False memory and delusions in Alzheimer’s disease

Thesis submitted for the degree of Doctor of Philosophy (PhD)

Dr Emma McLachlan

Primary Supervisor: Professor Robert Howard
Subsidiary Supervisors: Professor Neil Burgess and Professor Suzanne Reeves

May 2022

Funding: NIHR UCLH BRC Mental Health theme

Division of Psychiatry, Faculty of Brain Sciences, University College London
6th floor, Maple House, 149 Tottenham Court Rd, London W1T 7NF
Declaration

I, Emma M'Lachlan, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Emma M'Lachlan
(Signed electronically)
Abstract

Aims

This thesis aimed to investigate the relationship between memory errors and delusions in Alzheimer’s disease (AD), in order to further elucidate the mechanisms underlying delusion formation. This was achieved by undertaking narrative and systematic review of relevant literature, by exploring the relationship between performance on memory and metamemory tasks and delusions in AD patient populations and by investigating the neuroanatomical correlates of memory errors and delusions in AD patient populations.

Methods

I recruited 27 participants with and without delusions in AD and compared performance on measures of context memory, false memory and metamemory. I explored statistically significant behavioural findings further in the Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohort of participants with AD (n = 733). I then conducted hypothesis-driven region of interest and exploratory voxel-based morphometric analyses to determine the relationship between false memory and delusions and regional brain volume in the ADNI cohort. This informed similar analyses of neuroimaging data in my own participants (n = 8).

Results

In both samples, individuals with delusions in AD had higher false recognition rates on recognition memory tasks than those without delusions. False recognition was inversely correlated with volume of medial temporal lobe, ventral visual stream and prefrontal cortex in both samples. In the ADNI sample, false recognition was also inversely correlated with anterior cingulate cortex (ACC) volume bilaterally. Participants with delusions had reduced volume of right ACC and increased volume of right parahippocampal gyrus compared to the control group.

Conclusions

These two complementary studies provide evidence of specific memory impairments associated with both delusions and a distinct pattern of brain atrophy in AD. Simple cognitive interventions can reduce false recognition rates in AD. Given the significant risks associated with antipsychotic drug treatment of delusions, exploring how these non-pharmacological interventions potentially affect psychosis symptoms in AD is an important next step.
Impact Statement

The finding of this thesis of increased false recognition (incorrectly recognising a new item as previously seen) in those with delusions in Alzheimer’s disease (AD) compared to those without was consistent across two relatively simple and quick to administer recognition memory tests. This has potential cross-diagnostic clinical relevance and could have significant impact in terms of providing directions for future delusions research.

Firstly, these tests may have potential for use in early identification of those at risk of psychosis, both in AD and other psychiatric disorders. Early pharmacological intervention improves outcomes in schizophrenia, and while the impact of early intervention in AD is less clear, prompt assessment and intervention via non-pharmacological measures has the potential to limit both patient and carer distress.

Secondly, various interventions have demonstrated promising results in terms of reducing false memories, including in patients with hypoxic brain injury, psychosis in depression and in healthy older people. Given the high risk associated with pharmacological management of delusions in AD, with use of antipsychotics in this patient population leading to greater morbidity and mortality, I suggest future prospective studies should explore how relatively low risk cognitive interventions for false memory affect psychosis symptoms in AD.

In addition to this potential for clinical impact, my findings, of an inverse correlation between false recognition rate and volume of medial temporal lobe, ventral visual stream and prefrontal cortex, have potential for academic impact. These findings provide further evidence to support the role of these structures and their associated functional networks in false memory formation and will direct future research in this area.
Statement of Contributions

I applied for and was awarded a research post under the National Institute for Health Research University College London Hospitals Biomedical Research Centre Mental Health theme in order to complete the work reported in this thesis.

I designed both systematic reviews (described in Chapters 3 and 4). Ahmed Al-Shihabi, Justin Chan (JC), Professor Robert Howard, Dr Jonathan Huntley and Professor Suzanne Reeves (SRe) gave feedback during the design process. I completed literature searches for both reviews. I received support with data extraction and quality rating from Selina Rai (SRa) and JC. I conducted the narrative synthesis of results. Earlier iterations of the systematic reviews were submitted by Ahmed Al-Shihabi as a MBBS project (Chapter 3) and JC as an MSc project, for which I was secondary supervisor (Chapter 4). The review described in Chapter 3 was subsequently published (see M’Lachlan et al. (2020)). Specific contributions from the above researchers are described within the relevant chapters.

I designed my own patient-based study (described in Chapters 5 – 7, Chapter 12 and Chapter 13), and was responsible for participant recruitment (with support in local memory services from Sarah Boughetane, Dr Waleed Fawzi, Dr Lauren Huzzey, Dr Kathy Liu and SRe and for Join Dementia Research from Dr Dorothea Hammerer). I completed participant screening and behavioural testing. I accompanied participants to MRI scanning appointments at the Wellcome Centre for Human Neuroimaging, where scans were completed by the radiographers at the centre. The scanning protocol was programmed by Professor of MRI Physics, Martina Callaghan. Support for coding of computerised tasks is acknowledged within the relevant chapters. I conducted the data analysis.

I designed the ADNI cohort study (described in Chapters 8 – 11). I completed data extraction and cleaning. Support for coding of scripts for image preprocessing is acknowledged within the relevant chapters. I also received support in image preprocessing and analysis from Dilek Ocal at the Dementia Research Centre. I conducted the data analysis. Data used were obtained from the ADNI database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and provided data but did not participate in analysis or writing of this thesis. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Feedback on the whole thesis was provided by my supervisors Professor Robert Howard, Professor Neil Burgess (NB) and SRe.

I wrote the following thesis.
Acknowledgements

First and foremost, I would like to acknowledge those who participated in my patient-based study and their families, all of whom were extremely generous with their time. I am grateful to all of these individuals for their contributions, without which this thesis would not have been possible, and for so kindly welcoming me into their homes.

I would like to thank my supervisors Professors Robert Howard, Suzanne Reeves and Neil Burgess for their exceptional and thoughtful guidance and support, which has been unwavering throughout this PhD and is very much appreciated. Particular thanks to Rob and Suzanne, who have been providing me with encouragement and inspiration for the past nine years.

Thank you to Sarah Boughetane, Dr Waleed Fawzi, Dr Dorothea Hammerer, Dr Lauren Huzzey and Dr Kathy Liu for their support with recruitment. I am also grateful for the support of many of my academic colleagues at various stages of this PhD. At UCL: Dr James Bisby and Professor John King for coding computerised versions of memory tasks, Dr Daniel Bush for sharing his version of the associative inference task, Ahmed Al-Shihabi, Justin Chan, Dr Jonathan Huntley and Selina Rai for their help with the systematic reviews, Professor Martina Callaghan for support with the MRI protocol, FIL radiographers Megan Creasey and Curt Strangward, Amy Kirkwood and Dr Graham Wheeler for their statistics advice, Dr Ian Malone at the DRC for support with image preprocessing, and especially Dilek Ocal for her patience and support throughout the image analysis. At the University of Oxford: Professor Masud Husain for use of the object-location binding task and Maria Raquel Maio for her support with task interpretation. Data used in the preparation of this thesis were obtained from the ADNI database, led by Professor Michael W. Weiner. Thanks to all ADNI personnel, research institutions and participants. Thank you to Professor Tim Dornan for his early support of my academic career. With appreciation and thanks to the NIHR UCLH BRC for funding this PhD.

A final thank you to all of my family, McLachlans and Clarks, to all of my friends, in particular to those I haven’t been able to see as often as usual over the past six months, and to NJC: thank you for everything else.
# Table of Contents

Declaration................................................................................................................................. 2

Abstract........................................................................................................................................ 3

Aims ................................................................................................................................................ 3

Methods......................................................................................................................................... 3

Results........................................................................................................................................... 3

Conclusions ................................................................................................................................... 3

Impact Statement .......................................................................................................................... 4

Statement of Contributions ........................................................................................................... 5

Acknowledgements....................................................................................................................... 6

Table of Contents........................................................................................................................... 7

Table of Figures .............................................................................................................................. 16

Table of Tables ............................................................................................................................... 19

Abbreviations................................................................................................................................. 21

1. Introduction ............................................................................................................................... 25

   1.1 Context memory and delusions in Alzheimer’s disease ......................................................... 25

   1.2 Overview of thesis .................................................................................................................. 25

2. Background .................................................................................................................................. 26

   2.1 Dementia and Alzheimer’s disease ....................................................................................... 26

   2.2 Delusions in Alzheimer’s disease .......................................................................................... 28

       2.2.1 Epidemiology and phenomenology ................................................................................. 28

       2.2.2 Cognitive models of delusion formation ......................................................................... 29

       2.2.3 Neural networks implicated in delusions in Alzheimer’s disease ................................. 32

       2.2.4 Implications for management of psychosis in Alzheimer’s disease ......................... 33

   2.3 The relationship between false memory, confabulation and delusion ............................ 34

       2.3.1 Background .................................................................................................................... 34

       2.3.2 False memory and Alzheimer’s disease ......................................................................... 34

       2.3.3 False memory and delusions ......................................................................................... 36

   2.4 Context memory .................................................................................................................... 37
### 4. Neuropsychological Correlates of Spatial Context Memory in Alzheimer’s Disease: A Systematic Review

4.1 Abstract .......................................................................................................................... 56
4.2 Introduction ..................................................................................................................... 56
4.3 Methods ................................................................................................................................ 57
  4.3.1 Protocol and registration .............................................................................................. 57
  4.3.2 Literature search ......................................................................................................... 57
  4.3.3 Inclusion/exclusion criteria and screening ................................................................. 58
  4.3.4 Data extraction ............................................................................................................ 58
  4.3.5 Risk of bias/quality assessment ................................................................................... 58
4.4 Results ..................................................................................................................................... 59
  4.4.1 Study selection ............................................................................................................. 59
  4.4.2 Study characteristics ................................................................................................... 64
  4.4.3 Assessment of memory for spatial context ................................................................. 64
  4.4.4 Neuropsychological correlates .................................................................................... 65
  4.4.5 Central findings ........................................................................................................... 66
4.5 Discussion ................................................................................................................................ 74
  4.5.1 Methodological limitations of selected studies ......................................................... 74
  4.5.2 Limitations .................................................................................................................. 75
  4.5.3 Recommendations for future study .......................................................................... 75
  4.5.4 Conclusions ................................................................................................................ 76

### 5. Context Memory and Metamemory Study: Methods for Analyses of Behavioural Data

5.1 Study overview .................................................................................................................. 77
5.2 Research questions ............................................................................................................. 77
5.3 Recruitment ....................................................................................................................... 78
  5.3.1 Inclusion and exclusion criteria ................................................................................. 79
  5.3.2 Informed consent ....................................................................................................... 80
5.4 Study procedures ............................................................................................................... 80
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.4.1</td>
<td>Baseline assessments</td>
<td>80</td>
</tr>
<tr>
<td>5.4.2</td>
<td>Neuropsychological assessment</td>
<td>82</td>
</tr>
<tr>
<td>5.4.3</td>
<td>Measures of interest</td>
<td>85</td>
</tr>
<tr>
<td>5.4.4</td>
<td>Task adaptation</td>
<td>92</td>
</tr>
<tr>
<td>5.5</td>
<td>Statistical methodology</td>
<td>93</td>
</tr>
<tr>
<td>5.5.1</td>
<td>Group size</td>
<td>93</td>
</tr>
<tr>
<td>5.5.2</td>
<td>Internal pilot</td>
<td>93</td>
</tr>
<tr>
<td>5.5.3</td>
<td>Whole group behavioural analyses</td>
<td>94</td>
</tr>
<tr>
<td>5.6</td>
<td>Ethical approval</td>
<td>95</td>
</tr>
<tr>
<td>6.</td>
<td>Context Memory and Metamemory Study: Results of Internal Pilot</td>
<td>96</td>
</tr>
<tr>
<td>6.1</td>
<td>Recruitment</td>
<td>96</td>
</tr>
<tr>
<td>6.2</td>
<td>Review and adaptation of neuropsychological tasks</td>
<td>96</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Demographic and behavioural data</td>
<td>96</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Associative inference and relational memory</td>
<td>97</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Participant feedback</td>
<td>97</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Task adaptation</td>
<td>98</td>
</tr>
<tr>
<td>6.3</td>
<td>Preliminary behavioural results</td>
<td>99</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Demographic data</td>
<td>99</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Memory for spatial context – object-location binding</td>
<td>100</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Memory for temporal context – temporal context confusion</td>
<td>101</td>
</tr>
<tr>
<td>6.3.4</td>
<td>False memory and metamemory</td>
<td>102</td>
</tr>
<tr>
<td>6.4</td>
<td>Impact of pilot on plans for full study</td>
<td>105</td>
</tr>
<tr>
<td>6.4.1</td>
<td>Summary of behavioural data</td>
<td>105</td>
</tr>
<tr>
<td>6.4.2</td>
<td>Confirmed neuroimaging methodology</td>
<td>105</td>
</tr>
<tr>
<td>6.4.3</td>
<td>Sample size calculation</td>
<td>105</td>
</tr>
<tr>
<td>7.</td>
<td>Context Memory and Metamemory Study: Results of Analyses of Behavioural Data</td>
<td>107</td>
</tr>
<tr>
<td>7.1</td>
<td>Recruitment</td>
<td>107</td>
</tr>
<tr>
<td>7.1.1</td>
<td>Impact of COVID-19</td>
<td>107</td>
</tr>
<tr>
<td>7.2</td>
<td>Demographic details</td>
<td>108</td>
</tr>
</tbody>
</table>
7.3 Neuropsychological assessments ................................................................. 111

7.3.1 Global cognitive function ........................................................................... 111
7.3.2 Executive function ...................................................................................... 111
7.3.3 Attention .................................................................................................... 112
7.3.4 Perception and visuospatial function ........................................................ 112

7.4 Measures of interest ...................................................................................... 113

7.4.1 Memory for spatial context – object-location binding ............................... 113
7.4.2 Memory for temporal context – temporal context confusion .................... 113
7.4.3 False memory and metamemory .................................................................. 117

8. False Memories and Delusions in the ADNI Cohort: Methods for Analyses of Behavioural Data ........................................................................................................ 120

8.1 Background .................................................................................................... 120

8.1.1 Overview of ADNI ..................................................................................... 120

8.2 Research questions ....................................................................................... 123

8.3 Participant selection ..................................................................................... 123

8.4 Data selection ................................................................................................ 124

8.4.1 Demographic details ................................................................................. 124
8.4.2 Neuropsychological covariates/confounders .............................................. 124
8.4.3 False memory measures ............................................................................ 124

8.5 Statistical methodology ................................................................................ 125

8.5.1 Group size .................................................................................................. 125
8.5.2 Behavioural analyses ............................................................................... 126

9. False Memories and Delusions in the ADNI Cohort: Results of Analyses of Behavioural Data ........................................................................................................ 127

9.1 Participant selection ..................................................................................... 127

9.2 Demographic details ..................................................................................... 127

9.3 Neuropsychological profile .......................................................................... 133

9.3.1 Global cognitive function .......................................................................... 133
9.3.2 Executive function ..................................................................................... 133
14.2 False memories and delusional beliefs ............................................................... 240

14.3 Role of region-specific atrophy in false memory and delusion ......................... 243
  14.3.1 Medial temporal areas ............................................................................. 243
  14.3.2 Ventral visual stream ............................................................................ 244
  14.3.3 Prefrontal cortex .................................................................................. 245
  14.3.4 Anterior cingulate cortex ..................................................................... 247

14.4 Role of regions of greater volume in false memory and delusion .................... 248

14.5 Context memory, false memory and delusion .............................................. 249

14.6 Limitations ................................................................................................. 250

14.7 Clinical relevance and future directions ...................................................... 253

14.8 Conclusion .................................................................................................. 255

References ......................................................................................................... 256

Appendix 1. Modified EPHPP Tool for Systematic Review 1: Neuroimaging Correlates of False Memory in Alzheimer’s Disease ................................................................. 310

Appendix 2. Modified EPHPP Tool for Systematic Review 2: Neuropsychological Correlates of Spatial Context Memory in Alzheimer’s Disease ............................................ 312

Appendix 3. Context Memory and Metamemory Study Participant Information Sheet .... 314

Appendix 4. Wellcome Trust Centre for Neuroimaging MRI Exclusion Criteria ....................... 321

Appendix 5. Context Memory and Metamemory Study Consent Form ....................... 323

Appendix 6. Context Memory and Metamemory Study Case Report Form ....................... 325

Appendix 7. Temporal Context Confusion Task MATLAB Script ............................ 336

Appendix 8. Associative Inference Task MATLAB Script ...................................... 340

Appendix 9. DRM/Metamemory Task MATLAB Script ........................................ 348

Appendix 10. Analysis of the Association Between ‘State’ Delusion and False Recognition on ADAS-Cog-13 and RAVLT, Excluding Outliers Identified on Box Plots .......................... 354

Appendix 11. Edited CAT12 cat_defaults.m File ................................................... 355

Appendix 12. Shell Script to Run Image Preprocessing ......................................... 364

Appendix 13. ADNI False Memory Measures Region of Interest Analysis: Model Diagnostics (Pearson Residuals Plots) ............................................................... 365
Appendix 14. ADNI False Memory Measures Region of Interest Analysis: Model Diagnostics (Cook’s Distance Bar Charts) ........................................................................................................ 377

Appendix 15. Poisson Regression Models for False Recognition on RAVLT and ADAS-Cog 13 in the ADNI Cohort, With ROIs as Predictors: Excluding Outliers, Defined as Cook’s Distance > 4/(n – k – 1) ........................................................................................................ 389

Appendix 16. Context Memory and Metamemory Study, False Memory Measures Region of Interest Analysis: Model Diagnostics (Pearson Residuals Plots) ........................................... 392

Appendix 17. Context Memory and Metamemory Study False Memory Measures Region of Interest Analysis: Model Diagnostics (Cook’s Distance Bar Charts) ........................... 397

Appendix 18. Poisson Regression Models for False Recognition on TCC and DRM Tasks in my Patient-based Study, With ROIs as Predictors: Excluding Outliers, Defined as Cook’s Distance > 4/(n – k – 1) ........................................................................................................ 402
# Table of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Simple cognitive model of ordinary belief formation</td>
<td>29</td>
</tr>
<tr>
<td>2.2</td>
<td>Three-factor model of ordinary belief formation</td>
<td>29</td>
</tr>
<tr>
<td>2.3</td>
<td>Five-stage framework of ordinary belief formation</td>
<td>30</td>
</tr>
<tr>
<td>3.1</td>
<td>PRISMA flow diagram</td>
<td>44</td>
</tr>
<tr>
<td>4.1</td>
<td>PRISMA flow diagram</td>
<td>59</td>
</tr>
<tr>
<td>5.1</td>
<td>Grammatical reasoning task</td>
<td>83</td>
</tr>
<tr>
<td>5.2</td>
<td>Sustained attention to response task</td>
<td>84</td>
</tr>
<tr>
<td>5.3</td>
<td>Object-location binding task</td>
<td>86</td>
</tr>
<tr>
<td>5.4</td>
<td>Temporal context confusion task</td>
<td>88</td>
</tr>
<tr>
<td>5.5</td>
<td>Associative inference task</td>
<td>90</td>
</tr>
<tr>
<td>5.6</td>
<td>DRM/metamemory task</td>
<td>92</td>
</tr>
<tr>
<td>6.1</td>
<td>Scatter plots showing performance on DRM and metamemory tasks between groups</td>
<td>103</td>
</tr>
<tr>
<td>7.1</td>
<td>Participant selection flow chart</td>
<td>108</td>
</tr>
<tr>
<td>7.2</td>
<td>False recognition on temporal context confusion task</td>
<td>116</td>
</tr>
<tr>
<td>7.3</td>
<td>Proportion of high confidence responses by response type, DRM task</td>
<td>118</td>
</tr>
<tr>
<td>9.1</td>
<td>Participant selection flow chart</td>
<td>127</td>
</tr>
<tr>
<td>9.2</td>
<td>False recognition on ADAS-Cog 13 and RAVLT between ‘state’ delusion, ‘trait’ delusion and control groups</td>
<td>137</td>
</tr>
<tr>
<td>10.1</td>
<td>Design matrix for multiple regression model including intrusions on RAVLT as a covariate</td>
<td>150</td>
</tr>
<tr>
<td>10.2</td>
<td>Design matrix for multiple regression model including false recognition on RAVLT as a covariate</td>
<td>151</td>
</tr>
<tr>
<td>10.3</td>
<td>Design matrix for multiple regression model including false recognition on ADAS-Cog 13 as a covariate</td>
<td>152</td>
</tr>
<tr>
<td>10.4</td>
<td>Design matrix for ‘state’ delusion &lt; control t-test, with covariates</td>
<td>156</td>
</tr>
<tr>
<td>10.5</td>
<td>Design matrix for ‘trait’ delusion &lt; control t-test, with covariates</td>
<td>157</td>
</tr>
<tr>
<td>11.1</td>
<td>Participant selection flow chart</td>
<td>158</td>
</tr>
<tr>
<td>11.2</td>
<td>Data homogeneity - mean correlation of imaging data</td>
<td>165</td>
</tr>
<tr>
<td>11.3</td>
<td>Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of medial temporal lobe regions of interest</td>
<td>168</td>
</tr>
<tr>
<td>11.4</td>
<td>Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of ventral visual stream regions of interest</td>
<td>170</td>
</tr>
</tbody>
</table>
Figure 11.5 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of prefrontal cortex and anterior cingulate cortex regions of interest................................................................. 172
Figure 11.6 Statistical parametric map of negative correlation of grey matter volume and false recognition on ADAS-Cog 13 (t-contrast), displayed on whole glass brain view.... 179
Figure 11.7 Locations of the four significant clusters with negative effect of ADAS-Cog 13 false recognition ........................................................................................................ 180
Figure 11.8 Statistical parametric map of negative correlation of grey matter volume and intrusions on RAVLT (t-contrast), displayed on whole glass brain view ............... 181
Figure 11.9 Location of the right middle frontal gyrus cluster with negative effect of intrusions on RAVLT ........................................................................................................... 182
Figure 11.10 Unthresholded t-contrast effect size map for association between intrusions on RAVLT and grey matter volume ................................................................. 183
Figure 11.11 Statistical parametric map of negative correlation of grey matter volume and false recognition on RAVLT (t-contrast), displayed on whole glass brain view........... 184
Figure 11.12 Location of the left superior frontal gyrus clusters with negative effect of false recognition on RAVLT .................................................................................................... 185
Figure 11.13 Unthresholded t-contrast effect size map for association between false recognition on RAVLT and grey matter volume ..................................................................... 186
Figure 11.14 Statistical parametric map of negative correlation of grey matter volume and false recognition on ADAS-Cog 13 (t-contrast), whole glass brain view ................ 188
Figure 11.15 Location of bilateral medial temporal lobe and left fusiform clusters with negative effect of false recognition on ADAS-Cog 13 ................................................................. 189
Figure 11.16 Unthresholded t-contrast effect size map for association between false recognition on ADAS-Cog 13 and grey matter volume ................................................................. 190
Figure 11.17 Intrusions on RAVLT: design orthogonality.............................................. 191
Figure 11.18 False recognition on RAVLT: design orthogonality .................................. 192
Figure 11.19 False recognition on ADAS-Cog 13: design orthogonality ....................... 193
Figure 11.20 Statistical parametric map of ‘state’ delusion group < control t-contrast, displayed on whole glass brain view .................................................................................. 205
Figure 11.21 Location of left middle frontal gyrus cluster with lower volume in those with delusions at baseline (‘state’) compared to controls ............................................... 206
Figure 11.22 Unthresholded t-contrast effect size map for association between ‘state’ delusions and grey matter volume .................................................................................. 207
Figure 11.23 Statistical parametric map of ‘trait’ delusion group < control t-contrast, displayed on whole glass brain view .................................................................................. 208
Figure 11.24 Location of right middle temporal gyrus cluster with lower volume in those with delusions at any time point (‘trait’) compared to controls................................. 209
Figure 11.25 Unthresholded t-contrast effect size map for association between ‘trait’ delusions and grey matter volume ........................................................................... 210
Figure 11.26 T test of delusions at baseline (‘state’) vs control: design orthogonality ........ 211
Figure 11.27 T test of delusions at any time point (‘trait’) vs control: design orthogonality ... 212
Figure 13.1 Participant selection flow chart ................................................................................................. 218
Figure 13.2 Data homogeneity - mean correlation of imaging data......................................................... 225
Figure 13.3 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of medial temporal lobe regions of interest................................. 227
Figure 13.4 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of ventral visual stream regions of interest................................. 228
Figure 13.5 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of prefrontal cortex and anterior cingulate cortex regions of interest................................................................. 230
Table of Tables

Table 2.1 Diagnostic criteria for Alzheimer’s disease ................................................................. 27
Table 2.2 NPI psychosis symptom classification by subtype .......................................................... 28
Table 3.1 Summary of study characteristics and main findings .................................................... 49
Table 4.1 Summary of study characteristics and methodology .................................................... 60
Table 4.2 Summary of main findings by type of spatial context memory task ................................ 68
Table 6.1 Preliminary object-location binding task results .............................................................. 100
Table 6.2 Preliminary temporal context confusion task results ....................................................... 101
Table 6.3 Preliminary DRM/metamemory task results ................................................................. 102
Table 7.1 Demographic details and screening results ..................................................................... 109
Table 7.2 Breakdown and classification of psychosis symptoms at baseline (n = 10) ...................... 110
Table 7.3 Addenbrooke’s Cognitive Examination-III subdomain scores ....................................... 111
Table 7.4 Visual Object and Space Perception battery subdomain scores ..................................... 112
Table 7.5 Object-location binding task results .............................................................................. 113
Table 7.6 Temporal context confusion task results ....................................................................... 114
Table 7.7 DRM/metamemory task results ...................................................................................... 117
Table 8.1 Overview of ADNI phases ............................................................................................. 120
Table 9.1 Demographic details and screening results ..................................................................... 129
Table 9.2 Breakdown and classification of psychosis symptoms at baseline (n = 32) .................... 132
Table 9.3 Performance on measures of executive function ............................................................ 134
Table 9.4 Performance on measures of false memory ................................................................... 136
Table 9.5 Analysis of the association between ‘state’ delusion and false recognition on ADAS-Cog-13 and RAVLT .................................................................................................................... 139
Table 10.1 Acquisition protocols across ADNI phases ................................................................. 142
Table 10.2 Participant scan details (n = 729) ................................................................................. 143
Table 11.1 Demographic details and screening results ................................................................. 161
Table 11.2 Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest ................................................................................................................. 173
Table 11.3 Pearson correlation matrix for continuous variables (n = 728) ....................................... 178
Table 11.4 Comparison of volume in identified regions of interest between those with delusions at baseline (‘state’) and control ................................................................. 195
Table 11.5 Comparison of volume in identified regions of interest between those with delusions at any time point (‘trait’) and control ........................................................................ 197
Table 11.6 Binary logistic regression model of volume of ventral visual stream regions of interest (as proportion of TIV) as predictors of delusions at baseline (n = 589) .................... 200
Table 11.7 Comparison of parahippocampal gyri volume between those with delusions and control.......................................................................................................................................................................................... 201
Table 11.8 Binary logistic regression model of volume of the right parahippocampal gyrus (as proportion of TIV) as a predictor of delusions at baseline (n = 589).......................... 202
Table 12.1 Acquisition protocol ................................................................................................................................. 214
Table 13.1 Demographic details and screening results ............................................................................................... 220
Table 13.2 Temporal context confusion task results ............................................................................................... 222
Table 13.3 DRM/metamemory task results ............................................................................................................... 223
Table 13.4 Relationship between false recognition on the temporal context confusion and DRM/metamemory tasks and volume of regions of interest............................................. 231
Table 13.5 Comparison of volume in identified regions of interest between those with delusions and control .................................................................................................................................................................................. 234
**Abbreviations**

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>18FDG</td>
<td>18F-fluoro-2-deoxyglucose</td>
</tr>
<tr>
<td>ACC</td>
<td>Anterior cingulate cortex</td>
</tr>
<tr>
<td>ACE-III</td>
<td>Addenbrooke’s Cognitive Examination, third version</td>
</tr>
<tr>
<td>ACE-R</td>
<td>Addenbrooke’s Cognitive Examination-Revised</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>ADAS-Cog 13</td>
<td>Alzheimer’s Disease Assessment Scale–Cognitive subscale</td>
</tr>
<tr>
<td>ADNI</td>
<td>Alzheimer’s Disease Neuroimaging Initiative</td>
</tr>
<tr>
<td>aMCI</td>
<td>Amnestic mild cognitive impairment</td>
</tr>
<tr>
<td>ANOVA</td>
<td>Analysis of variance</td>
</tr>
<tr>
<td>APOE</td>
<td>Apolipoprotein E gene</td>
</tr>
<tr>
<td>B-ADL</td>
<td>Bayer Activities of Daily Living scale</td>
</tr>
<tr>
<td>CANTAB</td>
<td>Cambridge Neuropsychological Test Automated Battery</td>
</tr>
<tr>
<td>CAT12</td>
<td>Computational Anatomy Toolbox</td>
</tr>
<tr>
<td>CDR</td>
<td>Clinical Dementia Rating</td>
</tr>
<tr>
<td>CERAD</td>
<td>Consortium to Establish a Registry for Alzheimer’s Disease neurological battery</td>
</tr>
<tr>
<td>CMRG1c</td>
<td>Cerebral metabolic rate for glucose</td>
</tr>
<tr>
<td>CN</td>
<td>Cognitively normal</td>
</tr>
<tr>
<td>DLB</td>
<td>Lewy Body Dementia</td>
</tr>
<tr>
<td>dIPFC</td>
<td>Dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>DRM</td>
<td>Deese-Roediger-McDermott paradigm</td>
</tr>
<tr>
<td>DTI</td>
<td>Diffusion tensor imaging</td>
</tr>
<tr>
<td>EPHPP</td>
<td>Effective Public Health Practice Project tool</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>FFG</td>
<td>Fusiform gyrus</td>
</tr>
<tr>
<td>fMRI</td>
<td>Functional magnetic resonance imaging</td>
</tr>
<tr>
<td>FWE</td>
<td>Family wise error</td>
</tr>
<tr>
<td>FWHM</td>
<td>Full width half maximum</td>
</tr>
<tr>
<td>GCUT</td>
<td>Graph-cut</td>
</tr>
<tr>
<td>GDS-15</td>
<td>Geriatric Depression Scale, 15 item</td>
</tr>
<tr>
<td>GLE</td>
<td>Gross localisation error</td>
</tr>
<tr>
<td>GM</td>
<td>Grey matter</td>
</tr>
<tr>
<td>GS</td>
<td>Geodesic Shooting</td>
</tr>
<tr>
<td>HMSE</td>
<td>Hindi Mental State Examination</td>
</tr>
<tr>
<td>IFAG</td>
<td>Inferior frontal angular gyrus</td>
</tr>
<tr>
<td>IFG</td>
<td>Inferior frontal gyrus</td>
</tr>
<tr>
<td>IFOG</td>
<td>Inferior frontal orbital gyrus</td>
</tr>
<tr>
<td>IMP</td>
<td>123I-N-isopropyl-p-iodoamphetamine</td>
</tr>
<tr>
<td>IQR</td>
<td>Image quality rating</td>
</tr>
<tr>
<td>ITG</td>
<td>Inferior temporal gyrus</td>
</tr>
<tr>
<td>JLO</td>
<td>Judgement of line orientation</td>
</tr>
<tr>
<td>LOG</td>
<td>Lateral orbital gyrus</td>
</tr>
<tr>
<td>LTL</td>
<td>Lateral temporal lobe</td>
</tr>
<tr>
<td>MCI</td>
<td>Mild cognitive impairment</td>
</tr>
<tr>
<td>MFC</td>
<td>Medial frontal cerebrum</td>
</tr>
<tr>
<td>MFG</td>
<td>Middle frontal gyrus</td>
</tr>
<tr>
<td>MMSE</td>
<td>Mini Mental State Examination</td>
</tr>
<tr>
<td>MNI</td>
<td>Montreal Neurological Institute</td>
</tr>
<tr>
<td>Acronym</td>
<td>Full Form</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>MoCA</td>
<td>Montreal Cognitive Assessment</td>
</tr>
<tr>
<td>mPFC</td>
<td>Medial prefrontal cortex</td>
</tr>
<tr>
<td>MP-RAGE</td>
<td>Magnetisation-prepared rapid gradient-echo</td>
</tr>
<tr>
<td>MRI</td>
<td>Magnetic resonance imaging</td>
</tr>
<tr>
<td>M-RMT</td>
<td>Money’s Road Map Test</td>
</tr>
<tr>
<td>MTG</td>
<td>Middle temporal gyrus</td>
</tr>
<tr>
<td>MTL</td>
<td>Medial temporal lobe</td>
</tr>
<tr>
<td>NINCDS-ADRDA</td>
<td>National Institute of Neurological and Communicative Disease and Stroke and the Alzheimer’s Disease and Related Disorders Association criteria</td>
</tr>
<tr>
<td>NPI</td>
<td>Neuropsychiatric Inventory</td>
</tr>
<tr>
<td>OFC</td>
<td>Orbitofrontal cortex</td>
</tr>
<tr>
<td>PD</td>
<td>Parkinson’s disease</td>
</tr>
<tr>
<td>PET</td>
<td>Positron emission tomography</td>
</tr>
<tr>
<td>PFC</td>
<td>Prefrontal cortex</td>
</tr>
<tr>
<td>PHG</td>
<td>Parahippocampal gyrus</td>
</tr>
<tr>
<td>PRISMA</td>
<td>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</td>
</tr>
<tr>
<td>RAVLT</td>
<td>Rey Auditory Verbal Learning Test</td>
</tr>
<tr>
<td>rCBF</td>
<td>Regional cerebral blood flow</td>
</tr>
<tr>
<td>REC</td>
<td>Research ethics committee</td>
</tr>
<tr>
<td>ROCF</td>
<td>Rey-Osterrieth Complex Figure</td>
</tr>
<tr>
<td>ROI</td>
<td>Region of interest</td>
</tr>
<tr>
<td>RVP</td>
<td>Rapid Visual Processing test</td>
</tr>
<tr>
<td>SART</td>
<td>Sustained Attention to Response Task</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SFG</td>
<td>Superior frontal gyrus</td>
</tr>
<tr>
<td>SMFG</td>
<td>Superior medial frontal gyrus</td>
</tr>
<tr>
<td>sMMSE</td>
<td>Standardised Mini Mental State Examination</td>
</tr>
<tr>
<td>SPECT</td>
<td>Single photon emission computed tomography</td>
</tr>
<tr>
<td>SPM</td>
<td>Statistical parametric map</td>
</tr>
<tr>
<td>SPM12</td>
<td>Statistical Parametric Mapping software</td>
</tr>
<tr>
<td>SPSS</td>
<td>Statistical Package for the Social Sciences</td>
</tr>
<tr>
<td>STG</td>
<td>Superior temporal gyrus</td>
</tr>
<tr>
<td>T</td>
<td>Tesla</td>
</tr>
<tr>
<td>TCC</td>
<td>Temporal context confusion</td>
</tr>
<tr>
<td>TIV</td>
<td>Total intracranial volume</td>
</tr>
<tr>
<td>TPMs</td>
<td>Tissue probability maps</td>
</tr>
<tr>
<td>UPDRS</td>
<td>Unified Parkinson’s Disease Rating Scale</td>
</tr>
<tr>
<td>VBM</td>
<td>Voxel-based morphometry</td>
</tr>
<tr>
<td>VIF</td>
<td>Variance inflation factor</td>
</tr>
<tr>
<td>vmPFC</td>
<td>Ventromedial prefrontal cortex</td>
</tr>
<tr>
<td>VOSP</td>
<td>Visual Object and Space Perception battery</td>
</tr>
<tr>
<td>WAIS-R</td>
<td>Wechsler Adult Intelligence Scale-Revised</td>
</tr>
<tr>
<td>WMS</td>
<td>Wechsler Memory Scale</td>
</tr>
<tr>
<td>WMS-R</td>
<td>Wechsler Memory Scale-Revised</td>
</tr>
</tbody>
</table>
1. Introduction

1.1 Context memory and delusions in Alzheimer’s disease

I first became interested in delusions in Alzheimer’s disease (AD) during clinical work as a Core Trainee in Psychiatry with the Croydon Community Mental Health Team for Older Adults. I was struck by the impact of delusional beliefs on patients and their families. During my Academic Clinical Fellowship I completed a retrospective analysis of the neuroanatomical correlates of delusions in the AddNeuroMed cohort. My finding, that individuals with delusions in AD had lower left parahippocampal gyrus (PHG) volume than those without, led to the design of this thesis. I hypothesised that this finding was related to the role of the PHG in processing memory for the spatial and temporal context of visual information, and that errors in context memory processing may be involved in delusion formation in AD. This PhD was initially conceptualised as a prospective exploration of the relationship between memory error and delusion formation in AD. When the first wave of the coronavirus pandemic (COVID-19) curtailed recruitment in March 2020, I extended my investigation to include longitudinal data from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) cohort.

1.2 Overview of thesis

Chapter 2 is a narrative exploration of the literature relevant to this PhD, broadly covering AD, delusions and memory errors of interest. Two areas are taken forward to systematic reviews in Chapters 3 and 4: neuroimaging correlates of false memory and neuropsychological correlates of spatial context memory in AD. In Chapter 5 I describe the methodology of my prospective, patient-based study of context memory and metamemory in individuals with and without delusions in AD. Pilot results from this study are described in Chapter 6. These results informed behavioural testing in the cohort and the neuroimaging methodology. Behavioural results are described in Chapter 7. In Chapter 8 I describe the methodology of my retrospective study of false memory and delusions in the ADNI cohort, with results from behavioural testing described in Chapter 9. Methodology and results of my investigation of neuroanatomical correlates of false memory and delusions in the ADNI cohort are described in Chapters 10 and 11, with results informing a similar neuroimaging analysis in the sample from my patient-based study, which is described in Chapters 12 and 13. In Chapter 14 I summarise findings, relate these to previous work in this area, discuss conclusions and recommend future directions.
2. Background

2.1 Dementia and Alzheimer’s disease

Dementia, an umbrella term which includes a range of neurodegenerative disorders characterised by acquired and progressive impairments of cognitive and functional ability, is one of the most significant challenges to global health and is now the seventh leading cause of death globally.\(^3\)\(^-\)\(^6\) The number of people living with dementia worldwide is estimated at over 55 million\(^6\); across the world 10 million new cases of dementia are diagnosed every year, equivalent to a new case every three seconds.\(^7\) As populations age worldwide, the number of people living with dementia will increase to a predicted 135 million by 2050.\(^8\) The global distribution of dementia cases is slowly changing, with the proportion of people living with dementia in low and middle income countries due to increase over the same time period.\(^8\) This has a correspondingly large economic impact; in 2016 the total worldwide cost of dementia was estimated at $818 billion (1.1% of global gross domestic product)\(^7\),\(^9\) and this had increased to $1.3 trillion by 2019.\(^6\)

AD accounts for the majority (between 50% and 75%)\(^10\) of these dementia diagnoses and costs the United Kingdom (UK) economy over £25 billion a year.\(^{11-14}\) AD typically begins with short term episodic memory impairments, for example, misplacing items or forgetting appointments. Cognitive deficits accumulate as the disease progresses to include increasingly significant memory deficits alongside neuropsychiatric pathology and impairments affecting diverse neurocognitive domains including language and visuospatial and executive function.\(^{15, 16}\) The diagnostic criteria for AD that are most commonly used in clinical settings (from the International Classification of Diseases, currently in its tenth edition with an eleventh in development\(^17\)) and research settings (2011 revision of the National Institute of Neurological and Communicative Disease and Stroke and the Alzheimer’s Disease and Related Disorders Association criteria; NINCDS-ADRDA\(^18\)) are summarised in Table 2.1. With no immediate prospect of disease modifying treatments and numbers set to double in the next 30 years, symptomatic treatment is a priority area for research both in the UK and globally.\(^7, 14\)
### Table 2.1 Diagnostic criteria for Alzheimer’s disease

#### SUMMARY OF ICD-10 CRITERIA FOR F00 DEMENTIA IN ALZHEIMER’S DISEASE

<table>
<thead>
<tr>
<th>CRITERIA FOR ALL-CAUSE DEMENTIA</th>
<th>ADDITIONAL CRITERIA FOR AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Decline in both memory and thinking (including reasoning capacity and flow of ideas) sufficient to impair activities of daily living</td>
<td>• Insidious onset with slow deterioration</td>
</tr>
<tr>
<td>• Clear consciousness (in the absence of superimposed delirium)</td>
<td>• Other systemic/neurological diseases which can induce a dementia ruled out by investigations or absence of clinical evidence</td>
</tr>
<tr>
<td>• Symptoms for at least six months for confident clinical diagnosis</td>
<td>• No sudden, apoplectic onset, or neurological signs early in the illness</td>
</tr>
</tbody>
</table>

#### SUMMARY OF NINCDS-ADRDA CRITERIA FOR ALZHEIMER’S DISEASE

<table>
<thead>
<tr>
<th>CRITERIA FOR ALL-CAUSE DEMENTIA</th>
<th>ADDITIONAL CRITERIA FOR ‘PROBABLE’ AD</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Cognitive or behavioural symptoms that interfere with functional ability, represent a decline from previous levels of functioning and are not explained by delirium or major psychiatric disorder</td>
<td>• Insidious onset over months to years</td>
</tr>
<tr>
<td>• Cognitive impairment detected by history-taking from patient and informant and an objective cognitive assessment</td>
<td>• Clear-cut history of worsening cognition by report or observation</td>
</tr>
<tr>
<td>Impairment involves a minimum of two of:</td>
<td>• No evidence of substantial comorbid cerebrovascular disease or core features of Lewy Body dementia, frontotemporal dementia or primary progressive aphasia</td>
</tr>
<tr>
<td>• Impaired ability to acquire and remember new information</td>
<td>• No evidence for active neurological disease or alternative organic cause of cognitive symptoms</td>
</tr>
<tr>
<td>• Poor judgement and impaired reasoning and handling of complex tasks</td>
<td>Amnestic presentation:</td>
</tr>
<tr>
<td>• Impaired visuospatial abilities</td>
<td>• Impairment in learning and recall of recently learned information, plus impairment in at least one other cognitive domain</td>
</tr>
<tr>
<td>• Impaired language functions</td>
<td>Non-amnestic presentations (deficits in other cognitive domains should be present for all):</td>
</tr>
<tr>
<td>• Changes in personality or behaviour</td>
<td>• Language: word finding difficulties</td>
</tr>
<tr>
<td></td>
<td>• Visuospatial: deficits are in spatial cognition</td>
</tr>
<tr>
<td></td>
<td>• Executive dysfunction: impaired reasoning, judgment, and problem solving</td>
</tr>
</tbody>
</table>

Notes:
AD = Alzheimer’s disease; ICD-10 = International Classification of Diseases and Related Health Problems, 10th Revision; NINCDS-ADRDA = National Institute of Neurological and Communicative Disease and Stroke and the Alzheimer’s Disease and Related Disorders Association criteria.
2.2 Delusions in Alzheimer’s disease

2.2.1 Epidemiology and phenomenology

Psychosis symptoms (delusions and hallucinations) affect between 30% and 50% of people with AD, with an incidence rate of 10% per year. Delusions are more common than hallucinations and occur earlier in the disease course. Psychosis symptoms in AD are associated with increased agitation and aggression, more rapid cognitive and functional decline, earlier institutionalisation and substantial morbidity for patients and carers.

Increasing evidence suggests that psychosis symptoms in AD can be divided into two distinct subtypes (see Table 2.2): ‘paranoid’, which includes delusions of persecution and/or abandonment and ‘misidentification’, characterised by the presence of misidentification phenomena and/or hallucinations. The misidentification but not the paranoid subtype has been associated with increased hippocampal, parahippocampal and entorhinal pathology, greater global cognitive deficits and faster cognitive decline.

Table 2.2 NPI psychosis symptom classification by subtype

<table>
<thead>
<tr>
<th>CONTENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARANOID</td>
</tr>
<tr>
<td>A. Delusions</td>
</tr>
<tr>
<td>1. In danger/others are planning to hurt him/her</td>
</tr>
<tr>
<td>2. Others are stealing from him/her</td>
</tr>
<tr>
<td>3. Spouse is having an affair</td>
</tr>
<tr>
<td>7. Family members plan to abandon him/her</td>
</tr>
<tr>
<td>4. Unwelcome guests are staying in his/her house</td>
</tr>
<tr>
<td>5. His/her spouse or others are not who they claim to be</td>
</tr>
<tr>
<td>6. His/her house is not his/her own</td>
</tr>
<tr>
<td>8. Television/magazine figures are present in his/her home</td>
</tr>
<tr>
<td>MISIDENTIFICATION</td>
</tr>
<tr>
<td>A. Delusions</td>
</tr>
<tr>
<td>7. Family members plan to abandon him/her</td>
</tr>
<tr>
<td>4. Unwelcome guests are staying in his/her house</td>
</tr>
<tr>
<td>5. His/her spouse or others are not who they claim to be</td>
</tr>
<tr>
<td>6. His/her house is not his/her own</td>
</tr>
<tr>
<td>8. Television/magazine figures are present in his/her home</td>
</tr>
<tr>
<td>B. Hallucinations</td>
</tr>
<tr>
<td>1. He/she can hear voices</td>
</tr>
<tr>
<td>2. Talks to people who are not there</td>
</tr>
<tr>
<td>3. Seeing things not seen by others</td>
</tr>
</tbody>
</table>

Notes:
Letters and numbers refer to sections of the NPI.
NPI = Neuropsychiatric Inventory.
2.2.2  **Cognitive models of delusion formation**

Proposed mechanisms for how delusions form have largely been developed through clinical observation and research involving non-Alzheimer’s populations. Delusions can occur in a range of non-Alzheimer’s psychiatric disorders (for example, but not limited to, schizophrenia and bipolar affective disorder) and neurological disorders (again, for example, but not limited to, Parkinson’s disease (PD), brain injury, encephalopathy and epilepsy).

Delusions are generally accepted to be fixed false beliefs. Cognitive models of delusion formation are therefore often constructed in reference to models of ordinary belief formation. The majority of these are fundamentally conceptually similar, involving feedback loops of bottom-up sensory input, processing of which leads to belief formation or alteration, and top-down effects of existing beliefs on the monitoring of sensory input. Models include the simple heuristic model of ordinary belief formation proposed by Bentall (1990), see Figure 2.1, the three-factor model proposed by Langdon and Coltheart (2000), summarised in Figure 2.2, and the five-stage theory developed by Connors and Halligan (2017), see Figure 2.3.

**Figure 2.1 Simple cognitive model of ordinary belief formation**

```
New data → Attention and perception → Inference → Belief formation

Search for more data
```

Notes:
Heuristic model of ordinary belief acquisition, from Bentall (1990) and Bentall et al. (2001).

**Figure 2.2 Three-factor model of ordinary belief formation**

```
Sensory input → Input judged to require explanation → Hypotheses generated → Hypothesis adopted as belief

Sensory input monitoring processes  Rational belief evaluation processes
```

Notes:
Three-factor model of ordinary belief formation, summarised from Langdon and Coltheart (2000).
In their three-factor model of belief formation, Langdon and Coltheart (2000) described a model involving monitoring and evaluation processes. Their proposed sensory monitoring processes identify sensory information requiring explanation due to either heightened personal salience or discordance with previous experience, through conscious and preconscious monitoring. Rational belief evaluation processes involve assessing and prioritising hypotheses developed to explain sensory information, based on previous experience and other salient information (for example, opinions of others). If the hypothesis chosen after belief evaluation is adopted as a belief, this may influence expectations regarding sensory information in the future. The more recent model put forward by Connors and Halligan (2017) recognises similar stages to the three-factor model, but with more weight given to the effect of beliefs on all stages of belief generation, see Figure 2.3, described as a ‘tentative five-stage framework’.

Figure 2.3 Five-stage framework of ordinary belief formation

![Five-stage framework of ordinary belief formation](image)

Notes: Five-stage framework of ordinary belief formation, from Connors and Halligan (2017) and Connors and Halligan (2020).

Early cognitive models for delusion formation developed primarily by Brendan Maher focused on pathology affecting the first stage of belief formation (i.e. a single-factor model of delusion formation), proposing that delusional beliefs develop as rational attempts to explain primary perceptual abnormalities, for example unusually intense or vivid sensory experiences.\(^40\)\(^-\)\(^42\) This theory has been widely contradicted by findings that delusional beliefs can occur in the absence of abnormal perceptual experiences.\(^43\)\(^-\)\(^45\) However, over time this model was extended to acknowledge that various neuropsychological abnormalities could cause experiences for which delusions may form as rational explanations.\(^46\) Maher (1999) suggested that these neuropsychological abnormalities may include incorrect attribution of feelings of significance to sensory information, drug intoxication, abnormalities in attention leading to difficulty discriminating between relevant and irrelevant stimuli and impairments of inferential reasoning and metacognition. The two-factor theory of delusional belief expanded on this single-factor model by proposing that there are two key components of delusion formation: a
neuropsychological impairment that prompts the false belief and a disruption in belief evaluation processes such that the false belief is not rejected.\textsuperscript{47}

There have been recent critiques of the two-factor theory (see Corlett (2019) and Corlett and Fletcher (2021)), which note the overly reductive nature of models of both belief and delusion formation and describe them as ‘convenient but simplified’ approaches to delineating a likely far more complex relationship between perception and belief formation. Corlett and colleagues instead propose a computational model of delusion formation based on Bayesian predictive-processing theory.\textsuperscript{50-52} According to predictive-processing theory, existing beliefs are constantly being used to make top-down predictions about the world around us.\textsuperscript{52} A prediction error occurs when these predictions do not align with sensory input. In ordinary belief formation these prediction errors are processed with a degree of balance such that beliefs are not either held too rigidly or changed too easily.\textsuperscript{52} The prediction error based theory of delusion formation proposes that delusions result from abnormalities in prediction error processing causing imprecise (unreliable) prediction errors to be incorrectly signalled as precise (reliable), and leading to inappropriate belief revision and delusion generation.\textsuperscript{53-57}

While often presented as such, computational and two-factor models are not necessarily mutually exclusive, if disrupted prediction error processing is conceptualised as a ‘first factor’ neuropsychological impairment leading to the generation of false beliefs.\textsuperscript{57} Other potential ‘first factor’ neuropsychological impairments identified in those with delusions include:

- Jumping to conclusions, as widely replicated using the beads in a jar task.\textsuperscript{58, 59} In this task, participants are shown two jars holding coloured beads in different proportions. With the jars removed from view, the individual draws a bead at a time from one of the jars and chooses when to predict which jar the beads are coming from. Individuals with delusions tend to jump to conclusions, making their prediction after fewer beads have been seen. This has been investigated in various populations, including: those with current delusions,\textsuperscript{60} those with a history of experiencing delusions,\textsuperscript{61} people with delusional disorder,\textsuperscript{45} those at high risk of developing delusions (unaffected first degree relatives of people with schizophrenia),\textsuperscript{62} and in individuals with schizophrenia regardless of current presence or severity of delusions.\textsuperscript{63-65}

- Selectively attending to threatening stimuli.\textsuperscript{38} Early studies found that individuals with delusions were slower at processing threat-related or delusion-congruent words on the Stroop test, which involves reading out the colour a word is printed in rather than the word itself.\textsuperscript{66, 67} Those with persecutory delusions also selectively recall threat-related words on tasks involving free recall of words lists.\textsuperscript{68}
• Theory of mind deficits. Theory of mind is the cognitive ability to represent and understand the mental states of others. Impairments have been demonstrated in individuals with delusions\textsuperscript{69-72} and appear to resolve after remission of psychosis.\textsuperscript{73, 74}

• Attribution biases, with those who hold persecutory delusional beliefs being more likely to excessively attribute negative events to non-self (‘external’ causes) and positive events to themselves (‘internal’ causes), and being more likely to blame others for negative events.\textsuperscript{38, 75-78} This finding is not replicated in non-delusional individuals with paranoid traits, including individuals with subclinical persecutory ideation\textsuperscript{79} or those with autism spectrum disorder.\textsuperscript{80}

• Errors in memory processes, associative inference and metamemory as described in sections 2.3 – 2.6 below.

2.2.3 Neural networks implicated in delusions in Alzheimer’s disease

Combined data from neurobiological and neuropsychological studies implicate various functional neural networks in delusion formation in AD.\textsuperscript{30, 81}

The dopaminergic system has been a key focus of psychosis research since the dopamine D2/D3 receptor antagonist chlorpromazine was first found to improve psychosis symptoms in schizophrenia over 60 years ago. Similar to young adults with schizophrenia, people with psychosis in AD have increased striatal D2/3 receptor availability and poorer performance on the Rapid Visual Processing (RVP) test of sustained attention, implicating corticostriatal dopaminergic networks\textsuperscript{52, 82-85} involved in salience attribution.\textsuperscript{86-89} Aberrant salience attribution is thought to be a key cause of transdiagnostic delusion formation, with neuroimaging studies in AD finding atrophy and hypometabolism in brain regions of the salience attribution network, including the insula and anterior cingulate cortex (ACC), to be correlated with delusional beliefs.\textsuperscript{90-92} Aberrant salience processing has been related to the predictive error model of delusion formation described in section 2.2.2, with aberrant salience attribution potentially occurring following mislabelling of prediction errors as highly precise.\textsuperscript{55, 93, 94}

Dysfunction of the ventral visual pathway (commonly thought of as the ‘what’ pathway of visual processing) has also been associated with psychosis in AD, with participants with psychosis in AD performing more poorly on both the RVP test and the incomplete letters test of the Visual Object and Space Perception battery (VOSP), and having increased brain atrophy in ventral visual stream brain regions.\textsuperscript{2, 84} In addition to its role in object recognition, the ventral visual stream has a role in assigning affect to visual stimuli, including during facial recognition, via connected ventral limbic structures; it is thought to be through these connections that ventral visual stream dysfunction may contribute to misidentification delusions.\textsuperscript{95} Ventral visual stream dysfunction
occurs in schizophrenia,\underline{96,97} misidentification delusions following stroke\underline{98} and psychosis in PD.\underline{99,100}

Individuals with misidentification-type delusions in AD also have increased atrophy in the hippocampus, possibly reflecting more severe impairments in global cognitive function in this group.\underline{30,101}

Lastly, frontal cortical networks are linked to delusional symptoms in AD.\underline{31,102-105} Neuroimaging studies have identified asymmetrical structural and functional changes in frontal regions in participants with delusions in AD, including atrophy in right frontal lobe more generally\underline{31,102,106} and in bilateral middle frontal gyri (MFG),\underline{107,108} and hypometabolism in right dorsolateral prefrontal cortex (dPFC), right inferior frontal gyrus (IFG) and in the orbitofrontal cortex (OFC) bilaterally.\underline{109,110} The role of frontal cortical dysfunction in delusion generation remains uncertain, with recent speculation that it may lead to loss of control functions involved in reality monitoring.\underline{108} These patterns of frontal atrophy are also present in other psychosis syndromes, including schizophrenia and bipolar affective disorder.\underline{111}

Psychosis in AD shows familial aggregation,\underline{112-114} is highly heritable\underline{115} and is associated with both novel common genetic variants and common genetic variants associated with schizophrenia.\underline{116}

Carriers of the ε4 allele of the apolipoprotein E (APOE) gene (in addition to being at increased risk of AD) appear to be at greater risk of psychosis symptoms in AD,\underline{117,118} and gene x gene interactions between APOE and the methylene tetrahydrofolate reductase gene have been linked to increased psychosis severity.\underline{117} A duplicated region on chromosome 16p11.2, also found in both schizophrenia and autism, has been found in those with psychosis in AD but not those without.\underline{119} Using a polygenic risk score approach, delusions in AD in particular are found to share genetic liability with schizophrenia.\underline{120}

Despite these accumulating advances in knowledge, the mechanisms underlying delusion formation in AD remain relatively poorly understood.

\underline{2.2.4 Imlications for management of psychosis in Alzheimer’s disease}

Due to limited understanding of the pathology underlying delusions in AD, current drug treatment is based on extrapolating from treatment of psychotic illnesses (for example, schizophrenia) in young adults, using atypical antipsychotics such as risperidone. Although there is evidence of efficacy of antipsychotic medication in the treatment of psychosis in AD,\underline{121} these medications have a considerable side effect burden (parkinsonism, postural hypotension, sedation, falls, increased stroke risk) and lead to increased mortality.\underline{122-124} Research that aims to improve our mechanistic understanding of delusion formation in AD will help to identify and target both pharmacological and non-pharmacological treatment options.\underline{125}
2.3 The relationship between false memory, confabulation and delusion

2.3.1 Background

The standard definition of delusions, widely used both through medical education and clinical practice, is that they are fixed, firmly held beliefs, out of keeping with an individual’s cultural norms. This aligns with the definition of delusions within the Diagnostic and Statistical Manual of Mental Disorders as ‘false beliefs that are not amenable to change in light of conflicting evidence’, with delusions deemed to be ‘bizarre’ when they are ‘clearly implausible, not understandable to same-culture peers and not derived from ordinary life experiences’. However, there are various counter examples that call this definition into question. These include work by Myin-Germeys et al. (2005) that demonstrates how conviction in delusions can fluctuate, that a significant number of people with delusions are willing to accept the possibility that their delusional belief may be false, and the post-internet phenomenon of extreme communities online meaning that the cultural group an individual is part of (albeit online) may be more likely to share an unusual or extreme belief.

Given this uncertainty, it is possible that mechanisms for delusion formation may to some degree overlap with memory-related false beliefs (false memories) which are more common in AD, with some suggesting that they exist on a spectrum. While they differ in terms of how false beliefs relate to time (i.e. delusions being based in present experience and false memories in past experience), both conditions involve an individual endorsing distorted representations of reality. False memory itself is a broad term encompassing a range of phenomena, including confabulation (giving false information without being aware of it), intrusion errors, misremembering of word lists, false recognition of novel stimuli and distortions in autobiographical memory. False memories are extremely common in AD, occurring in up to 90% of individuals.

2.3.2 False memory and Alzheimer’s disease

As described in section 2.3.1, people with AD are highly susceptible to generating false memories. These may be associated with high risk behaviours, such as missing medication doses due to falsely remembering taking it, or increased vulnerability to exploitation due to falsely remembering that strangers are close friends. False memories in AD also reduce functional ability and increase caregiver distress.

The Deese-Roediger-McDermott (DRM) paradigm has been widely used in attempts to further understand susceptibility to false memory formation. This relatively simple task requires participants to distinguish previously presented words (for example, medicine, sick, nurse) from
highly related but not presented words (for example, doctor) and unrelated and not presented words. When the DRM is used in healthy young adults, they consistently falsely recall the highly related words but not the unrelated words, indicating that this effect is more than simply a failure of episodic memory. Results reported using the DRM in AD have not been wholly consistent, but have demonstrated a greater susceptibility to false memory generation in individuals with AD due to impaired recollection (item-specific memory), proportionately less impairment and therefore overreliance on familiarity (gist memory) and impaired memory verification and monitoring. AD participants consistently demonstrate more provoked confabulation than controls, with results from one paradigm suggesting that confabulation arises when over-learned information replaces poorly encoded information, possibly due to impaired monitoring.

In addition to demonstrating a greater tendency to DRM-provoked false memory, people with AD are also consistently found to have impaired discrimination (inability to distinguish between target items and distractors on recognition memory tests) and a more liberal response bias: that is, they have higher rates of false recognition on recognition memory tasks than those without AD as they are more likely to endorse ‘new’ and unseen stimuli as ‘old’ and seen before. Discrimination has been suggested as a possible marker of prodromal AD as it is more significantly impaired in individuals with amnestic mild cognitive impairment (aMCI) with cerebrospinal fluid biomarkers of AD (β-amyloid and tau) than those without.

There has been a relative lack of research into the relationship between both neuropsychological factors and the anatomical location of the neuropathology of AD and false memories. Semantic confabulation in AD (i.e. the proportion of confabulated responses to semantic memory questions) increases with reduced cognitive function as measured by the Mini Mental State Examination (MMSE), and is negatively correlated with measures of attention, orientation and constructional praxis. In healthy younger adults, susceptibility to false memories induced by a misinformation paradigm correlates with overall intelligence and cognitive function, visual perception and facial recognition abilities. In older adults, impairment in source monitoring and increased susceptibility to false recognition correlates with neuropsychological tasks associated with both executive function and overall cognitive function.

In non-AD populations, lesion and neuroimaging studies show various brain regions to be involved in distinguishing true from false memory, including medial temporal lobe (MTL) and ventral visual stream structures, with a combination of frontal dysfunction and some memory preservation required for false memory generation. Of note, both the ACC and the superior parietal lobule have been associated with attributing feelings of confidence and
doubt to DRM responses, and dysfunction in these areas may also contribute. Preserved functioning of the medial prefrontal cortex (mPFC), which is involved in complex decision making, source attribution and veracity monitoring, is also thought to be essential to false memory creation. Lesion studies have implicated a large number of regions in confabulation, including the prefrontal cortex (PFC), OFC, ACC, basal forebrain, anterior insula, capsular genu, amygdala, perirhinal cortex and hypothalamus. One relatively consistent finding is that, as in false memory, some degree of preserved hippocampal function is also required to generate confabulations. Individuals with MTL damage experience autobiographical memory loss and reduced false memories with the DRM, while those with PFC damage experience autobiographical confabulation and increased false memories on DRM testing. Given the presence of AD pathology in the MTL, it is hypothesised that false memories may result from reduced function within MTL and ventral visual structures with some preservation of mPFC leading to increased reliance on familiarity, and additional PFC dysfunction leading to impairment of memory veracity monitoring.

See systematic review in Chapter 3 for further details.

2.3.3 False memory and delusions

Few studies have explored the relationship between false memories and delusions in AD, although there is preliminary but limited evidence that confabulation may correlate with delusions. In healthy individuals, false memories (as generated by the DRM paradigm) also correlate with sub-clinical delusional ideation. Individuals with schizophrenia are widely known to be more liable to false recognition on memory testing, however this area remains relatively unexplored, with existing evidence being mixed. Some studies using the DRM paradigm have not found any increase in false memories. Others have clearly demonstrated an increase in false memories, particularly in those with current symptoms of psychosis, with intrusion and false recognition positively correlated with positive symptoms and negatively correlated with negative symptoms. It is possible that this discrepancy is due to overall low recognition rates on the DRM in this patient group. When individuals with schizophrenia falsely recognise words on memory testing during functional magnetic resonance imaging (fMRI), activation is reduced in PFC, ACC and posterior cingulate cortex compared to controls.
2.4  Context memory

2.4.1  Previous research work

As described in section 2.2.3, people with psychosis symptoms in AD have reduced performance accuracy on the incomplete letters task of the VOSP compared to those without psychosis symptoms. This difference is largely explained by poorer performance in those with misidentification-type symptoms.⁸⁴ Given this association between visuoperceptual functional deficits and psychosis in AD, in my previous work I compared regional cortical volume in the ventral visual stream in 104 participants with and without psychosis in AD from the AddNeuroMed cohort.² This revealed a significant main effect of psychosis subtype on ventral visual stream regional brain volume and participants with psychosis symptoms had significantly lower left parahippocampal volume, most marked in those experiencing misidentification-type phenomena.² This is consistent with the role of the PHG in processing spatial and temporal context of visual information (i.e. where or when something was seen before).¹⁹⁷-¹⁹⁹

In relation to the two-factor theory of delusional belief formation described in section 2.2.2, these findings suggest that one potential neuropsychological impairment (i.e. the first factor in the model) may be an impaired ability to correctly attribute context to visual information as a result of parahippocampal dysfunction, leading to the formation of misidentification delusions (for example, the delusion that one’s house is not ‘home’).²

2.4.2  Context memory in Alzheimer’s disease

AD is characterised by loss of episodic memory: memory for experienced events in a spatial or temporal context. Memory for both spatial and temporal context is impaired in AD, with poorer performance on a task of temporal context memory showing predictive power for a diagnosis of mild AD versus frontotemporal lobar degeneration and normal ageing.²⁰⁰-²⁰² This reflects the importance of hippocampal networks, which are crucial for processing both spatial and temporal contextual information²⁰³, ²⁰⁴ and are the site of significant early AD pathology and dysfunction.²⁰⁵, ²⁰⁶

No studies that have investigated context memory in AD have screened participants for the presence of psychosis symptoms which, given the findings described in section 2.4.1, is a potentially important confounding factor. To my knowledge, there has been no specific investigation of context memory in AD patients with and without delusions.

See systematic review in Chapter 4 for further details.
2.4.3 Context memory and delusions

Deficits in both spatial and temporal context memory are associated with delusional beliefs in schizophrenia. These impairments appear to remain stable over time and are unchanged by antipsychotic medication. Impairments in context memory are also found in unaffected first degree relatives of individuals with schizophrenia and in individuals with schizotypal personality disorder. Compared to healthy controls, impaired context memory in individuals with schizophrenia is associated with reduced activity in the dIPFC bilaterally. Context memory impairments are also present in delusional disorder, bipolar affective disorder, in individuals who experience hallucinations in PD and those with confabulation following brain injury.

2.5 Associative inference

2.5.1 Associative inference in Alzheimer’s disease

Spatial and temporal memory are fundamentally relational in nature and involve the encoding of relationships between an item and its place in space or time. In addition to being spatial or temporal in nature, relationships held in memory can be associative, such as the associations between unrelated items such as a name and a face. The ability to infer associative relationships between stimuli, or ‘associative inference’, involves constructing these relationships indirectly from previously acquired knowledge. For example, if one learns directly that A is related to B and B is related to C, using associative inference one can infer that A is related to C. Transitive inference is a form of logical reasoning which extends associative inference to include how items are related, for example – if A > B and B > C then one can infer that A > C. Both associative and transitive inference require preserved hippocampal and ventromedial PFC (vmPFC) functioning. Both associative and transitive inference have been found to be impaired in AD and amnesia, likely as a consequence of hippocampal dysfunction.

2.5.2 Associative inference and delusions

People with schizophrenia demonstrate an impaired ability to infer relationships in both associative and transitive inference tasks. When completing associative inference tasks during fMRI, individuals with schizophrenia have reduced activity in bilateral ACC and right parietal cortex compared to healthy controls. Impaired transitive inference is also observed in unaffected first degree relatives of people with schizophrenia and in individuals with psychosis in PD, although not in affective psychosis in bipolar affective disorder.
No studies have been identified that examine how associative or transitive inference relate to delusions in AD.

### 2.6 Metamemory

#### 2.6.1 Metamemory in Alzheimer’s disease

‘Metamemory’ is the ability to self-evaluate memory capabilities, for example, how accurately one can identify which items were correctly remembered in a memory test. This is a crucial aspect of normal cognitive function and includes the ability to recognise and correct false memories.\(^{185, 233}\) Metamemory can be conceptualised as two related categories: global metamemory, which is the ability to assess one’s overall memory capabilities (where a deficit would equate to anosognosia) and local, task-specific metamemory, which is the online monitoring of memory processes.\(^{234}\) Anosognosia is commonly associated with AD, with prevalence rates as high as 80%.\(^{235}\) Findings regarding local metamemory function in AD have been less consistent; some studies have found intact metamemory in AD participants,\(^{234, 236}\) while others have demonstrated significant impairments.\(^{237}\) Where metamemory impairments are demonstrated, these have been associated largely with frontotemporal dysfunction. Impaired metamemory correlates with executive dysfunction\(^{238}\) and functional neuroimaging has demonstrated an association between anosognosia and decreased functional activation of mPFC and anterior temporal lobe regions.\(^{239}\)

#### 2.6.2 Metamemory and delusions

Impaired task-specific metamemory is a postulated mechanism for delusion formation in schizophrenia.\(^{240}\) Patients with schizophrenia have higher confidence in memory errors than healthy controls, and those who are currently delusional have higher confidence in memory errors than those who are not.\(^{193, 241}\) This is also the case for healthy individuals, where those with higher rates of delusional ideation are more confident in items falsely recognised as ‘old’ on memory tests.\(^{242}\) People with schizophrenia are also found to be more confident in correct responses on memory tasks.\(^{193}\) A further metamemory phenomenon has been observed in individuals with schizophrenia known as the ‘confidence gap’, in which individuals have a combination of high confidence in incorrect answers and low confidence in correct answers.\(^{241, 243}\) In schizophrenia, reduced metacognitive accuracy is associated with reduced grey matter (GM) volume in the ventrolateral and right dlPFC,\(^{244}\) with the PFC put forward as a possible ‘hub’ for confidence assessment.\(^{245}\)

Delusions in AD have previously been associated with global metamemory impairment\(^{246}\) but no studies have explored how task-specific metamemory relates to delusions in AD.
2.7 Clinical relevance

Early intervention in dementia is a key priority of the Department of Health’s National Dementia Strategy. This thesis aimed to investigate how delusions form in AD. During the design of the thesis it was noted that if, as hypothesised (see section 5.2 and section 8.2), AD patients with delusions performed more poorly than those without on simple context memory or false memory tasks, such tasks could be used as predictive tools, which could allow pre-emptive treatment and input for patients and their families. Further, understanding the role of metamemory in delusion formation could provide a basis for non-pharmacological intervention, for example cognitive training, to reinforce this important cognitive function. Increased understanding of how delusions form by elucidating both neuropsychological and neuroanatomical correlates is important as it may ultimately allow identification of novel therapeutic targets for drug development to better treat delusions in AD.

2.8 Aims of thesis

The overall aim of this thesis was to further elucidate the mechanism underlying delusion belief formation in AD by:

- Undertaking systematic review of relevant literature relating to memory errors of interest in AD
- Exploring the relationship between performance on memory and metamemory tasks and delusions in AD patient populations
- Exploring neuroanatomical correlates of memory errors and delusions in AD patient populations

Specific research questions and hypotheses are described in detail in the relevant chapters.
3. Neuroimaging Correlates of False Memory in Alzheimer’s Disease: A Systematic Review

3.1 Introduction

The design of studies described later in this thesis was informed and guided by systematic reviews, which aimed to understand the existing research base for two of the areas of interest described in Chapter 2: false memory and memory for spatial context. The first systematic review, described in this chapter, explored neuroimaging studies of false memory in AD. The second, described in Chapter 4, explored the neuropsychological correlates of memory for spatial context in AD.

3.2 Abstract

AD is characterised by episodic memory impairment, but patients also experience memory distortions, including false memories, which can impact on safety and functional ability. For example, falsely remembering that medication was taken when it was not. There has been a relative absence of research into how the location of neuropathology in AD contributes to false memory generation. Understanding the neural networks that underpin false memories could help to guide the development of cognitive strategies to reduce memory errors. This chapter contains a systematic review of the methodology and outcomes of published studies that investigated the neuroimaging correlates of false memory in AD. Seven studies were identified, four using structural imaging and three using functional imaging. Studies were heterogenous in methodology and received mostly ‘weak’ quality assessment ratings. Combined, and consistent with neuroimaging findings in non-AD populations, results from reviewed studies provide preliminary support for a hypothesis that MTL and PFC dysfunction may lead to generation of false memories in AD. However, the small number of studies and significant heterogeneity within them means further study will be necessary to assess the reproducibility of results.

3.3 Background

As described in section 2.3.1, AD is a neurodegenerative disorder characterised not only by impaired short term episodic memory, but also by an increased susceptibility to generate false memories.\textsuperscript{137, 138} The term false memory is used to describe a broad range of phenomena, including misremembering of word lists, false recognition of novel stimuli and distortions in autobiographical memory.\textsuperscript{135} Confabulation is included in this spectrum and ranges in expression from intrusions to more elaborate false beliefs relating to imagined past experiences.\textsuperscript{154}
As described in Chapter 2, the DRM paradigm\textsuperscript{142} is the tool used most widely to study false memory\textsuperscript{142} and requires participants to distinguish previously presented words from highly related but not presented ‘lure’ words and unrelated and not presented words. Results reported using the DRM in AD have been inconsistent, with some studies reporting equivalent or increased false memories\textsuperscript{144, 145, 148} and some finding reduced false memories\textsuperscript{146, 147} in individuals with AD compared to healthy controls. As described in section 2.3.2, there has been a relative absence of research into the relationship between the anatomical location of neuropathology of AD and false memories. Based on lesion and neuroimaging studies in non-AD populations,\textsuperscript{165, 168, 177, 178} it is hypothesised that false memories may result from reduced function within MTL structures with some preservation of mPFC leading to increased reliance on familiarity\textsuperscript{150, 180, 181} and additional frontal dysfunction leading to impairment of memory veracity monitoring.

False memories in dementia impact on safety (for example, when an individual incorrectly remembers turning off the oven or taking medication), reduce functional ability and increase caregiver distress.\textsuperscript{140, 141, 248} Use of cognitive strategies to reduce false memories is being explored.\textsuperscript{156, 249} An increased understanding of the neural underpinnings of false memory could guide further development in this area. I therefore completed the following systematic review of existing studies investigating the neuroimaging correlates of false memory in AD.

3.4 Methods

No ethical approval was required. Data were gathered and findings were reported according to the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) systematic review guidelines.\textsuperscript{250}

3.4.1 Literature search

An electronic database search (Embase, Embase Classic, Medline, PsycINFO, Web of Science) was completed on December 18th 2018 using the search terms (AD OR Alzheimer*) AND (False memor* OR false recall OR false recognition OR false alarm* OR false positive* OR confab* OR intrusion* OR DRM OR Deese Roediger McDermott OR Deese-Roediger-McDermott) AND (Imaging OR neuroimaging OR MRI OR fMRI OR magnetic resonance OR functional magnetic resonance OR PET OR Positron emission tomography OR positron-emission tomography OR CT OR computed tomography OR computerised tomography).

3.4.2 Inclusion and exclusion criteria

Studies were included if they were published in peer reviewed papers, in English, carried out in human subjects with diagnosed AD, and included a measure of false memory (based on the broad definition described in section 3.3) in relation to a neuroimaging measure (brain structure,
functioning or metabolism). Studies were excluded if they contained no analysis of the relationship between false memory task performance and neuroimaging or included patients with other forms of dementia or mild cognitive impairment (MCI) in the sample. Case reports, poster abstracts and dissertations were excluded. Reviews were excluded but reference lists from relevant reviews were screened.

### 3.4.3 Data extraction

SRa and I independently screened titles and abstracts, followed by full article texts, and extracted data to a pre-specified excel spreadsheet. This included study characteristics (inclusion and exclusion criteria, imaging modality), demographics (sex, mean age, ethnicity, handedness, mean years of education, medication status), clinical characteristics (AD diagnostic criteria, mean cognitive score, mean years since diagnosis) and relevant outcomes (false memory task results, neuroimaging correlates). Discrepancies between raters were resolved through discussion, or in consultation with a third researcher (SRe). Demographic characteristics of the studies (when specified) were then combined as weighted pooled means and standard deviations and pooled proportions, using formulae recommended by the Cochrane Collaboration applied sequentially, see Higgins et al. (2022).

### 3.4.4 Quality assessment

SRa and I independently evaluated the quality of each study using a modified version of the Effective Public Health Practice Project (EPHPP) tool for quality assessment, see Appendix 1. Studies were rated in the following areas: selection bias, study design, false memory task, neuroimaging methodology and statistical analysis. Sections D (blinding), F (withdrawals and drop-outs) and G (intervention integrity) were excluded as they are specific for randomised controlled trials and inappropriate for the studies included in this review. Scores in each subsection were used to provide an overall rating of ‘strong’ (no weak ratings in any section), ‘moderate’ (only one weak rating), or ‘weak’ (two or more weak ratings), and an overall score was given (5 - 14 points) based on the rating of each category (strong = 1, weak = 3), with lower scores indicating better methodological quality. Any disagreements were resolved through discussion.

### 3.5 Results

#### 3.5.1 Literature search and study selection

Database searches identified 837 potential studies, with manual search of relevant review references yielding a further 32 results. Of the 617 studies remaining after duplicate removal, seven met the criteria for data extraction (see PRISMA flow chart, Figure 3.1). Inter-rater
reliability was high, with 97.8% agreement, \( \kappa = 0.76 \). Study characteristics, main findings and quality assessment are shown in Table 3.1 and summarised in section 3.5.2.

**Figure 3.1 PRISMA flow diagram**

3.5.2 Characteristics of included studies

The majority of studies (\( n = 4 \)) received a ‘weak’ rating on quality assessment, with the remaining three studies rated as ‘moderate’. Across scoring categories, five studies had a ‘weak’ rating for selection bias; two studies had a ‘weak’ rating for study design; no study received a ‘weak’ rating for either false memory task or neuroimaging methodology; three studies received a ‘weak’ rating for analysis.

Mean age of participants was 70.4 ± 8.7 years (reported in six studies), 55.3% of participants were female and 44.7% were male (reported in five studies), mean MMSE score was 21.7 ± 3.8 (reported in five studies) and participants had spent on average 11.7 ± 3.2 years in education (reported in three studies). Duration of AD and handedness were reported in two studies (mean
34.4 ± 25.9 months, 90.9% right handed) and race in one study,\textsuperscript{253} with the majority (85.1\%) of participants being white. Medication status was partially reported by four studies. Four studies did not report medical or psychiatric history of participants or include this in inclusion/exclusion criteria. Information on recruitment setting was included in four studies, two of whom recruited participants via outpatient clinics, one from hospitals, and one from both inpatient and community sites. None of the studies provided detail of whether or not participants were living independently. Six studies used a cross-sectional methodology.

Five studies investigated intrusion errors, using either the California Verbal Learning Test\textsuperscript{254} (used in three studies) or Grober and Buschke’s test\textsuperscript{255} (used in two studies). Two studies investigated more elaborate confabulatory responses using a specifically developed Modified Confabulation Battery,\textsuperscript{161} a prose memory test\textsuperscript{256} or verbal recall tests from the Consortium to Establish a Registry for Alzheimer’s Disease (CERAD).\textsuperscript{257} Two studies investigated false recognition on Grober and Buschke’s test or the Rey Auditory Verbal Learning Test (RAVLT).\textsuperscript{258}

One study investigated correlates of ‘visuospatial confabulation’ in response to the Rey-Osterrieth Complex Figure (ROCF).\textsuperscript{259, 260}

Four studies investigated structural neuroimaging correlates of a false memory measure, three using MRI\textsuperscript{161, 253, 261} and one using diffusion tensor imaging (DTI).\textsuperscript{262} Each of the three remaining identified studies employed a different functional neuroimaging technique including connectivity analysis using resting-state fMRI,\textsuperscript{263} positron emission tomography (PET) using the 18F-fluoro-2-deoxyglucose (18FDG) tracer\textsuperscript{264} and single photon emission computed tomography (SPECT) using the 123I-N-isopropyl-p-iodoamphetamine (IMP) tracer.\textsuperscript{265}

### Neuroimaging methodology

#### Structural magnetic resonance imaging

All three studies used T1 weighted coronal contiguous acquisition protocols. Deweer et al. (1995) provided no detail of magnetic field strength and little further detail of the imaging protocol used, but Lee et al. (2009) and Weiner et al. (2011) used 1.0 Tesla (T) and 1.5T respectively. Lee et al. (2009) used rapid acquisition and employed a three dimensional (3D), spoiled gradient recalled sequence. Three different scanners were used by Weiner et al. (2011), which used 3D magnetisation-prepared rapid gradient echo, spoiled gradient recalled echo and turbo field echo sequences. The imaging data were co-registered either to other study participants or to an anatomical template in two of the studies – the Montreal Neurological Institute-152 (MNI-152) template was used by Lee et al. (2009) and a hippocampal template developed by Csernansky et al. (1998) by Weiner et al. (2011). Only Lee et al. (2009) gave information on slice thickness, voxel size or spatial smoothing of images. Slice thickness was
1.5mm, with voxel size 1.02 x 1.02 x 1.50mm, and images were smoothed using a Gaussian kernel of 10mm full width half maximum (FWHM) prior to voxel-based morphometry (VBM) with SPM2 running in MATLAB, with a statistical threshold of $p < .001$ uncorrected for multiple comparisons, rejecting cluster sizes less than 27 x 3mm.

Two of the studies used a region of interest (ROI) analysis, and one study used VBM. Deweer et al. (1995) completed a ROI analysis, manually tracing the hippocampus, amygdala and caudate nucleus, although identifying the ventricles using a semi-automated approach. The volume of each structure was then calculated from surface area by multiplying by inter-slice thickness. Weiner et al. (2011) also completed a ROI analysis, calculating volume change of left and right hippocampus, ventricles and total brain volume between scans at baseline, 24 and 48 weeks. Scans were co-registered to each other to provide a measure of volume change in total brain volume and ventricular volume via the boundary shift interval. Hippocampal volume was measured using a semi-automated technique, using a high-dimensional brain mapping tool.

3.5.3.2 Diffusion tensor imaging

Flanagan et al. (2016) completed a DTI analysis. Two DTI-weighted sequences were acquired for each participant using 3T magnetic field strength, 32 gradient directions and 2.5 x 2.5 x 2.5mm voxel size. Images were co-registered to the MNI template. A ROI analysis was then conducted using a mask for the fornix, created using the Johns Hopkins University White-Matter Tractography Atlas.

3.5.3.3 Functional magnetic resonance imaging connectivity analysis

Data for the fMRI connectivity analysis conducted by Venneri et al. (2017) were acquired at two different sites. In Italy, 18 participants had two resting-state fMRI echoplanar sequences (voxel dimensions 3.28 x 3.28 x 6.0mm) and a complementary 3D anatomical T1-weighted scan (1.1 x 1.1 x 6.0mm voxels) acquired on a 1.5T scanner. In the UK, the remaining 18 participants underwent resting-state fMRI sequences (voxel dimensions 1.8 x 1.8 x 4.0mm) and 3D T1-weighted anatomical scan (0.9 x 0.9 x 1.0mm voxels) on a 3T scanner.

Analysis was completed with VBM using SPM12 running in MATLAB. Images were co-registered to MNI space and smoothed with a Gaussian kernel of 6mm FWHM. Connectivity maps were created using a main seed of interest in the inferior-dorsal portion of the right orbitolateral PFC, contralateral homologous seed in the left orbitolateral PFC as an anatomical control and a third seed in the calcarine cortex as a methodological control – all generated using the WFU PickAtlas toolbox. Only clusters surviving family wise error (FWE) correction at $p < .05$ were considered for interpretation.
3.5.3.4 Positron emission tomography

Desgranges et al. (2002) investigated resting-state brain glucose utilisation (cerebral metabolic rate for glucose; CMRG1c) using FDG-PET. Images were collected from 50 minutes post injection, with 63 septa out planes acquired, (slice thickness 2.43mm, lateral resolution 2.2 x 2.2mm) with slices positioned parallel to the canthomeatal line. 3 - 5mCi of 18FDG were injected via radial artery catheter and serial sampling was used to determine the time course of 18FDG in plasma and average plasma glucose. CMRG1c values were normalised using the cerebellar vermis as a control region. Images were processed using SPM99 in MATLAB, co-registered to a Talairach atlas and smoothed using a 3D Gaussian filter of 14mm. Only voxels with normalised CMRG1c values above 40% of the mean for the whole brain were selected for analysis. A hypothesis-driven ROI analysis was completed, including the right PFC and left hippocampal region, with clusters considered as significant at \( p < .05 \), corrected for multiple tests using the small volume corrections routine of SPM99. To test specificity of the results, a further whole brain analysis was completed, with significance threshold set to uncorrected \( p < .01 \).

3.5.3.5 Single photon emission computed tomography

Reed et al. (1993) used SPECT to measure regional cerebral blood flow (rCBF), using the IMP tracer. Images were collected from 10 minutes after injection of a 5mCi bolus of IMP. Two tomographic levels were studied for 20 minutes each, corresponding to 5cm and 7cm above and parallel to the orbitomeatal line. A ROI approach was used to investigate frontal cortex, OFC, parietal, temporal and occipital cortices. rCBF values (activity/mm²) were normalised using the occipital cortex as control. No information was provided regarding co-registration or smoothing.

3.5.4 Central findings

Measures of volume, connectivity, perfusion and metabolism of medial temporal structures all correlated negatively with measures of false memory, although there was no consensus regarding lateralisation. In terms of structural imaging, the volume of both hippocampal formation and amygdala correlated negatively with intrusions and false recognition respectively,\(^{261}\) and the volume of the bilateral MTL and right middle temporal gyrus (MTG) correlated negatively with confabulation.\(^{161}\) Functional imaging also revealed findings related to MTL regions, with reduced connectivity between right OFC and right hippocampus and uncus and between left OFC and left PHG and fusiform gyri (FFG) in those patients with confabulation.\(^{263}\) Further, hypoperfusion of the right temporal lobe correlated with false recognition\(^{265}\) and hypometabolism in left entorhinal and left perirhinal regions correlated with intrusion errors.\(^{264}\)

Three studies had findings related to prefrontal regions, with ACC volume correlating negatively with confabulation\(^{161}\) and hypometabolism in the right dIPFC correlating with intrusion errors.\(^{264}\)
Findings related to connectivity of the PFC varied according to laterality, with increased connectivity of the left OFC (with right PFC and ACC), and reduced connectivity of right OFC (with MTL) in those with confabulations.\textsuperscript{263}
<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>n</th>
<th>AGE (YEARS)</th>
<th>GENDER (FEMALE)</th>
<th>RIGHT HANDED</th>
<th>MMSE</th>
<th>DURATION AD (MONTHS)</th>
<th>MEDICATION</th>
<th>FM MEASURE</th>
<th>SUMMARY OF FINDINGS</th>
<th>QA (SCORE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deweer et al. (1995)</td>
<td>18</td>
<td>72.4 ± 1.5</td>
<td>NR</td>
<td>NR</td>
<td>22.3 ± 0.9</td>
<td>28.8 ± 3.6</td>
<td>NR</td>
<td>Grober and Buschke test intrusions and false recognition</td>
<td>HPC formation volume correlates negatively with number of extra list intrusions, percentage of intrusions at cued recall and total recall ($r = .60, p = .009$, slope = -16.8 ± 5.6; $r = .62, p = .006$, slope = -42.9 ± 13.5; $r = .63, p = .005$, slope = -43.8 ± 13.6 respectively). Amygdala volume correlates negatively with false recognition ($r = .52, p = .03$, slope = -14.4 ± 6.0).</td>
<td>Weak (10)</td>
</tr>
<tr>
<td>Lee et al. (2009)</td>
<td>22</td>
<td>80.4 ± 6.2†</td>
<td>19 (86.4)</td>
<td>NR</td>
<td>15.7 ± 4.6†</td>
<td>80.5% donepezil†</td>
<td>MCB</td>
<td>Volume of anterior cingulate, bilateral MTL and right MTG ($p &lt; .001$) correlate negatively with semantic confabulation score.</td>
<td>Moderate (7)</td>
<td></td>
</tr>
<tr>
<td>AUTHOR (YEAR)</td>
<td>n</td>
<td>AGE (YEARS)</td>
<td>GENDER (FEMALE)</td>
<td>RIGHT HANDED</td>
<td>MMSE</td>
<td>DURATION AD (MONTHS)</td>
<td>MEDICATION</td>
<td>FM MEASURE</td>
<td>SUMMARY OF FINDINGS</td>
<td>QA (SCORE)</td>
</tr>
<tr>
<td>---------------</td>
<td>---</td>
<td>-------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>------</td>
<td>----------------------</td>
<td>------------</td>
<td>-------------</td>
<td>-------------------</td>
<td>------------</td>
</tr>
<tr>
<td><strong>STRUCTURAL MRI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weiner et al. (2011)</td>
<td>40</td>
<td>75.3 ± 7.6†</td>
<td>30 (63.8)†</td>
<td>41 (87.2)†</td>
<td>19.1 ± 3.7</td>
<td>NR</td>
<td>All taking ChEI</td>
<td>CVLT intrusions</td>
<td>No significant correlation between change in HPC volume and change in intrusions on CVLT after 24 weeks of add-on Memantine treatment.</td>
<td>Moderate (7)</td>
</tr>
<tr>
<td><strong>DTI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flanagan et al. (2016)</td>
<td>37</td>
<td>64.1 ± 7.5</td>
<td>16 (43.2)</td>
<td>NR</td>
<td>24.0 ± 3.7</td>
<td>37.1 ± 31.2</td>
<td>NR</td>
<td>RAVLT false recognition</td>
<td>No significant correlation between fornix integrity and false recognition.</td>
<td>Moderate (8)</td>
</tr>
<tr>
<td><strong>fMRI CONNECTIVITY ANALYSIS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venneri et al. (2016)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Confabulators</td>
<td>18</td>
<td>68.7 ± 10.6</td>
<td>10 (55.6)</td>
<td>NR</td>
<td>21.5 ± 3.3</td>
<td>NR</td>
<td>NR</td>
<td>‘Confabulatory tendencies’ on any of: ROCF (delayed recall), paired associates test, prose memory test (immediate or delayed) and CERAD learning and verbal recall</td>
<td>When compared to non-confabulators, confabulators had: • Reduced connectivity between right OFC and MTA including right and left STG (p = .28, 156 voxels, Z = 4.13†, p &lt; .001, 321 voxels, Z = 4.09† respectively), right uncus and HPC (p = 0.28, 156 voxels, Z = 3.97 and 3.75† respectively) and right and left insula (p &lt; .001, 612 voxels, Z = 4.64† and p &lt; .001, 321 voxels, Z = 4.13† respectively)</td>
<td>Weak (8)</td>
</tr>
<tr>
<td>Non-Confabulators</td>
<td>18</td>
<td>66.9 ± 11.5</td>
<td>10 (55.6)</td>
<td>NR</td>
<td>22.3 ± 4.0</td>
<td>NR</td>
<td>NR</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 3.1 Cont. Summary of study characteristics and main findings

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>n</th>
<th>AGE (YEARS)</th>
<th>GENDER (FEMALE)</th>
<th>RIGHT HANDED</th>
<th>MMSE</th>
<th>DURATION AD (MONTHS)</th>
<th>MEDICATION</th>
<th>FM MEASURE</th>
<th>SUMMARY OF FINDINGS</th>
<th>QA (SCORE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>fMRI CONNECTIVITY ANALYSIS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Venneri et al. (2016) Cont.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PET; 18FDG tracer</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Desgranges et al. (2002)</td>
<td>19</td>
<td>72.9 ± 5.5</td>
<td>12 (63.2)</td>
<td>19 (100)</td>
<td>22.6 ± 2.4</td>
<td>NR</td>
<td>‘Unmedicated’ Grober and Buschke test free recall intrusions</td>
<td>Metabolism (CMRG1c) in right SFG (p &lt; .05, 445 voxels, Z = 2.58), and left STG (p &lt; .01 uncorrected, 10 voxels, Z = 2.36) correlates negatively with intrusions in free recall.</td>
<td>Weak (8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When compared to non-confabulators, confabulators had:
- Reduced connectivity between left OFC and MTA including right and left STG (p = .005, 230 voxels, Z = 4.23 and p < .001, 491 voxels, Z = 4.89 respectively) and left FFG and PHG (p < .001, 326 voxels, Z = 4.24 and 4.15 respectively).
- Increased connectivity between right OFC and bilateral prefrontal areas including right SFG (p < .001, 1088 voxels, Z = 4.37) and left SFG and MFG (p = .008, 200 voxels, Z = 4.03 and 3.74 respectively).
- Increased connectivity between left OFC and bilateral frontal areas including right SFG and left MFG (p < .001, 1329 voxels, Z = 4.51 and 4.48 respectively).
### Table 3.1 Cont. Summary of study characteristics and main findings

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>n</th>
<th>AGE (YEARS)</th>
<th>GENDER (FEMALE)</th>
<th>RIGHT HANDED</th>
<th>MMSE</th>
<th>DURATION AD (MONTHS)</th>
<th>MEDICATION</th>
<th>FM MEASURE</th>
<th>SUMMARY OF FINDINGS</th>
<th>QA (SCORE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reed et al. (1993)</td>
<td>20</td>
<td>71.0 ± 6.5</td>
<td>10 (50.0)</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Excluded if taking antidepressant/antipsychotic</td>
<td>CVLTm false recognition</td>
<td>Right temporal lobe perfusion (rCBF) correlates negatively with false recognition (p &lt; .05, r = -0.47).</td>
<td>Weak (10)</td>
</tr>
</tbody>
</table>

**Notes:**
- All included studies were cross-sectional. Values are expressed as mean ± SD; range or n (%).
- AD = Alzheimer’s disease; CERAD = Consortium to Establish a Registry for Alzheimer’s Disease; ChEI = cholinesterase inhibitor; CVLT = California Verbal Learning Test; DTI = diffusion tensor imaging; 18FDG=18F-fluoro-2-deoxyglucose; FFG = fusiform gyrus; FM = false memory; fMRI = functional magnetic resonance imaging; HPC = hippocampus; IMP = 123I N-isopropyl-p-iodoamphetamine; MCB = Modified Confabulation Battery; MFG = medial frontal gyrus; MRI = magnetic resonance imaging; MTA = medial temporal area; MTG = medial temporal gyrus; MTL = medial temporal lobe; NR = not reported; OFC = orbitofrontal cortex; PET = positron emission tomography; PHG = parahippocampal gyrus; QA = quality assessment; RAVLT = Rey Auditory Verbal Learning Test; ROCF = Rey-Osterrieth Complex Figure; SD = standard deviation; SFG = superior frontal gyrus; SPECT = single photon emission computed tomography; STG = superior temporal gyrus.
- † demographic details are unavailable for subgroup who received imaging, therefore details of larger sample are provided.
- ‡ peak Z score.
3.6 Discussion

The small number of studies identified by this review and their overall heterogeneity both in terms of false memory measure and imaging methodology utilised means that the results should be interpreted with caution. With these qualifications in mind however, there are some patterns in the data, and these findings overlap with neuroimaging data of false memory and confabulation in non-AD populations in that they implicate both MTL and PFC dysfunction in the generation of false memories. Neuroimaging studies that have investigated false memory in non-AD populations have shown reduced MTL activity and increased PFC activity (including right dIPFC and ACC) during retrieval of false versus true memories (see Dennis et al. (2015)). The findings of reduced metabolism and reduced connectivity of PFC regions in those with false memory superficially conflicts with the evidence base, but changes due to age-related structural and functional decline may be relevant here. Older adults similarly show reduced MTL function in false memory, with a more mixed picture of PFC involvement (see Devitt and Schacter (2016)). Indeed, similar to the findings of Venneri et al. (2017), older adults have been shown to be more prone to false memories (as measured by false recognition) when they fail to recruit frontoparietal networks. Reduced connectivity between right OFC and MTL could also lead to confabulation through dysfunction of the reality-monitoring network suggested by Schnider (2003), whereby the posterior OFC suppresses activated memories not related to ongoing reality. Clearly, both MTL and PFC are large anatomical areas containing multiple structures with their own distinct functions and connected neural networks. However, due to a lack of replicated findings related to any one of these regions, it is difficult to draw any more specific conclusions based on the studies reviewed here.

3.6.1 Methodological issues

There are some significant limitations within the reviewed studies which further impact on the strength of any conclusions. Over half of the papers received a ‘weak’ quality rating on risk of bias assessment. The majority of studies had a small sample size and may have been underpowered to detect differences and, significantly, the majority of studies did not report either medication status or medical or psychiatric history of participants. Of note, Weiner et al. (2011), the study with the largest sample size (n = 40), found no significant relationship between hippocampal volume and intrusions. Sample selection was an issue in three studies, as Flanagan et al. (2016) included largely early onset cases of AD, which may affect generalisability of results. Reed et al. (1993) included two patients with clinical depression, and Lee et al. (2009) included 13 patients with delusions. Neither Reed et al. (1993) nor Lee et al. (2009) recorded any control for these clear potential confounders during analysis. Indeed, only two studies controlled for any potential confounders (age, gender, general cognitive function, attention, executive
functioning or presence of psychosis symptoms). Not controlling for baseline measures of cognitive function in particular could lead to misinterpretation of the results of the structural imaging studies, where overall disease severity could be a significant factor related to both reduced medial temporal volume and increase in false memories. There is also potential cause for concern regarding interpretation of some of the reported $p$ values, as only three of the studies corrected for multiple comparisons, despite the fact that only four were hypothesis-driven.

Methodological issues were also identified in relation to neuroimaging protocols. All of the studies involved co-registration of images to a pre-existing template, apart from Deweer et al. (1995) and Reed et al. (1993) who both completed their scanning in the early 1990s, prior to significant developments in neuroimaging analysis. However, several different templates were used and none of the studies assessed the precision of co-registration. This is particularly relevant for measurement of small structures such as the PHG.

3.6.2 Limitations

The main limitation of this review, as discussed previously, is the small number of studies and their significant heterogeneity, which makes it difficult to meaningfully extrapolate from the results. It is also true that by using a broad definition of false memory (to include false recognition, intrusion and confabulation) I have introduced further heterogeneity. However, it is generally acknowledged that these phenomena share similar underlying mechanisms.

The current review focused solely on neuroimaging correlates, without including other studies of neural correlates (for example, electroencephalography) or neuropsychological correlates of false memory. While strict inclusion and exclusion criteria are a strength of this review, this led to the exclusion of studies in which AD participants were combined with MCI or control participants, which could have contained interesting insights. The exclusion of grey literature means that this review has missed unpublished results, which are more likely to have non-significant findings.

3.6.3 Recommendation for future study

Given the limited number of studies identified by this review, further investigation of the neural correlates of false memory in AD is clearly warranted to assess the replicability of the findings. This would ideally involve both exploratory voxel-based analyses and hypothesis-driven research in larger samples to provide greater power. This would enable clearer delineation of the likely overlapping neural networks involved in false recognition, false recall, confabulation and delusion. I would also recommend that further research explore the effect of, and consider
controlling for, age, gender and psychosis symptoms on false memory and neuroimaging measures in AD. The DRM false memory paradigm is conspicuous by its absence in the results of this review. I would suggest that given the wealth of literature surrounding the DRM in non-AD participants and outside of neuroimaging in AD participants that this would be an interesting tool to use to further explore this area. Lastly, I would recommend that any further study in this area should control for variation in attention and upstream memory processes.

3.6.4 Conclusions

Few studies have explored the neuroimaging correlates of false memory in AD. This review provides initial support for pre-existing hypotheses that the MTL-PFC network of recollection, familiarity and veracity monitoring is involved in false memory generation in AD. However, further study is necessary to assess replicability of results.
4. Neuropsychological Correlates of Spatial Context Memory in Alzheimer’s Disease: A Systematic Review

4.1 Abstract

Characteristic loss of episodic memory in AD includes memory for the spatial context of experienced events. Impaired spatial memory has been identified as accompanying conversion from MCI to AD and has also been associated with item misplacement, wandering behaviour and psychosis symptoms. Increased understanding of the cognitive processes that contribute to spatial memory impairment in AD is therefore clinically relevant. This chapter contains a systematic review of the methodology and outcomes of studies exploring the neuropsychological correlates of spatial memory impairment in AD. Twenty-two studies were identified. These had used heterogenous methodologies and the majority (n = 14 studies; 63.6%) received ‘weak’ quality assessment ratings. One consistent finding was a positive correlation between measures of memory for spatial context and measures of general cognitive function. Consistent with evidence that spatial memory is a sensitive marker of early cognitive decline and of conversion of MCI to AD, these results lend weight to studies identifying spatial memory tasks as candidates for screening for AD. Understanding spatial memory impairment in AD is important, and I would recommend that future studies use valid and reliable existing tests to facilitate direct comparison of results.

4.2 Introduction

As described in section 2.4, AD is characterised by loss of episodic memory: memory for personally-experienced events within a spatial and temporal context (i.e. where or when an event occurred). Spatial memory is a broad concept, encompassing a range of neuropsychological processes, which can be divided into three main areas: the ability to remember the location of objects, the ability to remember topographical aspects of an environment, and the ability to remember the spatial context of an autobiographical event.

In healthy adults, information about the spatial context of experiences is processed via the ventral and dorsal (‘what’ and ‘where’) visual pathways which terminate in temporal and parietal areas respectively. In the MTL, the hippocampus, parahippocampus and medial entorhinal cortex process and integrate spatial contextual information, with egocentric (relative to the self) and allocentric (relative to the environment) location information processed in parallel (see Byrne et al. (2007) for a review). Similarly, a temporoparietal pathway involving hippocampal, parahippocampal and right posterior parietal regions is implicated in spatial
Of these regions, the hippocampus in particular provides spatial context for stimuli (see Bird and Burgess (2008) for a review).

The hippocampus is an area of significant neuropathological burden in AD, and shows evidence of dysfunction early in the disease course. It is therefore unsurprising that people with AD have impaired spatial memory. Impaired memory for spatial context has been implicated in item misplacement, spatial disorientation (leading to wandering) and delusions in AD. These symptoms are common in AD and reduce quality of life for patients and their families. Delusions are associated with accelerated cognitive and functional decline, while wandering behaviour increases risk of injury alongside its similar association with poorer functional outcomes.

Impaired spatial memory is also observed in those with aMCI compared to healthy controls, meaning that spatial memory deficit could be a useful predictor of those who might benefit from intervention. In addition, impairments in spatial memory may function as a phenotypic marker of behavioural and psychological symptoms (including wandering and psychosis symptoms as detailed above). Increased understanding of the cognitive mechanisms underlying memory for spatial context in AD is therefore important. This chapter includes a systematic review of published studies investigating the neuropsychological correlates of spatial memory impairment in AD.

4.3 Methods

4.3.1 Protocol and registration

No ethical approval was required. Methodology and data reporting followed the PRISMA systematic review guidelines, PROSPERO registration number CRD42018104565.

4.3.2 Literature search

A search of the online literature databases Embase, Embase Classic, Medline, PsycINFO and Web of Science was completed up to 5th March 2019, using the following search terms: (context memor* OR spatial memor* OR visuospatial memor* OR visuo-spatial memor* OR source memor* OR spatial context* OR spatiotemporal context* OR visuospatial context* OR source monitoring OR source recognition OR object location) AND (Alzheimer* OR AD OR DAT) AND (neuropsych* OR cognitive).
4.3.3  **Inclusion/exclusion criteria and screening**

Studies were included if they: included participants with AD; included a measure of spatial memory (defined below) in relation to the outcome of any other neuropsychological test; and were peer-reviewed and published in English. Studies were excluded if healthy controls, or patients with MCI or other forms of dementia were included in the sample and analysed collectively. Case reports, conference papers, poster abstracts and dissertations were also excluded. A manual search was conducted of the reference lists of relevant reviews.

We defined a test of spatial memory as any measure of recall or recognition of the previous location of a stimulus or of the participant themselves. This included spatial working memory tasks, pattern recognition or reproduction, navigational tasks involving route memory and recall of the spatial context of autobiographical memories.

JC and I independently screened all titles and abstracts for eligibility using these inclusion and exclusion criteria, followed by full texts of studies thought to be potentially relevant. Discrepancies were resolved through discussion between JC, SR, and me.

4.3.4  **Data extraction**

JC and I extracted data from eligible studies to a pre-defined excel spreadsheet. This included information about study characteristics, sample characteristics and relevant outcomes (including details of neuropsychological tasks and correlation with spatial context memory). Results were considered significant if a $p$ value $< .05$ was reported. Demographic characteristics of the studies (when specified) were then combined as weighted pooled means and standard deviations and pooled proportions, as per section 3.4.3.251

4.3.5  **Risk of bias/quality assessment**

JC and I independently assessed the quality of studies using a modified version of the EPHPP tool,252 see Appendix 2. Studies were rated in the following areas: selection bias, study design, spatial context memory task, correlated neuropsychological task and statistical analysis. Sections D, F and G of the tool were excluded as they are specific for randomised controlled trial methodology. Scores in each subsection were used to provide an overall rating of ‘strong’ (no weak ratings in any section), ‘moderate’ (one weak rating), or ‘weak’ (more than one weak rating), and an overall score was given (from 5 - 14 points) based on the rating of each category (strong = 1, weak = 3), with lower scores indicating better methodological quality. Discrepancies in scoring were resolved through discussion.
4.4 Results

4.4.1 Study selection

3665 studies were identified through database and manual reference searches, with 2969 remaining after duplicate removal and 181 after title and abstract screening (see Figure 4.1). Of these, 22 were ultimately eligible for inclusion. Inter-rater agreement was high: 97.2%, $\kappa = 0.78$. Characteristics, methodology and quality assessment of the included studies are shown in Table 4.1 and summarised below.

*Figure 4.1 PRISMA flow diagram*

- Identification: Records identified through database searching ($n = 3628$) and manual search of review references ($n = 37$).
- Screening: Records after duplicates removed ($n = 2969$).
- Eligibility: Records excluded based on titles and abstracts ($n = 2788$), Full-text articles excluded ($n = 159$) for various reasons including not reporting neuropsychological correlates of spatial context memory, not investigating spatial context memory, not including a pure Alzheimer's disease sample, not being primary research studies, and not being available in English.
- Included: Full-text articles assessed for eligibility ($n = 181$), Studies included in qualitative synthesis ($n = 22$).
<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>n</th>
<th>AGE (YEARS)</th>
<th>GENDER (FEMALE)</th>
<th>EDUCATION (YEARS)</th>
<th>COGNITIVE ASSESSMENT</th>
<th>DURATION AD (YEARS)</th>
<th>MEDICATION</th>
<th>SPATIAL CONTEXT MEMORY MEASURE</th>
<th>NEUROPSYCHOLOGICAL TEST(S) INCLUDED IN ANALYSIS</th>
<th>STATISTICAL METHODOLOGY</th>
<th>QA (SCORE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benke et al. (2014)</td>
<td>37</td>
<td>78.9 ± 7.3</td>
<td>28 (75.7)</td>
<td>9.3 ± 1.9</td>
<td>MMSE: 23.4 ± 2.2</td>
<td>NR</td>
<td>NR</td>
<td>Naturalistic route/landmark learning task</td>
<td>CERAD: word list recall and recognition; constructional praxis; verbal fluency</td>
<td>Stepwise multiple regression</td>
<td>Moderate (8)</td>
</tr>
<tr>
<td>Bucks and Willison (1997)</td>
<td>18</td>
<td>70.4 ± 6.5</td>
<td>14 (73.7)</td>
<td>NR</td>
<td>MMSE: 19.5 ± 4.5</td>
<td>NR</td>
<td>NR</td>
<td>Location learning task</td>
<td>MMSE</td>
<td>NR</td>
<td>Weak (11)</td>
</tr>
<tr>
<td>Budolfson et al. (2015)</td>
<td>67</td>
<td>83.6 ± 6.4</td>
<td>29 (43.3)</td>
<td>14.2 ± 2.6</td>
<td>MMSE: 22 [19,27] (median [IQR])</td>
<td>NR</td>
<td>NR</td>
<td>BVMT-R</td>
<td>AQ</td>
<td>Spearman</td>
<td>Weak (10)</td>
</tr>
<tr>
<td>Cherrier et al. (2001)</td>
<td>16</td>
<td>76.0 ± 5.0</td>
<td>4 (25.0)</td>
<td>14.4 ± 3.0</td>
<td>DRS: 107 ± 18</td>
<td>3.8 ± 1.4</td>
<td>NR</td>
<td>Naturalistic route/landmark learning task</td>
<td>DRS; FMT-R; HVOT; JLO; M-RMT; NMT-R; VSLT-R</td>
<td>Stepwise linear regression</td>
<td>Moderate (8)</td>
</tr>
<tr>
<td>Efklides et al. (2002)</td>
<td>39</td>
<td>72.0 ± NR</td>
<td>28 (71.8)</td>
<td>7.2 ± NR</td>
<td>MMSE: range 12-23</td>
<td>NR</td>
<td>NR</td>
<td>RBMT: immediate and delayed route</td>
<td>WMS; EMQ</td>
<td>Pearson</td>
<td>Weak (9)</td>
</tr>
<tr>
<td>El Haj et al. (2017)</td>
<td>24</td>
<td>73.7 ± 6.5</td>
<td>17 (70.8)</td>
<td>8.9 ± 2.8</td>
<td>MMSE: 21.8 ± 1.5</td>
<td>NR</td>
<td>NR</td>
<td>Location learning task</td>
<td>Remote and recent event dating task</td>
<td>NR</td>
<td>Moderate (9)</td>
</tr>
<tr>
<td>El Haj and Antoine (2018)</td>
<td>31</td>
<td>72.9 ± 7.8</td>
<td>20 (64.5)</td>
<td>8.8 ± 2.8</td>
<td>MMSE: 21.8 ± 1.4</td>
<td>NR</td>
<td>NR</td>
<td>Memory for spatial context of autobiographical events</td>
<td>MMSE</td>
<td>NR</td>
<td>Weak (9)</td>
</tr>
<tr>
<td>AUTHOR (YEAR)</td>
<td>n</td>
<td>AGE (YEARS)</td>
<td>GENDER (FEMALE)</td>
<td>EDUCATION (YEARS)</td>
<td>COGNITIVE ASSESSMENT</td>
<td>DURATION AD (YEARS)</td>
<td>MEDICATION</td>
<td>SPATIAL CONTEXT MEMORY MEASURE</td>
<td>NEUROPSYCHOLOGICAL TEST(S) INCLUDED IN ANALYSIS</td>
<td>STATISTICAL METHODOLOGY</td>
<td>QA (SCORE)</td>
</tr>
<tr>
<td>---------------</td>
<td>---</td>
<td>-------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>--------------------</td>
<td>-------------</td>
<td>--------------------------------</td>
<td>-----------------------------------------------</td>
<td>------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Elgh et al. (2006)</td>
<td>16</td>
<td>75.3 ± 7.1</td>
<td>11 (68.8)</td>
<td>NR</td>
<td>MMSE: 20.5 ± 5.8</td>
<td>NR</td>
<td>NR</td>
<td>WAIS-R spatial span</td>
<td>MMSE; RMT; ADAS- word recall</td>
<td>Spearman</td>
<td>Moderate (8)</td>
</tr>
<tr>
<td>Foxe et al. (2013)</td>
<td>12</td>
<td>66.9 ± 8.4</td>
<td>4 (33.3)</td>
<td>11.5 ± 2.5</td>
<td>ACE-R: 58.6 ± 15.6</td>
<td>Age at onset: NR 62.1 ± 7.6 years</td>
<td>NR</td>
<td>NR</td>
<td>WMS-III spatial span</td>
<td>ACE-R; Doors Test A; Picture naming; ROCF; Sentence repetition; TMT; Word repetition</td>
<td>Spearman</td>
</tr>
<tr>
<td>Kéri (2014)</td>
<td>20</td>
<td>66.2 ± 7.9</td>
<td>8 (40.0)</td>
<td>11.5 ± 3.1</td>
<td>ACE: 84.2 ± 5.7; MMSE: 26.5 ± 3.4</td>
<td>NR</td>
<td>NR</td>
<td>Computerised and real-life PAL task†</td>
<td>RMET</td>
<td>Pearson</td>
<td>Weak (10)</td>
</tr>
<tr>
<td>Lee et al. (2014)</td>
<td>20</td>
<td>72.4 ± 5.6</td>
<td>10 (50.0)</td>
<td>7.0 ± 4.2</td>
<td>MMSE: 20.5 ± 3.4</td>
<td>NR</td>
<td>NR</td>
<td>Virtual radial arm maze‡</td>
<td>Spatial span; SRFT</td>
<td>Kendall’s tau-b</td>
<td>Weak (9)</td>
</tr>
<tr>
<td>MacPherson et al. (2007)</td>
<td>15</td>
<td>75.0 ± 8.2</td>
<td>10 (66.7)</td>
<td>12.1 ± 2.9</td>
<td>MMSE: 22.1 ± 1.8</td>
<td>NR</td>
<td>NR</td>
<td>Visual pattern span‡</td>
<td>Dual task: digit span plus articulatory suppression, visual patterns and tracking†</td>
<td>Spearman</td>
<td>Moderate (9)</td>
</tr>
<tr>
<td>Monacelli et al. (2003)</td>
<td>14</td>
<td>73.4 ± 5.9</td>
<td>NR</td>
<td>NR</td>
<td>MMSE: 23.1 ± 3.3</td>
<td>NR</td>
<td>NR</td>
<td>Naturalistic route/ landmark learning task‡</td>
<td>Benton: naming, facial recognition and JLO; M-RMT; WMS-R: figural memory and verbal PAL</td>
<td>NR</td>
<td>Weak (10)</td>
</tr>
<tr>
<td>Morganti et al. (2013)</td>
<td>26</td>
<td>81.0 ± 6.3</td>
<td>NR</td>
<td>NR</td>
<td>MMSE: 21.6 ± 2.5</td>
<td>NR</td>
<td>NR</td>
<td>Corsi spatial span</td>
<td>VR-maze test‡ and VR-road map test‡</td>
<td>Pearson</td>
<td>Weak (10)</td>
</tr>
<tr>
<td>Plancher et al. (2012)</td>
<td>15</td>
<td>76.5 ± 5.5</td>
<td>13 (86.7)</td>
<td>NR</td>
<td>MMSE: 19.3 ± 3.5</td>
<td>NR</td>
<td>NR</td>
<td>VR route/ landmark learning task‡</td>
<td>CDS</td>
<td>Pearson</td>
<td>Weak (11)</td>
</tr>
<tr>
<td>AUTHOR (YEAR)</td>
<td>n</td>
<td>AGE (YEARS)</td>
<td>GENDER (FEMALE)</td>
<td>EDUCATION (YEARS)</td>
<td>COGNITIVE ASSESSMENT</td>
<td>DURATION AD (YEARS)</td>
<td>MEDICATION</td>
<td>SPATIAL CONTEXT MEASURE</td>
<td>NEUROPSYCHOLOGICAL TEST(S) INCLUDED IN ANALYSIS</td>
<td>STATISTICAL METHODOLOGY</td>
<td>QA SCORE</td>
</tr>
<tr>
<td>----------------------</td>
<td>---</td>
<td>-------------</td>
<td>-----------------</td>
<td>-------------------</td>
<td>----------------------</td>
<td>---------------------</td>
<td>-------------</td>
<td>-------------------------</td>
<td>------------------------------------------------</td>
<td>-------------------------</td>
<td>----------</td>
</tr>
<tr>
<td>Sahgal et al. (1991)</td>
<td>14</td>
<td>73.7 ± 6.0</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>MMSE: 19.9 ± 4.8</td>
<td>NR</td>
<td>No psychotropic medication</td>
<td>CANTAB: PAL, delayed matching to sample, pattern recognition</td>
<td>MMSE</td>
<td>NR</td>
</tr>
<tr>
<td>Sahgal et al. (1992)</td>
<td>15</td>
<td>78.0 ± 2.3</td>
<td>11 (73.3)</td>
<td>NR</td>
<td>NR</td>
<td>MMSE: 19.3 ± 4.9</td>
<td>NR</td>
<td>No psychotropic medication</td>
<td>CANTAB: spatial recognition, spatial span, spatial working memory</td>
<td>MMSE</td>
<td>Bivariate regression analysis</td>
</tr>
<tr>
<td>Simone and Bayliss (1997)</td>
<td>13</td>
<td>79.9 ± 6.5</td>
<td>5 (38.5)</td>
<td>MMSE: 13.7 ± 3.3;</td>
<td>MMSE: 13.1 ± 2.9</td>
<td>NR</td>
<td>NR</td>
<td>Location learning task⁴</td>
<td>MMSE</td>
<td>NR</td>
<td>Moderate (9)</td>
</tr>
<tr>
<td>Stuart-Hamilton et al. (1988)</td>
<td>13</td>
<td>68.6 ± 5.3</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
<td>Spatial span¹</td>
<td>Object recognition¹</td>
<td>Kendall’s tau</td>
<td>Weak (12)</td>
</tr>
<tr>
<td>Tetewsky and Duffy (1999)</td>
<td>11</td>
<td>73.0 ± NR</td>
<td>NR</td>
<td>NR</td>
<td>MMSE: 21.9 ± 3.8</td>
<td>NR</td>
<td>NR</td>
<td>Naturalistic route/landmark learning task⁴</td>
<td>MMSE; shape, motion and self-movement discrimination⁴</td>
<td>Linear regression</td>
<td>Weak (10)</td>
</tr>
<tr>
<td>Tosto et al. (2015)</td>
<td>153</td>
<td>73.2 ± 8.1</td>
<td>72 (47.1)</td>
<td>10.7 ± 4.9</td>
<td>MMSE: 25.0 ± 2.8</td>
<td>2.5 ± 1.7</td>
<td>All taking ChEi/Memantine</td>
<td>ROCF</td>
<td>MMSE rate of decline</td>
<td>Cox regression models</td>
<td>Moderate (8)</td>
</tr>
</tbody>
</table>
**Table 4.1 Cont. Summary of study characteristics and methodology**

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>n</th>
<th>AGE (YEARS)</th>
<th>GENDER (FEMALE)</th>
<th>EDUCATION (YEARS)</th>
<th>COGNITIVE ASSESSMENT</th>
<th>DURATION AD (YEARS)</th>
<th>MEDICATION</th>
<th>SPATIAL CONTEXT MEMORY MEASURE</th>
<th>NEUROPSYCHOLOGICAL TEST(S) INCLUDED IN ANALYSIS</th>
<th>STATISTICAL METHODOLOGY</th>
<th>QA (SCORE)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tripathi et al. (2013)</td>
<td>39</td>
<td>66.0 ± 8.0</td>
<td>10 (25.6)</td>
<td>12.9 ± 4.6</td>
<td>HMSE: 23.7 ± 5.0</td>
<td>NR</td>
<td>NR</td>
<td>Corsi spatial span; stick construction recall</td>
<td>HMSE; CDR</td>
<td>Pearson</td>
<td>Weak (8)</td>
</tr>
</tbody>
</table>

Notes:
All included studies were cross-sectional. Values are expressed as mean ± SD; range or n (%) unless otherwise specified. Racial demographics were not recorded in any study.
ACE = Addenbrooke’s Cognitive Examination; ACE-R = Addenbrooke’s Cognitive Examination-Revised; AD = Alzheimer’s disease; ADAS = Alzheimer’s Disease Assessment Scale; AQ = Alzheimer’s questionnaire; BVMT-R = Brief Visuospatial Memory Test-revised; CANTAB = Cambridge Neuropsychological Test Automated Battery; CDR = Clinical Dementia Rating; CDS = Cognitive Difficulties Scale; CERAD = Consortium to Establish a Registry for Alzheimer’s disease neuropsychological battery; ChEI = cholinesterase inhibitor; DRS = Dementia Rating Scale; EMQ = Everyday Memory Questionnaire; HMSE = Hindi mental status examination; HVOT = Hooper Visual Organisation Test; IQR = interquartile range; JLO = Judgement of Line Orientation test; MMSE = Mini Mental State Examination; M-RMT = Money’s Road Map Test; NMT-R = New Map Test-Revised; NR = not recorded; PAL = paired associate learning; QA = quality assessment; RBMT = Rivermead Behavioural Memory Test; RMET = Reading the mind in the eyes; RMT = Rey 15 Item Memory Test; ROCF = Rey-Osterrieth Complex Figure; SD = standard deviation; SRFT = Simplified ROCF; TMT = Trail Making Test (B-A); VR = virtual reality; VSLT-R = Visuospatial Learning Test-Revised; WAIS-R = Wechsler Adult Intelligence Scale-Revised; WMS = Wechsler Memory Scale; WMS-R = Wechsler Memory Scale-Revised; WMS-III = Wechsler Memory Scale third edition.
† demographic details are unavailable for subgroup included in correlation analysis, therefore details of larger sample are provided.
‡ novel task designed by study authors.
4.4.2 Study characteristics

All included studies were cross-sectional, with a median sample size of 16 (range 11 - 153). Three studies recruited participants living independently in the community, two from ‘retirement homes’ and one recruited a mixed sample of participants living in the community and residential care. Information on setting was not reported by 16 studies. The majority of studies (n = 14 studies) were rated ‘weak’ on quality assessment and the remaining eight studies were rated ‘moderate’. Mean age of participants was 74.7 ± 8.4 years (reported with SD in n = 20 studies), 53.4% of participants were female and 46.6% were male (reported in n = 17 studies). Thirteen studies reported years of education (mean 11.3 ± 4.2). No studies reported the racial demographics of participants. Most studies (n = 15 studies) excluded participants with a history of other neurological or psychiatric illness, with the remaining seven studies not reporting this information. Three studies provided partial details of participants’ medication status. The MMSE was the most used measure of general cognitive function (n = 18 studies; mean score 22.5 ± 4.1). Duration of AD was reported by two studies (mean 2.6 ± 1.7 years). No studies included information regarding presence or absence of psychosis symptoms.

4.4.3 Assessment of memory for spatial context

The majority of studies (n = 15 studies) used lab-based neuropsychological assessments of spatial memory. The remaining seven studies employed a naturalistic/real-life testing methodology.

4.4.3.1 Spatial working memory
The most commonly used tasks to assess spatial memory were visuospatial working memory tests (n = 7 studies, see Table 4.2). The majority of these tasks were variations of the original Corsi block tapping test,\textsuperscript{289, 290} which requires the involvement of a combination of visuospatial sketchpad, controlled attention and central executive function.\textsuperscript{291} These were: spatial spans from various iterations of the Wechsler Memory Scale (WMS),\textsuperscript{292} with a 10 cube span\textsuperscript{293, 294}; a computerised version of the Corsi task as part of the Cambridge Neuropsychological Test Automated Battery (CANTAB)\textsuperscript{295, 296}; and an eight-item two dimensional spatial span with two conditions (black squares or coloured shapes).\textsuperscript{297} One study used the visual working memory test of pattern recall,\textsuperscript{298} which was developed by Della Sala et al. (1999) and has been proposed as a paradigm that can specifically test memory for visual appearance rather than spatial location.\textsuperscript{300}

4.4.3.2 Location/paired associate learning
Tasks requiring participants to learn object locations or object-location relationships were used by six studies. Three studies used tasks from the CANTAB; Sahgal et al. (1992) used the spatial
recognition task and Kéri (2014) and Sahgal et al. (1991) used the paired associate learning task, which involves attention and executive processes alongside associate learning. Three studies developed novel location learning tasks.

4.4.3.3 Visuospatial recall and recognition
One study used CANTAB tasks of visuospatial recognition, while four studies used visuospatial reproduction tasks to assess immediate and delayed visuospatial memory. These included the ROCF (n = 2 studies), which is known to involve elements of executive function (organisation and planning); the Brief Visuospatial Memory Test revised; and the National Institute of Mental Health and Neurosciences, India, Neuropsychological battery for the elderly stick construction task.

4.4.3.4 Landmark and route learning/navigation
Two studies developed virtual reality (VR) landmark/route learning and navigation tasks. A further four studies developed novel real-life route and landmark learning tasks, while Efklides et al. (2002) used the route recall subsections of the Rivermead Behavioural Memory Test.

4.4.3.5 Spatial context of autobiographical memory
El Haj and Antoine (2018) developed a novel measure for spatial context memory of autobiographical events. Participants were asked to remember three events in detail and prompted for details regarding their spatial context, which were clarified with family members to ascertain the number of correct location details.

4.4.4 Neuropsychological correlates

4.4.4.1 General Cognitive Function
The majority of studies (n = 14 studies) explored the correlation between spatial memory and a general cognitive assessment tool. Eight used the Mini Mental State Examination (MMSE); one used the Hindi MMSE equivalent (Hindi Mental State Examination, HMSE). The remaining five studies each used one of: the Addenbrooke’s Cognitive Examination-Revised (ACE-R), the Alzheimer’s questionnaire, the Clinical Dementia Rating (CDR), the Cognitive Difficulties Scale and the Dementia Rating Scale.

4.4.4.2 Memory
A wide range of memory subtypes were assessed for potential correlation with spatial memory. These included working memory (digit and spatial span; n = 3 studies) and facial recognition (Benton et al. (1983) and Everyday Memory Questionnaire; n = 2 studies). Five studies explored correlation with components of verbal memory, including: recognition (using the CERAD word list), immediate recall (CERAD and Alzheimer’s Disease Assessment Scale
(ADAS) word recall and WMS logical memory subtest) and WMS verbal paired associate learning. Visual memory components explored as potential correlates included recognition (Doors A test, WMS figural memory and a novel object recognition task; n = 3 studies), immediate recall (measured by Rey’s 15 item memory test, the New Map Test-Revised, ROCF and simplified ROCF and WMS visual reproduction; n = 5 studies) and delayed recall (measured by CERAD constructional praxis recall and ROCF/simplified ROCF; n = 3 studies). Cherrier et al. (2001) used the visuospatial learning test-revised, which combines immediate and delayed visual recall.

### 4.4.4.3 Visuospatial abilities

Foxe et al. (2013) used the visuospatial subscale of the ACE-R as a general measure of visuospatial ability. Visual discrimination was investigated by Cherrier et al. (2001), Monacelli et al. (2003) and Tetewsky and Duffy (1999) (using figure matching, judgement of line orientation (JLO) and novel tasks of shape, motion and self-movement discrimination). Three studies used the ROCF to explore the relationship between spatial memory and constructional praxis.

The most commonly used measure of visuospatial ability was Money’s Road Map Test (M-RMT; n = 3 studies), which assesses left-right discrimination and mental spatial rotation.

### 4.4.4.4 Executive function/other

Fewer studies assessed correlation between spatial memory measures and executive function. Three studies used measures of executive function including problem-solving (Hooper Visual Organisation Test; n = 1 study), set shifting (Trail Making Test; n = 1 study) and verbal fluency tasks (CERAD/ACE-R; n = 2 studies). One study assessed correlation between performance on spatial memory tasks and theory of mind (Reading the Mind in the Eyes test).

### 4.4.5 Central findings

There was significantly heterogeneity across studies, in terms of sample characteristics, spatial memory assessment tool used, and neuropsychological correlates explored, and the majority of studies were rated as having low methodological quality. However, with these caveats in mind, there are some patterns in the data as summarised in Table 4.2.

The most frequently replicated finding was that impaired spatial memory across various tasks was related to lower MMSE scores. MMSE (or HMSE) score was consistently positively correlated with visuospatial working memory measured by Corsi block tapping, Wechsler Adult Intelligence Scale-Revised (WAIS-R) and CANTAB and with object-location learning measures; two validated and one novel. Measures of visuospatial memory (by recall and recognition) were positively correlated with MMSE or HMSE score in two studies. Tosto et al. (2015) found no significant correlation, however their study differed significantly in methodology, as it
explored the potential relationship with rate of decline of MMSE score rather than outright MMSE score. General cognitive function as measured by MMSE also positively correlated with the number of spatial context details provided for autobiographical memories.\textsuperscript{321}

Naturalistic tasks involving route learning and landmark recall correlated with spatial orientation as assessed by M-RMT.\textsuperscript{316, 317} Verbal and visual immediate recall was also positively correlated with performance on route learning tasks.\textsuperscript{308, 315, 319} There was some indication that visual discrimination has a role in visuospatial working memory tasks and route learning, as object discrimination,\textsuperscript{297} self-movement discrimination\textsuperscript{318} and JLO test score\textsuperscript{317} were correlated with these measures. However, Cherrier et al. (2001), who unlike Stuart-Hamilton et al. (1988) and Monacelli et al. (2003) controlled for general cognitive function as a confounder, found no significant correlation between JLO score and route learning task performance.
| Table 4.2 Summary of main findings by type of spatial context memory task |

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>MMSE</th>
<th>Memory</th>
<th>Visuospatial</th>
<th>Language</th>
<th>Executive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Elgh et al. (2006)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Foxe et al. (2013)</td>
<td></td>
<td>×</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MacPherson et al. (2007)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morganti et al. (2013)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>↑</td>
</tr>
<tr>
<td>Saghal et al. (1992)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**MEMORY**
- Working memory
- Facial recognition
- Verbal IR
- Verbal PAL
- Visual recognition
- Visual IR
- Visual DR
- Visual discrimination
- Constructional praxis
- Spatial orientation
- Picture naming
- Verbal fluency

**NEUROPSYCHOLOGICAL DOMAINS**
- Global cognitive function
- Memory
- Visuospatial
- Language
- Executive function

**SUMMARY OF FINDINGS**
- **Visuospatial working memory:**
  - WAIS-R spatial span positively correlates with MMSE \( r = .60, p < .05 \). No correlation between spatial span and ADAS-word recall.
  - No correlation between WMS-III spatial span and Doors A, picture naming, ROCF or ACE-R verbal fluency.
  - Visual pattern span positively correlates with digit span plus interpolated tracking and articulatory suppression plus visual pattern span \( r = .40, p < .005; r = .27, p < .05; r = .41, p < .005 \).
  - Corsi’s supraspan positively correlates with VR-RMT \( r = .33, p < .049 \). No correlation between Corsi’s span and VR-RMT.
  - CANTAB spatial span positively correlates with MMSE \( r = .54, p = .03 \).
### Table 4.2 Cont. Summary of main findings by type of spatial context memory task

<table>
<thead>
<tr>
<th>Neuropsychological Domains</th>
<th>Memory</th>
<th>Visuospatial</th>
</tr>
</thead>
<tbody>
<tr>
<td>Global Cognitive Function</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Working memory</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Facial recognition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Verbal PAL</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual recognition</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual DR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visual discrimination</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constructional praxis</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spatial orientation</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Executive Function</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>MMSE</th>
<th>Digit span</th>
<th>EMQ/Benton</th>
<th>Word list/story recall</th>
<th>WMS</th>
<th>Object/pattern</th>
<th>Pattern/figure recall</th>
<th>JLO/Object/Shape/Mvmt</th>
<th>Pattern/figure recall</th>
<th>Pattern/figure copy</th>
<th>M-RMT/VR-RMT</th>
<th>Picture naming</th>
<th>Verbal fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stuart-Hamilton et al. (1988)</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tripathi et al. (2013)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sahgal et al. (1991)</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saghal et al. (1992)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Summary of Findings**

**Visuospatial Working Memory (Cont.):**
- Spatial span of black or coloured shapes negatively correlates with total errors on object recognition plus distractors ($\tau = -0.65; \tau = -0.62$).
- Corsi forward and backward span positively correlate with HMSE ($r = .70, p < .01; r = .64, p < .01$).

**Object-Location Learning:**
- No significant correlation between location learning task and MMSE.
- CANTAB PAL$^*$ negatively correlates with MMSE ($r = -.76, p < .01$).
- CANTAB spatial recognition positively correlates with MMSE ($r = .54, p = .03$).
<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>MMSE</th>
<th>Digit span</th>
<th>EMQ/Benton</th>
<th>Word list/story recall</th>
<th>WMS</th>
<th>Object/Pattern</th>
<th>Pattern/figure recall</th>
<th>JLO/Object/Shape/Mvt</th>
<th>Pattern/figure copy</th>
<th>M-RMT/VR-RMT</th>
<th>Picture naming</th>
<th>Verbal fluency</th>
<th>SUMMARY OF FINDINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simone and Bayliss (1997)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Object location learning correlates with MMSE when demands on attention reduced ($r = .76, p &lt; .01$).</td>
</tr>
<tr>
<td>Sahgal et al. (1991)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>CANTAB delayed match to sample positively correlates with MMSE ($r = .82, p &lt; .01$). No correlation between CANTAB pattern recognition and MMSE.</td>
</tr>
<tr>
<td>Tosto et al. (2015)</td>
<td>✗</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No correlation between ROCF immediate recall and MMSE.</td>
</tr>
<tr>
<td>Tripathi et al. (2013)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NNBE stick construction immediate recall positively correlates with HMSE ($r = .71, p &lt; .01$).</td>
</tr>
</tbody>
</table>
Table 4.2 Cont. Summary of main findings by type of spatial context memory task

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>MMSE</th>
<th>Digit span</th>
<th>EMQ/Benton</th>
<th>Word list/ story recall</th>
<th>WMS</th>
<th>Object/ pattern</th>
<th>Pattern/ figure recall</th>
<th>Pattern/ figure recall</th>
<th>JLO/ Object/ Shape/ Mvmt</th>
<th>Pattern/ figure recall</th>
<th>M-RMT/ VR-RMT</th>
<th>Picture naming</th>
<th>Verbal fluency</th>
<th>SUMMARY OF FINDINGS</th>
</tr>
</thead>
<tbody>
<tr>
<td>VISUAL DELAYED RECALL:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tosto et al. (2015)</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>No correlation between ROCF delayed recall and MMSE.</td>
</tr>
<tr>
<td>Tripathi et al. (2013)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NNB-E stick construction delayed recall positively correlates with HMSE (r = .61, p &lt; .01).</td>
</tr>
<tr>
<td>SPATIAL ORIENTATION (VR):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al. (2014)</td>
<td></td>
<td></td>
<td></td>
<td>↓</td>
<td>↓</td>
<td>↓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Navigation errors in VR-maze negatively correlate with SRFT copy, IR and DR (tb = -0.40; tb = -0.41; tb = -0.41, all p &lt; .05).</td>
</tr>
<tr>
<td>SPATIAL ORIENTATION (NATURALISTIC):</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>↓</td>
<td></td>
<td></td>
<td>Navigation errors on route learning negatively correlate with CERAD word list recall (B = -0.44, t = -4.88, p &lt; .001) and verbal fluency (B = -0.28, t = -3.11, p &lt; .002).</td>
</tr>
<tr>
<td>AUTHOR (YEAR)</td>
<td>MMSE</td>
<td>Digit span</td>
<td>EMQ/Benton</td>
<td>Word list/story recall</td>
<td>WMS</td>
<td>Object/ pattern</td>
<td>Pattern/ figure recall</td>
<td>Pattern/ figure recall</td>
<td>JLO/Object/ Shape/ Mvmt</td>
<td>Pattern/ figure copy</td>
<td>M-RMT/ VR-RMT</td>
<td>Picture naming</td>
<td>Verbal fluency</td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>------</td>
<td>------------</td>
<td>------------</td>
<td>------------------------</td>
<td>-----</td>
<td>----------------</td>
<td>----------------------</td>
<td>----------------------</td>
<td>-----------------------</td>
<td>---------------------</td>
<td>----------------</td>
<td>----------------</td>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>Cherrier et al. (2001)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efklides et al. (2002)</td>
<td>↑</td>
<td>↓</td>
<td>↑</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monacelli et al. (2003)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tetewsky and Duffy (1999)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### SUMMARY OF FINDINGS

**SPATIAL ORIENTATION (NATURALISTIC; CONT.):**

- **Route learning task performance positively correlates with M-RMT** ($r = .82, p < .01$); no correlation with VSLT-R or JLO.
- **RBMT route IR and DR positively correlate with WMS: digit span** ($r = .48$); logical memory ($r = .49$); visual reproduction ($r = .49; r = .53$) and verbal PAL ($r = .57; r = .61$). $p < .01$. Route IR negatively correlates with EMQ facial recognition $^*$ ($r = -.39, p < .05$).
- **Route learning task performance positively correlates with JLO** ($r = .67, p < .02$) and M-RMT ($r = .74, p < .004$). No correlation with PAL, figural or facial recognition.
- **Route learning negatively correlates with self-movement discrimination threshold $^*$** (slope = -0.78, $r = -.66, p < .05$). No correlation between route learning and MMSE, shape or motion discrimination.
Table 4.2 Cont. Summary of main findings by type of spatial context memory task

<table>
<thead>
<tr>
<th>AUTHOR (YEAR)</th>
<th>MMSE</th>
<th>Digit span</th>
<th>EMQ/Benton</th>
<th>Word list/story recall</th>
<th>WMS</th>
<th>Object/pattern</th>
<th>Pattern/figure recall</th>
<th>Pattern/figure recall</th>
<th>Pattern/figure recall</th>
<th>Pattern/figure recall</th>
<th>Pattern/figure recall</th>
<th>JLO/Object/Shape/Mvmt</th>
<th>Pattern/figure copy</th>
<th>M-RMT/VR-RMT</th>
<th>Picture naming</th>
<th>Verbal fluency</th>
</tr>
</thead>
<tbody>
<tr>
<td>El Haj and Antoine (2018)</td>
<td>↑</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

SUMMARY OF FINDINGS

Number of details provided in relation to location of autobiographical memory positively correlates with MMSE ($r = .49$, $p < .01$).

Notes:
Results included where relationship between spatial context memory and neuropsychological outcome was investigated by more than one study.

ACE-R = Addenbrooke’s Cognitive Examination-Revised; ADAS = Alzheimer’s Disease Assessment Scale; BVMT-R = Brief Visuospatial Memory Test-revised; CANTAB = Cambridge Neuropsychological Test Automated Battery; CERAD = Consortium to Establish a Registry for Alzheimer’s disease neuropsychological battery; DR = delayed recall; EMQ = Everyday Memory Questionnaire; HMSE = Hindi mental status examination; IR = immediate recall; JLO = Judgement of Line Orientation test; MMSE = Mini Mental State Examination; M-RMT = Money’s Road Map Test; Mvmt = Movement; NNB-E = NIMHANS Neuropsychological Battery for the Elderly; PAL = paired associate learning; RBMT = Rivermead Behavioural Memory Test; ROCF = Rey-Osterrieth Complex Figure; SRFT = Simplified = Rey-Osterrieth Complex Figure; VR = virtual reality; VSLT-R = Visuospatial Learning Test-Revised; WAIS-R = Wechsler Adult Intelligence Scale-Revised; WMS = Wechsler Memory Scale; WMS-R = Wechsler Memory Scale-Revised; WMS-III = Wechsler Memory Scale third edition.

↑ positive correlation observed, ↓ negative correlation observed; both $p < .05$. ✗ no significant correlation. † Lower score = better performance.
4.5 Discussion

The significant heterogeneity of the included studies and the methodological concerns, which are discussed in more detail below, mean that the results of this review should be interpreted with caution. However, the central finding that spatial memory impairment is consistently correlated with global cognitive impairment is in keeping with neuroimaging studies which have shown associations between spatial memory impairment and neuroimaging measures of hippocampal and parietal volume loss and reduced functional integrity in MCI and AD. In addition, previous studies have demonstrated that a variety of measures of spatial memory, including VR allocentric spatial memory testing, CANTAB visual paired associate learning and revised WMS (WMS-R) visual delayed recall, predict conversion of MCI to AD. Given the poor ability of the MMSE to predict conversion from MCI to AD, it is possible that spatial memory tests may be a more relevant early screening tool, and the findings of this review certainly support the need for further research in this area.

4.5.1 Methodological limitations of selected studies

There were significant issues with the methodology of the majority of included studies, and none were rated as high quality, which limits the generalisability of their findings. Over half of the studies (n = 13 studies) had a sample size of less than 20, with most of the study samples (n = 18 studies) rated as not being representative of an AD population in terms of age, gender and race. Criteria for AD diagnosis was unclear in three studies. Medication status and medical/psychiatric history were inconsistently reported, in three and fifteen studies respectively. As described in section 2.4.3, non-AD psychosis is consistently associated with impaired memory for spatial context. Despite this, none of the studies identified in this review screened (or controlled) for presence of psychosis symptoms. Benke et al. (2014) included participants with clinically significant depressive symptomatology and Cherrier et al. (2001) recruited from a veteran hospital, both of which further limit generalisability.

A wide range of tasks of spatial memory were used, with 23 different spatial memory tasks employed across the 22 studies. Of these, only nine had been validated in an AD population. Floor effects were seen with spatial memory tasks in five studies, particularly for measures of delayed visuospatial recall and in one naturalistic route learning task. Neither Efklides et al. (2002) nor Lee et al. (2014) provided spatial memory task results. Of note, Simone and Baylis (1997) found no correlation with MMSE score when participants’ attention was divided, by requiring participants to press and hold one lit button while also remembering the location of another lit button on a keypad. This suggests an important role for attentional
processes in object-location tasks, and potentially explains why Bucks and Willison (1997) found no significant correlation between their novel location-learning task and MMSE score.

There are also concerns with the statistical methodology of many of the studies. Apart from small sample sizes, in which outliers could have led to falsely exaggerated correlations, few studies (n = 6 studies) controlled for potential confounding variables, such as age, gender or overall cognitive function. Only three studies controlled for overall cognitive function, which could be particularly relevant given the finding of consistent correlation between spatial memory tests and MMSE score. Despite correlation analyses not being hypothesis driven in 20 studies, only five studies applied any correction for multiple comparisons in their statistical analyses.

4.5.2 Limitations

The main limitation of this review is the significant heterogeneity of the identified studies, which limits the meaningful interpretation of their findings. This is due in part to the broad nature of the concept and definition of spatial memory, which encompasses various elements of spatial cognition. This review also only included studies in participants with diagnosed AD, in whom impairments and pathology are more extensive, rather than pre-AD patients with biomarkers or retrospective studies of individuals with MCI who later converted to AD in whom cognitive impairment and pathology is more specific.

The nature of neuropsychological testing is such that many of the tasks require attentional and executive processing in addition to spatial memory and only five studies explored or adjusted for this. This review focused on studies which investigated the neuropsychological correlates of spatial memory. It will be important to similarly evaluate studies that have investigated the neurobiological correlates of these tasks, as this could further elucidate the neural underpinnings of memory for spatial context. The exclusion of grey literature means that this review has missed unpublished results, which are more likely to have non-significant findings.

4.5.3 Recommendations for future study

To reduce the heterogeneity identified by this review and allow more reliable comparison of results, I would recommend that future studies of spatial memory use existing, and ideally validated, measures. As discussed above, extending investigation to include pre-AD patients with biomarkers and those with aMCI would be informative, given the potential role of spatial memory tasks as screening tools. This review has identified several variables which should be considered for inclusion as confounders in future studies of spatial memory. Firstly, the finding that MMSE score is correlated with spatial memory suggests that it is important to consider the
roll of overall cognitive function in any analysis exploring measures of spatial memory. Similarly, in light of previous findings described in section 2.4.1, future analyses should consider screening and controlling for presence of psychosis symptoms. Whether or not to include these variables as confounders is fundamentally study-specific and dependent on the underlying research question. As always, care must be taken to ensure that such variables are not in fact intermediaries on the causal pathway, where their inclusion would be inappropriate.

4.5.4 Conclusions

Studies identified by this review explored the relationships between a wide range of tasks of spatial memory and an equally wide range of neuropsychological tasks. This significant heterogeneity makes it difficult to interpret results, with the only replicated finding being that impaired spatial memory is related to global cognitive impairment, as measured by the MMSE. This provides support for the pre-existing hypothesis that spatial memory impairment could be an important early marker of cognitive decline. Understanding spatial memory impairment in AD is important, and I would recommend that future studies use existing and validated testing paradigms to allow direct comparison of results.
5. Context Memory and Metamemory Study: Methods for Analyses of Behavioural Data

5.1 Study overview

The study was designed in two stages. The first was an internal acceptability and pilot study of neuropsychological tasks measuring context memory, associative inference and metamemory ability in AD patients with and without delusions. The aim of the pilot was to adapt neuropsychological tasks based on acceptability and participant performance; to obtain initial behavioural data results; to assess whether neuropsychological task performance was correlated with the presence of delusions; and to finalise the imaging protocol for the second stage of the study.

The second stage involved further neuropsychological testing of participants with tasks adapted following the pilot. This was followed by a structural neuroimaging exploration of the neural correlates of delusions in AD and the relationship between neurobiology and performance on the neuropsychological measures of interest.

5.2 Research questions

The principal research question posed by the full behavioural study was:

• Do AD participants with delusions have greater impairments in context memory (including spatial, temporal and associative memory organisation) and metamemory than those without delusions, when matched for overall cognitive impairment?

The secondary research question for the study was:

• Do participants with misidentification-type delusions have a greater impairment in memory for spatial and temporal context than those with mixed or paranoid-type delusions?

The a priori behavioural hypotheses, based on existing literature reviewed in Chapters 2 - 4, were that:

• Participants with delusions in AD would have greater impairments of memory for spatial and temporal context and associative inference ability than those without delusions
• Participants with delusions in AD would have higher confidence in memory errors than those without delusions
• Participants with misidentification-type delusions would have a greater impairment in memory for spatial and temporal context than those with mixed or paranoid-type delusions

5.3 Recruitment

Participants were recruited from two memory services in Camden and Islington National Health Service (NHS) Foundation Trust, three memory services in Barnet, Enfield and Haringey Mental Health NHS Trust, Mile End Hospital and three memory services in East London NHS Foundation Trust. Participants were also recruited from the Join Dementia Research register, a UK-wide Department of Health initiative, which enables people with dementia to register their interest in participation in research and aims to match them with suitable studies via an online database (www.joindementiaresearch.nihr.ac.uk).

No participants self-referred themselves for inclusion in the study. Memory clinics were contacted prior to recruitment and provided with a leaflet which contained study inclusion and exclusion criteria and participant information sheets (see Appendix 3). Potential participants from memory services were initially identified and approached by members of their clinical team and, if permission was granted, their details were then passed on to me. Potential participants from Join Dementia Research were identified by searching the online database using inclusion and exclusion criteria as described in section 5.3.1.

I provided potential participants with a detailed patient information sheet. This information sheet included a contact telephone number should the potential participant wish to ask me any further questions about the study and a self-addressed envelope to return a signed expression of interest form. A follow-up phone call was made on receipt of this signed form and, for those who remained interested in being involved in the study, a provisional date was arranged for consent, screening and assessment at the participant’s home.

Participants were recruited to two groups: the control group included participants with AD without delusions and the delusion group included participants with AD and delusions. The first five participants (regardless of delusion status) were recruited to an internal pilot to assess task acceptability and to make any adaptations to tasks and their presentation necessary for this patient group. An initial behavioural analysis was completed once five participants had been recruited to the delusion group, in order to inform plans for neuroimaging. All pilot participants remained involved in the full study. If any tasks were adapted, appointments were made to repeat testing with the adapted tasks for those five participants who had completed the tasks prior to this.
Recruitment began on 9th August 2018, with the fifth participant completing pilot testing by 13th November 2018, at which stage task acceptability was reviewed and adaptations made. Five participants had been recruited to the delusion group by 15th August 2019, at which stage an interim behavioural analysis was completed to inform the MRI protocol and neuroimaging hypotheses.

5.3.1 Inclusion and exclusion criteria

Inclusion criteria for the study were:

1) Diagnosis of AD made by clinical services, based on NINCDS-ADRDA criteria. A previous diagnosis of mixed Alzheimer’s and vascular dementia (i.e. made by memory service clinicians) was permitted, if following review of the participant’s general practitioner (GP) medical notes summary and on discussion with the participant their history of memory problems was consistent with AD, they had no history of CVA or TIA or significant vascular findings on previous imaging (minor/mild small vessel disease was permitted), and provided all other inclusion and exclusion criteria were met

2) Capacity to provide informed consent for inclusion

3) Aged 55 years or older

4) Standardised MMSE (sMMSE) score ≥ 22

5) Fluent in English, as translated versions of the tasks were unfortunately not available

Exclusion criteria were:

1) Current or past history of other major psychiatric or neurological illness

2) Any evidence of previous infarct or significant vascular disease on previous imaging (if available)

3) Presence of clinically relevant depressive symptoms, using the short form Geriatric Depression Scale (GDS-15), with a cut-off of > 5

4) Any medical illness that would interfere with completing assessments or impair the safety of the participant

5) Insufficient visual and auditory acuity to complete the tasks

6) Presence of parkinsonian symptoms or other features suggestive of a diagnosis of Lewy Body Dementia (DLB), including fluctuating conscious level, visual hallucinations or neuroimaging evidence, as in McKeith et al. (1996) and Ballard et al. (1997). Participants were excluded if they had a score > 8 on the modified Unified Parkinson's Disease Rating Scale (UPDRS)
7) Documented history of, or current issues with, alcohol or drug abuse
8) Contraindications for MRI, as assessed using exclusion criteria provided and approved by the Wellcome Trust Centre for Neuroimaging (see Appendix 4), including:
   a. Pacemaker
   b. Surgical Aneurysm clips
   c. Implanted metal prosthesis or pump
   d. Possibility of metal fragments in or near eyes or blood vessels

Any queries regarding participant safety for MRI were discussed in advance with the Wellcome Trust Functional Imaging Laboratory safety enquiries team.

5.3.2 Informed consent

Participants who expressed an interest in participating in the study were visited at home, to discuss the study processes and answer any questions that arose. Participants were assessed to determine whether they had capacity to provide informed consent for participation in the study and, if so, verbal and written consent was obtained (see Appendix 5). Contact was then made with their GP to confirm their safety to have an MRI, and their GP medical notes summary was reviewed. Participants also answered MRI safety screening questions, answers to which were confirmed with their carer if available. Provided there were no identified contradictions, participants completed the following assessments, split over two visits.

5.4 Study procedures

5.4.1 Baseline assessments

Assessments at baseline included a full history, which followed a standardised format (based on a template designed by SRe, see Appendix 6) and included demographic details (date of birth, gender, handedness, total years of education), past medical history, medication history and mental state examination. As described in section 5.3.1, participants were screened for eligibility using the sMMSE, GDS-15 and UPDRS.

The sMMSE is a brief, 30 item cognitive test, which provides a measure of global cognitive function, assessed over 11 domains including orientation, registration, attention and calculation, recall, naming, repetition, comprehension, writing and constructional praxis. Higher scores indicate better performance. While it was not developed to identify early-stage dementia, it is now widely used to screen for cognitive decline in older adults. A score of 26 or below is considered suggestive of dementia, with scores of 21 - 26 classified as ‘mild’, 11 - 20 ‘moderate’
and 0 - 10 'severe'. The sMMSE was chosen over the MMSE as it has improved reliability and there is no cost attached to its use.

The GDS-15 is a 15 item depression screening tool, which has been well validated in older adults. Fifteen questions regarding the participant’s mood over the past week are answered either ‘yes’ or ‘no’, and score either one point or zero depending on whether the positive or negative response is suggestive of depression. Higher scores indicate the presence of more depressive symptoms, and a score of greater than five suggests depression.

A modified version of the UPDRS was used to screen for motor symptoms of Parkinson’s disease, which would be indicative of dopaminergic disruption and a likely diagnosis of DLB rather than AD. On this scale, five motor parkinsonian symptoms are rated from zero (no symptoms) to four (severe symptoms), with a total possible score of 20. Symptoms assessed include bradykinesia, rigidity, facial expression, kinetic tremor of the hands and rest tremor. As described in section 5.3.1, participants were also screened for any further symptoms indicative of DLB.

With participants’ consent, two measures were completed by a family member, friend or carer: the neuropsychiatric inventory (NPI), which alongside mental state examination was used to screen for delusions (no cut-off) and the Bayer Activities of Daily Living scale (B-ADL), which was used as a measure of general functioning.

The NPI covers 12 areas of psychopathology: delusions, hallucinations, agitation, depression, anxiety, euphoria, apathy, disinhibition, irritability, aberrant motor behaviour, sleep and appetite. Questions relate to symptoms which have appeared since the onset of AD. Carers are first asked a ‘yes’ or ‘no’ question regarding presence or absence of the above symptoms at any time in the past, with subquestions if the symptom is present. Frequency is scored from zero (‘rarely’) to four (‘very often’) and severity is scored from one (‘mild’) to three (‘severe’), with NPI score calculated by multiplying frequency and severity scores. Higher scores indicate more neuropsychiatric symptoms are present. Scores of one or more in the delusion domain were used to define the presence of delusions, with participants receiving a score of one if they had experienced delusions at any time in the past month.

The B-ADL is a 25 item scale developed to assess deficits in performance of everyday activities. Carers are asked ‘does the person have difficulty...’ followed by an everyday task, with each item scored from one (‘never’) to 10 (‘always’), and options to reply ‘not applicable’ or ‘unknown’. The total score is the mean response, with higher scores corresponding to greater functional impairment.
5.4.2 Neuropsychological assessment

All participants completed a neuropsychological battery, including a baseline measure of global cognitive function, and tasks measuring performance across the domains of attention, executive function and visuospatial ability and perception. These tasks were chosen as individuals with delusions in AD and non-AD disorders demonstrate evidence of significantly different performance to those without delusions across these domains, which could potentially confound results for context memory and false memory tasks.\textsuperscript{84,103,359-362}

5.4.2.1 Global cognitive function

Global cognitive function was assessed using the sMMSE as described above and the Addenbrookes Cognitive Examination, (ACE-III) Version A.\textsuperscript{363} The ACE-III is another widely used cognitive screening tool, which is scored from a possible maximum of 100 points split over five subdomains: attention (maximum 18 points), memory (maximum 26), fluency (maximum 14), language (maximum 26) and visuospatial (maximum 16). Higher scores indicate better performance.
5.4.2.2 Executive function

Executive functioning and verbal reasoning were assessed using a computerised grammatical reasoning task. This is a 90 second test, in which the participant has to answer ‘true’ or ‘false’ to as many verbal reasoning questions relating to the relationship between a square and a circle as possible, see Figure 5.1. After each (true/false) response, I interfaced with the computer on behalf of the participant in order to minimise the impact of computer literacy on task performance. This task, and all following computer-based tasks unless otherwise stated were presented to participants on a Dell XPS 13, 13.3” screen laptop computer with a 1920 x 1080 pixel matrix, with participants sat approximately 40 cm from the screen. The task is scored as the total number of correct answers in 90 seconds, with higher scores indicating better performance.

Figure 5.1 Grammatical reasoning task

Notes:
Condensed example of task design. See text for full details.
5.4.2.3 **Attention**

Sustained attention and response inhibition were assessed using the computerised Sustained Attention to Response Task (SART)\(^\text{365}\). This is an eight minute test, in which the numbers one to nine appear in the centre of the screen, in a variety of sizes, for 500ms, with a 1000ms mask in between each trial, see Figure 5.2. There are 50 practice trials, followed by 270 test trials (containing 30 of each of the nine numbers, with the number three distributed throughout in a pre-fixed quasi-random manner). Participants are required to press the space bar for every number except the number three. Total errors are recorded, including both commission errors (pressing the space bar when three is presented) and omission errors (not pressing the space bar for numbers other than three).

*Figure 5.2 Sustained attention to response task*

![Figure 5.2 Sustained attention to response task](image)

**Notes:**
Condensed example of task design. See text for full details.

5.4.2.4 **Perception and visuospatial function**

Perception and visuospatial function were assessed using subtests from the VOSP\(^\text{366}\). Prior to undertaking these subtests, participants completed a shape detection screening test which involves identifying whether a degraded X is present in 20 speckled patterns, with one mark awarded for each correct yes/no decision. If participants scored less than 16, they were deemed not to have sufficient visuoperceptual abilities to complete the further tests, which were:

1. **Incomplete letters:** Participants identify 20 70% degraded letters, scored as number correct, total 20
2. **Number location:** Participants identify which number out of sixteen displayed in a box, corresponds to the position of a dot in a box below, scored as number correct, total 10
3. **Cube analysis:** Participants count the number of cubes displayed, scored as number correct, total 10
4. **Object decision:** Participants identify which of an array of four shapes corresponds to a silhouette drawing of a real object, scored as number correct, total 20

Each of these subtests has been well validated, in terms of ability to distinguish patients with right and left sided brain lesions and controls,\(^\text{366}\) in healthy older people\(^\text{367}\) and in AD.\(^\text{368}\)
5.4.3 Measures of interest

5.4.3.1 Memory for spatial context – object-location binding
This is a visuospatial short term memory paradigm measuring object-location binding, or recall of ‘what was where’, developed by Pertzov et al. (2012). Access to the task was provided online (https://omt.psy.ox.ac.uk/) by the Cognitive Neurology Lab at the University of Oxford. Participants were presented with two coloured ‘fractal’ shapes on a black screen and asked to remember both the fractal shapes and their locations. These shapes were displayed for 2000ms followed by a 4000ms delay. Two fractal shapes then appeared on the screen (one that was previously seen and one distractor) and the participant was asked to touch the shape that was seen before and identify where they previously saw it, see Figure 5.3. Participants performed three practice trials, followed by 20 test trials. In the pilot, I interfaced with the computer on behalf of the participant, choosing the shape they verbally specified (‘top’ or ‘bottom’) and dragging it to a location they indicated by touching the screen.
Figure 5.3 Object-location binding task

Fractal shapes appear on screen (2000ms)

Delay (4000ms)

Select shape seen at encoding (‘what’)

Drag shape to remembered location (‘where’)

Notes:
Condensed example of task design. See text for full details.

Performance was measured across several domains using calculations described by Pertzov et al. (2012):

- Identification performance (was the correct, i.e. previously seen, fractal shape selected?)
- Gross localisation error (GLE; how far, in Euclidean distance, the participant located the fractal shape from its correct location)
• ‘Nearest neighbour control’ (A correction of localisation performance to account for misbinding errors: localisation distance is calculated from the closest fractal shape to the end location, rather than the correct fractal shape)
• Misbinding errors (number of times fractal shapes were dragged by participants to be closer to the location of the alternative shape seen in the original array, described in more detail below)

GLE was calculated by comparing the x and y coordinates of the correct fractal shape location with the x and y coordinates of the end location of the fractal shape as placed by the participant. Coordinates were first converted from pixels to mm and then to visual degrees, using the formula

\[ \theta = 2 \times \text{atan} \left( \frac{x}{2(z)} \right) \]

where \( x \) is the coordinate (in mm) and \( z \) is the distance between participant and laptop screen (400mm). The difference between the points was then calculated in terms of Euclidean distance using the formula

\[ d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \].

Various options can be used as a cut-off to determine how close the correct fractal shape should be to the location of the alternative shape in order to be counted as a misbinding error. Early studies using this task used the more conservative cut-off of 1.5° of visual angle. However, given the overall poor gross localisation performance on the task by participants with AD, a previous study in familial AD used a more generous cut-off of 4.5° and ongoing studies using the task in the Cognitive Neurology Lab at the University of Oxford were using participants’ mean GLE as the misbinding cut-off. Mean GLE was therefore used as the cut-off for a misbinding error in this study, to account for overall poor localisation performance of participants.

Increased misbinding errors using this task have been demonstrated in patients with focal medial temporal lobe damage and in symptomatic and asymptomatic carriers of the genetic mutations for familial AD. Healthy older adults show reduced object localisation accuracy but intact object-location binding. In carriers of genetic mutations for familial AD object-location binding deficits are associated with reduced hippocampal volume.

5.4.3.2 Memory for temporal context – temporal context confusion
In the temporal context confusion (TCC) task participants were shown a run of 80 meaningful pictures of objects (adapted to be coloured rather than black and white versions of the Snodgrass and Vanderwart images) and verbally indicated any image that they had seen before during the run, see Figure 5.4. After the first run, participants were instructed to forget all the images they had seen, were immediately shown a new run of images and were then required to identify the temporal context of items (i.e. whether the image was shown in the current run or a previous run) by only saying ‘yes’ to repeated images from the current run. This
task was then repeated five minutes after the end of the second run. Participants responded by saying ‘yes’ when they saw a repeat image and I interfaced with the computer to record their response.

**Figure 5.4 Temporal context confusion task**

![Image of temporal context confusion task](image)

Notes:
Condensed example of task design. See text for full details.
Correct hits (i.e. items repeating within the run) are indicated by *.

There were four target (repeating) images in each run. Unknown to participants, the 80 images were split into eight blocks of 10, with the four target images randomly distributed within each block. As such, in each run there were 52 novel images. These images were chosen at random (within the MATLAB code) from a set of 260 at the start of the task. Each of the three runs contained the same 52 images, but with a different set of four repeating – as such, from the second run onwards the participants had seen all of the images before at least once. Each stimulus was onscreen for 2000ms. After 600ms a 100ms warning beep sounded before the next stimulus was presented. Stimuli presentation and response recording were handled by the MATLAB (MATLAB R2018a; https://www.mathworks.com; The Mathworks, Inc., Natick, MA) Cogent toolbox (Cogent; http://www.vislab.ucl.ac.uk/cogent/). In order to confirm their understanding of the task, prior to undertaking the testing, participants completed a brief trial (20 images split into five blocks of four, with three items repeating in each block) during which they saw different geometric shapes and indicated repeated images as described above.

The MATLAB code for this task is included in Appendix 7, written by Professor John King.

Degree of TCC was calculated as the relative increase in false alarms (corrected for overall number of positive responses) in subsequent runs compared to the first run: 

\[
TCC = \frac{FA_x}{H_x} - \frac{FA_1}{H_1}
\]
Where FA₁ and FAₓ are false alarm rates in the first and subsequent runs, and H₁ and Hₓ are hit rates in the first and subsequent runs. There were 28 possible hits and 52 possible false alarms in each run. In addition to calculating degree of TCC, measures of sensitivity and response bias were also calculated. Using Gaussian signal detection theory, discrimination (d’) was calculated as a measure of sensitivity and c as a measure of response bias. To account for zero values, a loglinear transformation was applied to hit and false alarm rates. At d’ = 0, signal is indistinguishable from noise, with increasing values indicating improving performance, with +∞ = perfect performance. c = 0 indicates no response bias, with negative values indicating a bias towards ‘yes’ and positive values indicating a bias towards ‘no’. These measures were only calculated for the first run, as the signal in later runs is confounded by repetition.

As in Macmillan (1993), d’ was calculated as (H = hit rate; FA = false alarm rate):

\[ d’ = \phi^{-1}(H) - \phi^{-1}(FA) \]

and c was calculated as:

\[ c = -\frac{\phi^{-1}(H) - \phi^{-1}(FA)}{2} \]

In non-dementia related amnesia, increased TCC is associated with an increased production of spontaneous confabulation and is linked to dysfunction of the OFC. In a similar task using words rather than pictures, El Haj et al. (2018) found AD participants to show increased TCC compared to controls, and that this correlated with a measure of hallucinations.

5.4.3.3 Associative inference and relational memory
The associative inference task was adapted from a version used previously by my supervisor NB’s research group. In the task, participants were shown a run of 24 pairs of words in two blocks of 12. These were drawn from 24 triples of person, place and object words (randomly combined for each participant from lists of 36 people, places and objects) so that the participant saw only two associations for each, for example place-object and person-object.

Participants were asked to imagine the two elements shown on screen interacting as vividly as possible. They were then shown a single item (either person, place or object) and asked to judge if the element was seen during encoding (i.e. ‘old’) or not (i.e. ‘new’). If the item was presented at encoding (regardless of whether the participant identified it as such), they were then shown six items (one correct and five distractors) and asked to indicate which of these the item was associated with during encoding. For each item this included the seen and unseen, or ‘inferred’, associations. For example, participants would be asked to infer a relationship between place-person in the example above, or between the basement and Paul McCartney in Figure 5.5.
Participants responded verbally and I recorded their response on the computer. Stimuli presentation and response recording were handled by the MATLAB Cogent toolbox. See Appendix 8 for the MATLAB code for this task, written by Dr Daniel Bush. Performance was measured as proportion correct for known and inferred pairs. Measures of old/new recognition, sensitivity and response bias were calculated as described above.

Names of famous individuals used for the ‘person’ element of this task were adapted from those used to run the task with working age adults (including for example, Beyonce and Lady Gaga) to names generated by and demonstrated to be familiar to older adults (age range 40 – 91; including for example, Cliff Richard and Nelson Mandela).

Figure 5.5 Associative inference task

Encoding: Two each of 24 (two blocks of 12) person-place/person-object/place-object from person-place-object triples seen during encoding. Displayed for 9000ms with 1000ms fixation period before each pair.

Test: Six new objects, and all three possible associations (two seen, one unseen) for each triple presented in random order. Initial old/new decision for each, followed by choice of six associations (including one target) for all items seen at encoding. Again, 1000ms cross fixation period (as seen above) prior to each old/new choice.

Notes:
Condensed example of task design. See text for full details.

In healthy young adults, this task has been associated with hippocampal and PFC processing using fMRI and MTL-mPFC theta coupling on magnetoencephalography. Impairments of associative inference are found in patients with schizophrenia. To my knowledge this task has not been used before with AD participants, however similar associative inference tasks have been used in healthy older adults and patients with non-dementia amnesia; both groups show evidence of impairment compared to healthy young subjects.
5.4.3.4 False memory and metamemory

A metamemory (memory confidence) task was used, which followed a computerised version of the DRM task (described briefly in sections 2.3.2 and 3.3). In the initial iteration of the computerised DRM, participants were shown six lists of 12 words (total 72) selected at random from 40 DRM lists of semantic associates,42 displayed sequentially on screen. Words were presented in the centre of the screen for 3000ms, followed by a 1000ms mask. Presentation time, number and length of word lists were based on previous studies using the DRM with AD participants.145, 150

List order was random, but within each list words were presented in descending order of association with the critical ‘lure’ word, which is not shown (for example, thread, eye, sewing, sharp, haystack; critical lure word ‘needle’). Participants were instructed to read the words aloud and were aware that they would subsequently have their recall tested.

Participants then completed a recognition task in which 42 words appeared on screen sequentially. This included six lure words, the three words most closely associated with the lure from each of the six lists and 18 unrelated words drawn from a pool of 160 words matched for average concreteness and frequency, with no evident semantic association to the DRM word lists.385

Each word remained on screen until the participant verbally responded either ‘old’ or ‘new’ and I recorded this answer. After each old/new decision, the participant was asked to complete a binary confidence rating: whether they were ‘sure’ or ‘unsure’ in their decision. Stimuli presentation and response recording were again handled by the MATLAB Cogent Toolbox. See Appendix 9 for the MATLAB code for this task, written by Dr James Bisby. Task design is summarised in Figure 5.6.
**Figure 5.6 DRM/metamemory task**

| Snow | Warm | Winter | Ice |

Encoding: Six lists of 12 words in descending order of association to the critical lure. Each word is on screen for 3000ms. Participant reads each word aloud.

| Winter | Star | Cold | Tiger |

LURE

* Test: Six lure words, 18 previously seen words, 18 unrelated words not previously seen. For each word participant is asked ‘did you read this word aloud?’ and ‘how sure are you?’ with a choice of ‘sure’ or ‘unsure’.

Notes:
Condensed example of task design. See text for full details.
Correct recognition marked by * above. Lure word is ‘cold’ in this example.

Task performance was measured by ‘true recognition’ (number of hits), false recognition of related lures (‘false memories’) and false recognition of unrelated words. Measures of sensitivity ($d'$) and bias (c) were calculated as described previously. Metamemory performance was measured as proportion of high confidence hits and false recognition.

As described in sections 2.3.2 and 3.3, use of the DRM in AD has demonstrated a greater susceptibility for false memory generation due to impaired recollection (item-specific memory),144, 146-148 a proportionately lesser impairment and therefore overreliance on familiarity (gist memory)145, 149 and impaired memory verification and monitoring.150 When people with schizophrenia completed a similar DRM/confidence rating paradigm they demonstrated increased confidence in false recognition compared to controls.240

5.4.4 Task adaptation

After the first five participants completed the tasks described above, they were asked to provide feedback on acceptability of the task in terms of level of difficulty, tolerability and length of time taken. If a family member or carer was present their opinion was also sought. Where possible task difficulty was then adjusted using a strategy previously used by my research group,386 by
making modifications to number of images/words presented, number of trials and length of time stimuli were shown on screen, see Chapter 6 for details.

5.5 Statistical methodology

Data analysis was carried out in Statistical Package for the Social Sciences (SPSS v27.0; www.spss.com). For all analyses, results were considered significant at a p value of < .05.

5.5.1 Group size

Prior to conduct of the study, necessary group size was estimated based on previous studies. Previous studies using these tasks in similar participant groups found significant differences with group sizes of less than 30. For example, one study using the spatial context memory task in patients with and without familial AD found significant results with a group size of eight, with 28 controls, and a study of metamemory comparing 16 participants with AD and six with frontotemporal dementia was able to detect significant differences between the groups. One similar study using fMRI to compare differences in activity during a spatial navigation task in those with and without MCI, which identified significant differences in functional activity of hippocampus and dIPFC with group numbers of 10 and eight (controls and patients respectively). The majority of previous studies comparing those with and without delusions in AD have used FDG-PET or SPECT.

No study was identified comparing context memory or metamemory between those with AD with and without psychosis symptoms. A power calculation based on a study examining TCC in AD without psychosis symptoms was completed using G*Power, and gave a sample size of ≥ 17 per group (p = .05, power = 0.8) to detect a difference between groups of one standard deviation (SD) from the mean. The planned sample size for this study was therefore set at 25 per group.

5.5.2 Internal pilot

As described in section 5.3, behavioural data were analysed for two pilot groups: the first five individuals recruited, and all participants recruited at the point at which five individuals had been recruited to the delusion group.

For the primary pilot, mean group performance on the tasks was calculated and participant feedback collated narratively.

For the secondary pilot, demographic details of the two groups were summarised. Task performance data were visualised as scatter and box plots and then assessed for normality using Q-Q plots and the Shapiro-Wilk test. Differences between the groups were compared using
independent-samples t-tests or Mann Whitney tests if assumptions of parametric data were violated.

5.5.3 **Whole group behavioural analyses**

Demographic and behavioural data were initially visualised as scatter plots and histograms and then assessed for normality using Q-Q plots and the Shapiro-Wilk test for non-normality. Demographic characteristics and task performance were compared between the control and delusion groups using independent-samples t-tests or Mann Whitney tests if assumptions of parametric data were violated, and chi-squared tests for categorical variables.

As the primary outcome measure was binary (i.e. does a participant have delusions or not), significant findings were further explored using binary logistic regression, with task performance measures on which there were significant differences between the groups included as predictors of control or delusion group in individual models. Potential confounding variables were chosen prior to data analysis. These were: age, gender, years of education, sMMSE score, cholinesterase inhibitor prescription and category fluency on ACE-III. False memories are known to increase with age, and older adults also perform more poorly on recognition tasks. Gender can influence response bias, with men being less likely to endorse items as previously seen. The same is true for education: those with a lower level of education have a more strict response bias. Level of education is also widely recognised to play a role in an individual’s degree of ‘cognitive reserve’ and to influence cognitive function in AD as well as the relationship between cognitive function and AD pathology. The sMMSE was used as a measure of severity of AD given that this would impact task performance (and regional brain volume in later neuroimaging analyses, see Chapters 10 and 12). Cholinesterase inhibitor prescription was included as this may influence both task performance (via improvements in executive function and visual attention) and degree of brain atrophy in AD participants. Category fluency for animals was chosen as a proxy of executive function (as well as language ability) as this was also used in the larger ADNI dataset (see Chapter 8) and was the executive function task completed by the greatest number of participants in this cohort. A measure of executive function was included as a covariate as poor performance on measures of executive function has been linked to an increase in false memory, with increased false recognition also found following frontal lobe injury.

Where binary logistic regression models were significant, model sensitivity, specificity and positive and negative predictive values were calculated. Further model diagnostics were run for models with significant results. Variance inflation factor (VIF) scores were calculated as a measure of collinearity in the data. Significant multi-collinearity is considered present at a VIF >
and a more conservative threshold of 2.5 was used for the purpose of the following analyses. The logistic regression model assumption of linearity between continuous variables and the logit transformation (log-odds) of the dependent variable was tested using the Box-Tidwell test, in which interaction terms between the continuous variables and their log are entered into each model. Non-significant findings indicate a linear relationship between the variable and the logit transformation of the outcome, and that this assumption is met.\textsuperscript{403} Model fit was further assessed by reviewing outcomes for the Hosmer and Lemeshow goodness of fit test, a chi-squared test that compares observed values to expected values using the model, meaning a significant result ($p < .05$) indicates a poor model fit.\textsuperscript{404}

5.6 Ethical approval

The study was reviewed and received approval from the University College London and University College London Hospital (UCL/UCLH) Joint Research Office (project ID 18/0038); Westminster Research Ethics Committee (REC; IRAS number 240572; REC reference 18/LO/0709) and the Health Research Authority.
6. Context Memory and Metamemory Study: Results of Internal Pilot

6.1 Recruitment

Recruitment to the study began on 9th August 2018 from Camden and Islington memory services. The recruitment rate was slower than expected, and as a result two substantial amendments were made to the study protocol (dated 3rd September 2018 and 9th January 2019) and received further ethical approval. The first amendment extended recruitment to include the three memory services in Barnet, Enfield and Haringey Mental Health NHS Trust and the ‘Join Dementia Research’ register. The second added Mile End Hospital and the three memory services in East London NHS Foundation Trust.

The fifth participant had completed pilot behavioural testing by 13th November 2018, at which stage evaluation of task difficulty and acceptability was undertaken. Preliminary behavioural results were analysed and a neuroimaging plan finalised when five participants with delusions had been recruited (15th August 2019). This was an internal pilot, with all participants remaining enrolled for the neuroimaging portion of the study. Participants who completed pilot versions of tasks that underwent any significant adaptations were invited to repeat the adapted tasks at a later date.

6.2 Review and adaptation of neuropsychological tasks

6.2.1 Demographic and behavioural data

The first five pilot participants were all from the control group (AD without delusions). Baseline demographic details of this group were: mean age 83.6 ± 8.2 years; two (40.0%) female and three (60.0%) male; all right handed; all white; mean years of education 11.8 ± 2.0; mean years since diagnosis 1.8 ± 1.1; mean sMMSE score 24.8 ± 2.6. All were prescribed cholinesterase inhibitors.

On the object-location binding task, participants had above-chance identification performance (mean hit rate of 0.69 ± 0.13), with a mean GLE of 9.2 ± 4.0° of visual angle. This is similar to findings in those with familial AD. Given the degree of localisation error it is perhaps unsurprising that none of the fractal shapes fell within the conservative 1.5° cut-off for misbinding. However, using each participant’s mean GLE, misbinding errors accounted for a small proportion (0.13 ± 0.11) of localisation performance.

Data from the TCC task indicated that participants were able to identify repeating items with high accuracy (mean run one hit rate 0.94 ± 0.09; mean false recognition rate 0.03 ± 0.05).
On the DRM/metamemory task, mean hit, false recognition and false memory rates (0.66 ± 0.30; 0.24 ± 0.21; 0.63 ± 0.18 respectively) were comparable to performance of participants with mild AD in other studies using this task. With the binary choice of ‘sure’ or ‘unsure’ for confidence ratings, one participant rated all items ‘sure’.

6.2.2 Associative inference and relational memory

The associative inference task was added to the study through a later amendment, as researchers within my supervisor NB’s group had recently found differences between participants with schizophrenia and healthy controls using this task; those with schizophrenia had increased false recognition and impaired memory for both direct and inferred associations. As this was added later, a different group of participants completed the initial pilot of this task (n = 8; four with delusions, four without). Baseline demographic details of this group were: mean age 83.9 ± 3.5 years; three (37.5%) female and five (62.5%) male; five (62.5%) right handed; six (75.0%) white and two (25.0%) black; mean years of education 14.2 ± 4.3; mean years since diagnosis 2.1 ± 1.2; mean sMMSE score of 25.1 ± 2.3. All were prescribed cholinesterase inhibitors.

The accuracy of old/new judgements was above chance (mean hit rate 0.67 ± 0.21). Associative inference performance was around the chance hit rate of 0.17 for both seen associations (mean hit rate 0.20 ± 0.05) and unseen ‘inferred’ associations (mean hit rate 0.18 ± 0.09).

6.2.3 Participant feedback

Participants tolerated the object-location binding, TCC and DRM/metamemory tasks well. The only comment provided on feedback sheets for these tasks (the same comment was made by two participants) was a request for fractal shapes in the object-location binding task to remain on screen for longer.

Participants struggled to tolerate the associative inference task. Four participants (50.0%) expressed their frustration with the task while completing it and, apart from one participant, all reported that they found the task too difficult and were guessing the answers. I observed that participants were finding the task challenging and appeared to be either guessing or becoming confused by expected associations, for example assuming that a sportsperson would be associated with a gym, or an actor with the cinema. One carer reported that the participant had later said that they had not enjoyed the task.
6.2.4 Task adaptation

6.2.4.5 Memory for spatial context – object-location binding
Participants in the internal pilot completed this task using a laptop screen with the researcher interfacing with the mouse to drag fractal shapes to a location specified by the participant by touching the screen. While interfacing with the laptop for the other tasks was deemed appropriate as response time was not being recorded and tasks did not include a spatial component, after the initial stage of the pilot it was decided that participants should ideally complete the object-location binding task themselves due to the precision measured by the task. As such, the remaining participants completed this task themselves using a tablet device (iPad; 6th generation).

6.2.4.6 Memory for temporal context – temporal context confusion
Minimal adaptations were made to this task. The size of the images was increased, and made uniform, in response to my observation that inconsistency in size might introduce unnecessary bias. The instructions displayed on screen were simplified – the initial version of the task used the language from the original Schnider and Ptak (1999) task ‘You will be presented with a series of pictures one at a time; ‘*Only respond to items repeated during this very run*’. This was found to be confusing for participants, and therefore changed to ‘Say YES when you see a repeat’, ‘**Try to forget the objects you saw in the previous test**’.

6.2.4.7 Associative inference and relational memory
Given the poor participant performance and levels of frustration experienced as a result of this task, it was not taken forward for use in the full participant cohort.

6.2.4.8 False memory and metamemory
On inspection of the lists presented by the initial MATLAB code it transpired that there had been an error in the coding of the random number generator which was rectified. Adaptations made for task use in the full participant cohort were:

- In order to increase power, list length and number were changed such that in the adapted task participants saw 15 lists of the four closest associated words at encoding and two words from each of the 15 lists, the 15 lure words and 30 matched unrelated words at test (as in Chadwick et al. (2016))
- Direction of presentation of lists was changed from decreasing to increasing association to the lure, following a finding by Evrard et al. (2018) that this presentation order increased false recognition of lure words in AD
- Confidence rating was changed from binary (sure/unsure) to a three-point Likert scale (complete guess – fairly confident – 100% confident) to allow for variation in confidence responses
• Wording during presentation of the task was clarified and simplified from ‘which of these words have you seen before?’, ‘did you see this word before?’, ‘Sure/Unsure that you did (not) see this word?’, to: ‘which of these words did you read aloud?’, ‘did you read this word aloud?’, ‘how sure are you?’

6.3 Preliminary behavioural results

6.3.1 Demographic data

The preliminary behavioural analysis included a total of 17 participants (five with delusions, 12 without). The mean age of the control group was 83.8 ± 7.1 years and the mean age of the group with delusions was 87.5 ± 3.7 years. Five (41.7%) of the control group were female (seven, 58.3% male) and three (60.0%) of the delusion group were female (two, 40.0% male). The majority (n = 14; 82.4%) were white, 11 (91.7%) of the control group and three (60.0%) of the delusion group. The majority (n = 12) were prescribed cholinesterase inhibitors, nine (75.0%) of the control group and three (60.0%) of the delusion group. The mean sMMSE score of the control group was 25.0 ± 2.6 and of the delusion group was 25.0 ± 2.0.

Perhaps unsurprisingly, given the small sample sizes, data were not normally distributed. While the small numbers also meant that results were interpreted with caution, Mann-Whitney tests were conducted to compare behavioural data between groups to help to identify any trends towards significance.
6.3.2 Memory for spatial context – object-location binding

Five participants with delusions and nine control participants used the iPad version of the object-location binding task. Participants in both groups performed similarly in relation to their ability to correctly identify the previously seen fractal shape. Participants with delusions performed significantly better in terms of their ability to correctly locate the fractal shape. They also had fewer misbinding errors, but this did not achieve statistical significance. See summary results in Table 6.1.

**Table 6.1 Preliminary object-location binding task results**

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 9)</th>
<th>DELUSIONS (n = 5)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HIT RATE</strong></td>
<td>0.71 ± 0.01</td>
<td>0.71 ± 0.08</td>
<td>U = 21.5, p = .898</td>
</tr>
<tr>
<td><strong>MEAN GLE</strong></td>
<td>7.0 ± 1.9°</td>
<td>4.8 ± 1.2°</td>
<td>U = 7.0, p = .042</td>
</tr>
<tr>
<td><strong>MISBINDING ERROR RATE</strong></td>
<td>0.13 ± 0.14</td>
<td>0.03 ± 0.04</td>
<td>U = 13.0, p = .240</td>
</tr>
</tbody>
</table>

Notes:
- Values expressed as mean ± SD.
- GLE = gross localisation error, in degrees of visual angle.
- Means compared by Mann Whitney U test, with significant p values shown in bold.
6.3.3 Memory for temporal context – temporal context confusion

There were no between group differences in number of hits on the temporal context confusion task, or in terms of the measure of temporal context confusion itself (i.e. increase in false recognition in later runs relative to the first run). Participants with delusions had increased false recognition across all three runs, although not at a level that reached statistical significance in this small sample, see Table 6.2.

Table 6.2 Preliminary temporal context confusion task results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 12)</th>
<th>DELUSIONS (n = 5)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HITS RUN 1</strong></td>
<td>24.9 ± 3.4</td>
<td>24.4 ± 5.3</td>
<td>U = 29.5, p = 1.000</td>
</tr>
<tr>
<td><strong>HITS RUN 2</strong></td>
<td>23.5 ± 3.4</td>
<td>23.2 ± 4.8</td>
<td>U = 27.5, p = .799</td>
</tr>
<tr>
<td><strong>HITS RUN 3</strong></td>
<td>24.6 ± 4.8</td>
<td>24.1 ± 5.6</td>
<td>U = 29.5, p = 1.000</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 1</strong></td>
<td>1.6 ± 2.1</td>
<td>5.7 ± 9.9</td>
<td>U = 20.0, p = .328</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 2</strong></td>
<td>8.3 ± 3.1</td>
<td>10.4 ± 12.0</td>
<td>U = 26.5, p = .721</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 3</strong></td>
<td>8.8 ± 4.2</td>
<td>13.0 ± 15.1</td>
<td>U = 28.5, p = .879</td>
</tr>
<tr>
<td><strong>TCC RUN 2</strong></td>
<td>0.24 ± 0.17</td>
<td>0.18 ± 0.15</td>
<td>U = 29.5, p = 1.000</td>
</tr>
<tr>
<td><strong>TCC RUN 3</strong></td>
<td>0.26 ± 0.22</td>
<td>0.25 ± 0.18</td>
<td>U = 29.0, p = 1.000</td>
</tr>
</tbody>
</table>

**Notes:**
Values expressed as mean ± SD.
TCC = temporal context confusion, calculated as $TCC = \frac{FA_{x}}{HR_{x}} - \frac{FA_{1}}{HR_{1}}$ where HR = hit rate, FA = false alarm rate in run x and run 1 respectively.
Maximum number hits per run is 28, maximum number false recognition per run is 52.
Means compared by Mann Whitney U test.
6.3.4 False memory and metamemory

Two trends were identified that did not achieve statistical significance: participants with delusions were more likely to endorse all items as previously seen and as such had higher hit, false recognition and false memory rates than control participants. They were also more highly confident in their responses, reporting more high confidence hits, false recognition and false memories than control participants, with the difference in proportion of high confidence hits reaching statistical significance ($p = .029$). These results are summarised in Table 6.3, and displayed as scatter plots in Figure 6.1.

**Table 6.3 Preliminary DRM/metamemory task results**

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 9)</th>
<th>DELUSIONS (n = 5)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HITS</strong></td>
<td>15.0 ± 10.5</td>
<td>20.1 ± 9.9</td>
<td>$U = 15.5, p = .364$</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION</strong></td>
<td>5.1 ± 5.1</td>
<td>11.7 ± 13.5</td>
<td>$U = 20.0, p = .797$</td>
</tr>
<tr>
<td><strong>FALSE MEMORIES</strong></td>
<td>4.4 ± 4.5</td>
<td>7.8 ± 7.5</td>
<td>$U = 18.0, p = .606$</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE HITS</strong></td>
<td>0.35 ± 0.31</td>
<td>0.77 ± 0.26</td>
<td>$U = 6.0, p = .029$</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE FR</strong></td>
<td>0.22 ± 0.29</td>
<td>0.57 ± 0.50</td>
<td>$U = 6.0, p = .279$</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE FM</strong></td>
<td>0.37 ± 0.36</td>
<td>0.83 ± 0.24</td>
<td>$U = 3.5, p = .117$</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. High confidence values as proportion of total hits, false recognition or false memories respectively.
DRM = Deese-Roediger-McDermott paradigm; FM = false memories; FR = false recognition.
Maximum number hits per run is 30, maximum number false recognition per run is 30, maximum number false memories per run is 15.
Means compared by Mann Whitney U test, significant p values shown in bold.
Figure 6.1 Scatter plots showing performance on DRM and metamemory tasks between groups
Figure 6.1 Cont. Scatter plots showing performance on DRM and metamemory tasks between groups
6.4  **Impact of pilot on plans for full study**

6.4.1  **Summary of behavioural data**

The trends that emerged from this data indicated differences in localisation performance on the object-location binding task, false recognition rates on both temporal context and DRM tasks, and in metamemory performance between the control and delusion groups. Although not significant, given that the false recognition finding was present across two tasks in a small sample, this finding informed planning for the second stage of the study.

6.4.2  **Confirmed neuroimaging methodology**

Consideration was given to the relative benefits of further exploring this finding using structural or functional MRI. Both the TCC and DRM/metamemory tasks were designed in a format to allow potential transition to use as a fMRI paradigm. However, neither had been previously used as fMRI paradigms in AD and, in order to optimise fMRI data, significant task development would be required prior to their use.

Given the relative complexity of the tasks, in terms of the likely involvement of overlapping functional networks (for example, attention, working memory and visuoperceptual networks alongside temporal context and metamemory judgements) it was decided that both tasks would need to be further refined through use in healthy populations to first understand areas of brain activation during these tasks in individuals without AD. Neither had participants completed these tasks independently, as I had recorded their verbal responses. The tasks would therefore need to be further piloted with the MRI-safe button box and in the scanner environment and adapted to optimise them for use in the scanner. In addition, the tasks would ideally be adjusted to match level of cognitive difficult between groups.

The extensive task piloting and development that would be required to maximise the quality of fMRI data was beyond the remit of the current study. It was therefore decided that structural imaging was a reasonable first step to explore correlations between memory false alarms, metamemory performance and patterns of neurodegeneration as indicated by differences in brain volume, combining hypothesis driven ROI and exploratory VBM approaches. See Chapters 10 and 12 for further details.

6.4.3  **Sample size calculation**

After neuroimaging methodology was confirmed a further, neuroimaging specific, power calculation was undertaken using G*Power. Based on my previous study in which a ROI analysis was completed in the AddNeuroMed cohort (described in more detail in section 2.4.1),

105
using difference in PHG volume between those with and without psychosis, this gave a sample size of ≥ 22 per group (power = 0.8) to detect a significant difference between groups (p < .05). This confirmed the minimum sample size suggested based on detecting difference in TCC task results (see section 5.5.1).
7. **Context Memory and Metamemory Study: Results of Analyses of Behavioural Data**

### 7.1 Recruitment

Participant recruitment began on 9\textsuperscript{th} August 2018, from identified memory services and Join Dementia Research as stated in Chapter 5.

#### 7.1.1 Impact of COVID-19

All face-to-face contact with participants, including recruitment, behavioural testing and scanning was stopped on 16\textsuperscript{th} March 2020 due to the coronavirus pandemic (COVID-19). Recruitment to the study was still in process at this point. Twenty-seven participants had undergone behavioural testing (10 with delusions and 17 without). Due to the uncertain nature of the situation, and the risks to participants of potential exposure to the virus, it was decided to end recruitment to the study on 20\textsuperscript{th} March 2020 and to analyse available data, while acknowledging the significant limitations introduced by the inability to recruit the planned sample size.

Participant screening and recruitment is summarised in Figure 7.1.
Demographic details

Demographic details and baseline screening assessments are summarised in Table 7.1. There were no significant differences between the delusion and control groups in terms of any of the demographic details (age, gender, race, handedness, years of education or numbers prescribed cognitive enhancers or antidepressants). No participants were prescribed any other psychotropic medication, including antipsychotics, anxiolytics or sedating hypnotics.

The groups did not differ in terms of baseline sMMSE, NPI or B-ADL scores or screening measures (GDS-15 and UPDRS). Scores on GDS-15 and UPDRS were well below cut-offs for exclusion from the study.
## Table 7.1 Demographic details and screening results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 17)</th>
<th>DELUSIONS (n = 10)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE (years)</strong></td>
<td>82.5 ± 8.0; 65 - 95</td>
<td>84.4 ± 4.7; 76 - 92</td>
<td>$t_{25} = -.693, p = .494$</td>
</tr>
<tr>
<td><strong>GENDER (female)</strong></td>
<td>4 (23.5)</td>
<td>6 (60.0)</td>
<td>$^*p = .101$</td>
</tr>
<tr>
<td><strong>RACE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>14 (82.4)</td>
<td>8 (80.0)</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>1 (5.9)</td>
<td>1 (10.0)</td>
<td>$\chi^2 (3, n = 17) = .881, p = .830$</td>
</tr>
<tr>
<td>Asian</td>
<td>1 (5.9)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Mixed race</td>
<td>1 (5.9)</td>
<td>1 (10.0)</td>
<td></td>
</tr>
<tr>
<td><strong>RIGHT HANDEDNESS</strong></td>
<td>14 (82.4)</td>
<td>7 (70.0)</td>
<td>$\chi^2 (2, n = 17) = .556, p = .757$</td>
</tr>
<tr>
<td><strong>EDUCATION (years)</strong></td>
<td>14.4 ± 4.4; 6 - 23</td>
<td>13.4 ± 4.7; 5 - 21</td>
<td>$t_{25} = .527, p = .603$</td>
</tr>
<tr>
<td><strong>MEDICATION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChEI/Memantine</td>
<td>15 (88.2)</td>
<td>7 (70.0)</td>
<td>$^*p = .326$</td>
</tr>
<tr>
<td>Antidepressant</td>
<td>2 (11.8)</td>
<td>1 (10.0)</td>
<td>$^*p = 1.000$</td>
</tr>
<tr>
<td><strong>sMMSE</strong></td>
<td>25.8 ± 2.6; 22 - 29</td>
<td>25.1 ± 2.1; 22 - 28</td>
<td>$U = 67.5, p = .386$</td>
</tr>
<tr>
<td><strong>NPI</strong></td>
<td>6.7 ± 9.5; 0 - 32</td>
<td>13.2 ± 12.2; 1 - 34</td>
<td>$U = 49.0, p = .074$</td>
</tr>
<tr>
<td><strong>B-ADL</strong></td>
<td>3.85 ± 2.13; 1.00 - 7.28</td>
<td>4.71 ± 1.65; 1.08 - 7.23</td>
<td>$t_{25} = -1.100, p = .282$</td>
</tr>
<tr>
<td><strong>GDS-15</strong></td>
<td>2.4 ± 1.8; 0 - 5</td>
<td>2.3 ± 1.3; 0 - 4</td>
<td>$U = 85.0, p = .1000$</td>
</tr>
<tr>
<td><strong>UPDRS</strong></td>
<td>1.8 ± 1.6; 0 - 5</td>
<td>2.7 ± 1.4; 0 - 5</td>
<td>$t_{25} = -1.526, p = .140$</td>
</tr>
</tbody>
</table>

Notes:
Values are expressed as mean ± SD; range or n (%). All measures at baseline.
B-ADL = the Bayer Activities of Daily Living scale; ChEI = cholinesterase inhibitor; GDS-15 = short form Geriatric Depression Scale; NPI = Neuropsychiatric Inventory; sMMSE = standardised MMSE; UPDRS = modified Unified Parkinson’s Disease Rating Scale.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data. Categorical comparisons by chi-squared unless otherwise specified.
$^*p$ value from Fisher’s exact test.
The mean NPI score for the delusion group was 13.2 ± 12.2, with all participants currently experiencing delusional symptoms. Two of the 10 participants in the delusion group were also experiencing hallucinations. Overall, psychosis symptoms were relatively mild: the mean delusion frequency x severity score was 1.7 ± 1.1 (n = 10), and the mean hallucination frequency x severity score was 1.0 ± 0.0 (n = 2). Seven of the 10 participants had paranoid-type psychosis symptoms, one had misidentification-type psychosis symptoms and two had mixed-type psychosis symptoms.

Breakdown of psychosis symptoms by subtype is shown in Table 7.2.

### Table 7.2 Breakdown and classification of psychosis symptoms at baseline (n = 10)

<table>
<thead>
<tr>
<th>CONTENT</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARANOID</td>
<td></td>
</tr>
<tr>
<td>Delusions</td>
<td>In danger/others are planning to hurt him/her</td>
</tr>
<tr>
<td></td>
<td>Others are stealing from him/her</td>
</tr>
<tr>
<td></td>
<td>Spouse is having an affair</td>
</tr>
<tr>
<td></td>
<td>Family members plan to abandon him/her</td>
</tr>
<tr>
<td>MISIDENTIFICATION</td>
<td></td>
</tr>
<tr>
<td>Delusions</td>
<td>Unwelcome guests are staying in his/her house</td>
</tr>
<tr>
<td></td>
<td>His/her spouse or others are not who they claim to be</td>
</tr>
<tr>
<td></td>
<td>His/her house is not his/her own</td>
</tr>
<tr>
<td></td>
<td>Television/magazine figures are present in his/her home</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>He/she can hear voices</td>
</tr>
<tr>
<td></td>
<td>Talks to people who are not there</td>
</tr>
<tr>
<td></td>
<td>Seeing things not seen by others</td>
</tr>
<tr>
<td>OTHER</td>
<td></td>
</tr>
<tr>
<td>Delusions</td>
<td>Any other unusual beliefs</td>
</tr>
<tr>
<td></td>
<td>- Family members having behaved cruelly towards him/her on specific past occasion</td>
</tr>
<tr>
<td></td>
<td>- People monitoring him/her</td>
</tr>
<tr>
<td></td>
<td>- Having won a prize</td>
</tr>
<tr>
<td>Hallucinations</td>
<td>Smells odours not smelled by others</td>
</tr>
<tr>
<td></td>
<td>Feels things on his/her skin</td>
</tr>
<tr>
<td></td>
<td>Tastes without known cause</td>
</tr>
<tr>
<td></td>
<td>Any other unusual sensory experiences</td>
</tr>
</tbody>
</table>
7.3 Neuropsychological assessments

7.3.1 Global cognitive function

As described previously, baseline sMMSE score did not differ between the groups (25.8 ± 2.6 for the control group and 25.1 ± 2.1 for the delusion group; \( U = 67.5, p = .386 \)). Similarly, groups did not differ on baseline ACE-III score (78.5 ± 15.0 for the control group and 71.70 ± 11.8 for the delusion group; \( t_{25} = 1.216, p = .235 \)). Group scores on ACE-III subdomains are shown in Table 7.3. Again, no differences were found between the control and delusion groups.

7.3.2 Executive function

See section 5.4.2.2 for full task description.

Participants in the control and delusion groups performed similarly on the computerised grammatical reasoning task, chosen as a measure of executive function (mean total correct responses in 90 seconds was 9.4 ± 4.0 for the control group and 7.3 ± 2.7 for the delusion group; \( t_{25} = 1.476, p = .152 \)).

<table>
<thead>
<tr>
<th>Table 7.3 Addenbrooke's Cognitive Examination-III subdomain scores</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>ATTENTION (18)</td>
</tr>
<tr>
<td>MEMORY (26)</td>
</tr>
<tr>
<td>FLUENCY (14)</td>
</tr>
<tr>
<td>LANGUAGE (26)</td>
</tr>
<tr>
<td>VISUOSPATIAL (16)</td>
</tr>
<tr>
<td>TOTAL (100)</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. All measures at baseline.
Maximum scores for each category provided in brackets.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.
7.3.3 Attention

See section 5.4.2.3 for full task description.

Participants also performed at similar levels on the SART. Number of total errors (including errors of both commission and omission, over 270 trials) was $24.6 \pm 18.7$ for the control group and $29.3 \pm 20.0$ for the delusion group ($t_{25} = -0.616, p = .543$).

7.3.4 Perception and visuospatial function

See section 5.4.2.4 for full task description.

All participants passed the shape detection screening test and so were considered to have sufficient visuoperceptual abilities to proceed with the full VOSP. No differences were observed between groups in performance across any of the VOSP subdomains, see Table 7.4.

**Table 7.4 Visual Object and Space Perception battery subdomain scores**

<table>
<thead>
<tr>
<th>Subdomain</th>
<th>CONTROL (n = 17)</th>
<th>DELUSIONS (n = 10)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>INCOMPLETE LETTERS (20)</td>
<td>18.4 ± 2.6</td>
<td>17.8 ± 1.7</td>
<td>$U = 53.5, p = .115$</td>
</tr>
<tr>
<td>NUMBER LOCATION (10)</td>
<td>8.4 ± 2.4</td>
<td>7.0 ± 2.5</td>
<td>$U = 53.0, p = .115$</td>
</tr>
<tr>
<td>CUBE ANALYSIS (10)</td>
<td>7.8 ± 2.5</td>
<td>8.8 ± 1.8</td>
<td>$U = 63.5, p = .286$</td>
</tr>
<tr>
<td>OBJECT DECISION (20)</td>
<td>16.1 ± 2.7</td>
<td>15.6 ± 2.3</td>
<td>$U = 68.0, p = .414$</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. All measures at baseline.
Maximum scores for each category provided in brackets.
Means compared by Mann Whitney U test for non-parametric data.
7.4 Measures of interest

7.4.1 Memory for spatial context – object-location binding

See section 5.4.3.1 for full task description.

Fourteen control participants and 10 participants with delusions completed the iPad version of the object-location binding task. In this larger group, in comparison to the preliminary analysis, there were no significant differences in performance in terms of ability to correctly identify which fractal shape participants had previously seen (hit rate), ability to locate the fractal shape in its previously seen location (mean GLE) or in rate of misbinding errors (locating the fractal shape closer to where the distractor shape had been located). There was also no difference when localisation performance was corrected for misbinding errors (nearest neighbour control). See Table 7.5.

Table 7.5 Object-location binding task results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 14)</th>
<th>DELUSIONS (n = 10)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>HIT RATE</td>
<td>0.74 ± 0.11</td>
<td>0.72 ± 0.11</td>
<td>t\textsubscript{22} = .533, p = .599</td>
</tr>
<tr>
<td>MEAN GLE</td>
<td>7.5 ± 3.4°</td>
<td>6.6 ± 2.6°</td>
<td>t\textsubscript{22} = .717, p = .481</td>
</tr>
<tr>
<td>NNC</td>
<td>6.5 ± 2.4°</td>
<td>5.7 ± 1.8°</td>
<td>t\textsubscript{22} = .905, p = .375</td>
</tr>
<tr>
<td>MISBINDING ERROR RATE</td>
<td>0.21 ± 0.25</td>
<td>0.14 ± 0.20</td>
<td>U = 58.5, p = .508</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. All measures at baseline.
Mean gross localisation error and nearest neighbour control given in degrees of visual angle.
GLE = gross localisation error; NNC = nearest neighbour control.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.

7.4.2 Memory for temporal context – temporal context confusion

See section 5.4.3.2 for full task description.

All 27 participants completed the TCC task. Participants with delusions had similar hit rates to those without delusions, but higher false recognition rates across the three runs. This difference was significant for run one, with the difference between groups also trending towards significance for run three, see Table 7.6. Discrimination performance (measured by \(d'\) as described in section 5.4.3.2) was significantly worse in the delusion group. Visualising the data in scatter plots indicated that one individual in the delusion group had outlying results, with high
rates of false recognition across the three runs (participant CM007; 24, 31 and 40 compared to means of $3.4 \pm 3.4$, $11.3 \pm 7.5$ and $7.2 \pm 8.0$ for the delusion group with this participant removed). The difference between the groups for run one false recognition and discrimination however, remained significant even when this individual’s data were excluded ($ps < .05$). Differences in false recognition between the delusion and control groups across the three runs are displayed in Figure 7.2.

**Table 7.6 Temporal context confusion task results**

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 17)</th>
<th>DELUSIONS (n = 10)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HITS RUN 1</strong></td>
<td>25.7 ± 3.0</td>
<td>24.7 ± 4.1</td>
<td>$U = 66.5, p = .359$</td>
</tr>
<tr>
<td><strong>HITS RUN 2</strong></td>
<td>24.4 ± 3.2</td>
<td>22.8 ± 5.5</td>
<td>$U = 81.5, p = .863$</td>
</tr>
<tr>
<td><strong>HITS RUN 3</strong></td>
<td>24.0 ± 5.0</td>
<td>24.8 ± 4.2</td>
<td>$U = 76.0, p = .675$</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 1</strong></td>
<td>$1.4 \pm 1.9$</td>
<td>$5.5 \pm 7.2$</td>
<td>$U = 36.5, p = .013$</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 2</strong></td>
<td>$8.3 \pm 4.6$</td>
<td>$13.3 \pm 9.4$</td>
<td>$t_{25} = -1.580, p = .141^*$</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 3</strong></td>
<td>$7.1 \pm 3.4$</td>
<td>$15.3 \pm 11.5$</td>
<td>$t_{25} = -2.178, p = .055^*$</td>
</tr>
<tr>
<td><strong>TCC RUN 2</strong></td>
<td>$0.25 \pm 0.18$</td>
<td>$0.36 \pm 0.30$</td>
<td>$t_{25} = -1.154, p = .260$</td>
</tr>
<tr>
<td><strong>TCC RUN 3</strong></td>
<td>$0.24 \pm 0.19$</td>
<td>$0.38 \pm 0.32$</td>
<td>$t_{25} = -1.439, p = .163$</td>
</tr>
<tr>
<td><strong>DISCRIMINATION (d')</strong></td>
<td>$3.51 \pm 0.64$</td>
<td>$2.72 \pm 0.86$</td>
<td>$t_{25} = 2.720, p = .012$</td>
</tr>
<tr>
<td><strong>RESPONSE BIAS (c)</strong></td>
<td>$0.22 \pm 0.43$</td>
<td>$0.04 \pm 0.44$</td>
<td>$t_{25} = 1.001, p = .326$</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. All measures at baseline.
Maximum number hits per run is 28, maximum number false recognition per run is 52.
TCC = temporal context confusion, calculated as $TCC = \frac{FA_x}{HR_x} - \frac{FA_1}{HR_1}$ where $HR =$ hit rate, $FA =$ false alarm rate in run $x$ and run one respectively.
Discrimination and response bias are calculated for run one.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.
*Levene’s test significant for false recognition run two and three ($p = .018, p = .001$ respectively), equal variances not assumed.
Significant findings were further explored through binary logistic regression with delusion group as the dependent variable. Given the relatively small sample size, potential confounding variables were identified a priori: age, gender, years of education, sMMSE score, cholinesterase inhibitor prescription and category fluency on ACE-III. Univariate analyses were then carried out for false recognition on run one or discrimination and each potential predictor value. Predictor values that reached the more liberal threshold of \( p < .10 \) were then included in the full model. Univariate models including false recognition on run one and discrimination were both significant (\( ps < .05 \)). Gender was the only potential covariate that had a \( p \) value below the specified threshold and was the only covariate included for multivariate analyses.

The model for false recognition on run one of the TCC task was significant \( (X^2 (2, n = 27) = 12.885, p = .002) \), and explained 51.8% (Nagelkerke \( R^2 \)) of the variance in delusion group. 81.5% of cases were correctly classified by the model, with 80.0% sensitivity and 82.4% specificity, a positive predictive value of 72.7% and negative predictive value of 87.5%. For each additional item falsely recognised on the first run of the TCC task, participants were 63.5% more likely to be in the ‘delusion’ group (\( \text{Exp(B)} 1.635, 95\% \text{CI} 1.017 – 2.628, p = .042 \)). The model for discrimination on the first run of the TCC task was also significant \( (X^2 (2, n = 27) = 12.538, p = .002) \), explained a similar amount of the variance in delusion group (Nagelkerke \( R^2 \) 50.7%), but only correctly classified 70.4% of cases with a 50.0% sensitivity and 82.3% specificity, a positive predictive value of 62.5% and a negative predictive value of 73.7%. Discrimination was significant in the model, with delusions less likely as discrimination performance improved (\( \text{Exp(B)} .125, 95\% \text{CI} .021 -.739, p = .022 \)).

Models were run excluding the outlying participant and remained significant (\( ps \leq .005 \)) with false recognition and discrimination remaining significant within the models (\( ps < .05 \)). VIF scores were all < 2.5, with a mean VIF of 1.0. The Box-Tidwell test was used to test the assumption of linearity between continuous variables and the logit transformation of the dependent variable, with both models meeting this assumption. Both models passed the Hosmer and Lemeshow test for goodness of fit (\( ps > .30 \)).
Figure 7.2 False recognition on temporal context confusion task

Notes:
○ = outlying value over 1.5 x interquartile range above third quartile; ✱ = outlying value over 3 x interquartile range above third quartile.
7.4.3  False memory and metamemory

See section 5.4.3.3 for full task description.

Fourteen control participants and 10 participants with delusions completed the adapted version of the DRM/metamemory task. There was a trend towards increased false recognition and false memories in the group with delusions, but this did not reach statistical significance. Participants with delusions were more likely to be highly confident in hits, false recognition and false memories (see Figure 7.3); this finding was significant for proportion of high confidence hits \( (p = .049) \). Results are summarised in Table 7.7.

**Table 7.7 DRM/metamemory task results**

<table>
<thead>
<tr>
<th></th>
<th>CONTROL ( (n = 14) )</th>
<th>DELUSIONS ( (n = 10) )</th>
<th>( P ) VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HITS</strong></td>
<td>15.4 ± 8.3</td>
<td>18.3 ± 8.6</td>
<td>( t_{22} = -.823, p = .420 )</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION</strong></td>
<td>3.4 ± 4.6</td>
<td>10.0 ± 10.3</td>
<td>( U = 45.0, p = .154 )</td>
</tr>
<tr>
<td><strong>FALSE MEMORIES</strong></td>
<td>3.7 ± 3.7</td>
<td>7.7 ± 6.4</td>
<td>( U = 50.0, p = .259 )</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE HITS</strong></td>
<td>0.44 ± 0.30</td>
<td>0.70 ± 0.31</td>
<td>( t_{22} = -2.085, p = .049 )</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE FR</strong></td>
<td>( (n = 10) ) 0.27 ± 0.37</td>
<td>( (n = 8) ) 0.40 ± 0.40</td>
<td>( U = 33.0, p = .573 )</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE FM</strong></td>
<td>( (n = 13) ) 0.40 ± 0.41</td>
<td>( (n = 7) ) 0.68 ± 0.30</td>
<td>( U = 26.0, p = .135 )</td>
</tr>
<tr>
<td><strong>Discrimination (d’)</strong></td>
<td>1.46 ± 0.84</td>
<td>0.95 ± 0.75</td>
<td>( t_{22} = 1.506, p = .146 )</td>
</tr>
<tr>
<td><strong>Response bias (c)</strong></td>
<td>0.66 ± 0.68</td>
<td>0.11 ± 1.08</td>
<td>( t_{22} = 1.538, p = .138 )</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. All measures at baseline.
Maximum number hits per run is 30, maximum number false recognition per run is 30, maximum number false memories per run is 15.
High confidence values as proportion of total hits, false recognition or false memories respectively.
DRM = Deese-Roediger-McDermott paradigm; FM = false memories; FR = false recognition.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.
Figure 7.3 Proportion of high confidence responses by response type, DRM task

Notes:

- o = outlying value over 1.5 x interquartile range above third quartile;
- ✱ = outlying value over 3 x interquartile range above third quartile.
The participant identified as an outlier for performance on the TCC task at baseline (CM007; see section 7.3.2) was also an outlier in terms of false recognition and false memory rates on the DRM/metamemory task, with high false recognition and false memory rates (30 and 15 respectively, compared to means of 7.8 ± 8.0, 3.7 ± 3.7 for the delusion group with this participant removed). With this individual’s data removed, false recognition rates remained greater in the delusion group than the control group, although there was no longer a difference in false memory rates between the groups. The delusion group continued to have poorer discrimination performance (mean d’ 1.10 ± 0.59).

Significant findings were further explored through binary logistic regression with delusion group as the dependent variable. Proportion of high confidence hits was transformed into a percentage score, to aid model interpretation. Given the relatively small sample size, potential confounding variables were identified a priori: age, gender, years of education, sMMSE score, cholinesterase inhibitor prescription and category fluency on ACE-III. Univariate analyses were then carried out for proportion of high confidence hits and each potential predictor value. Predictor values that reached the more liberal threshold of \( p < .10 \) were then included in the full model. The univariate model including number of high confidence hits was significant (\( p = .037 \)). Gender was the only potential covariate that had a \( p \) value below the specified threshold and was the only covariate included for multivariate analyses.

The multivariate model including gender was significant (\( \chi^2 (2, n = 24) = 11.656, p = .003 \)), and explained 51.8% (Nagelkerke \( R^2 \)) of the variance in delusion group. 87.5% of cases were correctly classified by the model, with 80.0% sensitivity and 92.9% specificity, a positive predictive value of 88.9% and negative predictive value of 86.7%. For each 1% increase in high confidence hits, participants were 5.4% more likely to be in the ‘delusion’ group (Exp(B) 1.054, 95% CI 1.007 – 1.105, \( p = .025 \)). The VIF scores for the model were all < 2.5, with a mean VIF of 1.0. The Box-Tidwell test was used to test the assumption of linearity and the model met this assumption. The model passed the Hosmer and Lemeshow test for goodness of fit (\( p = .307 \)).

Results were not further explored by subtype of psychosis symptoms as group numbers were considered too small, with only one individual experiencing misidentification-type phenomena.
8. False Memories and Delusions in the ADNI Cohort: Methods for Analyses of Behavioural Data

8.1 Background

In my patient-based study of memory error and delusions in AD (see Chapters 5 - 7), those participants with delusions in AD had higher rates of false recognition than those without delusions on two different behavioural tasks. That is, they were more likely to incorrectly endorse a new stimulus (word or picture) as seen before. This finding reached statistical significance for the TCC task and survived once confounders (gender) were included in regression modelling. However, the numbers in the sample were small (27 participants with AD in total, 10 with delusions and 17 without).

The study described in this chapter was designed to explore this finding further in a larger sample of people with AD, available as part of a publicly available data source: the ADNI cohort.

8.1.1 Overview of ADNI

ADNI is a North American multicenter, longitudinal study, initiated by a group of United States (US) public health bodies, private pharmaceutical companies and nonprofit organisations. ADNI launched in 2004 and continues to recruit subjects from 59 sites across the US and Canada. Over the past 17 years the study has amassed an extensive data archive of behavioural and neuroimaging data for individuals with AD, aMCI and healthy older adults. Data collection has been split into four phases (ADNI1, ADNI-GO, ADNI2, ADNI3), see Table 8.1.

<table>
<thead>
<tr>
<th>ADNI PHASE</th>
<th>YEARS RUNNING</th>
<th>PRIMARY STUDY GOAL</th>
<th>COHORT ROLLOVER</th>
<th>NEW</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADNI1</td>
<td>2004 - 2010</td>
<td>Investigation of biomarkers to develop more sensitive biomarkers for early detection and tracking of AD</td>
<td>N/A</td>
<td>200 AD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>400 MCI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>200 CN</td>
</tr>
<tr>
<td>ADNI-GO</td>
<td>2009 - 2011</td>
<td>Investigation of biomarkers at an earlier stage of disease</td>
<td>500 MCI and CN</td>
<td>200 MCI</td>
</tr>
<tr>
<td>ADNI2</td>
<td>2011 - 2017</td>
<td>As above, plus amyloid PET with Florbetapir included, and added participants with ‘significant memory concern’ but without MCI diagnosis</td>
<td>650 - 700 MCI and CN</td>
<td>150 AD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>250 MCI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>150 CN</td>
</tr>
<tr>
<td>ADNI3</td>
<td>2017 - 2022</td>
<td>Investigation of relationship between biomarker characteristics across Alzheimer’s disease iterations</td>
<td>130 - 150 AD</td>
<td>85 - 185 AD</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>275 - 320 MCI</td>
<td>150 - 515 MCI</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>295 - 330 CN</td>
<td>135 - 500 CN</td>
</tr>
</tbody>
</table>

Notes:
AD = Alzheimer’s disease; ADNI = Alzheimer’s Disease Neuroimaging Initiative; CN = cognitively normal, healthy elderly controls; MCI = mild cognitive impairment.
8.1.1.1  **Inclusion and exclusion criteria**

**Inclusion criteria for ADNI participation are:**

1) Age 55 - 90  
2) For AD participants: meets the NINCDS/ADRDA criteria for probable AD  
3) MMSE score of 24 - 30 for cognitively normal (CN) and mild cognitive impairment (MCI) participants and MMSE 20 - 26 for AD participants  
4) CDR score of 0 for CN, 0.5 for MCI and 0.5, 1.0 for AD participants  
5) GDS-15 < 6  
6) Hachinski Ischaemia Score of < 5  
7) Has ‘study partner’ with whom they have > 10 hours contact with per week, who can accompany participant on visits  
8) Four weeks of stability on permitted medications  
9) Adequate visual and auditory ability to complete neuropsychological testing  
10) Fluent speaker of English or Spanish  
11) Completed six grades of education, or have the equivalent work history  
12) No contraindications to MRI  
13) Committed to involvement in a 2 - 3 year study, and agreeing to blood, urine, DNA testing and neuroimaging  
14) Not enrolled in any other trials or studies  

**Exclusion criteria for ADNI participation for participants with AD are:**

1) Significant neurological disease (including PD, vascular dementia, Huntington’s disease)  
2) Significant psychiatric illness (including major depression, schizophrenia, bipolar affective disorder)  
3) Unstable medical illness such that would prevent a participant from following study protocol  
4) History of alcohol or substance abuse  
5) Resident in a nursing home  
6) Taking any of the following medications:
   a. Antidepressants with anticholinergic properties (e.g. – amitriptyline)  
   b. Narcotic analgesics (> 2 doses/week, e.g. oxycodone)  
   c. Other anticholinergic agents either with neuroleptic or other central nervous system activity (e.g. – amantadine). NB – This excludes those taking the antipsychotics chlorpromazine, fluphenazine, loxapine, perphenazine, thioridazine, thiothixene, trifluoperazine, clozapine and haloperidol.
Aripiprazole, olanzapine, quetiapine, risperidone and ziprasidone are permitted.
d. Antiparkinsonian medication
e. Selected benzodiazepines/sedatives (chlordiazepoxide, clonazepam, diazepam, flurazepam, meprobamate, triazolam)

7) Screening 1.5T MRI scan fails to pass quality control due to either:
a. A significant neurological abnormality being detected (for example, evidence of infection, infarction or other lesions)
b. Motion due to participant non-compliance with scanning, due to either chronic illness or deteriorated cognitive ability

8.1.1.2 Screening and baseline assessment
Demographic details are collected from the ADNI cohort at baseline, which include age, gender, race, years of education, occupation, year of onset of AD (where appropriate), medications and medical history.

As part of their screening visit, participants complete the CDR, the Modified Hachinski Ischaemia Scale and the GDS. Behavioural and psychological symptoms are assessed using the NPI or the briefer, cross-validated NPI-Q, at baseline and follow-up visits. The NPI-Q was primarily used for ADNI1 and ADNI-GO, with ADNI2 and ADNI3 using the NPI-Q for telephone follow-up only. In person follow-up occurs at six months and annually thereafter.

8.1.1.3 Neuropsychological battery
Global cognitive function is assessed in ADNI using the MMSE and the 13-item ADAS–Cognitive subscale (ADAS-Cog 13). Participants also complete a battery of neuropsychological tests assessing premorbid intelligence (the American National Adult Reading Test); confrontation naming (Boston Naming Test); verbal memory (RAVLT and the Logical Memory Test I and II – Immediate and Delayed Recall from the WMS-R; associative learning (Digit Symbol Substitution Test from the WAIS-R); attention and working memory (forward and backward digit spans from the WMS-R); language and executive function (category fluency for both animals and vegetables); attention and executive function (Trail Making Tests A and B); visuospatial ability, praxis and executive function (clock drawing and clock copying).

From ADNI-GO onwards the Montreal Cognitive Assessment (MoCA) was added to the battery, and the Multilingual Naming Test has replaced the Boston Naming Test in ADNI3.
8.1.4 Ethical approval

ADNI participants provide written informed consent for their involvement. Each institution recruiting to the study has received ethical approval for the study from the relevant institutional Research Ethics Committee.

8.2 Research questions

The primary research question was:

- Do AD participants with delusions at baseline show evidence of more false memories than those without delusions, when matched for overall cognitive impairment as measured by MMSE score?

The secondary research question was:

- Do AD participants who experience delusions at any time during ADNI follow-up show evidence of more false memories at baseline than those who do not, when matched for overall cognitive impairment as measured by MMSE score?

The a priori hypothesis, based on findings from analysis of my own data as above and described in detail in Chapter 7, was that those individuals with delusions in AD at baseline or any later time point would have increased rates of false recognition compared to those without delusions.

8.3 Participant selection

Participants were selected from within the ADNI cohort in accordance with inclusion/exclusion criteria as below. As such, participants also met all ADNI inclusion/exclusion criteria laid out above in section 8.1.1.1.

Inclusion criteria:

1) Diagnosis of possible or probable AD based on NINCDS-ADRDA criteria at any ADNI timepoint
2) Recruited to either ADNI 1, GO, 2 or 3
3) Completed NPI or NPI-Q at baseline

Exclusion criteria:

1) Inconsistent diagnosis of AD through ADNI follow-up (if participant reverted to a diagnosis of aMCI at a later visit)

The above criteria were chosen to maximise sample size. While an individual may be nominally ‘cognitively normal’ on recruitment to ADNI, this includes individuals with MMSE scores of 24
and over, and as such those who later meet criteria for a diagnosis of AD are assumed for the purpose of this study to have AD pathology present at baseline. A subgroup analysis excluding those who were CN at baseline will be run to confirm that this does not significantly alter findings.

8.4 Data selection

8.4.1 Demographic details

Demographic data for all participants were downloaded from the ADNI data archive. This included: age, gender, race, primary language, handedness, years of education, years since diagnosis and any psychotropic medication prescribed (cholinesterase inhibitor/memantine; antidepressant; antipsychotic), measured at baseline.

8.4.2 Neuropsychological covariates/confounders

NPI and NPI-Q data from all available timepoints were used to determine if participants developed delusions (or hallucinations) at any study time point. As ADNI phase determined whether NPI or NPI-Q was completed, baseline NPI-Q scores were used as a measure of overall severity of behavioural and psychological symptoms of dementia, with NPI scores converted to NPI-Q scores if the NPI-Q was not available. Delusions were considered present if a participant scored one or more on the delusions item of the NPI-Q or answered ‘yes’ to the delusions domain of the full NPI (NPI section A). Participants were considered to have a delusion ‘state’ if this was positive at baseline, and to have a delusion ‘trait’ if this was the case at baseline or any other time point during ADNI follow-up. Participants with a delusion ‘trait’ but without delusions at baseline were not included in the control group for the ‘state’ group. The delusion ‘trait’ group includes those with ‘state’ delusions: these two groups were therefore never compared to each other, only to the control group.

MMSE score was used as a measure of overall cognitive function. Available measures of executive function were also explored, including MoCA, clock drawing, category fluency and Trails B score.

8.4.3 False memory measures

The ADAS-Cog is widely used as an outcome measure of cognitive function in AD.\textsuperscript{332} The 13 item ADAS-Cog includes word recall (immediate and delayed), following commands, constructional praxis, naming objects and fingers, ideational praxis, orientation, word recognition, number cancellation, remembering instructions, comprehension, word finding difficulty and spoken language ability. The first nine items are scored as ‘number incorrect’, and scores on the final
four items are based on the clinician’s impression of the participant’s degree of impairment in these domains, ranging from zero (none) to five (severe). Lower scores on the ADAS-Cog 13 indicate better performance.

In the word recognition element of the ADAS-Cog 13 participants are shown a list of 12 words printed on white cards, which they are asked to read aloud and ‘try to remember’. They then complete a word recognition task, in which they are shown a further set of words including the 12 words previously seen and 12 distractors. Positive responses to distractors (range 0 - 12, higher score indicates more false recognition) are used as one index of false memory.

The RAVLT consists of five trials of learning 15 words, read aloud by the examiner. For each trial, participants immediately repeat the words they can remember, with the number of words correct and number of intrusions (number of words recalled that were not in the word list) recorded. A distractor list B is read, following which participants are asked to recall words from the original list, with the number of correct words and intrusions recorded. From the point at which list B is read, there is a 30 minute delay before the participant is asked to recall the words again. Finally, participants are shown a word recognition sheet, containing the original 15 words and 15 distractors, for which the words correctly recalled and those that are falsely recognised are recorded.

From the RAVLT, the total number of intrusions across the initial five trials (intrusions in immediate recall, range 0 - 75), total number of intrusions after the distractor list (range 0 - 15) and total number of intrusions after 30 minutes (range 0 - 15) were calculated. Both the sum of total intrusions across the task (range 0 - 105) and number of positive responses to distractors on word recognition (range 0 - 15) were used as measures of false memory.

Discrimination and response bias were calculated for both ADAS-Cog 13 and RAVLT using formulae as described in section 5.4.3.2.

8.5 Statistical methodology

Behavioural data analysis was carried out in SPSS. For all analyses, results were considered significant at a $p$ value of $< .05$.

8.5.1 Group size

Prior to initiation of the study, necessary group size was estimated based on previous studies. This analysis was hypothesis driven, and a power calculation in G*Power was completed based on the findings described in Chapter 7. Using the results for false recognition on run one of the TCC task in those with and without delusions gave a sample size of $\geq 22$ per group (power = 0.8)
to detect a significant difference in false recognition between groups \( (p < .05) \), numbers which were far exceeded by this dataset: as a result, correction for multiple comparisons was not performed.

### 8.5.2 Behavioural analyses

Data were first viewed in the form of scatter plots, histograms and Q-Q plots to assess whether they followed a normal distribution. The assumption of normality was further tested using the Kolmogorov-Smirnov test for non-normality. Baseline characteristics of the population and task performance were compared between those with delusions at baseline and the control group, and between those with delusions at any time point and the control group using independent-samples t-tests or Mann-Whitney non-parametric tests if parametric data assumptions were violated, or chi-squared tests for categorical variables.

As the primary outcome measures were binary (i.e. does a participant have ‘state’ delusions or not; does a participant have ‘trait’ delusions or not), binary logistic regression was used to further explore significant findings. Given the large sample size, potential confounding variables were chosen a priori. These were: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription and category fluency, for reasons as described in section 5.5.3.

Potential outlying data points were identified by visual inspection of scatter plots, with the data point then reviewed to exclude data entry error. Models were then re-run excluding any outlying values.

Further model diagnostics were run for models with significant results. As described in section 5.5.3, VIF scores were calculated as a measure of collinearity in the data, with multi-collinearity considered present at the more conservative VIF threshold of 2.5. The logistic regression model assumption of linearity was tested using the Box-Tidwell test. Model fit was assessed using the Hosmer and Lemeshow goodness of fit test, as described in section 5.5.3.

Given that ‘state’ delusion was a rare event in the group (see section 9.2), Firth’s logistic regression models were also run. Firth’s logistic regression reduces bias introduced by small samples, and corrects for separation issues that can occur when the event studied is rare (as is the case for ‘state’ delusion in the large ADNI sample).
9. False Memories and Delusions in the ADNI Cohort: Results of Analyses of Behavioural Data

As described in Chapter 8, demographic and behavioural data were downloaded from the ADNI data archive (http://adni.loni.usc.edu/data-samples/access-data/) on 24th June 2020, with participant diagnostic information downloaded for screening and selection on 26th March 2020.

9.1 Participant selection

Of the 2248 participants in ADNI, 783 had received an AD diagnosis at any point during follow-up. As described in Chapter 8, participants whose diagnosis was subsequently changed back to CN or aMCI were excluded. Participants whose dementia diagnoses were recorded as ‘suspected not due to AD aetiology’ were excluded, as were those with PD or with a GDS-15 > 5. A total of 733 participants were included in the behavioural analysis, see Figure 9.1

![Figure 9.1 Participant selection flow chart](image)

Notes:
AD = Alzheimer’s disease; ADNI = Alzheimer’s Disease Neuroimaging Initiative; CN = cognitively normal, healthy elderly controls; GDS-15 = short form Geriatric Depression Scale; MCI = mild cognitive impairment; PD = Parkinson’s disease.

9.2 Demographic details

Demographic details and baseline screening assessments are summarised in Table 9.1.

Participants in the sample had a mean age of 74.8 ± 7.5 years, 318 (43.4%) were female, 415 (56.6%) were male and the majority (n = 686; 93.6%) were white. Of note, the ADNI protocol
states an aim to recruit 12.0% of study participants from black and minority groups, to match their estimate of the 14.0% aged minority population of the US.407 387 (52.8%) had an AD diagnosis at baseline and the majority of participants (n = 589, 80.4%) had received an AD diagnosis by the end of the second year of follow-up.

The sample was split into those with ‘state’ delusions (delusions at baseline, n = 42) and those with ‘trait’ delusions (to include the ‘state’ group, and those with delusions at any later time point, n = 179). Both of these groups were compared to a control group of participants who did not develop delusions at any time point (n = 554).

For the ‘state’ delusion group, there were no significant differences from the control group in terms of age, gender, primary language, handedness, years of education, cholinesterase inhibitor or antidepressant prescriptions or GDS-15 score. However, the ‘state’ delusion group did differ significantly from the control group in terms of race, diagnosis at baseline, years since diagnosis, antipsychotic prescribing, MMSE, NPI and NPI-Q scores, and presence of hallucinations. The control group included a lower proportion of individuals with AD at baseline and had on average fewer years since diagnosis. Individuals in the control group were prescribed fewer antipsychotics, had higher scores on the MMSE and had a lower burden of neuropsychiatric symptoms as measured by total NPI and NPI-Q scores.

The ‘trait’ delusion and control groups did not differ significantly in terms of age, race, primary language, handedness, years of education, diagnosis at baseline, years since diagnosis, cholinesterase inhibitor or antidepressant prescriptions, MMSE or GDS-15 score. The ‘trait’ group differed significantly from the control group in terms of gender and, similar to the ‘state’ group, had higher rates of antipsychotic prescribing and neuropsychiatric symptoms, including hallucinations, on the NPI and NPI-Q.
**Table 9.1 Demographic details and screening results**

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 554)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th><strong>P VALUE</strong> ('STATE' / CONTROL)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th><strong>P VALUE</strong> ('TRAIT’ / CONTROL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE (years)</strong></td>
<td>74.8 ± 7.6; 55 - 90</td>
<td>76.0 ± 6.4; 64 - 90</td>
<td>$t_{594} = -.985, p = .325$</td>
<td>75.0 ± 7.2; 55 - 91</td>
<td>$t_{731} = -.275, p = .783$</td>
</tr>
<tr>
<td><strong>GENDER (female)</strong></td>
<td>225 (40.6)</td>
<td>18 (42.9)</td>
<td>$X^2 (1, n = 596) = .081, p = .871$</td>
<td>93 (52.0)</td>
<td>$X^2 (1, n = 733) = 7.085, p = .008$</td>
</tr>
<tr>
<td><strong>RACE</strong></td>
<td></td>
<td></td>
<td>$^p = .036$</td>
<td></td>
<td>$^p = .054$</td>
</tr>
<tr>
<td>White</td>
<td>524 (94.6)</td>
<td>36 (85.7)</td>
<td>162 (90.5)</td>
<td></td>
<td>12 (6.7)</td>
</tr>
<tr>
<td>Black/African</td>
<td>13 (2.3)</td>
<td>4 (9.5)</td>
<td>4 (2.2)</td>
<td></td>
<td>1 (0.6)</td>
</tr>
<tr>
<td>American</td>
<td></td>
<td></td>
<td>$^p = .670$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>12 (2.2)</td>
<td>2 (4.8)</td>
<td>4 (2.2)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mixed race</td>
<td>5 (0.9)</td>
<td>0 (0.0)</td>
<td>1 (0.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>PRIMARY LANGUAGE</strong></td>
<td></td>
<td></td>
<td>$^p = .766$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>540 (97.5)</td>
<td>41 (97.6)</td>
<td>175 (97.8)</td>
<td></td>
<td>4 (2.2)</td>
</tr>
<tr>
<td>Spanish</td>
<td>9 (1.6)</td>
<td>1 (2.4)</td>
<td>0 (0.0)</td>
<td></td>
<td>$^p = .441$</td>
</tr>
<tr>
<td>Other</td>
<td>5 (0.9)</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>RIGHT HANDEDNESS</strong></td>
<td></td>
<td></td>
<td>$^p = .766$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 595/732)</td>
<td>509 (92.0)</td>
<td>38 (90.5)</td>
<td>169 (94.4)</td>
<td></td>
<td>$X^2 (1, n = 732) = 1.112, p = .292$</td>
</tr>
<tr>
<td><strong>EDUCATION (years)</strong></td>
<td>15.6 ± 2.8; 6 - 20</td>
<td>15.6 ± 2.8; 8 - 20</td>
<td>$t_{594} = -.119, p = .905$</td>
<td>15.4 ± 2.9; 4 - 20</td>
<td>$t_{731} = .889, p = .374$</td>
</tr>
<tr>
<td><strong>DIAGNOSIS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>21 (37.9)</td>
<td>0 (0.0)</td>
<td>$X^2 (2, n = 596) = 11.151, p = .005$</td>
<td>6 (3.4)</td>
<td>$X^2 (2, n = 733) = 2.914, p = .233$</td>
</tr>
<tr>
<td>aMCI</td>
<td>229 (41.3)</td>
<td>8 (19.0)</td>
<td>87 (48.6)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>304 (54.9)</td>
<td>34 (81.0)</td>
<td>86 (48.0)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>YEARS SINCE DIAGNOSIS</strong> (n = 576/712)</td>
<td>0.3 ± 4.0; -13.0 - 12.3</td>
<td>2.2 ± 2.6; -4.0 - 7.8</td>
<td>$t_{54,513} = -2.960, p = .000$</td>
<td>0.3 ± 4.1; -12.0 - 13.8</td>
<td>$t_{710} = .040, p = .968$</td>
</tr>
</tbody>
</table>
### Table 9.1 Cont. Demographic details and screening result

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 554)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>P VALUE (‘STATE’ / CONTROL)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th>P VALUE (‘TRAIT’ / CONTROL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MEDICATION (n = 595/732)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChEI/Memantine</td>
<td>404 (73.1)</td>
<td>34 (81.0)</td>
<td>$X^2 (1, n = 595) = 1.253, p = .284$</td>
<td>123 (68.7)</td>
<td>$X^2 (1, n = 732) = 1.264, p = .261$</td>
</tr>
<tr>
<td>Antidepressant</td>
<td>181 (32.7)</td>
<td>17 (40.5)</td>
<td>$X^2 (1, n = 595) = 1.055, p = .312$</td>
<td>72 (40.2)</td>
<td>$X^2 (1, n = 732) = 3.357, p = .067$</td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>4 (0.1)</td>
<td>4 (9.5)</td>
<td>$^* p = .001$</td>
<td>6 (3.4)</td>
<td>$^* p = .017$</td>
</tr>
<tr>
<td><strong>NPI-Q</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.4 ± 2.7; 0 - 15</td>
<td>7.9 ± 4.9; 1 - 24</td>
<td>$U = 3086.5, p = .000$</td>
<td></td>
<td>4.3 ± 4.3; 0 - 24</td>
<td></td>
</tr>
<tr>
<td><strong>NPI (n = 260/306)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.1 ± 7.5; 0 - 35</td>
<td>19.5 ± 12.1; 3 - 46</td>
<td>$U = 732.5, p = .000$</td>
<td></td>
<td>11.2 ± 11.4; 0 - 46</td>
<td>$U = 5887.5, p = .001$</td>
</tr>
<tr>
<td><strong>HALLUCINATIONS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline</td>
<td>9 (1.6)</td>
<td>7 (16.7)</td>
<td>$^* p = .000$</td>
<td>16 (8.9)</td>
<td></td>
</tr>
<tr>
<td>Any time point</td>
<td>46 (8.3)</td>
<td>16 (38.1)</td>
<td>$^* p = .000$</td>
<td>70 (39.1)</td>
<td></td>
</tr>
<tr>
<td><strong>MMSE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25.1 ± 2.9; 18 - 30</td>
<td>23.4 ± 2.7; 17 - 29</td>
<td>$t_{594} = 3.533, p = .000$</td>
<td></td>
<td>25.1 ± 2.9; 17 - 30</td>
<td>$t_{731} = -.082, p = .935$</td>
</tr>
<tr>
<td><strong>GDS-15</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.6 ± 1.4; 0 - 5</td>
<td>1.7 ± 1.5; 0 - 5</td>
<td>$U = 11072.5, p = .590$</td>
<td></td>
<td>1.6 ± 1.4; 0 - 5</td>
<td>$U = 48939.0, p = .787$</td>
</tr>
</tbody>
</table>
Table 9.1 Cont. Demographic details and screening result

Notes:
Values are expressed as mean ± SD; range or n (%). All measures at baseline unless otherwise specified.
AD = Alzheimer’s disease; aMCI = amnestic mild cognitive impairment; ChEI = cholinesterase inhibitor; CN = cognitively normal; GDS-15 = short form Geriatric Depression Scale; MMSE = Mini Mental State Examination; NPI = Neuropsychiatric Inventory.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data. Categorical comparisons by chi-squared unless otherwise specified.
†p value from Fisher’s exact test.
The NPI-Q was completed for all participants. The mean NPI-Q score for the ‘state’ delusion group was 7.9 ± 4.9, for the ‘trait’ delusion group 4.3 ± 4.3, and for the control group 2.4 ± 2.7. For the subset of participants for whom a full NPI was completed, the mean NPI score for the ‘state’ delusion group was 19.5 ± 12.1, for the ‘trait’ delusion group was 11.2 ± 11.4, and for the control group was 6.1 ± 7.5. In the ‘state’ group mean delusion severity on the NPI-Q at baseline was 1.6 ± 0.7. The NPI was completed for 21 of the ‘state’ group at baseline, with mean delusion frequency x severity score of 4.0 ± 3.2. Twenty-five participants were experiencing hallucinations at baseline, with a further 91 developing hallucinations at a later timepoint.

When psychosis symptoms were broken down by subtype, 32 participants had psychosis symptoms at baseline, with seven experiencing paranoid-type symptoms, nine misidentification-type symptoms, eight mixed-type symptoms (i.e. a combination of paranoid and misidentification symptoms) and the remaining eight other psychotic phenomena. Including participants who had an NPI at any timepoint (n = 439), 103 experienced psychosis symptoms, with 27 experiencing only paranoid-type symptoms, 39 misidentification-type symptoms and 37 a mixed picture. Breakdown of baseline psychosis symptoms by subtype is shown in more detail in Table 9.2.

**Table 9.2 Breakdown and classification of psychosis symptoms at baseline (n = 32)**

<table>
<thead>
<tr>
<th>CONTENT</th>
<th>n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delusions</td>
<td></td>
</tr>
<tr>
<td>PARANOID</td>
<td></td>
</tr>
<tr>
<td>In danger/others are planning to hurt him/her</td>
<td>5 (15.6)</td>
</tr>
<tr>
<td>Others are stealing from him/her</td>
<td>12 (37.5)</td>
</tr>
<tr>
<td>Spouse is having an affair</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>Family members plan to abandon him/her</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Unwelcome guests are staying in his/her house</td>
<td>4 (12.5)</td>
</tr>
<tr>
<td>His/her spouse or others are not who they claim to be</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>His/her house is not his/her own</td>
<td>2 (6.3)</td>
</tr>
<tr>
<td>Television/magazine figures are present in his/her home</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>Hallucinations</td>
<td></td>
</tr>
<tr>
<td>He/she can hear voices</td>
<td>5 (15.6)</td>
</tr>
<tr>
<td>Talks to people who are not there</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Seeing things not seen by others</td>
<td>7 (21.9)</td>
</tr>
<tr>
<td>Delusions</td>
<td></td>
</tr>
<tr>
<td>Any other unusual beliefs</td>
<td>11 (34.4)</td>
</tr>
<tr>
<td>Smells odours not smelled by others</td>
<td>-</td>
</tr>
<tr>
<td>Feels things on his/her skin</td>
<td>3 (9.4)</td>
</tr>
<tr>
<td>Tastes without known cause</td>
<td>1 (3.1)</td>
</tr>
<tr>
<td>Any other unusual sensory experiences</td>
<td>4 (12.5)</td>
</tr>
</tbody>
</table>
9.3 Neuropsychological profile

9.3.1 Global cognitive function

The ‘state’ delusion group had significantly greater impairments on both measures of global cognitive function than the control group (MMSE scores 23.4 ± 2.7 and 25.1 ± 2.9 respectively, \( t_{594} = 3.533, p = .000 \); ADAS-Cog 13 scores 30.9 ± 6.8 and 25.2 ± 9.2 respectively, \( t_{588} = -3.931, p = .000 \)). The ‘trait’ delusion group and the control group were more closely matched, with no significant difference in either MMSE or ADAS Cog-13 scores (MMSE scores 25.1 ± 2.9 and 25.1 ± 2.9 respectively, \( t_{731} = -.082, p = .935 \); ADAS-Cog 13 scores 25.3 ± 9.2 and 25.2 ± 9.2 respectively, \( t_{723} = -.029, p = .977 \)).

9.3.2 Executive function

A similar pattern was seen in measures of executive function, with significant differences in performance between ‘state’ delusion group and control group, and no significant difference in performance between ‘trait’ delusion group and control, see Table 9.3.
### Table 9.3 Performance on measures of executive function

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 554)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>P VALUE ('STATE' / CONTROL)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th>P VALUE ('TRAIT' / CONTROL)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MoCA (EXECUTIVE FUNCTION)</strong> (n = 271/320)</td>
<td>3.2 ± 1.4</td>
<td>2.7 ± 1.5</td>
<td>U = 2183.5, p = .106</td>
<td>3.1 ± 1.3</td>
<td>U = 8448.0, p = .559</td>
</tr>
<tr>
<td><strong>CLOCK DRAWING</strong> (n = 595/732)</td>
<td>3.8 ± 1.3</td>
<td>3.2 ± 1.5</td>
<td>U = 9036.5, p = .013</td>
<td>3.7 ± 1.3</td>
<td>U = 48359.0, p = .631</td>
</tr>
<tr>
<td><strong>CATEGORY FLUENCY</strong> (n = 595/732)</td>
<td>14.0 ± 5.3</td>
<td>11.5 ± 4.5</td>
<td>t_{593} = 2.890, p = .004</td>
<td>13.8 ± 4.9</td>
<td>t_{730} = .342, p = .733</td>
</tr>
<tr>
<td><strong>TRAILS B SCORE</strong> (n = 560/693)</td>
<td>165.8 ± 86.9</td>
<td>197.6 ± 82.8</td>
<td>U = 7376.0, p = .015</td>
<td>169.0 ± 90.5</td>
<td>U = 44063.0, p = .862</td>
</tr>
</tbody>
</table>

**Notes:**
Values are expressed as mean ± SD. All measures at baseline.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.
MoCA = Montreal Cognitive Assessment, executive function subdomain only. Category fluency for animals only.
9.4 False memory measures

The ‘state’ delusion group had higher rates of false recognition on both the ADAS-Cog 13 and RAVLT. This finding was statistically significant for ADAS-Cog 13 false recognition, see Table 9.4 and Figure 9.2. On RAVLT intrusion measures, there were no significant differences between the groups, see Table 9.4. There were no significant differences between ‘trait’ delusion and control groups on any of the false memory measures.

Discrimination performance (measured by d’ as described in section 5.4.3.2) was worse for word recognition on both ADAS-Cog 13 and RAVLT in the ‘state’ delusion group compared to the control group. This was statistically significant for both measures. Both ‘state’ delusion group and control group had negative response bias, with no statistically significant difference between the groups. There were no significant differences between ‘trait’ delusion and control groups for either discrimination or response bias.

Potential outliers were identified by visual inspection of scatter plots and box plots of false-recognition data, see Figure 9.2. These were first inspected to exclude data entry error. Findings that were significant were explored excluding potential outliers. Comparison of false recognition on ADAS-Cog 13 between the ‘state’ delusion and control groups remained significant ($U = 9125.0, p = .025$). When participants who were CN at baseline were removed from comparison of false recognition on ADAS-Cog 13 between the ‘state’ delusion and control groups the finding was no longer significant ($U = 9308.5, p = .065$). However, when eight data points identified as outliers were excluded from this analysis the comparison remained significant ($U = 8978.5, p = .042$).
Table 9.4 Performance on measures of false memory

<table>
<thead>
<tr>
<th>FALSE RECOGNITION</th>
<th>CONTROL (n = 554)</th>
<th>'STATE' DELUSIONS (n = 42)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADAS-Cog 13 (12)</td>
<td>2.6 ± 2.7</td>
<td>3.5 ± 3.0</td>
<td>U = 9455.0, p = .040</td>
</tr>
<tr>
<td>d'</td>
<td>1.47 ± 0.83</td>
<td>1.21 ± 0.61</td>
<td>U = 9511.5, p = .048</td>
</tr>
<tr>
<td>c</td>
<td>-0.14 ± 0.58</td>
<td>-0.02 ± 0.68</td>
<td>t_{594} = -1.287, p = .199</td>
</tr>
<tr>
<td>RAVLT (15) (n = 595)</td>
<td>2.5 ± 2.4</td>
<td>3.0 ± 2.5</td>
<td>U = 9907.0, p = .107</td>
</tr>
<tr>
<td>d'</td>
<td>1.17 ± 0.91</td>
<td>0.82 ± 0.63</td>
<td>U = 8842.0, p = .010</td>
</tr>
<tr>
<td>c</td>
<td>-0.46 ± 0.58</td>
<td>-0.48 ± 0.60</td>
<td>U = 11180.0, p = .687</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>INTRUSIONS</th>
<th>CONTROL (n = 554)</th>
<th>'TRAIT' DELUSIONS (n = 179)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Immediate recall (90) (n = 590)</td>
<td>3.7 ± 4.9</td>
<td>3.4 ± 3.3</td>
<td>U = 10883.5, p = .721</td>
</tr>
<tr>
<td>After distraction (15) (n = 594)</td>
<td>1.3 ± 1.5</td>
<td>1.7 ± 1.6</td>
<td>U = 9804.5, p = .132</td>
</tr>
<tr>
<td>Delayed recall (15) (n = 595)</td>
<td>1.2 ± 1.9</td>
<td>1.0 ± 1.5</td>
<td>U = 11182.5, p = .663</td>
</tr>
<tr>
<td>Total (120) (n = 588)</td>
<td>6.3 ± 6.7</td>
<td>6.0 ± 4.8</td>
<td>U = 10286.5, p = .515</td>
</tr>
</tbody>
</table>

Notes:
Values are expressed as mean ± SD. All measures at baseline.
Maximum scores for each category provided in brackets. All intrusion measures are from RAVLT.
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; c = response bias; d' = discrimination; RAVLT = Rey’s Auditory Verbal Learning Test.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.
**Figure 9.2** False recognition on ADAS-Cog 13 and RAVLT between ‘state’ delusion, ‘trait’ delusion and control groups

Notes:
- o = outlying value over 1.5 x interquartile range above third quartile; ✱ = outlying value over 3 x interquartile range above third quartile.
- ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; RAVLT = Rey’s Auditory Verbal Learning Test.
Binary logistic regression modelling was used to further explore the relationship between ‘state’ and ‘trait’ delusions and performance on measures of false recognition. Four models were run in total. Two models predicted whether participants would be in ‘state’ delusion or control group, one including false recognition on RAVLT as a predictor and one including false recognition on ADAS-Cog 13 as a predictor. Two models predicted whether participants would be in ‘trait’ delusion or control group, again with one model including false recognition on RAVLT and one including false recognition on ADAS-Cog 13. Each model also adjusted for confounding variables identified a priori: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription and category fluency for animals.

For ‘state’ delusions, the ADAS-Cog 13 false recognition model was statistically significant ($X^2 (7, n = 594) = 17.220, p = .016$), although only 7.1% (Nagelkerke $R^2$) of the variance in group (‘state’ delusions or control) was explained by the predictors overall. The RAVLT false recognition model was also statistically significant ($X^2 (7, n = 593) = 16.546, p = .000$), and explained only 6.9% (Nagelkerke $R^2$) of the variance in delusion group. False recognition was not a significant predictor in either model. Both models were unable to identify participants with delusions; they had 0.0% positive predictive value and classified all cases as controls.

For ‘trait’ delusions, neither model was statistically significant ($ps > .05$).

Significant models were re-run excluding those participants who were CN at baseline, with results remaining unchanged; both binary logistic regression models predicting presence of ‘state’ delusion remained significant overall ($ps < .05$). Neither false recognition on ADAS-Cog 13 or RAVLT were significant predictors in the models. Potential outliers were identified by visual inspection of scatter plots and box plots of false-recognition data ($n = 19$), see Figure 9.2. These were first inspected to exclude data entry error. Models were then re-run excluding these outliers. Both models remained significant ($ps < .05$), see Appendix 10.

For further detail of ‘state’ delusion model results, see Table 9.5.

### 9.4.1 Model diagnostics

Further model diagnostics were run for the two models which were statistically significant (i.e. the two models for ‘state’ delusion group).

VIF scores were all $< 2.5$, with a mean VIF of 1.2. The Box-Tidwell test was used to test the assumption of linearity between continuous variables and the logit transformation of the dependent variable, with all four models meeting this assumption. Both models also passed the Hosmer and Lemeshow test for goodness of fit ($ps > .05$).
Firth’s logistic regression, run to address the issue of ‘state’ delusions being a rare event in the group (n = 42 of total n = 594, 7.1%) yielded similar results; both models were significant (both ps < .05), with false recognition not reaching significance in either model.

Table 9.5 Analysis of the association between ‘state’ delusion and false recognition on ADAS-Cog-13 and RAVLT

<table>
<thead>
<tr>
<th>PREDICTOR</th>
<th>EXP(β) (95% CI)</th>
<th>P VALUE</th>
<th>EXP(β) (95% CI)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>FALSE RECOGNITION</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAS-Cog 13</td>
<td>1.062 (.952, 1.185)</td>
<td>.281</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RAVLT</td>
<td>-</td>
<td>-</td>
<td>1.047 (.929, 1.180)</td>
<td>.452</td>
</tr>
<tr>
<td>AGE</td>
<td>1.024 (.979, 1.070)</td>
<td>.304</td>
<td>1.022 (.978, 1.068)</td>
<td>.341</td>
</tr>
<tr>
<td>GENDER</td>
<td>.879 (.445, 1.735)</td>
<td>.710</td>
<td>.853 (.434, 1.677)</td>
<td>.645</td>
</tr>
<tr>
<td>YEARS OF EDUCATION</td>
<td>1.060 (.937, 1.199)</td>
<td>.355</td>
<td>1.058 (.936, 1.197)</td>
<td>.365</td>
</tr>
<tr>
<td>ChEI</td>
<td>1.011 (.439, 2.331)</td>
<td>.979</td>
<td>1.025 (.445, 2.362)</td>
<td>.954</td>
</tr>
<tr>
<td>MMSE</td>
<td>.863 (.756, .986)</td>
<td><strong>.030</strong></td>
<td>.858 (.752, .979)</td>
<td><strong>.023</strong></td>
</tr>
<tr>
<td>CATEGORY FLUENCY</td>
<td>.951 (.885, 1.022)</td>
<td>.170</td>
<td>.950 (.884, 1.020)</td>
<td>.158</td>
</tr>
</tbody>
</table>

Notes:
All measures at baseline.
Model one: ‘State’ delusion with false recognition on ADAS-Cog 13; Model two: ‘State’ delusion with false recognition on RAVLT.
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; RAVLT = Rey’s Auditory Verbal Learning Test. Category fluency for animals only.
10. False Memories and Delusions in the ADNI Cohort: Neuroimaging Methods

10.1 Background

As described in Chapter 7, recruitment and data collection for the clinical study stopped on 16th March 2020 due to the COVID-19 pandemic. Eight participants had completed MRI scanning by this point. Given the larger sample size of the ADNI cohort, I made the decision to analyse neuroimaging data from ADNI first, to inform the analysis for the patient-based study.

As described in section 8.2.4, ADNI has ethical approval through all involved institutional Research Ethics Committees.

10.2 Research questions

The primary research questions were:

- Do false memories correlate with specific patterns of lower regional brain volume across all AD participants (regardless of delusion group membership)?
- Do AD participants with either delusions at baseline (‘state’ delusions) or delusions at any time point (‘trait’ delusions) during ADNI follow-up show different patterns of lower regional brain volume than those without delusions?

The secondary research questions were:

- Do AD participants with delusions (either ‘state’ or ‘trait’) show evidence of lower GM volume in areas associated with false memory for the whole participant group?
- Do participants with different subtypes of psychosis symptoms have different patterns of reduced GM volume?

My a priori hypotheses, developed from the existing literature regarding delusion and false memory generation, as discussed in Chapters 2 and 3, were as follows:

- Across ADNI participants, regardless of delusion group membership, increased false recognition would correlate with reduced volume of MTL structures, ventral visual stream structures (entorhinal cortices, PHG, FFG, lingual gyri), PFC structures, ACC and superior parietal lobules
- Hippocampal, entorhinal, mPFC and dIPFC volumes would be reduced in those with delusions
- Volumes of ventral visual stream structures (as listed above) would be reduced in those with delusions
• Participants with misidentification-type symptoms would have reduced GM volume in the ventral visual stream than those with mixed or paranoid-type psychosis symptoms

10.3 Participant selection

Participants were selected from within the ADNI cohort in accordance with inclusion/exclusion criteria listed below. Participants also met all ADNI inclusion/exclusion criteria laid out in section 8.1.1.1.

Inclusion criteria:

4) Diagnosis of possible or probable AD based on NINCDS-ADRDA criteria at any ADNI timepoint

5) Recruited to either ADNI 1, GO, 2 or 3

6) Completed baseline NPI or NPI-Q

7) Completed baseline MRI

Exclusion criteria:

2) Inconsistent diagnosis of AD (if participant reverted to a diagnosis of CN or aMCI at a later visit)

The above criteria were chosen to maximise sample size. While an individual may be nominally ‘cognitively normal’ on recruitment to ADNI, this includes individuals with MMSE scores of 24 and over, and as such, those who later meet criteria for a diagnosis of AD were assumed for the purpose of this study to have AD pathology present at baseline. A subgroup analysis removing those who were CN at baseline was carried out to confirm that this did not significantly alter any observed findings.

10.4 MRI protocol

As described in section 8.2, ADNI recruited from 59 sites across the US and Canada in four phases (ADNI1, ADNI-GO, ADNI2 and ADNI3). Scanner manufacturers and hardware (for example, head coils) differed between sites, with a large variety of scanner platforms included in the dataset. The MR protocol was also changed between ADNI phases. All participants recruited to ADNI1 had 1.5T imaging, with 25% also having the same protocol at 3T field strength. In ADNI-GO, ADNI2 and ADNI3, all imaging was done on 3T scanners, see acquisition protocols in Table 10.1.\textsuperscript{420}
Table 10.1 Acquisition protocols across ADNI phases

<table>
<thead>
<tr>
<th>SCAN #</th>
<th>ADNI1 (1.5T SCANNER)</th>
<th>ADNI-GO/ADNI2 (3T SCANNER)</th>
<th>ADNI 3 (3T SCANNER)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Localiser</td>
<td>Localiser</td>
<td>Localiser</td>
</tr>
<tr>
<td>2</td>
<td>MP-RAGE</td>
<td>MP-RAGE/IR-SPGR</td>
<td>Accelerated MP-RAGE</td>
</tr>
<tr>
<td>3</td>
<td>MP-RAGE repeat</td>
<td>Accelerated MP-RAGE/IR-SPGR</td>
<td>Sagittal 3D FLAIR</td>
</tr>
<tr>
<td>4</td>
<td>B1 Calibration – Head coil</td>
<td>Resting state fMRI</td>
<td>Axial T2-star</td>
</tr>
<tr>
<td>5</td>
<td>B1 Calibration – Body coil</td>
<td>Axial T2 FLAIR</td>
<td>Axial 3D PASL</td>
</tr>
<tr>
<td>6</td>
<td>T2 dual echo</td>
<td>Axial T2-star</td>
<td>Axial DTI</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>Axial ASL Perfusion</td>
<td>Field mapping</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>Axial DTI</td>
<td>Axial rsfMRI</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>HighResHippocampus</td>
</tr>
</tbody>
</table>

Notes:
ADNI = Alzheimer’s Disease Neuroimaging Institute; ASL = arterial spin labelling; DTI = diffusion tensor imaging; FLAIR = fluid-attenuated inversion recovery; fMRI = functional magnetic resonance imaging; IR-SPGR = inversion recovery spoiled gradient-echo; MP-RAGE = magnetisation-prepared rapid gradient-echo; PASL = pulsed arterial spin labelling; rsfMRI = resting state functional magnetic resonance imaging.

I decided to include structural MRI data from participants across all sites and phases, in order to maximise the number of participants with delusions at baseline and increase statistical power. Although this introduced heterogeneity due to differences in scanning protocols, previous VBM analyses using structural MR imaging from AD participants across multiple sites and scanners found only minimal confounding of results, with significant effect of scanner and field strength found only in cerebellum, precentral gyri and thalamus.421, 422

ADNI provides image data either in raw form or following ADNI specific preprocessing. Raw data were chosen as they were more comparable with data from my patient-based study.

All ADNI1 participants had repeated imaging at each visit, see Table 10.1. For consistency, I chose the first of these two scans unless it was judged of insufficient quality due to evident head motion on subjective review during scan reorientation or on subsequent automated image quality assessment using the Computational Anatomy Toolbox (CAT12, version 7 (r1725); http://www.neuro.uni-jena.de/cat/), see section 10.5.4. In this case, data from repeated imaging were used. Fully sampled images were chosen where available, with accelerated imaging used if not.
Structural MRI scans completed at baseline for ADNI participants were 3D T1 weighted sagittal images using either magnetisation-prepared rapid gradient-echo (MP-RAGE) or inversion recovery spoiled gradient-echo techniques. Baseline MRI was available for 729 (99.5%) of the 733 participants. Table 10.2 summarises scan details for participants across the different study phases.

**Table 10.2 Participant scan details (n = 729)**

<table>
<thead>
<tr>
<th>PROTOCOL</th>
<th>FIELD STRENGTH</th>
<th>SAMPLING PATTERN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.5T</td>
<td>3T</td>
</tr>
<tr>
<td>MP-RAGE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR-SPGR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADNI1 (n = 412)</td>
<td>412 (56.5)</td>
<td>-</td>
</tr>
<tr>
<td>ADNI-GO/ADNI2 (n = 260)</td>
<td>186 (25.5)</td>
<td>75 (10.3)</td>
</tr>
<tr>
<td>ADNI3 (n = 57)</td>
<td>40 (5.5)</td>
<td>17 (2.3)</td>
</tr>
</tbody>
</table>

Notes:
Values are expressed as n (%).
ADNI = Alzheimer’s Disease Neuroimaging Institute; IR-SPGR = inversion recovery spoiled gradient-echo; MP-RAGE = magnetisation-prepared rapid gradient-echo.

Variability between protocols included scanner manufacturer (Siemens, Philips or GE systems) and hardware. Slice thickness varied from 1.0 – 1.2mm, the repetition time varied from 6.5 – 2300ms, the echo time from 2.8 – 4.4ms and the flip angle from 8 – 11 degrees.

### 10.5 Image preprocessing

#### 10.5.1 Setting the origin

The origin of the ADNI structural images, i.e. the point at which x, y and z dimensions are all zero, was set as the centre of the scanner. Preprocessing involved normalising images to MNI space by mapping each brain to the generic MNI template. This reduced variation in size and orientation such that equivalent anatomical regions in each brain image were within the same voxels to allow comparison between images. The origin of the MNI template is set to the anterior commissure. Prior to image preprocessing the origin for each ADNI scan was manually reset to the anterior commissure using the ‘Display’ function in Statistical Parametric Mapping software (SPM12; [https://www.fil.ion.ucl.ac.uk/spm/software/spm12/]), to reduce error during normalisation.
10.5.2 Manual image quality control

As setting the origin involved visual review of all scans, this was used as an initial image quality control. Two scans were identified as having high degrees of likely movement artifact, and for these participants the repeat MP-RAGE scans were used.

10.5.3 Preprocessing

Preprocessing was completed using the CAT12 toolbox for SPM12, run through MATLAB. The scans for the first 10 participants (by ADNI participant number) were preprocessed using CAT12 default settings, with the Geodesic Shooting (GS) approach for image registration as this method is thought to be potentially superior to the Dartel approach for computational anatomy studies. Any variations from CAT12 defaults are discussed with reasoning in sections 10.5.3.1 - 10.5.3.12.

Once segmentation and skull stripping appeared to be optimised on visual inspection of the preprocessed subgroup of scans, preprocessing was run for the whole participant group using a batch script (the generic cat_batch_cat.sh, calling on an edited cat_defaults.m file, see Appendix 11). Due to the large number of scans the script was run on the UCL Computer Science High Performance Computing cluster. An additional shell script was written in Bash to run the job as an array job on the cluster, see Appendix 12. This original script was written by Dr Ian Malone.

The preprocessing pathway involved the following stages:

10.5.3.1 Noise reduction
The first preprocessing stage in CAT12 is noise reduction using a spatial-adaptive Non-Local Means (SANLM) denoising filter, which removes noise from the images but maintains edges.

10.5.3.2 Internal resampling
Internal resampling then occurs, which standardises the voxel size and thereby accommodates variety in image resolution. Internal resampling was kept to the CAT12 default of ‘optimal’, which finds an optimal resolution based on median voxel size.

10.5.3.3 Bias correction
The bias correction algorithm removes bias artefact introduced by transmit/receive inhomogeneities of the scanner, which would otherwise impede automated tissue segmentation and registration. The algorithm was changed from default to ‘full’, to optimise registration for images with a high degree of atrophy. The ‘full’ bias correction uses an iterative process, followed by a final maximum-based filter to remove remaining local inhomogeneities.
10.5.3.4 **Affine registration**
Affine registration uses mutual information,\(^\text{429}\) which brings images into the same anatomical space by aligning them to a template consisting of a smoothed average of the T1 scans of 152 healthy European adults (the International Consortium for Brain Mapping template).

10.5.3.5 **Unified segmentation**
Standard SPM12 ‘unified segmentation’ segments voxels into GM, white matter or cerebrospinal fluid based on voxel intensity, and creates tissue probability maps (TPMs) to determine which tissue class a voxel is likely to belong to, based on its position in space.

10.5.3.6 **Skull stripping**
Skull stripping (to remove the skull from images) was initially run using the default ‘SPM’ skull stripping method. However, on visual inspection of the resulting images in the initial 10 scan subgroup this option was too aggressive and led to incorrect removal of GM. As a result, a graph-cut (GCUT) skull stripping approach was used, with the default of 0.5 strength. The GCUT approach uses intensity thresholding alongside an approach making cuts using a graph theoretic framework.\(^\text{430}\)

10.5.3.7 **Local intensity transformation**
GM intensity varies by region for anatomical reasons (for example, regional iron content or myelinisation), and depends on MRI protocol. To correct for this, CAT12 applies an intensity transformation to correct for variability in intensity across all tissue classes.\(^\text{431}\)

10.5.3.8 **Refined segmentation**
A further segmentation process then occurs, based on an Adaptive Maximum A Posterior technique,\(^\text{432}\) in which an adaptive segmentation process occurs which does not rely on TPMs and therefore accounts for variations in intensity. This further segmentation process also includes a partial volume correction, which estimates the fraction of each pure tissue type present in each voxel – acknowledging that due to size each voxel is unlikely to contain only one tissue type.\(^\text{433}\)

10.5.3.9 **Spatial registration**
Images are then aligned to a common reference template using either Dartel or the GS approach,\(^\text{427}\) with the GS approach chosen for the reasons described in section 10.5.3. It is not possible to complete further ROI analysis in CAT12 if a sample-specific template is used at this point, and therefore the existing GS template (in MNI space) was used, which was developed from 555 healthy control subjects in the IXI-database.\(^\text{434}\) This template is considered appropriate for all subjects apart from children and has been previously used in ROI and whole brain VBM analyses in AD participants.\(^\text{431, 435, 436}\) Modulation of images then occurs using the Jacobian
determinant. This process preserves the original voxel size and therefore the original GM volume prior to shrinking or expanding to match the template.

10.5.3.10  Brain parcellation
Tissue volumes (i.e. for white matter, GM and cerebrospinal fluid) in ml were estimated for ROIs using the volume-based ‘Neuromorphometrics’ atlas. This is a maximum probability map provided by Neuromorphometrics, Inc. under academic supervision.\(^ {437} \) This atlas was developed by neuroanatomical technicians, who manually traced ROIs on structural MR images of 30 healthy control subjects from the OASIS database and registered them to MNI space.\(^ {437-439} \) The atlas includes 140 ROIs in GM only.

10.5.3.11  Output
Modulated, warped and partial volume segmented image files were saved for each tissue class. GM image files were then smoothed, to compensate for variability in structural characteristics and imperfect co-registration and to reduce the number of independent comparisons in the VBM analysis. Images were smoothed using the ‘Smooth’ function in SPM12 at default (8mm FWHM of the Gaussian smoothing kernel) for use in the whole brain VBM analysis described below.

10.5.4  Further quality check
A further image quality check was then completed using the ‘Check sample’ function in CAT12 to assess data homogeneity and by checking the image quality measures which are included in output following image preprocessing in CAT12.

10.6  Statistical analyses

10.6.1  Behavioural data and normality testing
Demographic details, neuropsychological covariates and false memory measures were downloaded and processed as per the behavioural analysis (see section 8.5). Behavioural data analyses and ROI analyses were carried out in SPSS. Whole brain VBM analyses were carried out in SPM12. For all analyses results were considered significant at a \( p \) value of < .05 unless otherwise stated.

Behavioural and ROI data were viewed in the form of scatter plots, histograms and Q-Q plots to assess whether they followed a normal distribution. The assumption of normality was further tested using the Kolmogorov-Smirnov test for non-normality.

Baseline characteristics of the population were compared between those with delusions at baseline (‘state’ delusion) and the control group, and between those with delusions at any time
point (‘trait’ delusion) and the control group using independent-samples t-tests or Mann-Whitney non-parametric tests for parametric and non-parametric data respectively and chi-squared tests for categorical variables.

10.6.2 False memory measures

10.6.2.1 Regions of interest analyses

ROIs were selected from those available in the Neuromorphometric atlas based on hypotheses described in section 10.2 and included (bilateral): hippocampi, entorhinal cortices, PHG, FFG, lingual, MFG, superior frontal (SFG), IFG, inferior frontal orbital (IFOG), inferior frontal angular (IFAG), superior medial frontal (SMFG), lateral orbital (LOG) and anterior cingulate gyri, medial frontal cerebra (MFC) and superior parietal lobules. ROIs were corrected for global brain volume, by calculating each ROI as a proportion of total intracranial volume (TIV). The relationships between each of the three measures of false memory (total intrusions on the RAVLT, false recognition on the RAVLT and the ADAS-Cog 13, see section 8.5.3) and the TIV-corrected volume of each individual ROI were visualised using scatter plots to assess for any correlation (either linear or monotonic) between these variables.

Correlations between variables were then further explored using generalised linear regression modelling. Potential confounding variables were chosen a priori. These were: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription and category fluency for animals (see section 5.5.3) and MRI field strength, as this is known to bias the sensitivity of MRI to atrophy. In order to allow a more meaningful interpretation of regression coefficients, in regression analyses ROIs were further transformed by multiplying by a factor of $10^3$. This transformation meant that a one unit change in ROI volume in the regression model was the equivalent to a change in ROI volume of 0.01% of TIV.

The primary outcome false memory measures (intrusions on RAVLT, false recognition on RAVLT and false recognition on ADAS-Cog 13) are numeric count data with a Poisson distribution (histograms that were unimodal with a positive skew). Individual Poisson regression models were run for each of the three measures with individual ROIs included as independent variables, and potential confounders included as covariates. Covariates were assessed for collinearity using VIF scores, with scores above 10.0 suggestive of a high degree of collinearity. A more conservative threshold of 2.5 was used for the purpose of the following analyses.

The model fit was assessed using the Pearson $X^2/df$ value as a measure of overdispersion (where one represents equidispersion). A Pearson $X^2/df$ value $> 1.5$ was considered to indicate a degree of overdispersion at which use of an unadjusted Poisson model was likely to lead to greater error. For models with values above this cutoff a Poisson model accounting for
overdispersion or a negative binomial model were considered as alternatives. Log-likelihood values were used to compare the fit between different models, with less negative log-likelihood values indicating better model fit.

Model diagnostics were run for any ROIs with significant results. Residuals were examined by creating scatter plots of standardised Pearson residuals against predicted values. If there was a constant spread of residuals across predicted values and the majority of standardised Pearson residuals fell between -2.0 and 2.0, the models were considered to be a good fit.

To assess the impact of outlying values, for any ROIs with significant results, Cook’s distance, which combines leverage and residual values to provide a measure of influence of each data point on the model, was calculated and visualised in bar charts. Various thresholds for Cook’s distance can be used, with no one universally accepted cut-off: these range from one to the more conservative $4/(n-k-1)$, where $n$ = number of participants and $k$ = number of variables in the model. Cook’s distance was checked and values that exceeded the more liberal threshold of one were checked for user error on data entry. If no error was identified, the models were re-run excluding these values, to assess their impact on results. The process was repeated using the more conservative threshold of $4/(n-k-1)$.

Finally, a subgroup analysis explored the impact of including CN participants. For any ROIs with significant findings, models were re-run excluding those participants who were CN at the time of MRI.

A power calculation was carried out using G*Power, based on a previous study which explored the relationship between hippocampal atrophy (volume of hippocampal formation corrected for TIV) and intrusions on the California Verbal Learning Test and found a significant correlation between the two ($r = 0.54$, $p = .02$, slope = -18.5 (7.3), no covariates). This gave a sample size of 39 (power = 0.8) to detect a significant ($p < .05$) relationship between the false memory outcome and eight covariates, a number which is exceeded by this dataset. As a result, correction for multiple comparisons was not performed.

10.6.2.2 Whole brain analysis
An exploratory whole brain analysis was completed by VBM using SPM12, to investigate the relationship between regional volume and false memory performance measures.

Multiple regression models were carried out for each false memory measure, using a between-subject second-level analysis in SPM12. Each model included the covariates specified in section 10.6.2.1, with the addition of TIV to control for overall participant brain volume. Continuous variables were centred around the mean to reduce multicollinearity. Collinearity was assessed
by creating a Pearson correlation matrix for continuous variables, with correlation coefficients < .80 considered acceptable, and using VIFs with a conservative threshold of < 2.5.

Explicit threshold masking was applied to reduce the number of multiple comparisons and remove voxels with a high likelihood of tissue type misclassification. A threshold mask was created using the ‘make_majority_mask’ function for SPM12 developed by Ridgway et al. (2009), to ensure that all GM was included for this participant group as areas of atrophy can be incorrectly excluded by the default SPM12 masking procedure. For each model a mask was created from the smoothed GM images, creating a majority mask to include voxels for which 80% of images had an intensity above 0.2. After creation, each mask was inspected using the SPM12 ‘Check Reg’ function to confirm it was not overly stringent and was a good fit for the data.

T-contrasts were run to explore both positive and negative correlation between performance on false memory measures and brain volume, see Figure 10.1, Figure 10.2 and Figure 10.3.

Multiple regression models were first run with a FWE correction applied, with a significance threshold of $p_{FWE} < .05$ and an extent threshold of $k = 10$ voxels. Exploratory models were run as a secondary analysis, with an uncorrected $p$ value of < .001 and the same extent threshold of $k = 10$ voxels. Significant results were displayed on warped, skull-stripped, bias and noise-corrected average brains, created using the SPM12 ‘ImCalc’ function. Unthresholded effect size maps were also created for each contrast applied.

Design orthogonality matrices were created and viewed for each model to further assess the degree of collinearity between variables as a proxy of model-fit.
Figure 10.1 Design matrix for multiple regression model including intrusions on RAVLT as a covariate

Notes:
* mean centered variables.
ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; RAVLT = Rey’s Auditory Verbal Learning Test; TIV = total intracranial volume.
Figure 10.2 Design matrix for multiple regression model including false recognition on RAVLT as a covariate

Notes:
* mean centered variables.
ChEI = cholineresterase inhibitor prescription; MMSE = Mini Mental State Examination; RAVLT = Rey's Auditory Verbal Learning Test; TIV = total intracranial volume.
**Figure 10.3** Design matrix for multiple regression model including false recognition on ADAS-Cog 13 as a covariate

Notes:
* mean centered variables.
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; TIV = total intracranial volume.
10.6.3 Delusion group

10.6.3.1 Regions of interest analyses

ROIs were selected from those available in the Neuromorphometric atlas based on hypotheses described in section 10.2. These were as described in section 10.6.2.1, with the exception of IFG, IFOG, IFAG, LOG and superior parietal lobules. Prior to analyses, ROIs were corrected for TIV as described in section 10.6.2.1. The relationships between ROI volumes (as a proportion of TIV) and delusion group (‘state’ or ‘trait’ vs control) were explored using independent-samples t-tests. As described in section 8.4.2, the control group was the same for both delusion groups; those who developed delusions at a later timepoint were removed from the control group for analyses of delusions at baseline.

To further explore this relationship, logistic regression models were run with each ROI as a predictor of delusion group (‘state’ or ‘trait’ vs control), alongside covariates defined a priori (age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency and MRI field strength) for reasons as described in section 5.5.3.

Binary logistic regression models were used to predict group status for both ‘state’ delusions or control and ‘trait’ delusions or control groups and included each individual ROI as the predictor variable and previously described confounding factors as covariates.

For the purposes of comparison with my previous finding, the relationship between delusion group and ventral visual stream ROIs was explored in a separate analysis, which included ventral visual stream ROIs bilaterally and the covariates described above. Relationships of bilateral PHG to delusion groups were explored using independent-samples t-tests and binary logistic regression.

Covariates were assessed for collinearity using VIFs, using the conservative threshold of 2.5 to indicate high collinearity. The logistic regression model assumption of linearity was tested using the Box-Tidwell test and model fit was assessed by reviewing outcomes for the Hosmer and Lemeshow goodness of fit test, as described in section 5.5.3. Firth’s logistic regression models were run to confirm that reducing the bias caused by the ‘state’ delusion outcome being rare did not alter the results, as described in section 8.5.

Significant findings in the relationships between delusion group and ROIs from the above analyses were interrogated by delusion subtype (paranoid-type, misidentification-type or mixed, see description in section 2.2.1). One-way analyses of variance (ANOVAs) were used to investigate the relationship between delusion subtype (at both baseline and at any timepoint) and ROI volumes. In addition to confirming that data met the ANOVA assumption of normality
(confirmed as described in section 10.6.1), the assumption of homogeneity of variance (i.e. that variance of the ROI volume was equal across the delusion subtypes) was tested using Levene’s test on which significant findings indicate that this assumption is violated. If Levene’s test was significant, Welch’s ANOVA was used, to reduce type one error rates. Multinomial regression models were also run, with significant ROIs as predictors of delusion subtype, including the covariates described above.

Finally, a subgroup analysis was run to explore the impact of including CN participants. For any ROIs with significant findings, models were re-run excluding those participants who were CN at the time of MRI.

A power calculation (completed using G*Power), based on the difference in PHG volume observed between those with and without psychosis in an ROI analysis of the AddNeuroMed cohort, suggested a sample size of ≥ 22 per group (power = 0.8) would detect a significant difference between groups ($p < .05$). As this sample size was larger, for both ‘state’ and ‘trait’ delusion group comparisons with control, correction for multiple comparisons was not performed.

10.6.3.2 Whole brain analysis
An exploratory whole brain VBM analysis was carried out in SPM12, to further identify any associations between regional volume and delusion group.

Independent-samples t-tests were run for both ‘state’ delusion and ‘trait’ delusion groups compared to control group as between-subject second-level analyses in SPM12. These were first run without covariates, and then run to include the same covariates specified in section 10.6.3.1, with the addition of TIV to control for overall participant brain volume. As described in section 10.6.2.2, continuous variables were mean centred to reduce multicollinearity in the model. The degree of collinearity in the models was assessed by creating a Pearson correlation matrix for continuous variables, with correlation coefficients < .80 considered acceptable and using VIFs with a conservative threshold of < 2.5.

Explicit threshold masking was applied during the analysis, with a different mask created for both ‘state’ and ‘trait’ delusion group comparisons, using the same process as described in section 10.6.2.2. After creation, each mask was inspected using the SPM12 ‘Check Reg’ function to confirm it was not overly stringent and was a good fit for the data.

T-contrasts were run to identify areas of reduced volume in those with delusions compared to control: ‘state’ delusion group < control and ‘trait’ delusion group < control, see Figure 10.4 and Figure 10.5.
Multiple regression models were run with a FWE correction applied, with a significance threshold of $p_{FWE} < .05$ and an extent threshold of $k = 10$ voxels. More exploratory models were run as a secondary analysis, with an uncorrected $p$ value of $< .001$ and the same extent threshold of $k = 10$ voxels. Any significant results were displayed on warped, skull-stripped, bias and noise-corrected average brains, created using the SPM12 ‘ImCalc’ function.

Design orthogonality matrices were created and viewed for each model to further assess the degree of collinearity between variables as a proxy of model-fit.
**Figure 10.4** Design matrix for 'state' delusion < control t-test, with covariates

**Notes:**
* mean centered variables.
Group₁ = 'state' delusions; Group₂ = control.
ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; TIV = total intracranial volume.
Figure 10.5 Design matrix for ‘trait’ delusion < control t-test, with covariates

Notes:
* mean centered variables.
Group 1 = ‘trait’ delusions; Group 2 = control.
ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; TIV = total intracranial volume.
11. False Memories and Delusions in the ADNI Cohort: Results of Analyses of Neuroimaging Data

As described in Chapter 10, all demographic and behavioural data were downloaded from the ADNI data archive (http://adni.loni.usc.edu/data-samples/access-data/) on 24th June 2020. Participant diagnostic information had been downloaded for screening and selection on 26th March 2020. Participant MRI scans (MP-RAGE, prior to preprocessing through ADNI pipelines) were downloaded on 15th December 2020.

11.1 Participant selection

Of the 733 participants included in the behavioural analyses, 728 had baseline MRI scans available.

Figure 11.1 Participant selection flow chart

Notes:
AD = Alzheimer’s disease; ADNI = Alzheimer’s Disease Neuroimaging Initiative; CN = cognitively normal, healthy elderly controls; GDS-15 = short form Geriatric Depression Scale; MCI = mild cognitive impairment; MRI = magnetic resonance imaging; PD = Parkinson’s disease; ROI = region of interest; VBM = voxel-based morphometry.
11.2 Demographic details

Participants in the sample had a mean age of 74.8 ± 7.4 years, 317 (43.5%) were female and 411 (56.5%) were male and the majority (n = 681, 93.5%) were white. 386 (53.0%) had an AD diagnosis at baseline and the majority of participants (n = 585, 80.4%) had received an AD diagnosis by the end of the second year of follow-up. The mean baseline MMSE score for the group was 25.1 ± 2.9.

Similar to the behavioural data analyses described in Chapter 10, the sample was split into those with ‘state’ delusions (delusions present at baseline, n = 42) and those with ‘trait’ delusions (to include the ‘state’ group and those with delusions present at any later time point, n = 179). Both of these groups were compared to a control group of participants who did not develop delusions at any time point during the ADNI study period (n = 549).

Of the five participants who did not have baseline neuroimaging data, the mean age was 76.9 ± 13.1 years, one participant (20.0%) was female and four (80.0%) were male and the majority (n = 5, 100.0%) were white. Compared to the total participant population, a higher proportion (n = 4, 80.0%) had a diagnosis of AD at the baseline visit, but they had a similar mean baseline MMSE score of 25.6 ± 4.6. None of the participants without baseline imaging had delusions at baseline or at any later timepoint.

Demographic details and baseline screening assessments for the three groups are summarised in Table 11.1.

For the ‘state’ delusion group, there were no significant differences from the control group in terms of age, gender, primary language, handedness, years of education, use of cholinesterase inhibitor or antidepressant medication or GDS-15 score. However, the ‘state’ delusion group did differ significantly from the control group in race, diagnosis at baseline, years since diagnosis, antipsychotic prescribing, NPI and NPI-Q score, presence of hallucinations, MMSE score and category fluency for animals. The control group included a lower proportion of individuals with a diagnosis of AD at baseline and had fewer years since diagnosis. Individuals in the control group were prescribed fewer antipsychotics, had higher scores on the MMSE and had a lower burden of neuropsychiatric symptoms as measured by the NPI and NPI-Q.

The ‘trait’ delusion and control groups did not differ significantly in terms of age, race, primary language, handedness, years of education, diagnosis at baseline, years since diagnosis, cholinesterase inhibitor or antidepressant prescriptions, MMSE or GDS-15 score. There was a higher proportion of females in the ‘trait’ delusion compared to the control group, and similar to the ‘state’ group there were higher rates of antipsychotic prescribing and neuropsychiatric
symptoms, including hallucinations, on the NPI and NPI-Q in the ‘trait’ delusion compared to the control group.
Table 11.1 Demographic details and screening results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 549)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>P VALUE (‘STATE’ / CONTROL)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th>P VALUE (‘TRAIT’ / CONTROL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>74.8 ± 7.5;</td>
<td>76.0 ± 6.4;</td>
<td>t_{589} = 1.007, p = .314</td>
<td>75.0 ± 7.2;</td>
<td>t_{726} = -.306, p = .760</td>
</tr>
<tr>
<td></td>
<td>55 - 90</td>
<td>64 - 90</td>
<td></td>
<td>55 - 91</td>
<td></td>
</tr>
<tr>
<td>GENDER (female)</td>
<td>224 (40.8)</td>
<td>18 (42.9)</td>
<td>X^2 (1, n = 591) = .068, p = .794</td>
<td>93 (52.0)</td>
<td>X^2 (1, n = 728) = 6.831, p = .009</td>
</tr>
<tr>
<td>RACE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>519 (94.5)</td>
<td>36 (85.7)</td>
<td></td>
<td>162 (90.5)</td>
<td></td>
</tr>
<tr>
<td>Black/African American</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>13 (2.4)</td>
<td>4 (9.5)</td>
<td>† p = .037</td>
<td>12 (6.7)</td>
<td>† p = .058</td>
</tr>
<tr>
<td>Mixed race</td>
<td>12 (2.2)</td>
<td>2 (4.8)</td>
<td></td>
<td>4 (2.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5 (0.9)</td>
<td>0 (0.0)</td>
<td></td>
<td>1 (0.6)</td>
<td></td>
</tr>
<tr>
<td>PRIMARY LANGUAGE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>535 (97.4)</td>
<td>41 (97.6)</td>
<td>† p = .674</td>
<td>175 (97.8)</td>
<td>† p = .443</td>
</tr>
<tr>
<td>Spanish</td>
<td>9 (1.6)</td>
<td>1 (2.4)</td>
<td></td>
<td>4 (2.2)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>5 (0.9)</td>
<td>0 (0.0)</td>
<td></td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>RIGHT HANDEDNESS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 590/727)</td>
<td>504 (92.0)</td>
<td>38 (90.5)</td>
<td>† p = .767</td>
<td>169 (94.4)</td>
<td>X^2 (1, n = 727) = 1.171, p = .327</td>
</tr>
<tr>
<td>EDUCATION (years)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15.6 ± 2.8;</td>
<td>15.6 ± 2.8;</td>
<td>t_{589} = .115, p = .908</td>
<td>15.4 ± 2.9;</td>
<td>t_{726} = .892, p = .372</td>
</tr>
<tr>
<td></td>
<td>6 - 20</td>
<td>8 - 20</td>
<td></td>
<td>4 - 20</td>
<td></td>
</tr>
<tr>
<td>DIAGNOSIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CN</td>
<td>21 (38.3)</td>
<td>0 (0.0)</td>
<td>† p = .003</td>
<td>6 (3.4)</td>
<td>X^2 (2, n = 728) = 2.751, p = .253</td>
</tr>
<tr>
<td>aMCI</td>
<td>228 (41.5)</td>
<td>8 (19.0)</td>
<td></td>
<td>87 (48.6)</td>
<td></td>
</tr>
<tr>
<td>AD</td>
<td>300 (54.6)</td>
<td>34 (81.0)</td>
<td></td>
<td>86 (48.0)</td>
<td></td>
</tr>
<tr>
<td>YEARS SINCE DIAGNOSIS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 572/708)</td>
<td>0.3 ± 4.0;</td>
<td>2.2 ± 2.6;</td>
<td>t_{54,668} = 4.320, p = .000</td>
<td>0.3 ± 4.1;</td>
<td>t_{706} = -.055, p = .956</td>
</tr>
<tr>
<td></td>
<td>-13.0 – 12.3</td>
<td>-4.0 - 7.8</td>
<td></td>
<td>-12.0 - 13.8</td>
<td></td>
</tr>
</tbody>
</table>
Table 11.1 Cont. Demographic details and screening result

<table>
<thead>
<tr>
<th>MEDICATION (n = 590/727)</th>
<th>CONTROL (n = 549)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>P VALUE (‘STATE’/CONTROL)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th>P VALUE (‘TRAIT’ / CONTROL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ChEI/Memantine</td>
<td>400 (73.0)</td>
<td>34 (81.0)</td>
<td>$X^2 (1, n = 590) = 1.271$, $p = .283$</td>
<td>123 (68.7)</td>
<td>$X^2 (1, n = 727) = 1.223$, $p = .269$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antidepressant</td>
<td>179 (32.7)</td>
<td>17 (40.5)</td>
<td>$X^2 (1, n = 590) = 1.073$, $p = .311$</td>
<td>72 (40.2)</td>
<td>$X^2 (1, n = 727) = 3.411$, $p = .065$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>4 (0.1)</td>
<td>4 (9.5)</td>
<td>$^p = .001$</td>
<td>6 (3.4)</td>
<td>$^p = .017$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPI-Q</td>
<td>2.4 ± 2.7; 0 - 15</td>
<td>7.9 ± 4.9; 1 - 24</td>
<td>$U = 3032.0, p = .000$</td>
<td>4.3 ± 4.3; 0 - 24</td>
<td>$U = 35917.0, p = .000$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NPI (n = 256/302)</td>
<td>6.1 ± 7.5; 0 - 35</td>
<td>19.5 ± 12.1; 3 - 46</td>
<td>$U = 712.5, p = .000$</td>
<td>11.2 ± 11.4; 0 - 46</td>
<td>$U = 5734.5, p = .001$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HALLUCINATIONS</td>
<td>Baseline</td>
<td>8 (1.5)</td>
<td>$^p = .000$</td>
<td>16 (8.9)</td>
<td>$X^2 (1, n = 728) = 23.699, p = .000$</td>
</tr>
<tr>
<td></td>
<td>Any time point</td>
<td>45 (8.2)</td>
<td>$^p = .000$</td>
<td>70 (39.1)</td>
<td>$X^2 (1, n = 728) = 96.957, p = .000$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>25.0 ± 2.9; 18 - 30</td>
<td>23.4 ± 2.7; 17 - 29</td>
<td>$t_{589} = -3.537, p = .000$</td>
<td>25.1 ± 2.9; 17 - 30</td>
<td>$t_{726} = -.102, p = .919$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDS-15</td>
<td>1.5 ± 1.4; 0 - 5</td>
<td>1.7 ± 1.5; 0 - 5</td>
<td>$U = 10948.0, p = .574$</td>
<td>1.6 ± 1.4; 0 - 5</td>
<td>$U = 48397.0, p = .755$</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CATEGORY FLUENCY (n = 590/727)</td>
<td>14.0 ± 5.4; 0 - 35</td>
<td>11.5 ± 4.5; 1 - 22</td>
<td>$t_{588} = 2.919, p = .004$</td>
<td>13.8 ± 4.9; 1 - 26</td>
<td>$t_{725} = .409, p = .683$</td>
</tr>
</tbody>
</table>
Table 11.1 Cont. Demographic details and screening result

Notes:
Values are expressed as mean ± SD; range or n (%). All measures at baseline unless otherwise specified.
ChEI = cholinesterase inhibitor; CN = cognitively normal; GDS-15 = short form Geriatric Depression Scale; MMSE = Mini Mental State Examination; NPI = Neuropsychiatric Inventory.
Category fluency is for animals only.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data. Categorical comparisons by chi-squared unless otherwise specified.
† p value from Fisher’s exact test.
11.3 Image quality

Homogeneity in the sample was reasonable, with the majority of scans having a mean correlation coefficient of > .80 (see Figure 11.2). Thirty-three images were flagged as having a mean overall correlation over two SD from the mean, although only five of these had correlation coefficients < .80. All of these images were checked visually for artefact. One scan, which had the lowest correlation coefficient of .75, had a high degree of movement artefact. All other scans were considered to be of acceptable quality.

Scans also received quality ratings for noise (noise contrast ratio), bias (inhomogeneity contrast ratio) and resolution, which were combined as a weighted average image quality rating (image quality rating, IQR). IQRs are expressed as a grade from A-F and a percentage, with a rating of F equivalent to a score of < 50.0%. A low IQR is found to lead to underestimation of GM. All images received IQR grades between B and C+, and a mean score of 83.1 ± 2.3%, apart from one scan which was rated F, and scored 42.3%. This was the same scan that had notable movement artefact on inspection and given these significant quality concerns this scan was removed from further analysis.
Figure 11.2 Data homogeneity - mean correlation of imaging data

Notes:
Output from CAT12 'VBM Data Homogeneity' function.
11.4 False memory measures

11.4.1 Regions of interest analysis

The relationships between volume of previously specified ROIs corrected for TIV and measures of false memory (total intrusions on the RAVLT, false recognition on the RAVLT and the ADAS-Cog 13) were visualised using scatter plots. No direct (linear or monotonic) relationships were identified. The relationships were therefore further explored using regression modelling. Prior to running the models, the volumes of ROIs (as a proportion of TIV) were multiplied by $10^3$ to allow for meaningful interpretation of the regression coefficients (exponentiated $\beta$ values). This transformation meant that a one unit increase in ROI volume in the regression model equated to an increase in ROI volume of 0.01% of TIV. For all ROIs explored in the analysis, mean volume as a percentage of TIV was $0.32 \pm 0.26\%$; mean left hippocampal volume was $0.14 \pm 0.03\%$. Transforming by a factor of $10^3$ meant that a one unit change in the model was less than one SD from the mean ROI volume and was equivalent to (for example) a 7.1% change in mean left hippocampal volume.

Confounding variables were defined a priori (age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency and MRI field strength) and assessed for collinearity in each model, with transformed, TIV-corrected left hippocampal volume as a representative measure of ROI. The VIF scores were all $< 2.5$, with a mean VIF of 1.2.

Reviewing the distribution of results for all three false memory measures in histogram plots revealed a likely Poisson distribution. Data for modelling predictors of total intrusions on the RAVLT were overdispersed (mean Pearson $X^2$/df value of $6.6 \pm 0.0$, where one represents equidispersion). Poisson modelling was therefore run with correction for overdispersion, using the Pearson $X^2$/df value as the scale parameter in the model as described by McCullagh and Nelder (1989). Negative binomial modelling was considered but showed a worse fit for the data with more negative log-likelihood values. Poisson modelling accounting for overdispersion using the method described by McCullagh and Nelder (1989) was also used for false recognition on both the RAVLT and the ADAS-Cog 13 (mean Pearson $X^2$/df values of $2.2 \pm 0.0$ and $2.7 \pm 0.0$ respectively). Negative binomial modelling was explored, but again led to more negative log-likelihood values.

All of the models that assessed the relationship between ROIs and intrusions on the RAVLT were significant ($ps < .01$), but no ROI reached individual significance within any model. Similarly, all of the models that assessed the relationship between ROIs and false recognition on both the RAVLT and the ADAS-Cog 13 were significant ($ps < .005$). Increased volume of several ROIs was
linked to fewer words falsely recognised on both RAVLT and ADAS-Cog 13. See Table 11.2 for full details.

11.4.1.1 Medial temporal lobe regions of interest
Four MTL ROIs were significantly related to false recognition performance on both the RAVLT and the ADAS-Cog 13: bilateral hippocampi and bilateral entorhinal cortices. ROI volume was inversely related to number of words falsely recognised, see Figure 11.3 and Table 11.2.

For every 0.01% increase in volume (as a proportion of TIV) of the left hippocampus there was a 4.0% reduction in false recognition on the RAVLT and 9.1% on the ADAS-Cog 13 (Exp(β) .960, 95% CI .932 - .989, p = .008; Exp(β) .909, 95% CI .882 - .937, p = .000 respectively); right hippocampus there was a 6.0% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .940, 95% CI .915 - .965, p = .000); left entorhinal cortex there was a 3.2% reduction in false recognition on the RAVLT and 6.1% on the ADAS-Cog 13 (Exp(β) .968, 95% CI .940 - .997, p = .030; Exp(β) .939, 95% CI .911 - .969, p = .000 respectively); right entorhinal cortex there was a 5.2% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .948, 95% CI .920 - .978, p = .001).
Figure 11.3 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of medial temporal lobe regions of interest

Notes:
Data from Poisson regression modelling of volume of individual ROIs (as proportion of TIV) as predictors of false recognition, including covariates: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency for animals and MRI field strength.
Percentage change in false recognition is per unit increase in ROI volume (one unit = 0.01% of TIV).
Error bars indicate 95% confidence intervals. Only significant results displayed, \( p < .05 \).
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; EC = entorhinal cortex; HPC = hippocampus; L = left; R = right; RAVLT = Rey’s Auditory Verbal Learning Test.
11.4.1.2 **Ventral visual stream regions of interest**

Four ventral visual stream ROIs were significantly related to false recognition performance on either the RAVLT or the ADAS-Cog 13: bilateral PHG and FFG. Again, volume was inversely related to number of words falsely recognised, see Figure 11.4 and Table 11.2.

For every 0.01% increase in volume (as a proportion of TIV) of the left PHG there was a 3.6% reduction in false recognition on the RAVLT and 6.6% on the ADAS-Cog 13 (Exp(\(\beta\)) .964, 95% CI .936 - .993, \(p = .016\) and Exp(\(\beta\)) .934, 95% CI .905 - .963, \(p = .000\) respectively); right PHG there was a 3.1% reduction in false recognition on the RAVLT and 3.7% on the ADAS-Cog 13 (Exp(\(\beta\)) .969, 95% CI .943 - .997 \(p = .031\) and Exp(\(\beta\)) .963, 95% CI .935 - .992, \(p = .012\) respectively); left FFG there was a 2.6% reduction in false recognition on the ADAS-Cog 13 (Exp(\(\beta\)) .974, 95% CI .960 - .989 \(p = .000\)); right FFG there was a 1.9% reduction in false recognition on the ADAS-Cog 13 (Exp(\(\beta\)) .981, 95% CI .967 - .996, \(p = .011\)).
Figure 11.4 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of ventral visual stream regions of interest

Notes:
Data from Poisson regression modelling of volume of individual ROIs (as proportion of TIV) as predictors of false recognition, including covariates: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency for animals and MRI field strength.
Percentage change in false recognition is per unit increase in ROI volume (one unit = 0.01% of TIV). Error bars indicate 95% confidence intervals. Only significant results displayed, \( p < .05 \).
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; FFG = fusiform gyrus; L = left; PHG = parahippocampal gyrus; R = right; RAVLT = Rey’s Auditory Verbal Learning Test.
Several PFC ROIs were significantly related to false recognition performance on either the RAVLT or the ADAS-Cog 13. The ROI volume was again inversely related to number of words falsely recognised. These ROIs included bilateral MFG and bilateral SFG, the left IFG, bilateral SMFG, the right MFC and bilateral anterior cingulate gyri, see Figure 11.5 and Table 11.2.

For every 0.01% increase in volume (as a proportion of TIV) of the left MFG there was a reduction in false recognition on the ADAS-Cog 13 of 0.9% (Exp(β) .991, 95% CI .984 - .998, $p = .014$); right MFG there was a reduction in false recognition on the ADAS-Cog 13 of 0.9% (Exp(β) .991, 95% CI .984 - .997, $p = .008$); left SFG there was a 1.2% reduction in false recognition on the RAVLT (Exp(β) .988, 95% CI .980 - .997, $p = .010$); right SFG there was a 1.1% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .989, 95% CI .980 - .998, $p = .020$); left IFG there was a 4.0% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .960, 95% CI .932 - .989, $p = .007$); left SMFG there was a 2.0% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .980, 95% CI .962 - .999, $p = .036$; Exp(β) .973, 95% CI .954 - .993, $p = .009$ respectively); right SMFG there was a 1.6% reduction in false recognition on the RAVLT and 1.8% on the ADAS-Cog 13 (Exp(β) .984, 95% CI .971 - .998 $p = .023$; Exp(β) .982, 95% CI .968 - .996, $p = .015$ respectively); right MFC there was a 5.4% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .946, 95% CI .901 - .992, $p = .023$); left anterior cingulate gyri there was a 2.3% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .977, 95% CI .957 - .998, $p = .034$); right anterior cingulate gyri there was a 2.6% reduction in false recognition on the ADAS-Cog 13 (Exp(β) .974, 95% CI .953 - .995 $p = .014$).
Figure 11.5 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of prefrontal cortex and anterior cingulate cortex regions of interest

Notes:
Data from Poisson regression modelling of volume of individual ROIs (as proportion of TIV) as predictors of false recognition, including covariates: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency for animals and MRI field strength.
Percentage change in false recognition is per unit increase in ROI volume (one unit = 0.01% of TIV). Error bars indicate 95% confidence intervals. Only significant results displayed, $p < .05$.
ACC = anterior cingulate cortex (anterior cingulate gyri regions of interest); ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; IFG = inferior frontal gyrus; L = left; MFC = medial frontal cerebrum; MFG = middle frontal gyrus; R = right; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus; RAVLT = Rey’s Auditory Verbal Learning Test.
Table 11.2 Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest

<table>
<thead>
<tr>
<th></th>
<th>RAVLT FALSE RECOGNITION MODELS</th>
<th>ADAS-COG 13 FALSE RECOGNITION MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL LIKELIHOOD RATIO</td>
<td>EXP(β)</td>
</tr>
<tr>
<td></td>
<td>(df = 8, n = 725, All ps ≤ .002)</td>
<td></td>
</tr>
<tr>
<td>LEFT HIPPOCAMPUS</td>
<td>$X^2 = 32.002$</td>
<td>.960</td>
</tr>
<tr>
<td>LEFT ENTORHINAL CORTEX</td>
<td>$X^2 = 29.307$</td>
<td>.968</td>
</tr>
<tr>
<td>RIGHT ENTORHINAL CORTEX</td>
<td>$X^2 = 28.558$</td>
<td>.971</td>
</tr>
<tr>
<td>Left PHG</td>
<td>$X^2 = 30.440$</td>
<td>.964</td>
</tr>
<tr>
<td>Right PHG</td>
<td>$X^2 = 29.352$</td>
<td>.969</td>
</tr>
<tr>
<td>Left FFG</td>
<td>$X^2 = 25.444$</td>
<td>.994</td>
</tr>
<tr>
<td>Right FFG</td>
<td>$X^2 = 27.464$</td>
<td>.988</td>
</tr>
<tr>
<td>Left LING</td>
<td>$X^2 = 24.729$</td>
<td>.998</td>
</tr>
<tr>
<td>Right LING</td>
<td>$X^2 = 25.240$</td>
<td>.995</td>
</tr>
</tbody>
</table>
Table 11.2 Cont. Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest

<table>
<thead>
<tr>
<th></th>
<th>RAVLT FALSE RECOGNITION MODELS</th>
<th>ADAS-COG 13 FALSE RECOGNITION MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL LIKELIHOOD RATIO (df = 8, n = 725, All ps ≤ .002)</td>
<td>MODEL LIKELIHOOD RATIO (df = 8, n = 725, All ps &lt; .001)</td>
</tr>
<tr>
<td></td>
<td>EXP(β)</td>
<td>95% CI EXP(β)</td>
</tr>
<tr>
<td>Left MFG</td>
<td>$X^2 = 25.795$</td>
<td>.996</td>
</tr>
<tr>
<td>Right MFG</td>
<td>$X^2 = 26.750$</td>
<td>.995</td>
</tr>
<tr>
<td>Left SFG</td>
<td>$X^2 = 31.486$</td>
<td>.988</td>
</tr>
<tr>
<td>Right SFG</td>
<td>$X^2 = 28.094$</td>
<td>.992</td>
</tr>
<tr>
<td>Left IFG</td>
<td>$X^2 = 27.947$</td>
<td>.975</td>
</tr>
<tr>
<td>Right IFG</td>
<td>$X^2 = 26.113$</td>
<td>.983</td>
</tr>
<tr>
<td>Left IFOG</td>
<td>$X^2 = 24.623$</td>
<td>.998</td>
</tr>
<tr>
<td>Right IFOG</td>
<td>$X^2 = 24.665$</td>
<td>1.006</td>
</tr>
<tr>
<td>Left IFAG</td>
<td>$X^2 = 24.932$</td>
<td>1.008</td>
</tr>
<tr>
<td>Right IFAG</td>
<td>$X^2 = 24.637$</td>
<td>.998</td>
</tr>
<tr>
<td>M-PFC</td>
<td>$X^2 = 29.163$</td>
<td>.980</td>
</tr>
</tbody>
</table>
Table 11.2 Cont. Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest

<table>
<thead>
<tr>
<th>MODEL-Cog 13 FALSE RECOGNITION MODELS</th>
<th>RAVLT FALSE RECOGNITION MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td>MODEL LIKELIHOOD RATIO (df = 8, n = 725, All ps ≤ .002)</td>
<td>Likelihood Ratio</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------------------------</td>
</tr>
<tr>
<td>M-PFC Right SMFG</td>
<td>$X^2 = 29.995$</td>
</tr>
<tr>
<td>M-PFC Left MFC</td>
<td>$X^2 = 27.126$</td>
</tr>
<tr>
<td>M-PFC Right MFC</td>
<td>$X^2 = 27.225$</td>
</tr>
<tr>
<td>OFC Left LOG</td>
<td>$X^2 = 24.644$</td>
</tr>
<tr>
<td>OFC Right LOG</td>
<td>$X^2 = 24.823$</td>
</tr>
<tr>
<td>LEFT ANTERIOR CINGULATE GYRUS</td>
<td>$X^2 = 26.970$</td>
</tr>
<tr>
<td>RIGHT ANTERIOR CINGULATE GYRUS</td>
<td>$X^2 = 27.961$</td>
</tr>
<tr>
<td>LEFT SUPERIOR PARIETAL LOBULE</td>
<td>$X^2 = 25.553$</td>
</tr>
<tr>
<td>RIGHT SUPERIOR PARIETAL LOBULE</td>
<td>$X^2 = 24.618$</td>
</tr>
</tbody>
</table>

| MODEL LIKELIHOOD RATIO (df = 8, n = 725, All ps ≤ .001) | Likelihood Ratio | EXP(β) | 95% CI | P VALUE | FR REDUCED |
|--------------------------------------|--------------------------------|
|                                                | $X^2 = 71.408$ | .982 | .968, .996 | .015 | 1.8% |
|                                                | $X^2 = 69.088$ | .954 | .908, 1.001 | .057 | - |
|                                                | $X^2 = 70.905$ | .946 | .901, .992 | .023 | 5.4% |
|                                                | $X^2 = 65.708$ | .979 | .931, 1.030 | .411 | - |
|                                                | $X^2 = 64.971$ | .997 | .951, 1.046 | .916 | - |
|                                                | $X^2 = 69.869$ | .977 | .957, .998 | .034 | 2.3% |
|                                                | $X^2 = 71.553$ | .974 | .953, .995 | .014 | 2.6% |
|                                                | $X^2 = 67.975$ | .991 | .980, 1.002 | .108 | - |
|                                                | $X^2 = 66.478$ | .994 | .982, 1.005 | .270 | - |
**Table 11.2 Cont. Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest**

Notes:
Reduction is % reduction in false recognition per unit increase in ROI volume (one unit = 0.01% of TIV).
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; DL-PFC = dorsolateral prefrontal cortex; FFG = fusiform gyrus; FR = false recognition; IFAG = inferior frontal angular gyrus; IFG = inferior frontal gyrus; IFOG = inferior frontal orbital gyrus; LING = lingual gyrus; LOG = lateral orbital gyrus; MFC = medial frontal cerebral; MFG = middle frontal gyrus; M-PFC = medial prefrontal cortex; OFC = orbitofrontal prefrontal cortex; PHG = parahippocampal gyrus; RAVLT = Rey’s Auditory Verbal Learning Test; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus; VL-PFC = ventrolateral prefrontal cortex.

Poisson regression models include covariates: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency for animals and MRI field strength.
11.4.1.4 Model diagnostics

Further model diagnostics were run for those ROIs with significant results. Residuals were examined by creating scatter plots of standardised Pearson residuals fitted values, all of which demonstrated a constant spread of residuals and indicated a good model fit, with the majority of standardised Pearson residuals falling between -2.0 and 2.0 (see Appendix 13).

Similarly, for those ROIs with significant results, Cook’s distance was calculated and visualised in bar charts (see Appendix 14). No values exceeded one. Using the more conservative calculation for threshold of Cook’s distance to identify outliers of $4/(n - k - 1)$, several possible outliers were identified across all ROIs. The mean number of values exceeding this threshold was $31.4 \pm 1.8$, 4.3% (range 29 - 34) for models of false recognition on RAVLT and brain volume, and $40.1 \pm 2.7$, 5.5% (range 38 - 45) for models of false recognition on ADAS-Cog 13 and brain volume. All data points flagged by this method were checked for user error on entry, without any error identified. All models were then run again, excluding values above the $4/(n - k - 1)$ threshold (0.006 for this data set). All models for false recognition on both the RAVLT and ADAS-Cog 13 retained both overall significance ($ps < .001$) and significance for the individual ROIs within them (all $ps < .05$), see Appendix 15.

11.4.1.5 Subgroup analysis

Significant findings at the ROI level in the whole group were further explored in the subgroup of patients who had a diagnosis of either AD or aMCI at the time of their baseline MRI (i.e. excluding the 27 participants who were CN at baseline).

For false recognition on the RAVLT, all models retained overall significance ($ps < .05$), and all individual ROIs within them retained significance ($ps < .05$), apart from the right SFG ($\text{Exp}(\beta) = .992$, 95% CI .984 – 1.001, $p = .083$). For false recognition on the ADAS-Cog 13, all models retained significance ($ps < .001$) and the individual ROIs within them retained significance ($ps < .05$).

11.4.2 Whole brain analysis

A whole brain analysis by VBM was completed to further identify any brain regions that had volume correlating with performance on the false memory measures. Covariates were included as previously specified (age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency, TIV and MRI field strength). Continuous measures (age, years of education and TIV) were mean centered to reduce multicollinearity. The correlation matrix for continuous independent variables is shown in Table 11.3. There were significant correlations between half of the variables, particularly affecting years of education and category fluency, however all Pearson correlation coefficients were < .80 and VIFs were all < 2.5, mean 1.3.
Two control participants, for whom data were missing on cholinesterase inhibitor use and animal category fluency score respectively, were removed from the VBM analysis.

Nine participants with missing scores for total intrusions on RAVLT, one participant with no score for total false recognition on RAVLT and one participant with no score for total false recognition on ADAS-Cog 13 were excluded from the respective regression models.

Multiple regression models were run with a FWE correction applied, with a significance threshold of $p_{\text{FWE}} < .05$ and an extent threshold of $k = 10$ voxels. More exploratory models were run as a secondary analysis, with an uncorrected $p$ value of < .001 and the same extent threshold of $k = 10$ voxels.

Table 11.3 Pearson correlation matrix for continuous variables ($n = 728$)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. AGE (years, mean centered)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. EDUCATION (years, mean centered)</td>
<td>-.041</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. MMSE</td>
<td>-.063</td>
<td>.195$^b$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. CATEGORY FLUENCY (Animals)</td>
<td>-.084$^a$</td>
<td>.202$^b$</td>
<td>.456$^b$</td>
<td></td>
</tr>
<tr>
<td>5. TIV (ml, mean centered)</td>
<td>.061</td>
<td>.176$^b$</td>
<td>.054</td>
<td>.062</td>
</tr>
</tbody>
</table>

Notes:
MMSE = Mini Mental State Examination; TIV = total intracranial volume.

$^a p < .05.$

$^b p < .01.$

11.4.2.1 Family wise error corrected results
On multiple regression analyses with t-contrasts run for both positive and negative effects of intrusions and false recognition on RAVLT and positive effects of false recognition on ADAS-Cog 13 with GM intensity, no clusters reached cluster or peak-level significance at $p_{\text{FWE}} < .05$.

However, examination of the statistical parametric map (SPM) for the multiple regression analysis with t-contrast for negative effect of false recognition on ADAS-Cog 13 revealed significant reduction in GM volume in two areas of the right hippocampus ($x = 17, y = -11, z = -14$ and $x = 29, y = -29, z = -9$), the left FFG and the left middle temporal gyrus (MTG; $x = -27, y = -29, z = -15$ and $x = -35, y = 9, z = -33$ respectively), significant at the cluster-level ($p_{\text{FWE}} < .002$), see Figure 11.6 and Figure 11.7.
**Figure 11.6** Statistical parametric map of negative correlation of grey matter volume and false recognition on ADAS-Cog 13 (t-contrast), displayed on whole glass brain view.

**SPM results:**
Height threshold $T = 4.442411 \ (p<0.05 \ \text{FWE})$
Extent threshold $k = 10 \ \text{voxels}$

**Statistics:** $p$-values adjusted for search volume

<table>
<thead>
<tr>
<th>set-level</th>
<th>cluster-level</th>
<th>peak-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$c$</td>
<td>$p_{\text{FWE}}$</td>
</tr>
<tr>
<td>0.000</td>
<td>4</td>
<td>0.000</td>
</tr>
<tr>
<td>0.010</td>
<td>0.667</td>
<td>76</td>
</tr>
<tr>
<td>0.007</td>
<td>0.667</td>
<td>100</td>
</tr>
<tr>
<td>0.018</td>
<td>0.878</td>
<td>40</td>
</tr>
</tbody>
</table>

Notes:
FDR = false discovery rate; FWE = family wise error; FWHM = full width half maximum.
Design matrix covariates: 1 = mean, 2 = false recognition on ADAS-Cog 13, 3 = age (years; mean centered), 4 = gender, 5 = education (years; mean centered), 6 = cholinesterase inhibitor prescription, 7 = MMSE score, 8 = category fluency for animals, 9 = TIV (ml; mean centered), 10 = MRI field strength.
**Figure 11.7** Locations of the four significant clusters with negative effect of ADAS-Cog 13 false recognition

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes:
Significant clusters displayed on warped, skull-stripped, bias and noise-corrected average brain. 
$p_{FWE} < .05$, $k = 10$ voxels.

11.4.2.2 Exploratory analyses

On multiple regression analyses, with t-contrasts for negative effect of intrusions on RAVLT on GM intensity, seven clusters had peak-level significance at the more liberal threshold of uncorrected $p < .001$. The SPM revealed significant reduction in bilateral MFG ($x = -47$, $y = 44$, $z = 18$ and $x = 50$, $y = 6$, $z = 48$), right SFG ($x = 8$, $y = 63$, $z = 29$), the left fornix ($x = -2$, $y = 2$, $z = 2$), the right cuneus ($x = 17$, $y = -69$, $z = 14$) and in the cerebellum bilaterally ($x = -5$, $y = -56$, $z = -5$ and $x = 5$, $y = -62$, $z = -6$), see Figure 11.8 and Figure 11.9 and unthresholded effect size map in Figure 11.10.
Figure 11.8 Statistical parametric map of negative correlation of grey matter volume and intrusions on RAVLT (t-contrast), displayed on whole glass brain view.

Notes:
FDR = false discovery rate; FWE = family wise error; FWHM = full width half maximum.
Design matrix covariates: 1 = mean, 2 = intrusions on RAVLT, 3 = age (years; mean centered), 4 = gender, 5 = education (years; mean centered), 6 = cholinesterase inhibitor prescription, 7 = MMSE score, 8 = category fluency for animals, 9 = TIV (ml; mean centered), 10 = MRI field strength.
Figure 11.9 Location of the right middle frontal gyrus cluster with negative effect of intrusions on RAVLT

Notes:
Significant clusters displayed on warped, skull-stripped, bias and noise-corrected average brain. $p < .001$, $k = 10$ voxels.
Figure 11.10 Unthresholded t-contrast effect size map for association between intrusions on RAVLT and grey matter volume

Notes:
Red indicates stronger association between increased intrusions on RAVLT and lower grey matter volume. Blue indicates the reverse contrast.
Examination of the SPM for the multiple regression analysis with t-contrast for negative effect of false recognition on RAVLT using the more liberal uncorrected $p$ threshold of $<.001$ revealed significant reduction in GM volume in bilateral hippocampi ($x = -14, y = -11, z = -17$ and $x = 18, y = -17, z = -15$), the left inferior temporal gyrus (ITG; $x = -39, y = -15, z = -23$) the right MFG ($x = 36, y = 8, z = 48$) and two locations within the left SFG ($x = -15, y = 29, z = 50$ and $x = -15, y = 15, z = 59$), significant at the peak-level, see Figure 11.11 and Figure 11.12 and unthresholded effect size map in Figure 11.13.

**Figure 11.11 Statistical parametric map of negative correlation of grey matter volume and false recognition on RAVLT (t-contrast), displayed on whole glass brain view**
**Figure 11.12** Location of the left superior frontal gyrus clusters with negative effect of false recognition on RAVLT

Notes:
Significant clusters displayed on warped, skull-stripped, bias and noise-corrected average brain. $p < .001$, $k = 10$ voxels.
Figure 11.13 Unthresholded t-contrast effect size map for association between false recognition on RAVLT and grey matter volume

Notes:
Red indicates stronger association between increased intrusions on RAVLT and lower grey matter volume. Blue indicates the reverse contrast.
Multiple regression analyses with t-contrasts run for negative effects of false recognition on ADAS-Cog 13 with GM intensity revealed 22 regions with peak-level significance at the more liberal threshold of uncorrected $p < .001$. Regions with peak-level significance included eight regions explored in the ROI analysis: the right hippocampus ($x = 17, y = -11, z = -14$), the left FFG ($x = -27, y = -29, z = -15$), bilateral MFG (three left-sided clusters: $x = -30, y = 45, z = 17$, $x = -42, y = 30, z = 17$, $x = -35, y = 12, z = 32$ and one right: $x = 35, y = 48, z = 12$), bilateral SFG (two left-sided clusters: $x = -12, y = 59, z = 14$, $x = -17, y = 17, z = 56$ and one right: $x = 20, y = 50, z = 26$), the right ACC ($x = 0, y = 30, z = -12$) and the left superior parietal lobule ($x = -18, y = -62, z = 54$). Additionally, the exploratory VBM analysis identified significant clusters in the left precentral gyrus ($x = -33, y = 3, z = 14$), the left posterior cingulate ($x = -5, y = -35, z = 27$), the right posterior insula ($x = 33, y = -18, z = 15$), the left superior temporal gyrus (STG; $x = -32, y = -21, z = 9$), the right MTG ($x = 63, y = -54, z = -12$) and the right ITG (two clusters: $x = 59, y = -29, z = -23$ and $x = 44, y = -3, z = -36$), bilateral supramarginal gyrri ($x = -51, y = -42, z = 35$ and $x = 53, y = -36, z = 35$) and the left inferior occipital gyrus ($x = -51, y = -60, z = -6$). See Figure 11.14 and Figure 11.15 and unthresholded effect size map in Figure 11.16.
Figure 11.14 Statistical parametric map of negative correlation of grey matter volume and false recognition on ADAS-Cog 13 (t-contrast), whole glass brain view.
**Figure 11.14 Cont.** Statistical parametric map of negative correlation of grey matter volume and false recognition on ADAS-Cog 13 (t-contrast), whole glass brain view

Notes:
FDR = false discovery rate; FWE = family wise error; FWHM = full width half maximum.
Design matrix covariates: 1 = mean, 2 = false recognition on ADAS-Cog 13, 3 = age (years; mean centered), 4 = gender, 5 = education (years; mean centered), 6 = cholinesterase inhibitor prescription, 7 = MMSE score, 8 = category fluency for animals, 9 = TIV (ml; mean centered), 10 = MRI field strength.

**Figure 11.15** Location of bilateral medial temporal lobe and left fusiform clusters with negative effect of false recognition on ADAS-Cog 13

Notes:
Significant clusters displayed on warped, skull-stripped, bias and noise-corrected average brain. $p < .001$, k = 10 voxels.
Figure 11.16 Unthresholded t-contrast effect size map for association between false recognition on ADAS-Cog 13 and grey matter volume

Notes:
Red indicates stronger association between increased intrusions on RAVLT and lower grey matter volume. Blue indicates the reverse contrast.
11.4.2.3 Model diagnostics

The design orthogonality matrix showed that the majority of covariates were orthogonal to each other, but there was a degree of collinearity between independent variables, see Figure 11.17, Figure 11.18 and Figure 11.19.

**Figure 11.17 Intrusions on RAVLT: design orthogonality**

<table>
<thead>
<tr>
<th>Measure</th>
<th>Scale</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Note</strong>:</td>
<td></td>
</tr>
<tr>
<td>* mean centered variables.</td>
<td></td>
</tr>
<tr>
<td>ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; RAVLT = Rey’s Auditory Verbal Learning Test; TIV = total intracranial volume.</td>
<td></td>
</tr>
</tbody>
</table>
Figure 11.18 False recognition on RAVLT: design orthogonality

Notes:
* mean centered variables.

ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; RAVLT = Rey's Auditory Verbal Learning Test; TIV = total intracranial volume.
**Figure 11.19** False recognition on ADAS-Cog 13: design orthogonality

Notes:

* mean centered variables.

ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; TIV = total intracranial volume.
11.5 Delusion group

11.5.1 Regions of interest analysis

The relationships between volume of previously specified ROIs (as a proportion of TIV) and delusion group (‘state’ or ‘trait’) were first explored using independent-samples t-tests, and subsequently through binary logistic regression modelling, see Table 11.4 and 11.5. Volumes of ROIs (as a proportion of TIV) were again transformed by multiplying by a factor of $10^3$, to allow meaningful interpretation of odds ratios. Prior to running the models, potential confounding variables defined a priori (age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency and MRI field strength) were assessed for collinearity in each model, with transformed, TIV-corrected left hippocampal volume as a representative measure of ROI. The VIF scores were all $< 2.5$, with a mean VIF of 1.3. The Box-Tidwell test was used to test the assumption of linearity between continuous variables and the logit transformation of the dependent variable, with the majority meeting this assumption, the only exceptions being volume of the right anterior cingulate gyrus in the model for ‘trait’ delusions (log interaction $p = .014$) and volume of the right entorhinal cortex in the ‘trait’ delusion/ventral visual stream model (log interaction $p = .022$).

In the ‘state’ delusion group, the right anterior cingulate gyrus was significantly smaller in TIV corrected volume than in the control group ($t_{589} = 2.147$, $p = .032$). However, this finding was no longer significant when confounding factors were included in the binary logistic regression model ($p = .054$). All of the binary logistic regression models for ‘state’ delusion were significant, however no individual ROIs achieved significance within the models, see Table 11.4. None of the binary logistic regression models for ‘trait’ delusion reached significance, see Table 11.5. The right anterior cingulate gyrus was significant within the model for ‘trait’ delusions ($\text{Exp}(\beta) .947$, 95% CI .901 - .996, $p = .035$), but as described above this model was in violation of the assumption of linearity. All of the models passed the Hosmer and Lemeshow test for goodness of fit.

Firth’s logistic regression, run to address the issue of ‘state’ delusions being a rare event in the group ($n = 42$ of total $n = 589, 7.1\%$) yielded similar results; all models were significant (all $ps < .05$), but no individual ROIs achieved significance in any models.
Table 11.4 Comparison of volume in identified regions of interest between those with delusions at baseline (‘state’) and control

<table>
<thead>
<tr>
<th>ROI</th>
<th>CONTROL (n = 546)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>T TEST</th>
<th>LOGISTIC REGRESSION MODEL (df = 8, n = 589, All ps &lt; .05)</th>
<th>LOGISTIC REGRESSION FIT (df = 8, n = 589, All ps &gt; .05)</th>
<th>ROI P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEFT HIPPOCAMPUS</td>
<td>1.40 ± 0.27</td>
<td>1.36 ± 0.24</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = .869, p = .385</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.584</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 6.773</td>
<td>.717</td>
</tr>
<tr>
<td>RIGHT HIPPOCAMPUS</td>
<td>1.60 ± 0.32</td>
<td>1.56 ± 0.29</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = .853, p = .394</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.453</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 6.144</td>
<td>.979</td>
</tr>
<tr>
<td>LEFT ENTORHINAL CORTEX</td>
<td>1.29 ± 0.28</td>
<td>1.32 ± 0.29</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = -.486, p = .627</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 19.429</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 4.954</td>
<td>.085</td>
</tr>
<tr>
<td>RIGHT ENTORHINAL CORTEX</td>
<td>1.34 ± 0.28</td>
<td>1.35 ± 0.28</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = -.332, p = .740</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 18.254</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 5.870</td>
<td>.182</td>
</tr>
<tr>
<td>DL-PFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left MFG</td>
<td>9.78 ± 1.08</td>
<td>9.58 ± 1.29</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = 1.084, p = .279</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.806</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 6.399</td>
<td>.554</td>
</tr>
<tr>
<td>Right MFG</td>
<td>9.72 ± 1.09</td>
<td>9.53 ± 1.07</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = 1.092, p = .275</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.856</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 11.571</td>
<td>.526</td>
</tr>
<tr>
<td>Left SFG</td>
<td>7.85 ± 0.85</td>
<td>7.68 ± 0.76</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = 1.274, p = .203</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 17.464</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 4.550</td>
<td>.320</td>
</tr>
<tr>
<td>Right SFG</td>
<td>7.88 ± 0.85</td>
<td>7.79 ± 0.83</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = 0.615, p = .539</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.548</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 9.499</td>
<td>.758</td>
</tr>
<tr>
<td>M-PFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left SMFG</td>
<td>3.38 ± 0.39</td>
<td>3.30 ± 0.39</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = 1.319, p = .188</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 17.027</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 8.977</td>
<td>.450</td>
</tr>
<tr>
<td>Right SMFG</td>
<td>4.11 ± 0.53</td>
<td>4.02 ± 0.56</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = .990, p = .323</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.543</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 4.834</td>
<td>.763</td>
</tr>
<tr>
<td>Left MFC</td>
<td>0.92 ± 0.16</td>
<td>0.88 ± 0.15</td>
<td>t&lt;sub&gt;589&lt;/sub&gt; = 1.395, p = .164</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 16.809</td>
<td>X&lt;sup&gt;2&lt;/sup&gt; = 4.602</td>
<td>.551</td>
</tr>
</tbody>
</table>
Table 11.4 Cont. Comparison of volume in identified regions of interest between those with delusions at baseline (‘state’) and control

<table>
<thead>
<tr>
<th>ROI</th>
<th>CONTROL (n = 546)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>T TEST</th>
<th>LOGISTIC REGRESSION MODEL (df = 8, n = 589, All ps &lt; .05)</th>
<th>LOGISTIC REGRESSION FIT (df = 8, n = 589, All ps &gt; .05)</th>
<th>ROI P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>M-PFC</td>
<td>Right MFC</td>
<td>0.97 ± 0.16</td>
<td>0.93 ± 0.17</td>
<td>t_{589} = 1.568, p = .117</td>
<td>X^2 = 16.929</td>
<td>X^2 = 4.378</td>
</tr>
<tr>
<td>LEFT ANTERIOR CINGULATE GYRUS</td>
<td>2.63 ± 0.38</td>
<td>2.55 ± 0.43</td>
<td>t_{589} = 1.302, p = .193</td>
<td>X^2 = 16.710</td>
<td>X^2 = 9.520</td>
<td>.612</td>
</tr>
<tr>
<td>RIGHT ANTERIOR CINGULATE GYRUS</td>
<td>1.99 ± 0.35</td>
<td>1.87 ± 0.39</td>
<td>t_{589} = 2.147, p = .032</td>
<td>X^2 = 20.205</td>
<td>X^2 = 9.379</td>
<td>.054</td>
</tr>
</tbody>
</table>

Notes:
Values are (mean ± SD) x 10^-3. All measures are mean brain volume for specified region of interest, as proportion of total intracranial volume, at baseline.
DL-PFC = dorsolateral prefrontal cortex; MFC = medial frontal cerebrum; MFG = middle frontal gyrus; M-PFC = medial prefrontal cortex; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus.
Logistic regression included covariates: age, gender, years of education, MMSE, cholinesterase inhibitor prescription and MRI field strength.
Model fit assessed by Hosmer and Lemeshow test, p > .05 indicates good model fit.
Table 11.5 Comparison of volume in identified regions of interest between those with delusions at any time point ('trait') and control

<table>
<thead>
<tr>
<th>ROI</th>
<th>CONTROL (n = 546)</th>
<th>'TRAIT' DELUSIONS (n = 179)</th>
<th>T TEST</th>
<th>LOGISTIC REGRESSION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MODEL (df = 8, n = 726, All ps &gt; .05)</td>
</tr>
<tr>
<td>LEFT HIPPOCAMPUS</td>
<td>1.40 ± 0.27</td>
<td>1.40 ± 0.27</td>
<td>$t_{726} = - .304, p = .761$</td>
<td>$X^2 = 9.167$</td>
</tr>
<tr>
<td>RIGHT HIPPOCAMPUS</td>
<td>1.60 ± 0.32</td>
<td>1.59 ± 0.28</td>
<td>$t_{726} = .436, p = .663$</td>
<td>$X^2 = 9.505$</td>
</tr>
<tr>
<td>LEFT ENTORHINAL CORTEX</td>
<td>1.29 ± 0.28</td>
<td>1.29 ± 0.26</td>
<td>$t_{726} = .190, p = .850$</td>
<td>$X^2 = 9.240$</td>
</tr>
<tr>
<td>RIGHT ENTORHINAL CORTEX</td>
<td>1.34 ± 0.28</td>
<td>1.32 ± 0.25</td>
<td>$t_{726} = .796, p = .426$</td>
<td>$X^2 = 9.206$</td>
</tr>
<tr>
<td>DL-PFC</td>
<td>9.78 ± 1.08</td>
<td>9.82 ± 0.12</td>
<td>$t_{726} = -.492, p = .623$</td>
<td>$X^2 = 9.315$</td>
</tr>
<tr>
<td>Left MFG</td>
<td>9.72 ± 1.09</td>
<td>9.69 ± 0.11</td>
<td>$t_{726} = .334, p = .738$</td>
<td>$X^2 = 9.321$</td>
</tr>
<tr>
<td>Right MFG</td>
<td>7.85 ± 0.85</td>
<td>7.81 ± 0.83</td>
<td>$t_{726} = .613, p = .540$</td>
<td>$X^2 = 9.718$</td>
</tr>
<tr>
<td>Left SFG</td>
<td>7.88 ± 0.85</td>
<td>7.84 ± 0.82</td>
<td>$t_{726} = .478, p = .633$</td>
<td>$X^2 = 9.661$</td>
</tr>
<tr>
<td>Right SFG</td>
<td>3.38 ± 0.39</td>
<td>3.36 ± 0.43</td>
<td>$t_{726} = .475, p = .635$</td>
<td>$X^2 = 9.656$</td>
</tr>
<tr>
<td>M-PFC</td>
<td>4.11 ± 0.53</td>
<td>4.13 ± 0.54</td>
<td>$t_{726} = -.579, p = .562$</td>
<td>$X^2 = 9.423$</td>
</tr>
<tr>
<td>Left SMFG</td>
<td>0.92 ± 0.16</td>
<td>0.90 ± 0.16</td>
<td>$t_{726} = 1.073, p = .284$</td>
<td>$X^2 = 10.745$</td>
</tr>
</tbody>
</table>

197
Table 11.5 Cont. Comparison of volume in identified regions of interest between those with delusions at any time point (‘trait’) and control

<table>
<thead>
<tr>
<th>ROI</th>
<th>CONTROL (n = 546)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th>T TEST</th>
<th>LOGISTIC REGRESSION</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>MODEL (df = 8, n = 726, All ps &gt; .05)</td>
</tr>
<tr>
<td><strong>M-PFC</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Right MFC</td>
<td>0.97 ± 0.16</td>
<td>0.97 ± 0.17</td>
<td>(t_{726} = .283, p = .778)</td>
<td>(X^2 = 9.364)</td>
</tr>
<tr>
<td><strong>LEFT ANTERIOR CINGULATE GYRUS</strong></td>
<td>2.63 ± 0.38</td>
<td>2.62 ± 0.39</td>
<td>(t_{726} = .166, p = .868)</td>
<td>(X^2 = 9.244)</td>
</tr>
<tr>
<td><strong>RIGHT ANTERIOR CINGULATE GYRUS</strong></td>
<td>1.99 ± 0.35</td>
<td>1.93 ± 0.39</td>
<td>(t_{726} = 1.749, p = .081)</td>
<td>(X^2 = 13.637)</td>
</tr>
</tbody>
</table>

**Notes:**
Values are (mean ± SD) \(\times 10^{-3}\). All measures are mean brain volume for specified region of interest, as proportion of total intracranial volume, at baseline.

DL-PFC = dorsolateral prefrontal cortex; MFC = medial frontal cerebrum; MFG = middle frontal gyrus; M-PFC = medial prefrontal cortex; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus.

Logistic regression included covariates: age, gender, years of education, MMSE, cholinesterase inhibitor prescription and MRI field strength.

Model fit assessed by Hosmer and Lemeshow test, \(p > .05\) indicates good model fit.
An exploration of the relationship between delusions and ventral visual stream ROIs was undertaken separately, guided by my previous findings, see section 10.6.3.1. A binary logistic regression model including all ventral visual stream ROIs bilaterally (entorhinal cortex, PHG, FFG and lingual gyrus) and the same covariates as described above, was significant for the ‘state’ delusion group ($\chi^2 (15, n = 589) = 31.485, p = .008$) and explained 13.0% (Nagelkerke $R^2$) of the variance in delusion group. 92.9% of cases were correctly classified by the model, but with all cases classified as control (i.e. 0% sensitivity, 100% specificity). There was evidence of high collinearity between volumes of ventral visual stream regions within the model, with VIFs for ventral visual stream regions all > 2.5 and mean VIF 3.3 for the model. The model for ‘trait’ delusion group was not significant.

The right PHG was the only ROI that was individually significant in the model for ‘state’ delusion, see Table 11.6. For every 0.01% increase in volume of right PHG (as a proportion of TIV) participants were 40.3% more likely to be in the ‘state’ delusion group ($\text{Exp}(\beta) = 1.403, 95\% \text{ CI } 1.042 – 1.891, p = .026$). No ventral visual stream ROIs were individually significant in the model for ‘trait’ delusion.

Due to previous findings indicating a significant difference in parahippocampal volume between delusion and no-delusion groups, this ROI was also explored separately, with no significant difference found between ‘state’ delusion and control groups when compared by independent-samples t-tests, see Table 11.7. While the volume of the right PHG appeared to be greater than the left PHG for the ‘state’ delusion group, this finding did not reach significance on a paired-sample t-test ($t_{41} = -1.269, p = .212$).
Table 11.6 Binary logistic regression model of volume of ventral visual stream regions of interest (as proportion of TIV) as predictors of delusions at baseline (n = 589)

<table>
<thead>
<tr>
<th>Predictor</th>
<th>P VALUE</th>
<th>EXP(β)</th>
<th>95% CI EXP(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AGE (years)</td>
<td>.175</td>
<td>1.034</td>
<td>.985, 1.085</td>
</tr>
<tr>
<td>GENDER</td>
<td>.476</td>
<td>.767</td>
<td>.370, 1.590</td>
</tr>
<tr>
<td>EDUCATION (years)</td>
<td>.243</td>
<td>1.079</td>
<td>.950, 1.226</td>
</tr>
<tr>
<td>ChEI</td>
<td>.714</td>
<td>1.175</td>
<td>.496, 2.787</td>
</tr>
<tr>
<td>MMSE</td>
<td>.008</td>
<td>.827</td>
<td>.719, .951</td>
</tr>
<tr>
<td>CATEGORY FLUENCY</td>
<td>.030</td>
<td>.917</td>
<td>.849, .991</td>
</tr>
<tr>
<td>MRI FIELD STRENGTH</td>
<td>.736</td>
<td>.864</td>
<td>.371, 2.016</td>
</tr>
<tr>
<td>LEFT EC</td>
<td>.269</td>
<td>1.172</td>
<td>.884, 1.554</td>
</tr>
<tr>
<td>RIGHT EC</td>
<td>.483</td>
<td>.905</td>
<td>.685, 1.196</td>
</tr>
<tr>
<td>LEFT PHG</td>
<td>.060</td>
<td>.749</td>
<td>.555, 1.012</td>
</tr>
<tr>
<td>RIGHT PHG</td>
<td>.026</td>
<td>1.403</td>
<td>1.042, 1.891</td>
</tr>
<tr>
<td>LEFT FFG</td>
<td>.059</td>
<td>1.106</td>
<td>.996, 1.228</td>
</tr>
<tr>
<td>RIGHT FFG</td>
<td>.419</td>
<td>.957</td>
<td>.861, 1.065</td>
</tr>
<tr>
<td>LEFT LING</td>
<td>.254</td>
<td>1.058</td>
<td>.960, 1.167</td>
</tr>
<tr>
<td>RIGHT LING</td>
<td>.483</td>
<td>.905</td>
<td>.685, 1.196</td>
</tr>
</tbody>
</table>

Notes:
All measures at baseline unless otherwise specified.
ChEI = cholinesterase inhibitor; EC = entorhinal cortex; FFG = fusiform gyrus; LING = lingual gyrus; MMSE = Mini Mental State Examination; MRI = magnetic resonance imaging; PHG = parahippocampal gyrus.
Category fluency is for animals only.
Overall model significant, $X^2(15, n = 589) = 31.485, p = .008$. 
Table 11.7 Comparison of parahippocampal gyri volume between those with delusions and control

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 549)</th>
<th>‘STATE’ DELUSIONS (n = 42)</th>
<th>T TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LEFT PHG</strong></td>
<td>1.71 ± 0.29</td>
<td>1.72 ± 0.31</td>
<td>$t_{589} = -206, p = .837$</td>
</tr>
<tr>
<td><strong>RIGHT PHG</strong></td>
<td>1.70 ± 0.30</td>
<td>1.75 ± 0.34</td>
<td>$t_{589} = .259, p = .259$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 549)</th>
<th>‘TRAIT’ DELUSIONS (n = 179)</th>
<th>T TEST</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LEFT PHG</strong></td>
<td>1.71 ± 0.29</td>
<td>1.71 ± 0.29</td>
<td>$t_{726} = -.033, p = .974$</td>
</tr>
<tr>
<td><strong>RIGHT PHG</strong></td>
<td>1.70 ± 0.30</td>
<td>1.69 ± 0.30</td>
<td>$t_{726} = .308, p = .758$</td>
</tr>
</tbody>
</table>

Notes:
Values are (mean ± SD) x 10^{-3}. All measures are mean brain volume for specified region of interest, as proportion of total intracranial volume, at baseline.
PHG = parahippocampal gyrus.

Using binary logistic regression modelling, while the overall model for ‘state’ delusion with left PHG as a predictor was significant ($p = .023$) the left PHG was not significant within the model.
The model including the right PHG as a predictor of ‘state’ delusions was also significant ($p = .009$), with the right PHG significant within the model, see Table 11.8. For every 0.01% increase in volume of right PHG (as a proportion of TIV) participants were 14.6% more likely to have delusions at baseline ($\text{Exp}(\beta) = 1.146, 95\% \text{ CI} 1.002 – 1.311, p = .047$). This model explained 8.5% (Nagelkerke $R^2$) of the variance in delusion group. While 92.9% of cases were correctly classified by the model, all cases were again classified as control. All VIFs for this model were < 2.5, with a mean VIF of 1.3. Models of ‘trait’ delusion including bilateral PHG ROIs were not significant.

Hosmer and Lemeshow tests indicated all ventral visual stream models had good fit, all $ps > .05$. 
**Table 11.8** Binary logistic regression model of volume of the right parahippocampal gyrus (as proportion of TIV) as a predictor of delusions at baseline (n = 589)

<table>
<thead>
<tr>
<th></th>
<th>P VALUE</th>
<th>EXP(β)</th>
<th>95% CI EXP(β)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE (years)</strong></td>
<td>.174</td>
<td>1.032</td>
<td>.986, 1.079</td>
</tr>
<tr>
<td><strong>GENDER</strong></td>
<td>.531</td>
<td>.805</td>
<td>.408, 1.588</td>
</tr>
<tr>
<td><strong>EDUCATION (years)</strong></td>
<td>.338</td>
<td>1.062</td>
<td>.939, 1.202</td>
</tr>
<tr>
<td><strong>ChEI</strong></td>
<td>.690</td>
<td>1.187</td>
<td>.510, 2.763</td>
</tr>
<tr>
<td><strong>MMSE</strong></td>
<td>.010</td>
<td>.835</td>
<td>.729, .958</td>
</tr>
<tr>
<td><strong>CATEGORY FLUENCY</strong></td>
<td>.129</td>
<td>.947</td>
<td>.882, 1.016</td>
</tr>
<tr>
<td><strong>MRI FIELD STRENGTH</strong></td>
<td>.524</td>
<td>.774</td>
<td>.351, 1.705</td>
</tr>
<tr>
<td><strong>RIGHT PHG</strong></td>
<td>.047</td>
<td>1.146</td>
<td>1.002, 1.311</td>
</tr>
</tbody>
</table>

Notes:
- All measures at baseline unless otherwise specified.
- ChEI = cholinesterase inhibitor; MMSE = Mini Mental State Examination; MRI = magnetic resonance imaging; PHG = parahippocampal gyrus.
- Category fluency is for animals only.
- Overall model significant, $X^2$ (8, n = 589) = 20.475, $p = .009$

11.5.1.4 **Delusion subtype analysis**

Significant findings described above were further explored by subtype of psychosis symptoms.

Of the 256 participants in the ‘state’ delusion group or control who had the NPI completed at baseline (excluding those who developed delusions at a later timepoint), 232 (90.6%) had no delusions, seven (2.7%) had paranoid-type symptoms, five (2.0%) had misidentification-type symptoms and eight (3.1%) had a mixed picture. A total of 425 participants in the ‘trait’ or control group had the NPI completed. Of these individuals, 323 (76.0%) had no delusions at any time point, 27 (6.4%) had paranoid-type symptoms, 38 (8.9%) had misidentification-type symptoms and 37 (8.7%) had mixed-type symptoms.

One-way ANOVAs were run to compare volume of right anterior cingulate gyrus and right PHG between both ‘state’ and ‘trait’ delusion subtype (and control) groups. Multinomial regression models were then run to include covariates (age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency and MRI field strength).
For the ‘state’ subtype groups, none of the one-way ANOVAs or multinomial regression models were significant. For the ‘trait’ subtype group and right anterior cingulate gyrus volume, Welch’s test was performed as Levene’s test indicated that the assumption of homogeneity of variance was not met. This was significant \( (F_{3,62.616} = 3.214, \ p = .029) \) and driven by a significant difference in volume of this ROI between the paranoid-type and control groups \( (\ p = .017) \), mean volume of right anterior cingulate gyrus as a proportion of TIV was significantly lower in paranoid-type vs control group - \( 1.8 \times 10^{-3} \pm 0.4 \times 10^{-3} \) and \( 2.0 \times 10^{-3} \pm 0.3 \times 10^{-3} \) respectively). Only the multinomial regression model for the right anterior cingulate gyrus was significant \( (p = .023) \), with none of the ROIs significant within the models.

### 11.5.1.5 Subgroup analysis
Significant findings at the ROI level in the whole group were further explored in the subgroup of patients who had a diagnosis of either AD or aMCI at the time of their baseline MRI (i.e. excluding the 27 participants who were CN at baseline).

The right anterior cingulate gyrus remained significantly smaller in volume for the ‘state’ delusion group compared to the control group \( (t_{568} = 2.172, \ p = .030) \). The logistic regression model exploring the relationship between ‘state’ delusions and ROIs of ventral visual stream areas remained significant \( (p = .015) \), with the right PHG a significant predictor of ‘state’ delusion \( (p = .029) \). The one-way ANOVA for difference in volume of right anterior cingulate gyrus between ‘trait’ delusion subtype was significance \( (p = .010) \), again driven by the paranoid-type/control group differences \( (p = .003) \). The multinomial regression model for ‘trait’ delusion also remained significant \( (p = .005) \); for every 0.01% increase in right anterior cingulate volume (as a proportion of TIV) participants were 15.5% less likely to be in the paranoid-type group compared to the control group \( (\text{Exp}(\beta) .845, 95\% \ CI .722 - .988, \ p = .035) \).

### 11.5.2 Whole brain analysis
A whole brain analysis by VBM was completed to further identify any brain regions that had volume correlating with delusion group. Covariates were included as previously specified (age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency, TIV and MRI field strength). Continuous measures (age, years of education and TIV) were mean centered to reduce multicollinearity. The correlation matrix for continuous independent variables is shown in Table 11.3. While there were significant correlations between some variables, all Pearson correlation coefficients were < .80 and VIFs were all < 2.5, mean 1.3.

Two control participants, for whom data were missing on cholinesterase inhibitor use and animal category fluency score respectively, were removed from the VBM analysis.
T-tests were used to compare ‘state’ and ‘trait’ delusion groups with the control group. These were first run with no covariates, and then with the covariates as above. A FWE correction was applied, with a significance threshold of \( p_{FWE} < .05 \) and an extent threshold of \( k = 10 \) voxels. More exploratory models were run as a secondary analysis for t-tests with covariates included, and an uncorrected \( p \) value of < .001 with the same extent threshold of \( k = 10 \) voxels.

The contrasts of interest were ‘state’ delusion group < control and ‘trait’ delusion group < control. No clusters reached cluster or peak-level FWE corrected significance on SPMs.

11.5.2.1 Exploratory analyses
For the t-test contrast of ‘state’ delusion group < control, the left MFG (\( x = -24, y = 50, z = 38 \)) reached peak-level significance (uncorrected \( p < .001 \)), see Figure 11.20 and Figure 11.21 and unthresholded effect size map in Figure 11.22.
Figure 11.20 Statistical parametric map of ‘state’ delusion group < control t-contrast, displayed on whole glass brain view

Notes:
FDR = false discovery rate; FWE = family wise error; FWHM = full width half maximum.
Design matrix covariates: 1 = ‘state’ delusion group; 2 = control group; 3 = age (years; mean centered), 4 = gender, 5 = education (years; mean centered), 6 = cholinesterase inhibitor prescription, 7 = MMSE score, 8 = category fluency for animals, 9 = TIV (ml; mean centered), 10 = MRI field strength.
Figure 11.21 Location of left middle frontal gyrus cluster with lower volume in those with delusions at baseline ('state') compared to controls

Notes:
Significant clusters displayed on warped, skull-stripped, bias and noise-corrected average brain $p < .001$, $k = 10$ voxels.
Figure 11.22 Unthresholded t-contrast effect size map for association between ‘state’ delusions and grey matter volume

Notes:
Red indicates areas of lower grey matter volume in those with ‘state’ delusions compared to control. Blue indicates the reverse contrast.
For the t-test contrast of ‘trait’ delusion group < control the left SFG (x = -2, y = 65, z = 17) and right MTG (x = 68, y = -24, z = -14) reached peak-level significance at the more liberal significance threshold of $p < .001$, see Figure 11.23 and Figure 11.24 and unthresholded effect size map in Figure 11.25.

**Figure 11.23** Statistical parametric map of ‘trait’ delusion group < control t-contrast, displayed on whole glass brain view

![Statistical parametric map](image)

**Statistics:** p-values adjusted for search volume

<table>
<thead>
<tr>
<th></th>
<th>set-level</th>
<th>cluster-level</th>
<th>peak-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p$</td>
<td>$c$</td>
<td>$\alpha_{\text{FWE-coor}}$</td>
<td>$\alpha_{\text{FDR-coor}}$</td>
</tr>
<tr>
<td>0.793</td>
<td>2</td>
<td>0.816</td>
<td>0.729</td>
</tr>
<tr>
<td>0.769</td>
<td>7.89</td>
<td>0.729</td>
<td>11.3</td>
</tr>
</tbody>
</table>

Table shows 3 local maxima more than 6.0mm apart

Height threshold: $T = 3.10, p = 0.001$ (unc.)

Extent threshold: $k = 10$ voxels, $p = 0.035$ (unc.)

Expected voxels per cluster, $<k > = 173.052$

Expected number of clusters, $<c > = 2.95$

FWE: 4.408, FDR: Inf, FWE6: Inf, FDR6: Inf

Note:

FDR = false discovery rate; FWE = family wise error; FWHM = full width half maximum.

Design matrix covariates:
1. 'trait' delusion group
2. control group
3. age (years; mean centered)
4. gender
5. education (years; mean centered)
6. cholinesterase inhibitor prescription
7. MMSE score
8. category fluency for animals
9. TIV (ml; mean centered)
10. MRI field strength
**Figure 11.24** Location of right middle temporal gyrus cluster with lower volume in those with delusions at any time point (‘trait’) compared to controls

Notes:
Significant clusters displayed on warped, skull-stripped, bias and noise-corrected average brain.

\( p < .001, k = 10 \) voxels.
Figure 11.25 Unthresholded t-contrast effect size map for association between ‘trait’ delusions and grey matter volume

Notes:
Red indicates areas of lower grey matter volume in those with ‘trait’ delusions compared to control. Blue indicates the reverse contrast.
11.5.2.2 Model diagnostics

The design orthogonality matrix showed that the majority of covariates were orthogonal to each other, but there was a degree of collinearity between independent variables, see Figure 11.26 and Figure 11.27.

**Figure 11.26** T test of delusions at baseline (‘state’) vs control: design orthogonality

Notes:
* mean centered variables.

Group\(_1\) = ‘state’ delusion group; Group\(_2\) = control group.

ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; TIV = total intracranial volume.
Figure 11.27 T test of delusions at any time point ('trait') vs control: design orthogonality

Notes:
* mean centered variables.
Group1 = ‘trait’ delusion group; Group2 = control group.
ChEI = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; TIV = total intracranial volume.
12. Context Memory and Metamemory Study: Neuroimaging Methods

12.1 Background

As described in section 10.1, when participant recruitment stopped in March 2020, I decided to use the findings from the larger ADNI cohort to guide the analysis of the neuroimaging data from my patient-based study, as described below.

This study was reviewed and received ethical approval from the UCL/UCLH Joint Research Office (project ID 18/0038); Westminster REC (IRAS number 240572; REC reference 18/LO/0709) and the Health Research Authority.

12.2 Research questions

The primary research questions were:

- Do false memories correlate with specific patterns of lower regional brain volume across all AD participants (regardless of delusion group)?
- Do AD participants with delusions show different patterns of lower regional brain volume compared to those without?

The a priori hypotheses, refined by neuroimaging analyses in the ADNI cohort and previous work by my research group, were as follows:

- Regardless of delusion group, increased false recognition (on TCC or DRM/metamemory tasks) would correlate with reduced volume of hippocampi, entorhinal cortices, MTG, ventral visual stream structures (PHG, FFG), PFC structures and anterior cingulate gyri
- Volume of left PHG and anterior cingulate gyri would be reduced in those with delusions compared to control participants
- Volume of the right PHG would be greater in those with delusions compared to control participants

12.3 Participant selection

Participants were recruited as described in section 5.3, according to inclusion/exclusion criteria as laid out in section 5.3.1.
12.4 MRI protocol

The MRI protocol was finalised once pilot behavioural data were analysed, see Chapter 6. Structural MRI scans (3D T1 weighted sagittal images) were completed following the ADNI3 MP-RAGE protocol, to allow a direct comparison to the ADNI data analysis. All scans were performed at the Wellcome Centre for Neuroimaging at University College London, on the same Siemens MAGNETOM Prisma 3T scanner with a 64-channel headcoil. A 13 second three plane localiser scan was followed by an approximately five minute MP-RAGE scan, with parameters as per Table 12.1, leading to an isotropic voxel size of 1.0mm$^3$.

**Table 12.1 Acquisition protocol**

<table>
<thead>
<tr>
<th>SCAN #</th>
<th>SCAN TYPE</th>
<th>SLICE THICKNESS</th>
<th>TR/TE</th>
<th>FLIP ANGLE</th>
<th>FOV</th>
<th>SLICES/SLAB</th>
<th>DURATION (mm:ss)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Localiser</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>00:13</td>
</tr>
<tr>
<td>2</td>
<td>MP-RAGE</td>
<td>1.0mm</td>
<td>2300/2.96ms</td>
<td>9°</td>
<td>240 x 256mm</td>
<td>208</td>
<td>05:12</td>
</tr>
</tbody>
</table>

Notes:

FOV = field of view; MP-RAGE = magnetisation-prepared rapid gradient-echo; TR = repetition time; TE = time to echo.

12.4.1 Additional behavioural testing

Due to delays in recruitment and finalising the scanning protocol, there had been a significant delay between baseline visit and MRI visit for some participants. Behavioural measures of interest (the object-location binding, TCC and DRM/metamemory tasks, described in more detail in section 5.4.3) were therefore repeated prior to MRI scanning for any participant whose MRI was scheduled for more than one month after their baseline assessment.

12.5 Image preprocessing

Image preprocessing followed the same pathway as described in section 10.5 but, due to the smaller number of scans, was run on a Dell XPS 13 9370 laptop rather than through the UCL Computer Science High Performance Computing cluster.

12.6 Statistical analyses

12.6.1 Behavioural data and normality testing

Behavioural and ROI analyses were carried out in SPSS and included false memory measures that were most comparable with ADNI: false recognition on run one of the TCC task, and false
recognition on the DRM/metamemory task. For task descriptions, see section 5.4.3. For all analyses results were considered significant at a \( p \) value of <.05 unless otherwise stated.

Behavioural and ROI data were first viewed in the form of scatter plots, histograms and Q-Q plots to assess whether they followed a normal distribution. The assumption of normality was further tested using the Shapiro-Wilk test for non-normality.

Baseline characteristics and task performance at the time of MRI were compared between those with delusions and the control group using independent-samples t-tests or Mann-Whitney non-parametric tests if parametric data assumptions were violated.

### 12.6.2 False memory measures

#### 12.6.2.1 Regions of interest analyses

ROIs were selected from those available in the Neuromorphometric atlas based on hypotheses described in section 12.2 and include (bilateral) hippocampi, entorhinal cortices, MTG, PHG, FFG, MFG, SFG, IFG, SMFG, MFC and anterior cingulate gyri. ROIs were corrected for the degree of overall brain atrophy as described in section 10.6.2.

The relationships between false memory measures (false recognition on the TCC task and the DRM) and the TIV-corrected volume of each individual ROI were visualised using scatter plots to assess for any linear or monotonic relationships. Relationships were further explored using generalised linear regression modelling. Due to the small sample size, no covariates were included in the models. In order to allow a more meaningful interpretation of regression coefficients, in regression analyses ROIs were further transformed by multiplying by a factor of \( 10^3 \). This transformation meant that a one unit change in ROI volume in the regression model was the equivalent of a change in ROI volume of 0.01% of TIV.

As in the ADNI analysis, false recognition scores on both the TCC task and the DRM showed a Poisson distribution on visual inspection and Poisson regression models were run for each false memory measure, with each individual ROI as a predictor variable in a separate model.

Model fit was assessed by using the Pearson \( \chi^2/df \) value as a measure of overdispersion (where one represents equidispersion). A Pearson \( \chi^2/df \) value > 1.5 was considered to indicate a degree of overdispersion at which use of an unadjusted Poisson model was likely to lead to greater error.\(^{442}\) If this was the case, a Poisson model accounting for overdispersion\(^{443}\) or a negative binomial model were considered as alternatives. Log-likelihood values were used to evaluate model fit, with less negative log-likelihood values indicating better model fit.
Model diagnostics were run for any ROIs with significant results. Residuals were examined by creating scatter plots of standardised Pearson residuals against predicted values. If there was a constant spread of residuals across predicted values and the majority of standardised Pearson residuals were between -2.0 and 2.0, the model was considered a good fit.

To assess the impact of outlying values for any ROIs with significant results, as in section 10.6.2.1, Cook’s distance was calculated and visualised in bar charts. Cook’s distance was checked and if any values exceeded the more liberal threshold of one, further checks were carried out to rule out user error on data entry. If no error was identified, the model was re-run after excluding these values to determine their impact on results. This analysis was repeated using the more conservative calculation for threshold of Cook’s distance of \(4/(n - k - 1)\), see section 10.6.2.1.

12.6.2.2  **Whole brain analysis**

Exploratory whole brain analysis was not appropriate due to the small sample size.

12.6.3  **Delusion group**

12.6.3.1  **Regions of interest analyses**

ROIs were selected from those available in the Neuromorphometric atlas based on hypotheses described in section 12.2. These were (bilateral) PHG and anterior cingulate gyri. Prior to analysis ROIs were corrected for TIV as described in section 10.6.2.1.

The relationships between volume of these previously specified ROIs (as proportion of TIV) and delusion group were explored using independent-samples t-tests. As no significant results were found, no further regression analyses were performed.

12.6.3.2  **Whole brain analysis**

Exploratory whole brain analysis was not appropriate due to the small sample size.
13. Context Memory and Metamemory Study: Results of Analyses of Neuroimaging Data

As discussed in Chapter 12, participant recruitment occurred from memory services and from the Join Dementia Research register between 9th August 2018 and 16th March 2020.

13.1 Participant selection

Of the 27 participants included in the behavioural analyses, eight (four with delusions and four without) had completed MRI scanning by the point at which recruitment stopped because of COVID-19.

Neuroimaging was carried out between 19th December 2019 and 12th March 2020. Of the eight participants, four participants had MRI scans within a month of their baseline visit, with the remaining four having scans within six and 13 months (mean 10.5 ± 3.1 months) due to delays in recruitment and in finalising the scanning protocol. Behavioural measures of interest were repeated for the four participants whose MRI was scheduled for more than one month after their baseline assessment.
13.2 Demographic details

Participants in the sample had a mean age of 83.8 ± 8.9 (range 65 – 96) years, two (25.0%) were female and six (75.0%) were male and the majority (n = 7, 87.5%) were white. The mean sMMSE score for the group was 25.9 ± 3.0 (range 22 – 29). Four participants (50.0%) were experiencing delusions, identified from mental state examination and the NPI, and four (50.0%) were in the control group (AD, no delusions).

For the remaining 19 participants who did not have neuroimaging data collected, the mean age was 83.1 ± 6.4 years, eight (42.1%) were female and 11 were male (57.9%) and the majority (n = 15, 78.9%) were white. They had a lower mean sMMSE score of 25.2 ± 2.5.

Demographic details and baseline screening assessments for the delusion and control groups are summarised in Table 13.1. There were no significant differences between the delusion and
control group in terms of age, gender, race, handedness, years since diagnosis, cholinesterase inhibitor, antidepressant or antipsychotic prescriptions, baseline NPI score, presence of hallucinations, sMMSE score, baseline GDS-15 score or category fluency scores on baseline ACE-III. The groups only differed significantly in terms of years of education. However, given the small sample size, it less likely that differences between groups would reach significance and it is worth noting that there was a 3.8 point difference in sMMSE score between the groups (27.8 ± 2.5 in the control group compared to 24.0 ± 2.2 in the delusion group; $U = 1.5, p = .057$).

Performance of the groups on behavioural tasks repeated at time of MRI scan are summarised in Table 13.2. Note that the participant identified as an outlier for performance on the TCC task in section 7.3.2 was one of the participants in the delusion group who had imaging completed.
Table 13.1 Demographic details and screening results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 4)</th>
<th>DELUSIONS (n = 4)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGE (years)</strong></td>
<td>82.3 ± 12.8; 65 - 96</td>
<td>85.3 ± 3.8; 81 - 90</td>
<td>t6 = -.448, p = .670</td>
</tr>
<tr>
<td><strong>GENDER (female)</strong></td>
<td>0 (0.0)</td>
<td>2 (50.0)</td>
<td>†p = .429</td>
</tr>
<tr>
<td><strong>RACE</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>3 (75.0)</td>
<td>4 (100.0)</td>
<td>†p = 1.000</td>
</tr>
<tr>
<td>Black</td>
<td>1 (25.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td>Mixed race</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td></td>
</tr>
<tr>
<td><strong>RIGHT HANDEDNESS</strong></td>
<td>3 (75.0)</td>
<td>3 (75.0)</td>
<td>†p = 1.000</td>
</tr>
<tr>
<td><strong>EDUCATION (years)</strong></td>
<td>16.5 ± 3.1; 12 - 19</td>
<td>11.3 ± 1.7; 9 - 13</td>
<td>t6 = 2.960, p = .025</td>
</tr>
<tr>
<td><strong>YEARS SINCE DIAGNOSIS</strong></td>
<td>2.3 ± 1.5; 0.2 - 3.7</td>
<td>2.5 ± 1.2; 1.1 - 3.9</td>
<td>U = 8.0, p = 1.000</td>
</tr>
<tr>
<td><strong>MEDICATION</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ChEI/Memantine</td>
<td>4 (100.0)</td>
<td>2 (50.0)</td>
<td>†p = .429</td>
</tr>
<tr>
<td>Antidepressant</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>-</td>
</tr>
<tr>
<td>Antipsychotic</td>
<td>0 (0.0)</td>
<td>0 (0.0)</td>
<td>-</td>
</tr>
<tr>
<td><strong>NPI (baseline)</strong></td>
<td>15.0 ± 16.4; 0 - 32</td>
<td>19.0 ± 13.9; 1 - 34</td>
<td>U = 7.0, p = 1.000</td>
</tr>
<tr>
<td><strong>HALLUCINATIONS</strong></td>
<td>0 (0.0)</td>
<td>1 (25.0)</td>
<td>†p = 1.000</td>
</tr>
<tr>
<td><strong>sMMSE</strong></td>
<td>27.8 ± 2.5; 24 - 29</td>
<td>24.0 ± 2.2; 22 - 27</td>
<td>U = 1.5, p = .057</td>
</tr>
<tr>
<td><strong>GDS-15 (baseline)</strong></td>
<td>3.0 ± 2.5; 0 - 5</td>
<td>1.8 ± 1.3; 0 - 3</td>
<td>t6 = .908, p = .399</td>
</tr>
<tr>
<td><strong>CATEGORY FLUENCY</strong></td>
<td>11.3 ± 3.1; 7 - 14</td>
<td>9.0 ± 2.5; 7 - 12</td>
<td>t6 = 1.140, p = .298</td>
</tr>
</tbody>
</table>

Notes:
Values are expressed as mean ± SD; range or n (%). All measures at time of MRI unless otherwise specified.
ACE-III = Addenbrooke’s Cognitive Examination-III; ChEI = cholinesterase inhibitor; GDS-15 = short form Geriatric Depression Scale; NPI = Neuropsychiatric Inventory; sMMSE = standardised Mini Mental State Examination.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data. Categorical comparisons by chi-squared unless otherwise specified.
†p value from Fisher’s exact test.
13.3 Behavioural task performance

Performance of the groups on the TCC task at the time of MRI scanning is summarised in Table 13.2. The participant identified as an outlier for performance on the TCC task at baseline (CM007; see section 7.3.2) was again an outlier when the task was repeated prior to MRI, having high false recognition across the three runs (27, 36 and 41 compared to means of 5.0 ± 4.6, 17.0 ± 8.5 and 17.0 ± 9.6 for run one, two and three respectively for the delusion group with this participant removed). Removing this individual from comparison of the behavioural results between groups did not affect the outcome and it is notable that even with this individual removed, false recognition was greater in the delusion group than in the control group.

While no results reached significance, performance on the TCC task in this subgroup was representative of results in the larger group, with the delusion group having higher rates of false recognition across the runs, a greater degree of TCC, poorer discrimination and a response bias towards endorsing items as previously seen.
Table 13.2 Temporal context confusion task results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 4)</th>
<th>DELUSIONS (n = 4)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HITS RUN 1</strong></td>
<td>26.0 ± 3.4</td>
<td>26.3 ± 1.7</td>
<td>U = 6.5, p = .686</td>
</tr>
<tr>
<td><strong>HITS RUN 2</strong></td>
<td>25.3 ± 2.99</td>
<td>26.3 ± 1.7</td>
<td>U = 6.5, p = .686</td>
</tr>
<tr>
<td><strong>HITS RUN 3</strong></td>
<td>23.8 ± 7.23</td>
<td>24.3 ± 2.2</td>
<td>U = 5.0, p = .486</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 1</strong></td>
<td>0.0 ± 0.0</td>
<td>10.5 ± 11.6</td>
<td>U = 2.0, p = .114</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 2</strong></td>
<td>10.0 ± 6.2</td>
<td>21.8 ± 11.8</td>
<td>U = 2.5, p = .114</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION RUN 3</strong></td>
<td>6.3 ± 1.7</td>
<td>23.0 ± 14.4</td>
<td>U = 2.5, p = .114</td>
</tr>
<tr>
<td><strong>TCC RUN 2</strong></td>
<td>0.39 ± 0.23</td>
<td>0.42 ± 0.23</td>
<td>U = 6.0, p = .686</td>
</tr>
<tr>
<td><strong>TCC RUN 3</strong></td>
<td>0.29 ± 0.13</td>
<td>0.56 ± 0.25</td>
<td>tₜ = -1.941, p = .100</td>
</tr>
<tr>
<td><strong>DISCRIMINATION (d')</strong></td>
<td>3.97 ± 0.69</td>
<td>2.63 ± 1.29</td>
<td>U = 3.0, p = .200</td>
</tr>
<tr>
<td><strong>RESPONSE BIAS (c)</strong>*</td>
<td>0.36 ± 0.35</td>
<td>-0.22 ± 0.42</td>
<td>tₜ = 2.116, p = .079</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD.
TCC = temporal context confusion, calculated as TCC = FA x H⁰₁ / (H x FA), where HR = hit rate, FA = false alarm rate in run x and run one respectively.
Discrimination and response bias are calculated for run one.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.

Performance of the groups on the DRM/metamemory task at the time of the MRI scan is summarised in Table 13.3. The participant identified as an outlier for performance on the TCC task at baseline (CM007; see section 7.3.2) was also an outlier in terms of false recognition and false memory rates on the DRM/metamemory task prior to MRI, with high false recognition and false memory rates (29 and 15 respectively, compared to means of 9.0 ± 6.9, 7.0 ± 4.4 for the delusion group with this participant removed). Again, it is notable that even with this individual removed, false recognition and false memory rates were greater in the delusion group than in the control group. With this individual removed, the difference between the groups in discrimination was no longer significant (U = .0, p = .057), although the delusion group continued to have poorer discrimination performance (mean d’ 0.86 ± 0.20). For all other measures, the outcomes remained unchanged.
Performance on the DRM/metamemory task in this subgroup was largely representative of results in the larger group, with the delusion group having more hits, false recognition and false memories, poorer discrimination and a response bias towards endorsing items as previously seen. Performance differed in terms of the proportion of high confidence responses, which in this subgroup were greater in the control group than the delusion group. It is worth noting that for the control group there were only two participants who falsely recognised any items as previously seen – each of these participants only made one false recognition error during the run, but both were highly confident in their responses.

Table 13.3 DRM/metamemory task results

<table>
<thead>
<tr>
<th></th>
<th>CONTROL (n = 4)</th>
<th>DELUSIONS (n = 4)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>HITS</strong></td>
<td>18.8 ± 9.5</td>
<td>20.5 ± 9.2</td>
<td>U = 7.5, p = 1.000</td>
</tr>
<tr>
<td><strong>FALSE RECOGNITION</strong></td>
<td>0.5 ± 0.6</td>
<td>14.0 ± 11.5</td>
<td>U = 1.0, p = .057</td>
</tr>
<tr>
<td><strong>FALSE MEMORIES</strong></td>
<td>3.8 ± 4.2</td>
<td>9.0 ± 5.4</td>
<td>U = 3.5, p = .200</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE HITS</strong></td>
<td>0.55 ± 0.43</td>
<td>0.50 ± 0.37</td>
<td>U = 8.0, p = 1.000</td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE FR</strong></td>
<td>1.00 ± 0.00</td>
<td>0.28 ± 0.28</td>
<td>U = .0, p = .133</td>
</tr>
<tr>
<td>(n = 2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>HIGH CONFIDENCE FM</strong></td>
<td>0.43 ± 0.51</td>
<td>0.36 ± 0.45</td>
<td>U = 7.5, p = 1.000</td>
</tr>
<tr>
<td><strong>DISCRIMINATION (d’)</strong></td>
<td>2.35 ± 0.92</td>
<td>0.77 ± 0.25</td>
<td>U = .0, p = .029</td>
</tr>
<tr>
<td><strong>RESPONSE BIAS (c)</strong></td>
<td>0.73 ± 0.57</td>
<td>-0.30 ± 1.24</td>
<td>t_6 = 1.508, p = .182</td>
</tr>
</tbody>
</table>

Notes:
Values expressed as mean ± SD. High confidence values as proportion of total hits, false recognition or false memories respectively.
Maximum number hits per run is 30, maximum number false recognition per run is 30, maximum number false memories per run is 15.
DRM = Deese-Roediger-McDermott paradigm; FM = false memories; FR = false recognition.
Means compared by independent-samples t-test for parametric and Mann Whitney U test for non-parametric data.
13.4 Image quality

Homogeneity in the sample was reasonable, with scans having a mean correlation coefficient of > .80, see Figure 13.2. No images were flagged as having a mean overall correlation over two SD from the mean. Due to the small sample size, all images were checked visually for artefact, with no significant movement artefact visible. All scans were considered to be of acceptable quality.

As described in section 10.5.4, scans also received an overall weighted average IQR. All images received IQR grades of either B or B-, and a mean score of 83.2 ± 1.0%.
Figure 13.2 Data homogeneity - mean correlation of imaging data

Notes:
Output from CAT12 ‘VBM Data Homogeneity’ function.⁴²³
13.5 False memory measures

13.5.1 Regions of interest analysis

The relationships between volumes of previously specified ROIs corrected for TIV and measures of false memory (false recognition on run one of the TCC task and false recognition on the DRM/metamemory task) were visualised using scatter plots. No direct (linear or monotonic) relationships were identified. The relationships were therefore further explored using regression modelling. As in the ADNI analysis, prior to running the models, measures of volume of ROIs (as proportion of TIV) were transformed by multiplying by a factor of $10^3$.

Due to the small sample size, models were run without covariates.

Reviewing the distribution of results for both false memory measures in histogram plots revealed a likely Poisson distribution. Data for modelling predictors of false recognition in run one of the TCC task were overdispersed (mean Pearson $X^2/df$ value of $12.9 \pm 5.4$, where one represents equidispersion; for models that reached significance, mean Pearson $X^2/df$ value was $7.3 \pm 2.0$). Poisson modelling was therefore run with correction for overdispersion, using the Pearson $X^2/df$ value as the scale parameter in the model as described by McCullagh and Nelder (1989). Negative binomial modelling was considered but showed a worse fit for the data with more negative log-likelihood values.

Poisson modelling accounting for overdispersion using the method described by McCullagh and Nelder (1989) was also used for false recognition on the DRM (mean Pearson $X^2/df$ value of $13.7 \pm 3.6$; for models that reached significance, mean Pearson $X^2/df$ value was $7.4 \pm 4.0$). Negative binomial modelling was explored, but again led to more negative log-likelihood values.

Nine of the models that assessed the relationship between ROIs and false recognition on the TCC task were significant ($ps < .05$), with eight ROIs reaching individual significance within these models. Two models that assessed the relationship between false recognition on the DRM were significant ($ps < .05$), with both ROIs reaching individual significance within the models. Increased volume of several ROIs was linked to fewer words falsely recognised on either the TCC task or the DRM. For one model the opposite was the case: increased volume of the right anterior cingulate gyrus was linked to more words falsely recognised on the TCC task. See Table 13.4 for full details.
13.5.1.1 Medial temporal lobe regions of interest

For every 0.01% increase in volume (as a proportion of TIV) of the left entorhinal cortex there was a 45.4% reduction in false recognition on the TCC task (Exp(β) .546, 95% CI .322 - .924, p = .024); right entorhinal cortex there was a 41.3% reduction in false recognition on the TCC task (Exp(β) .587, 95% CI .345 - .997, p = .049). See Figure 13.3.

**Figure 13.3** Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of medial temporal lobe regions of interest

---

Notes:

Data from Poisson regression modelling of volume of individual ROIs (as proportion of TIV) as predictors of false recognition, without covariates.

Percentage change in false recognition is per unit increase in ROI volume (one unit = 0.01% of TIV).

Error bars indicate 95% confidence intervals. Only significant results displayed, p < .05.

EC = entorhinal cortex; L = left; R = right; TCC = temporal context confusion task.
13.5.1.2 Ventral visual stream regions of interest
For every 0.01% increase in volume (as a proportion of TIV) of the left and right PHG there was a reduction in false recognition on the TCC task of 51.9% and 45.2% respectively (Exp(β) .481, 95% Cl .276 - .836, p = .010; Exp(β) .548, 95% Cl .351 - .855, p = .008); left FFG there was a 16.5% reduction in false recognition on the TCC task (Exp(β) .835, 95% Cl .756 - .922, p = .000). See Figure 13.4.

Figure 13.4 Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of ventral visual stream regions of interest

![Bar chart showing percentage change in false recognition per 0.01% increase in volume for PHG and FFG.](image)

Notes:
Data from Poisson regression modelling of volume of individual ROIs (as proportion of TIV) as predictors of false recognition, without covariates.
Percentage change in false recognition is per unit increase in ROI volume (one unit = 0.01% of TIV).
Error bars indicate 95% confidence intervals. Only significant results displayed, p < .05.
L = left; R = right; PHG = parahippocampal gyrus; FFG = fusiform gyrus; TCC = temporal context confusion task.

13.5.1.3 Prefrontal cortex and anterior cingulate cortex regions of interest
For every 0.01% increase in volume (as a proportion of TIV) of the left and right SFG there was a reduction in false recognition on the TCC task of 14.0% and 26.3% respectively (Exp(β) .860, 95% Cl .762 - .970, p = .014; Exp(β) .737, 95% Cl .623 - .872, p = .000); left and right MFC there was a reduction in false recognition on the DRM task of 62.8% and 39.7% respectively (Exp(β) .372, 95% Cl .194 - .713, p = .003; Exp(β) .603, 95% Cl .370 - .982, p = .042); right anterior cingulate gyrus there was a 78.4% increase in false recognition on the TCC task (Exp(β) 1.784, 95% Cl 1.038 – 3.065, p = 0.036). See Figure 13.5.
Figure 13.5 Bar chart of percentage change in false recognition per 0.01% increase in volume (proportion of TIV) of prefrontal and anterior cingulate cortex regions of interest.
Figure 13.5 Cont. Bar chart of percentage change in false recognition per 0.01% increase in volume (as proportion of TIV) of prefrontal cortex and anterior cingulate cortex regions of interest

Notes:
Data from Poisson regression modelling of volume of individual ROIs (as proportion of TIV) as predictors of false recognition, without covariates.
Percentage change in false recognition is per unit increase in ROI volume (one unit = 0.01% of TIV). Error bars indicate 95% confidence intervals. Only significant results displayed, p < .05.
ACC = anterior cingulate cortex (anterior cingulate gyri regions of interest); DRM = Deese-Roediger-McDermott paradigm; L = left; MFC = medial frontal cerebrum; R = right; SFG = superior frontal gyrus; TCC = temporal context confusion task.

13.5.1.4 Model diagnostics
Further model diagnostics were run for those ROIs with significant results. Residuals were examined by creating scatter plots of standardised Pearson residuals fitted values, all of which demonstrated a reasonable model fit, with the majority of standardised Pearson residuals falling between -2.0 and 2.0. Reviewing the spread of the residuals in the plot indicated that three (75%) of the participants in the delusion group (participant identification numbers CM007, CM013 and CM022) had apparently outlying residuals across all ROIs (see Appendix 16).

Similarly, for those ROIs with significant results, Cook’s distance was calculated and visualised in bar charts (see Appendix 17). Values exceeded one on all but two ROIs, the majority of these were due to the participant clearly identified as an outlier on behavioural testing (CM007). Using the more conservative calculation for threshold of Cook’s distance to identify outliers of 4/(n – k - 1), several possible outliers were identified across all ROIs. The mean number of values exceeding this threshold was 1.1, 14.1% (range 1 - 2), for models of false recognition on the TCC task and brain volume, and 1.5, 18.8% (range 1 - 2) for models of false recognition on the DRM and brain volume. All data points flagged by this method were checked for user error on entry, without any error identified. All models were then run again, excluding values above the 4/(n – k - 1) threshold (0.67 for this data set).

All models for false recognition on the TCC task were no longer significant. Both models for false recognition on the DRM remained significant, with the right MFC being the only ROI to remain significant within the models (Exp(β) .228, 95% CI .122 - .425, p = .000) (see Appendix 18).
Table 13.4 Relationship between false recognition on the temporal context confusion and DRM/metamemory tasks and volume of regions of interest

<table>
<thead>
<tr>
<th>Region</th>
<th>TCC FALSE RECOGNITION MODELS</th>
<th>DRM FALSE RECOGNITION MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL (df = 1, n = 8)</td>
<td>EXP(β)</td>
</tr>
<tr>
<td>LEFT HIPPOCAMPUS</td>
<td>$X^2 = 3.152, p = .076$</td>
<td>-</td>
</tr>
<tr>
<td>RIGHT HIPPOCAMPUS</td>
<td>$X^2 = .053, p = .817$</td>
<td>-</td>
</tr>
<tr>
<td>LEFT ENTORHINAL CORTEX</td>
<td>$X^2 = 6.650, p = .010$</td>
<td>.546</td>
</tr>
<tr>
<td>RIGHT ENTORHINAL CORTEX</td>
<td>$X^2 = 4.853, p = .028$</td>
<td>.587</td>
</tr>
<tr>
<td>LEFT MIDDLE TEMPORAL GYRUS</td>
<td>$X^2 = 4.670, p = .031$</td>
<td>.911</td>
</tr>
<tr>
<td>RIGHT MIDDLE TEMPORAL GYRUS</td>
<td>$X^2 = 1.510, p = .219$</td>
<td>-</td>
</tr>
<tr>
<td>VENTRAL VISUAL STREAM</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left PHG</td>
<td>$X^2 = 9.281, p = .002$</td>
<td>.481</td>
</tr>
<tr>
<td>Right PHG</td>
<td>$X^2 = 10.693, p = .001$</td>
<td>.548</td>
</tr>
<tr>
<td>Left FFG</td>
<td>$X^2 = 14.667, p = .000$</td>
<td>.835</td>
</tr>
<tr>
<td>Right FFG</td>
<td>$X^2 = 1.749, p = .186$</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>TCC FALSE RECOGNITION MODELS</td>
<td></td>
</tr>
<tr>
<td>-------------------</td>
<td>------------------------------</td>
<td>-----------------------</td>
</tr>
<tr>
<td></td>
<td>MODEL</td>
<td>EXP(β)</td>
</tr>
<tr>
<td></td>
<td>(df = 1, n = 8)</td>
<td></td>
</tr>
<tr>
<td><strong>DL-PFC</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left MFG</td>
<td>$X^2 = .001, \ p = .981$</td>
<td>-</td>
</tr>
<tr>
<td>Right MFG</td>
<td>$X^2 = .192, \ p = .661$</td>
<td>-</td>
</tr>
<tr>
<td>Left SFG</td>
<td>$X^2 = 6.306, \ p = .012$</td>
<td>.860</td>
</tr>
<tr>
<td>Right SFG</td>
<td>$X^2 = 14.801, \ p = .000$</td>
<td>.737</td>
</tr>
<tr>
<td><strong>VL-PFC</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left IFG</td>
<td>$X^2 = 1.332, \ p = .248$</td>
<td>-</td>
</tr>
<tr>
<td>Right IFG</td>
<td>$X^2 = .022, \ p = .881$</td>
<td>-</td>
</tr>
<tr>
<td><strong>M-PFC</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left SMFG</td>
<td>$X^2 = .059, \ p = .807$</td>
<td>-</td>
</tr>
<tr>
<td>Right SMFG</td>
<td>$X^2 = .239, \ p = .625$</td>
<td>-</td>
</tr>
<tr>
<td>Left MFC</td>
<td>$X^2 = .819, \ p = .365$</td>
<td>-</td>
</tr>
<tr>
<td>Right MFC</td>
<td>$X^2 = 1.987, \ p = .159$</td>
<td>-</td>
</tr>
<tr>
<td><strong>LEFT ANTERIOR CINGULATE GYRUS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X^2 = .008, \ p = .928$</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>RIGHT ANTERIOR CINGULATE GYRUS</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 13.4 Cont. Relationship between false recognition on the temporal context confusion and DRM/metamemory tasks and volume of regions of interest

| Notes: | DL-PFC = dorsolateral prefrontal cortex; DRM = Deese-Roediger-McDermott paradigm; FFG = fusiform gyrus; IFG = inferior frontal gyrus; MFC = medial frontal cerebrum; MFG = middle frontal gyrus; M-PFC = medial prefrontal cortex; PHG = parahippocampal gyrus; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus; TCC = temporal context confusion task; VL-PFC = ventrolateral prefrontal cortex. Poisson regression models run without covariates. |
13.6 Delusion group

13.6.1 Regions of interest analysis

The relationships between volume of previously specified ROIs (as a proportion of TIV) and delusion group (delusions or control) were explored using independent-samples t-tests, see Table 13.5. No significant differences were identified.

Table 13.5 Comparison of volume in identified regions of interest between those with delusions and control

<table>
<thead>
<tr>
<th>Region</th>
<th>Control (n = 4)</th>
<th>Delusions (n = 4)</th>
<th>T Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEFT PARAHIPPOCAMPAL GYRUS</td>
<td>1.86 ± 0.22</td>
<td>1.69 ± 0.22</td>
<td>t_6 = 1.160, p = .290</td>
</tr>
<tr>
<td>RIGHT PARAHIPPOCAMPAL GYRUS</td>
<td>1.83 ± 0.28</td>
<td>1.63 ± 0.21</td>
<td>t_6 = 1.181, p = .282</td>
</tr>
<tr>
<td>LEFT ANTERIOR CINGULATE GYRUS</td>
<td>2.65 ± 0.20</td>
<td>2.55 ± 0.35</td>
<td>t_6 = 0.487, p = .643</td>
</tr>
<tr>
<td>RIGHT ANTERIOR CINGULATE GYRUS</td>
<td>1.48 ± 0.24</td>
<td>1.78 ± 0.28</td>
<td>t_6 = -1.639, p = .152</td>
</tr>
</tbody>
</table>

Notes:
Values are (mean ± SD) x 10^-3
All measures are mean brain volume for specified region of interest, as proportion of total intracranial volume.
14. Discussion

14.1 Principal findings

14.1.1 Behavioural research questions and hypotheses: patient-based study

The principal research question posed by the behavioural patient-based study was: Do AD participants with delusions have greater impairments in context memory (including spatial, temporal and associative memory organisation) and metamemory than those without delusions, when matched for overall cognitive impairment? The secondary research question for the study was: Do participants with misidentification-type delusions have a greater impairment in memory for spatial and temporal context than those with mixed or paranoid-type delusions? A priori behavioural hypotheses for the patient-based study were that participants with delusions in AD would have greater impairments in memory for spatial and temporal context and associative inference ability than those without delusions, that participants with delusions in AD would have higher confidence in memory errors than those without delusions and that individuals with misidentification-type delusions would have greater impairment in memory for spatial and temporal context than those with mixed or paranoid-type delusions.

14.1.2 Principal behavioural findings: patient-based study

There were no significant differences between those with and without delusions in AD in terms of demographic variables, global cognitive impairment, reasoning ability, attention or space and object perception.

No significant differences were observed between the groups in memory for context on the object-location binding task. The associative inference and relational memory task was not taken forward for use in the full participant cohort.

Across all three runs of the TCC task, participants with delusions had similar hit rates to the control group, but higher false recognition rates (this reached statistical significance for the first run only). Participants with delusions had greater impairment in memory for temporal context than those without, a finding which did not reach statistical significance. Participants with delusions had a more liberal response bias (being more likely to endorse all items as previously seen) and poorer discrimination between targets and distractors, with the difference between groups in terms of discrimination reaching statistical significance. In a binary logistic model which included gender as a covariate, both first run false recognition and discrimination remained statistically significant predictors of whether a participant had delusions, correctly predicting the presence of delusions in 80.0% and 50.0% of cases respectively.
On the DRM, there were no statistically significant differences between participants with delusions and those without in terms of either hits, false memories or false recognition. Participants with delusions were more likely to be highly confident in memory responses across all response types (hits, false recognition and false memories) than the control group. This difference reached statistical significance for high confidence responses to hits. In a binary logistic model with gender as a covariate, high confidence hits remained a statistically significant predictor of whether a participant had delusions, correctly predicting the presence of delusions in 80.0% of cases.

In terms of my primary research question, while no statistically significant differences were observed between the groups on measures of spatial or temporal context, group numbers were relatively small and results from this study are therefore not sufficient to rule out such a relationship. This is also the case for the finding of impaired metamemory in the delusion group, which did not reach statistical significance. Unfortunately, it was not possible to explore my secondary research question in this cohort due to the small sample size. Non-significant trends in the data for both memory for temporal context and metamemory were in line with my a priori hypotheses. The trend in terms of memory for spatial context was not.

14.1.3 Behavioural research questions and hypothesis: ADNI cohort study

In light of the results described in section 14.1.2, the primary research question posed by the behavioural study in the ADNI cohort was: Do AD participants with delusions at baseline (‘state’ delusions) show evidence of more false memories than those without delusions, when matched for overall cognitive impairment as measured by MMSE score? The secondary research question was: Do AD participants who experience delusions at any time during ADNI follow-up (‘trait’ delusions) show evidence of more false memories at baseline than those who do not, when matched for overall cognitive impairment as measured by MMSE score? The a priori hypothesis was that individuals with either ‘state’ or ‘trait’ delusions in AD would have increased rates of false recognition compared to those without delusions.

14.1.4 Principal behavioural findings: ADNI cohort study

Participants with ‘state’ delusions were significantly different from the control group across multiple demographic variables (race, diagnosis, years since diagnosis, prescription of antipsychotic medication and burden of neuropsychiatric symptoms). The ‘trait’ delusion group were more evenly matched with the control group, but differed statistically in terms of gender, prescription of antipsychotic medication and burden of neuropsychiatric symptoms. Participants with ‘state’ delusions were more cognitively impaired than the control group, and performed more poorly on measures of executive function, while those with ‘trait’ delusions did not
statistically differ from controls in terms of their degree of cognitive impairment or performance on executive function tasks.

The ‘state’ delusion group had higher rates of false recognition on both the ADAS-Cog 13 and RAVLT. This finding was statistically significant for false recognition on the ADAS-Cog 13. When those who were CN were excluded from this analysis the finding was no longer significant. However this appeared to be driven by outlying data points in the control group (representing 1.5% of the group), and the finding remained significant when these data points were excluded. Those in the ‘state’ delusion group had significantly poorer discrimination than controls on both the RAVLT and ADAS-Cog 13.

In multivariate binary logistic models for prediction of ‘state’ delusion group or control group, which included age, gender, years of education, prescription of cholinesterase inhibitors, MMSE score and category fluency as covariates, inclusion of false recognition on both ADAS-Cog 13 and RAVLT as independent variables led to statistically significant models. However, neither was a statistically significant predictor within the models, and the models had poor positive predictive value for delusion group overall. The only variable that was a statistically significant predictor within these models was MMSE score. This is not unexpected, given that the MMSE is a measure of disease severity, and delusions become more common as AD progresses. My findings of increased false recognition in those with delusions in my patient-based study (who were well matched with the control group in terms of sMMSE score) and of specific patterns of brain atrophy associated with false recognition when controlling for MMSE score support the fact that there is a relationship between false recognition and delusions, after accounting for the potential confounding effects of disease severity.

In terms of the primary research question, and in keeping with my a priori hypothesis, individuals with delusions at baseline did show evidence of more false memories than those without delusions. However, this finding was not sustained when overall degree of cognitive impairment was included as a confounder. In terms of the secondary research question, there were no statistically significant differences between those with ‘trait’ delusions and the control group on any of the false memory measures.

14.1.5 Neuroimaging research questions and hypotheses: ADNI cohort study

There were two principal research questions for the ADNI neuroimaging analyses. These were: Do false memories correlate with specific patterns of lower regional brain volume across all AD participants (regardless of delusion group)? Do AD participants with either ‘state’ delusions or ‘trait’ delusions show different patterns of lower regional brain volume than those without? The secondary research questions were: Do AD participants with delusions (either ‘state’ or ‘trait’)

237
show evidence of reduced GM volume in areas associated with false memory for the whole participant group? Do participants with different subtypes of psychosis symptoms have different patterns of reduced GM volume?

My a priori hypotheses were that across ADNI participants, regardless of delusion group, increased false recognition would correlate with reduced volume of MTL, ventral visual stream (entorhinal cortices, PHG, FFG and lingual gyri) and PFC structures, ACC and superior parietal lobules, that hippocampal, entorhinal, mPFC, dIPFC and regional ventral visual stream volume would be reduced in those with delusions and that participants with misidentification-type symptoms would have reduced GM volume in the ventral visual stream than those with mixed or paranoid-type psychosis symptoms.

14.1.6 Principal neuroimaging findings: ADNI cohort study

False recognition (as measured on either the RAVLT or ADAS-Cog 13) was inversely correlated with volume of MTL structures (bilateral hippocampi and entorhinal cortices), ventral visual stream structures (bilateral PHG and FFG), PFC structures in dIPFC (bilateral MFG and SFG), vIPFC (left IFG), mPFC (bilateral SMFG and right MFC) and bilateral ACC to a statistically significant degree, when confounding variables were included on overdispersion-corrected Poisson modelling. A FWE corrected whole brain VBM analysis including confounders additionally revealed a significant inverse relationship between false recognition on ADAS-Cog 13 and GM volume in the left MTG (as well as MTL structures also identified on ROI analysis – the right hippocampus and left FFG).

Using a more liberal significance threshold without FWE correction, significant inverse relationships were identified between intrusions on RAVLT and GM volume in dIPFC structures (bilateral MFG and right SFG), left fornix, right cuneus and bilateral cerebellum, between false recognition on RAVLT and GM volume in left ITG (as well as MTL and dIPFC structures identified on ROI analysis – bilateral hippocampi, right MFG and left SFG) and between false recognition on ADAS-Cog 13 and GM volume in left precentral gyrus, the left posterior cingulate cortex, the right posterior insula, lateral temporal lobe (LTL) structures (left STG, right MTG and right ITG), the left superior parietal lobule and bilateral supramarginal gyrus and the left inferior occipital gyrus (as well as MTL structures, dIPFC structures and ACC as also identified on ROI analysis – right hippocampus, left FFG, bilateral MFG, bilateral SFG and right ACC).

Participants in both delusion groups had smaller right ACC than the control group, a finding which reached statistical significance for the 'state' delusion group. This finding did not remain statistically significant when confounding variables were included in binary logistic modelling.
On exploration of this finding by psychosis subtype, there was no significant relationship between right ACC volume and psychosis subtype for the ‘state’ delusion group. However, for the ‘trait’ delusion group this appeared to be driven by lower right ACC volume in participants with the paranoid-type trait. Participants in the ‘state’ delusion group had larger right PHG than those in the control group, and this was also a statistically significant predictor of delusion group in a binary logistic model including confounding variables. Those with ‘state’ delusions also had larger right PHG compared to left PHG, although this finding did not reach statistical significance. A FWE corrected whole brain VBM analysis including confounders revealed no significant findings.

Using a more liberal significance threshold without FWE correction, the left MFG was found to be significantly lower in volume in the ‘state’ delusion compared to the control group, with the right MTG significantly lower in volume in the ‘trait’ delusion group compared to the control group.

In terms of how my findings relate to my research questions and hypotheses, the study showed an association between false memories and lower GM volume in multiple regions, which were generally in keeping with a priori hypotheses, including structures in MTL, ventral visual stream, PFC and ACC alongside the left MTG. As expected, using a more liberal threshold for significance showed lower GM volume across a more diverse array of brain regions. Presence of ‘state’ delusions were associated with lower volume of the right ACC, and individuals with different ‘trait’ psychosis subtypes had differing right ACC volumes. In terms of ‘state’ delusion, a possible role of lower left PHG volume and greater right PHG volume was also identified. Using a more liberal threshold for significance revealed a relationship between lower GM volume in dLPFC (left MFG) and ‘state’ delusion group, and lower LTL volume (right MTG) and ‘trait’ delusion group. Participants with delusions did have patterns of lower GM volume that overlapped with areas associated with false memory, including the right ACC and left MFG. False recognition was associated with lower left MTG volume while ‘trait’ delusion group was associated with lower right MTG volume.

14.1.7 Neuroimaging research questions and hypotheses: patient-based study

Informed by the results from analyses in the ADNI cohort described in section 14.1.6, the primary research questions were: Do false memories correlate with specific patterns of lower regional brain volume across all AD participants (regardless of delusion group)? Do AD participants with delusions show different patterns of lower regional brain volume compared to those without? The a priori hypotheses were that regardless of delusion group, increased false recognition (on the TCC or DRM/metamemory tasks) would correlate with reduced volume of
hippocampi, entorhinal cortices, MTG, ventral visual stream structures (PHG and FFG), PFC structures and ACC, that volume of left PHG and bilateral ACC would be reduced in those with delusions compared to control participants and that volume of the right PHG would be greater in those with delusions compared to control participants.

14.1.8 Principal neuroimaging findings: patient-based study

In the small sample available for my patient-based study (n = 8), false recognition (as measured on TCC task run one or the DRM/metamemory task) was inversely correlated with volume of MTL structures (bilateral entorhinal cortices), ventral visual stream structures (bilateral PHG and left FFG), PFC structures in dIPFC (bilateral SFG) and mPFC (bilateral MFC) to a statistically significant degree on overdispersion-corrected Poisson modelling. There was a positive correlation between false recognition on the TCC task and right ACC volume. Of note, once data with high degrees of influence as measured by Cook’s distance were removed, only the correlation between false recognition on the DRM/metamemory task and right MFC volume remained statistically significant.

There were no significant differences between delusion and control group in terms of PHG or ACC volume.

The small sample size means that this patient-based study lends only limited support to findings reported in section 14.1.6.

14.2 False memories and delusional beliefs

False memories (as measured by false recognition of both words and images on memory testing) were significantly associated with current delusional beliefs for two different cohorts of research participants with AD – both those recruited to my own study and those recruited to ADNI. In both groups, this finding was linked to significantly poorer discrimination ability in those with current delusional beliefs compared to those with AD without delusions. While participant groups in my own study were well matched in terms of overall cognitive impairment, ADNI participant groups were not, and false recognition was not a significant predictor of delusional beliefs when included in a model with confounders.

Attempts to explain how recognition memory errors occur involve two broader theories of memory: dual-process theory\(^{450, 451}\) and fuzzy-trace theory\(^{452, 453}\). In both models, new items are falsely recognised as old due to high feelings of familiarity (in fuzzy-trace theory this is increased by overreliance on gist memory), which are not appropriately overridden by recollection (due to reduced ability to retrieve verbatim traces of old stimuli).\(^{453, 454}\) This ability to correctly reject
new items that appear familiar by using specific details of old stimuli is known as ‘recollec-
tion’. While findings relating to familiarity are less consistent, both recollection and fa-
miliarity are thought to be reduced in AD compared to healthy ageing, with a cor-
responding reduction in false recognition.

My finding of increased false recognition in those with delusions in AD is in keeping with previous findings that false memories (as measured by confabulation) correlate with delusions in AD. It also supports findings that individuals with schizophrenia, particularly those who are currently experiencing delusions, have higher rates of false recognition and reduced discrimination ability, something which is also the case for otherwise-healthy individuals with high levels of positive schizotypy. There is however some clear inconsistency in the evidence base in this regard, as described in section 2.3.3, with some studies finding no difference, or indeed a reduction, in false recognition in individuals with schizophrenia, schizotypy and delusion-like ideation compared to controls. Few studies completed in individuals with schizophrenia or schizotypy have included overall cognitive function as a covariate, and it is possible that this goes some way to explain varying results. In this study, the finding of increased false recognition in those with delusions in the ADNI cohort was no longer significant when MMSE score was included in regression models alongside other potential confounding variables. While VIFs indicated no significant concerns regarding collinearity in these models, it appears likely that there is some relationship between the three variables. Delusions become more common as severity of AD increases and MMSE score reduces, with MMSE score the only significant predictor of delusion group in the ADNI cohort on binary logistic regression modelling. While false recognition is reduced in those with AD compared to healthy older adults, and therefore one might assume would be positively correlated with MMSE score, no studies have been identified that examine this relationship and more elaborate false memories (such as individuals with AD reporting that they have paid a bill when they haven’t) increase as MMSE score declines.

While no measure of memory confidence was available in the ADNI dataset, in my sample participants with delusions were more highly confident than those without delusions across all of their responses, a finding that reached statistical significance for high confidence hits. Individuals with schizophrenia are also found to be more highly confident in memory errors than controls, as are those with positive schizotypy and delusion-like ideation. However, this is not a consistent finding, with Bhatt et al. (2010) finding no overall difference in high confidence false recognition between those with and without delusions in schizophrenia but that those with delusions were more highly confident in false negative responses (i.e. incorrectly endorsing old items as new). Similar to the findings of the current study, Evans et al. (2019) found that
individuals with delusion-like ideation are more highly confident in correct responses. However, this finding is the exception in metamemory research in schizophrenia, with the majority of existing studies finding individuals with schizophrenia to have less confidence in correct responses than controls. Of note, confidence rating scales used to assess metamemory vary widely between studies, from binary judgements to Likert scales with various different wordings and ranges and continuous measures, for example, length of a button press. Impaired metamemory in schizophrenia and schizotypy has also been linked to a more liberal response bias, or a tendency for individuals to endorse items as previously seen. While there was a trend towards this for individuals with delusions on both the TCC and DRM tasks in my patient-based study, this did not reach significance, and there was no significant difference in response bias between either ‘state’ or ‘trait’ delusion group and control group in the ADNI cohort.

This study is the first to examine the relationship between both false recognition and metamemory and delusions in AD, and while the sample size is small (including ten participants with current delusions in my patient-based study and 42 in the ADNI cohort), it provides preliminary evidence to support common mechanisms for delusion formation across psychiatric disorders. Across both cohorts, individuals with delusions were more likely to falsely recognise new items as old, as per dual-trace theory, possibly due to increased feelings of familiarity. Impaired familiarity perception has been identified as a likely mechanism for lesion-induced delusional misidentification on meta-analysis of fMRI studies and in déjà vu experiences. In terms of how these findings relate to current cognitive models of delusion formation, Whittlesea and Williams (1998) describe how feelings of both false and true familiarity may arise when a mismatch between expectation and outcome (i.e. a prediction error) causes feelings of surprise. As described by Corlett et al. (2009), the experience of multiple surprising events requiring explanation due to aberrant prediction error processing may both lead to an increase in false recognition and predispose to inappropriate belief formation. Increased false recognition could also represent a mechanistic step in delusional belief formation by contributing to cognitive biases associated with delusional beliefs, for example, through impaired novelty detection leading to a bias against contradictory evidence. While metamemory findings were less robust, they contribute to accumulating evidence of metamemory impairment as an important factor in formation and maintenance of delusional beliefs, with metamemory impairment also associated with a bias against disconfirmatory information as demonstrated by Koller and Cannon (2021) and Rollwage et al. (2018).
14.3 Role of region-specific atrophy in false memory and delusion

14.3.1 Medial temporal areas

False recognition (as measured on either the RAVLT or the ADAS-Cog 13) was inversely correlated with volume of bilateral hippocampi and entorhinal cortices. FWE corrected whole brain VBM analysis including confounders confirmed the relationship between increased false recognition on ADAS-Cog 13 and reduced GM volume in the right hippocampus, as well as identifying a relationship between false recognition and lower LTL volume. While numbers in my patient-based study were small, a similar relationship was observed between false recognition on the TCC task and lower volume in bilateral entorhinal cortices.

Given that impaired recollection leads to increased false recognition, this finding is in keeping with the role of the hippocampus in cognitive function with gross hippocampal volume clearly associated with memory performance, and both hippocampus and entorhinal cortex crucial for recognition memory. Individuals with hippocampal lesions demonstrate the combination of impaired recollection with preserved familiarity thought to predispose to false recognition. In addition to their role in recollection, an important function of the hippocampus and entorhinal cortex (in conjunction with PFC and PHG) is encoding context to episodic memory, including spatial temporal order and more general contextual information including associative information (such as face-name pairings) and source (for example, if an item was seen as an image or a word). Of note, reduced left hippocampal volume has previously been associated with increased false recognition on the RAVLT in a smaller group (n = 77) of individuals with AD. Lower right hippocampal volume has been consistently linked to psychosis symptoms in AD with left entorhinal atrophy in AD recently found to increase risk of psychosis. Lower hippocampal volume is also a consistent finding in individuals with schizophrenia, with volume loss potentially preceding disease onset and correlated with severity of psychosis symptoms. The entorhinal cortex is reduced in volume in individuals with schizophrenia, but with less robust findings regarding correlation with severity of psychosis symptoms. The MTL is part of a network involved in encoding and retrieval of memory for spatial and temporal context, mediated via its functional connections to the PFC. It has therefore been hypothesised that MTL volume loss, and subsequent dysfunction within this network, leads to psychosis as a result of impaired ability to attribute context to sensory input, causing increased prediction errors and delusion formation.

In terms of the lower LTL volume observed on whole brain analysis (lower GM volume in the left MTG correlating with false recognition on ADAS-Cog 13), previous research has identified lower
right MTG volume in individuals with psychosis in AD, particularly in those with misidentification-type symptoms.\textsuperscript{90, 502} While psychosis-associated atrophy in AD has been found to be right-sided, left-sided change is in keeping with findings in schizophrenia. LTL volume loss is part of a pattern identified in very-early onset schizophrenia and correlates with psychosis symptom severity in these individuals\textsuperscript{503} with reduced left MTG volume observed in individuals with chronic schizophrenia diagnoses\textsuperscript{504} and those with newly diagnosed schizophrenia and their unaffected siblings.\textsuperscript{505, 506} It is thought that left-sided MTG atrophy may be related to impaired language processing,\textsuperscript{504} which may explain why left MTG atrophy was related to increased false recognition of words in the current study.

14.3.2 Ventral visual stream

False recognition (as measured on either the RAVLT or the ADAS-Cog 13) was inversely correlated with bilateral PHG and FFG volume, with the FWE corrected whole brain VBM analysis confirming the inverse relationship between false recognition on the ADAS-Cog 13 and GM volume in the left FFG. My patient-based study provided additional support for this relationship, with false recognition on the TCC task inversely related to volume of bilateral PHG and left FFG. If, as discussed in section 14.2, false memories represent a mechanistic step in delusion formation, this is in keeping with my previous finding of reduced volume of ventral visual stream structures in individuals with delusions in AD, particularly explained by lower left PHG volume.\textsuperscript{2}

The PHG and FFG are part of the hippocampal-PFC network of context encoding and retrieval, with the PHG involved in processing both spatial and temporal context of memory for visual information\textsuperscript{198, 199} and the FFG involved in object and face recognition,\textsuperscript{507} with the left FFG selectively responsive to written words.\textsuperscript{508} PHG and left FFG activity are observed during correct recognition on recognition memory testing\textsuperscript{165, 509, 510} and both the PHG and left FFG have therefore been hypothesised as having a role in distinguishing true from false memory.\textsuperscript{165, 510} Volume of the left PHG has previously been shown to be inversely related to false recognition on the RAVLT in AD.\textsuperscript{262} This study supports this hypothesis, as PHG and left FFG atrophy was linked to an increase in false memories.

In addition to my own previous findings of a relationship between left PHG atrophy and delusions in AD, a prospective study found lower PHG bilaterally in those who went on to develop delusions in AD,\textsuperscript{511} and right FFG atrophy has also been associated with AD psychosis.\textsuperscript{90} While there was no relationship between delusional beliefs and atrophy in these regions in the current study, it is notable that these previous findings were linked primarily to misidentification-type psychosis symptoms, while the majority of individuals in both ADNI and my own patient cohort had paranoid-type or mixed-type symptoms. However, Nakaaki et al.
(2013) did find bilateral PHG loss compared to controls in a group of AD participants (n = 18) who primarily had paranoid-type delusions (n = 13, 72.2%). While this study used a false discovery rate correction for multiple comparisons (which is less conservative when compared to the FWE correction used in this study), in the current study no ventral visual stream areas had significantly different GM volume between those with and without delusions even when the FWE correction was removed.

Alongside hippocampal changes, parahippocampal atrophy is a consistent finding in individuals with schizophrenia, and has been confirmed by various meta-analyses. Reduced left PHG volume is found in patients with first episode schizophrenia, with PHG volume negatively correlated with positive symptom scores and lower PHG volume predictive of poorer clinical outcomes. Reduced left PHG volume is also predictive of psychosis onset in high risk groups.

Atrophy in the FFG bilaterally has been found in individuals with both treatment-naïve and chronic schizophrenia in whom FFG volume correlates negatively with symptom severity and degree of insight, in individuals with psychosis in autism spectrum disorder and in those with schizotypal symptoms. Changes in PHG and FFG function have been linked to recognition memory impairments in non-AD psychosis; individuals with schizophrenia have reduced PHG activity while completing a recognition memory task during fMRI compared to controls and individuals with first episode psychosis have smaller differences in right FFG activity between true and false recognition compared to controls. Similar to how hippocampal pathology is thought to relate to psychosis as described in section 14.3.1, PHG atrophy is thought to impair context appraisal, leading to prediction errors due to disruption in hippocampal comparisons of internal representations against reality. FFG atrophy has been linked to delusional misidentification due to its role in facial recognition, with lower left FFG volume recently found to be associated with impaired recognition of emotion and emotional intensity in the faces of others. This potential first factor neuropsychological impairment is hypothesised to lead to delusional belief formation in combination with known biases towards negative interpretation of the feelings and intentions of others.

14.3.3 Prefrontal cortex

False recognition (as measured on either the RAVLT or ADAS-Cog 13) was inversely correlated with multiple PFC structures: diPFC (bilateral MFG and SFG), viPFC (left IFG) and mPFC (bilateral SMFG and right MFC). In the sample from my patient-based study, false recognition (as measured on either TCC task run one or the DRM/metamemory task) was also inversely correlated with PFC structures in both diPFC (bilateral SFG) and mPFC (bilateral MFC). While findings from exploratory whole brain analyses without FWE correction should be interpreted
with some caution, individuals with ‘state’ delusions in the ADNI cohort were found to have reduced GM volume in dIPFC (left MFG) compared to controls.

These findings are in keeping with the role of the PFC in cognitive control, including decision making, selective attention and response inhibition and goal-directed behaviour. Relevant to the current study, a region within the right dIPFC is specifically activated when prediction errors occur and is thought to be involved in re-evaluating and updating learned relationships in response to these errors. Findings are also in keeping with observations that individuals with lesions in dIPFC have increased false recognition on memory testing. Intact PFC function is thought to be important for monitoring memory retrieval to limit false recognition rates with bilateral PFC activity greater for true than for false recognition. Specifically, the MFG is recruited bilaterally (on fMRI) when making context memory judgements, with right MFG preferentially involved in recognition memory. Older adults have increased false recognition compared to younger healthy controls, associated with reduced PFC activity. As described in Chapter 3, false memories (as measured by intrusion errors) in AD correlate with hypometabolism in right dIPFC.

As discussed in section 2.2.3, both right-sided generalised PFC atrophy and bilateral MFG atrophy are associated with delusions in AD, alongside hypometabolism in right dIPFC, vIPFC and bilateral OFC, and increased dIPFC tau phosphorylation. PFC, and particularly dIPFC, dysfunction has been consistently implicated as a key component of schizophrenia neuropathology. dIPFC atrophy is present in individuals with chronic schizophrenia and schizotypal disorder, and is associated with delusional beliefs in bipolar affective disorder and onset of psychosis in individuals at risk mental state. As described in section 2.3.3, when individuals with schizophrenia falsely recognise words on memory testing during fMRI, PFC activity is reduced compared to controls and similar to findings for FFG activity described in section 14.3.2, the difference in left dIPFC (left MFG) activity between true and false recognition is reduced in individuals with first episode psychosis compared to controls. Aberrant right PFC activation is observed on fMRI during a prediction error task in a ketamine model of delusional symptoms and in individuals with first episode psychosis, in whom it correlates with the extent of delusion symptoms. PFC dysfunction is thought to lead to delusional beliefs due to impairments in reality monitoring and belief evaluation, with right dIPFC dysfunction specifically linked to delusion formation through abnormal prediction error signalling. Changes found in the current study of bilateral PFC atrophy associated with false recognition in AD are in keeping with these findings. While left sided MFG atrophy alone was associated with delusional beliefs, which appears at odds with
previous right-sided findings, this was uncorrected for error and should therefore be interpreted with caution.

14.3.4 Anterior cingulate cortex

False recognition (as measured on the ADAS-Cog 13) was inversely correlated with GM volume in ACC bilaterally. My own smaller sample contradicted this finding, with a positive correlation between false recognition on the TCC task and right ACC volume. In the ADNI cohort, participants in both ‘state’ and ‘trait’ delusion groups had smaller right ACC than the control group, a finding which reached statistical significance for the ‘state’ delusion group. On exploration of this finding by psychosis subtype, in the ‘trait’ delusion group this appeared to be driven by the paranoid-type subtype. This finding did not remain statistically significant when confounding variables were included in binary logistic modelling, and there was no significant relationship between right ACC volume and ‘state’ psychosis subtype.

The ACC has heterogeneous functionality, which remains relatively poorly understood. Its ventral region has a role in emotion regulation and monitoring social prediction error via functional connections to OFC and amygdala.\textsuperscript{554-557} Dorsally, the ACC appears involved in cognitive control, through conflict and error monitoring, motor planning and action execution\textsuperscript{558-561}, as such the ACC has also been identified as important to prediction error signalling.\textsuperscript{562-565} Results regarding ACC activity during false recognition are mixed, with some finding reduced activity during false compared to true recognition\textsuperscript{165, 538} and others the opposite,\textsuperscript{510, 566} possibly influenced by differing behavioural paradigms used between studies. ACC activity during false recognition is consistently found to be reduced in older adults (with greater false recognition rates) compared to younger adults.\textsuperscript{567, 568} As described in Chapter 3, in individuals with AD, ACC volume is inversely correlated with false memories (as measured by degree of confabulation),\textsuperscript{161} and increased connectivity is seen between right PFC and ACC and OFC in those with confabulations compared to those without.\textsuperscript{263}

The finding of reduced right ACC volume in those with ‘state’ delusions is in keeping with previous findings of bilateral ACC atrophy\textsuperscript{80} and generalised right frontal atrophy associated with psychosis in AD\textsuperscript{31, 105, 106}; bilateral ACC hypoperfusion is related to paranoid-type but not misidentification-type delusions in AD\textsuperscript{541} and correlates with delusion severity.\textsuperscript{110} ACC atrophy is also a well-replicated finding across the spectrum of psychosis diagnoses, including those with schizophrenia with and without current delusions,\textsuperscript{569} individuals with the at risk mental state who go on to develop psychosis\textsuperscript{570, 571} and affective psychosis.\textsuperscript{572} Compared to healthy controls, schizophrenia patients demonstrate reduced activation in the dorsal ACC on memory test errors
during fMRI. ACC dysfunction is hypothesised to lead to delusion formation through aberrant salience and prediction error processing.

Of note, conflicting results regarding the relationship between ACC atrophy and false recognition were found in the current study. While it is most likely that this is explained by the small sample size in the non-ADNI cohort, it worth noting that the ACC is known to be a region with interindividual anatomical variability and additional gender differences. Women are found to have larger overall ACC volume and greater asymmetry in ACC volume (right greater than left) compared to men, meaning that it is perhaps also relevant that in my smaller sample half of the group with delusions were female compared to none of the control group.

14.4 Role of regions of greater volume in false memory and delusion

Participants in the ‘state’ delusion group had larger right PHG volumes than those in the control group. This finding was particularly affected by confounding variables. While an independent-samples t-test comparing right PHG volume between ‘state’ delusion and control group was null, when standard confounding variables (including MMSE score) were included in the binary logistic regression model the right PHG was revealed as a significant predictor of delusions, with a 0.01% volume increase meaning participants were 14.6% more likely to be in the ‘state’ delusion group \( (p = .047) \). When additional ventral visual stream regions were included as covariates, the right PHG remained significant, with a 0.01% volume increase leading to a 40.3% increase in likelihood of participants being in the ‘state’ delusion group \( (p = .026) \).

This could be interpreted as an indication that the volumes of other ventral visual stream areas also play a role in predicting delusion group membership. Of note, both left PHG and left FFG were close to statistical significance within the model \( (p \text{ values of .060 and .059 respectively}) \). Comparing Nagelkerke \( R^2 \) values confirms that this larger ventral visual stream model explains a greater proportion of the variance in delusion group, but collinearity between the volumes of ventral visual stream structures (as indicated by the high VIFs, see section 11.5.1) would reduce the statistical power of the model. The regression coefficient from the binary logistic regression model that was restricted to right PHG volume is thus likely more reliable.

The finding that individuals with delusions in AD have greater right PHG volume alongside increased false recognition rates is potentially contradictory to my finding of an inverse relationship between volume of MTL structures (bilateral hippocampi, entorhinal cortices and PHG) and false recognition on the RAVLT and ADAS-Cog 13. It is possible that voxel clusters within bilateral MTL identified by VBM as inversely related to false recognition on ADAS-Cog 13 may have included PHG to some extent (see Figure 11.7). On VBM this lower MTL volume is
predominantly left-sided. This could suggest that it is asymmetry, rather than the overall volume of PHG that is predictive of delusions. It is also important to note that the participants included in the analysis of false recognition include both ‘state’ and ‘trait’ delusion groups and right PHG volume was not predictive of ‘trait’ delusions.

Hemispheric asymmetry is a well-known and extensively documented finding across a wide range of neurological and psychiatric disorders, see Ocklenburg and Gunturkun (2017) and Mundorf and Ocklenburg (2021) for recent in-depth reviews. Relative right-sided GM preservation and loss or reversal of brain asymmetry have been demonstrated in schizophrenia and proposed as a mechanistic step in psychosis in schizophrenia due to abnormalities in distinguishing between internal and external speech. Specifically, reversal of the left-greater-than-right PHG asymmetry of healthy controls has been observed previously in schizophrenia and schizotypal personality disorder.

This observation of rightward asymmetry of PHG volume is also found when comparing individuals with AD to those with MCI and healthy older adults, and it is therefore possible that this is a marker of disease severity rather than being a causal factor in delusion formation. However, limited existing literature suggests a possible role for right PHG overactivity in delusion formation in schizophrenia; Surguladze et al. (2006) finding that individuals with schizophrenia had reduced right PHG activity in response to fearful faces and increased activity in response to neutral faces, compared to controls, with increased activity correlating with degree of reality distortion and Kirino et al. (2019) observed that right PHG activity on fMRI correlated with severity of positive symptoms.

I therefore further explored the finding of greater right PHG volume in those with delusions to determine if within-group asymmetry was present in the sample. While those with ‘state’ delusions had a larger right compared to left PHG this finding did not reach statistical significance on a paired-samples t-test.

### 14.5 Context memory, false memory and delusion

As described in Chapter 1, the initial motivation for this thesis was to explore how memory for context relates to delusions in AD, with hypotheses as described in section 14.1.1 based on my previous research work. While there were no significant findings on specific context memory tasks in my patient-based study, and no specific context memory tasks were completed in the ADNI cohort, in both studies individuals with delusions in AD had increased false recognition on memory tasks. This included false recognition of both words (the ADAS-Cog 13) and pictures of objects (the TCC task). While not specifically designed to assess memory for context, false
recognition on these tasks can be conceptualised as a failure of context memory due to an inability to determine if these common words or objects were previously seen in the context of the task or not.\textsuperscript{592} In support of this, false recognition can be reduced by increasing contextual associations at encoding,\textsuperscript{592, 593} with degree of TCC known to correlate with degree of confabulation.\textsuperscript{594} That context memory is impaired in individuals with delusions in AD is also supported by the neuroimaging findings of this study, in which this increase in false recognition is inversely correlated with GM volume in brain regions known to be involved in processing memory for context: the hippocampus, PHG and PFC.\textsuperscript{278, 595}

14.6 Limitations

Several limitations mean that the results discussed in this chapter should be interpreted with caution. Firstly, the impact of the COVID-19 pandemic, as discussed in section 7.1.1, meant that recruitment for my patient-based study ended prematurely and numbers in my prospective patient-based study were therefore ultimately too low to give the study sufficient power based on a priori power calculations. While the ADNI sample was larger, the number of individuals with ‘state’ delusions remained relatively small, although above the number considered sufficient for power for both behavioural (see section 8.5.1) and neuroimaging analyses (see section 10.6.2). In both groups, small numbers reduced my ability to explore secondary hypotheses relating to delusional subtypes. Consistent findings across the groups in terms of behavioural results (increased false recognition associated with delusions) are reassuring. Small numbers in the delusion groups did not affect analysis of neuroimaging correlates of false memories, as this was completed for the entire cohort, with significant results as previously discussed in this chapter. However, relatively few areas of atrophy were associated with delusional beliefs on neuroimaging analysis. Given that regional volume loss in schizophrenia has been found to correlate with severity of psychosis symptoms, in both the patient-based sample (who had relatively mild delusional symptoms, with a mean frequency x severity score of 1.7 ± 1.1) and ADNI sample (for whom severity of delusional symptoms was unknown for half of the group) it is possible that relatively mild symptoms also reduced the power of the studies.

Both studies may also have been affected by limitations relating to accurate diagnosis, both of AD itself and of delusions. While both cohorts included only individuals with diagnosed AD as per NINCDS-ADRDA criteria,\textsuperscript{15} both cohorts also included individuals with relatively mild AD (the maximum sMMSE score for my patient-based sample was 29, and for the ADNI cohort was 30). Overlap in cognitive function between those with MCI and early AD is acknowledged, making correctly diagnosing these individuals challenging.\textsuperscript{596} In order to maximise sample size, and because the ADNI cut-off for ‘cognitively normal’ is an MMSE score of 24, I made the decision to
include individuals who received an AD diagnosis at any timepoint in the ADNI sample. This introduced further diagnostic heterogeneity. However, cognitive function (as measured by the MMSE) was included as a covariate in analyses of ADNI data, while my patient-based sample groups were well matched in terms of sMMSE score, which may go some way to controlling for this. Similarly, it cannot be completely ruled out that a proportion of participants in both cohorts had undiagnosed DLB, particularly given that some were experiencing hallucinations (n = 2 for my patient-based sample and n = 25 for the ADNI cohort at baseline) at a relatively early stage of disease. The high degree of correlation between delusions and hallucinations prevented inclusion of hallucinations as a confounder in regression analyses. There is further possible diagnostic overlap with posterior cortical atrophy (PCA). In my patient-based cohort I took a thorough history of onset of cognitive difficulties to confirm likely non-PCA AD diagnosis, and a clinical history is also taken as part of ADNI recruitment procedures. In both studies, the inclusion criteria of sufficient visual ability for relatively complex neuropsychological testing may also have reduced the chance of including participants with PCA diagnoses. In support of this, analysis of atrophy in posterior cortex and visuospatial deficits in ADNI1, ADNI-GO and ADNI2 participants did not reveal the deterioration over time that would be expected with PCA diagnosis. Delusions themselves present further diagnostic challenges, as explored in detail in section 2.3.1. Care is required in assessment to ensure that unusual beliefs are not in fact based in reality, and that they do not represent comorbid psychopathologies such as late onset schizophrenia, affective psychoses, delirium or substance misuse. Use of a carer-rated NPI in addition to me completing a mental state exam was intended to minimise these difficulties for my patient-based sample, as carers should be better placed to determine if unusual beliefs are in fact based in truth while I confirmed their presence. Jeste and Finkel (2000) propose that seven criteria should be met to diagnose psychosis in AD, for which participants in my study consistently met six, but did not all meet their severity criteria (‘severe enough to cause some disruption in patients’ and/or others’ functioning’), with the inclusion of individuals with relatively mild severity of psychosis symptoms chosen to increase sample size in both my patient-based and ADNI study, in line with previous work by my research group. Participants did not meet any of the exclusion criteria for dementia-related psychosis more recently proposed by Fischer et al. (2020). Relying on the carer-rated NPI for determining presence of absence of delusions itself has some limitations, as not all carers were living with participants so would not necessarily be aware of infrequently experienced unusual beliefs. In addition, the nature of longitudinal research with set follow-up time points as in the ADNI study means that some occurrences of delusions could have been missed. It therefore remains possible that some individuals with delusions may have been incorrectly classified as controls.
While the small sample size in my patient-based study is acknowledged as potentially limiting the reliability of data, a further issue that may impair comparisons between this sample and the ADNI cohort is the difference between the groups in terms of antipsychotic prescribing. None of the individuals with delusions were prescribed antipsychotics, compared to four (9.5%) in the ADNI ‘state’ delusion group. Given that antipsychotics are found to have a relationship with atrophy in schizophrenia (albeit with mixed results regarding direction of correlation\textsuperscript{599, 600}), this further reduces direct comparability of results.

Due to its impact on recruitment and sample size for the prospective study initially planned for this thesis, the COVID-19 pandemic necessitated inclusion of the retrospective ADNI cohort analysis. The use of retrospective data has inherent limitations, as it is necessary to work within the parameters of available neuropsychological testing and neuroimaging. This meant that it was unfortunately not possible to further explore findings relating to metamemory impairment in my patient-based cohort, as there were no measures of memory confidence completed for ADNI participants.

While all ADNI neuropsychological testing occurred in the relatively controlled clinic environment, for my patient-based study baseline assessment and neuropsychological testing was completed in participants’ homes, in order to reduce the barrier to access for the study for older people and therefore aid recruitment. The nature of testing in a home environment meant that there were some unpredictable (and unavoidable) interruptions to testing and assessment, which included telephones ringing and distraction by members of the family and pets. These were however rare and, if they occurred during timed tasks, the task was repeated. This affected only one assessment of a measure of interest: for one individual in the control group the TCC task had to be repeated due to a telephone ringing during the task. This could potentially have increased the rate of false recognition for this individual; to minimise the potential for this bias, repetition of the task was delayed until the end of the testing session, over 30 minutes after the interruption. Results from the TCC task in healthy controls indicate that any initial increase in false recognition on the task is no longer present after a 30 minute delay.\textsuperscript{217}

There are also potential limitations relating to the statistical methodology of the study. Firstly, for my own patient-based study, it was not possible to complete blinded testing and data analysis due to the nature of only one researcher (me) being responsible for all elements of the study. Both my patient-based study and the ADNI study involve multiple statistical comparisons, which increases the chance of type one errors. However, results in the study were generally consistent with hypotheses defined a priori and with previous findings as described earlier in this chapter. Given the acknowledged limitations of the small samples in terms of overall statistical power, further correcting for multiple comparisons was considered likely to increase
the risk of type two errors. There are potential limitations related to the use of regression analyses; the most significant of these as relates to the current study is the possibility of their output being affected by collinearity of variables. However, correlation between variables was assessed by calculating VIFs, which were below the less conservative threshold of 10.0 for all analyses. VIFs were also below the more conservative threshold of 2.5 for all analyses except those involving multiple ROIs (as discussed in section 14.4). For the ADNI dataset the highest correlation between continuous variables was a Pearson correlation coefficient < .80, see Table 11.3.

Lastly, various methodological decisions in the neuroimaging analysis, both in choice of ADNI image data and in the image analysis pipeline chosen for this study have the potential to introduce bias. For example, while including participants from the entire ADNI database allowed a large sample size, particularly important to increase the sample of individuals with delusions, the inclusion of data acquired from different imaging protocols (as described in section 10.4) was a significant source of heterogeneity. However, as described in section 10.4, previous VBM analyses using structural MRI data from AD participants across multiple sites and scanners have reported only minimal confounding of results. Magnetic field strength was also included as a confounding variable in ADNI neuroimaging analysis. The image analysis pipeline used for this study involved coregistration of images to an existing MNI space template developed from 555 healthy control subjects in the IXI-database rather than a sample-specific template. This was necessary as otherwise it would not have been possible to complete ROI analysis in CAT12. While the healthy control template is considered appropriate for all subjects apart from children and has previously been used in ROI and whole brain VBM analyses in AD participants, it remains possible that this impacted the warping process and could have reduced accuracy of ROI volume measurements for the more atrophied brains in the samples. More general limitations that are the case for all research using structural MRI include the inability to directly infer causality from patterns of atrophy, the poor reproducibility of MRI findings across studies, the possibility for atrophy to be produced from other factors leading to spurious results, the lack of molecular specificity and the lack of ability to assess function (reviewed in detail by Johnson et al. (2012) and recently discussed by Weinberger and Radulescu (2021)).

14.7 Clinical relevance and future directions

The finding of increased false recognition in those with delusions compared to controls across two relatively simple and quick to administer recognition memory tests has potential cross-diagnostic clinical relevance. This finding builds on existing research into markers of psychosis proneness in schizophrenia, where the importance of a shift towards cognitive markers of risk
has recently been emphasised.\textsuperscript{603} This is particularly important in schizophrenia, where evidence that early intervention (with programmes combining personalised medication management, individual therapy and family psychoeducation) for those at high risk improves outcomes\textsuperscript{604} and reduces cost burden for health services.\textsuperscript{605} The impact of early intervention for psychosis in AD, where options for pharmacological management are more limited and come with significant risks,\textsuperscript{122-124} is less well understood.\textsuperscript{606} This reflects current UK National Institute for Health and Care Excellence guidance, which places an emphasis on physical screening and avoiding antipsychotic drug use unless individuals are considered ‘at risk of harming themselves’ or in ‘severe distress’, but offers no guidance on psychological or psychosocial interventions.\textsuperscript{607} Early identification of those at risk of psychosis in AD, and prompt assessment and intervention via non-pharmacological means, has the potential to limit both patient and carer distress. This is an important area of investigation for future research. While findings suggest the possibility of behavioural and neuroimaging markers of delusion-proneness in AD, this study was not designed to explore this. Prospective, longitudinal studies are needed to further explore predictive markers of emergent delusions in AD, including false memory, metamemory and neuroimaging findings.

Studies that have investigated the effects of a variety of non-drug interventions in reducing false memories have shown promising results in terms of reducing false memories, including psychological interventions (including diary keeping, feedback-based and psychotherapy interventions) for confabulation following hypoxic brain injury,\textsuperscript{608} cognitive programs (including autobiographical reminiscence therapy and metacognitive training) and transcranial magnetic stimulation for impaired context memory in psychosis,\textsuperscript{609, 610} metacognitive training for false memories in depression,\textsuperscript{249} and direct transcranial current stimulation for false recognition in healthy older adults.\textsuperscript{611} Given the high risk associated with pharmacological management of delusions, exploring how these relatively low risk interventions for false memory relate to psychosis symptoms is an important next step.

This study provided tentative support for a link between impaired metamemory and delusions in AD, albeit in a small sample. This finding needs to be replicated and further explored prospectively in a larger sample, as it has implications for treatment. A recent study has demonstrated that 10 weeks of weekly 90 minute metamemory training improves cognitive performance (measured by delayed recall, category fluency and Boston Naming Test) in older adults with subjective memory complaints.\textsuperscript{612} Improvement of cognitive ability was correlated with increased prefrontal cortical thickness on MRI.\textsuperscript{613} Furthermore, as this training program can be administered by appropriately trained non-clinical staff, it is a relatively low cost, low risk and potentially relevant intervention for those with delusions. In addition, given the equivocal
findings regarding metamemory in psychosis research, future studies should endeavour to use consistent confidence-rating scales.

Further neuroimaging research in this area may be complementary to exploration of cognitive interventions for false memory and metamemory and their impact on psychosis in AD. Such future studies would ideally be prospective such that consistent imaging protocols are used between participants. For example, fMRI could be used to investigate whether brain activity during false memory paradigms maps to the structural correlates of false memory identified in this study, whether these patterns are different for individuals with delusions and whether cognitive interventions for false memory and metamemory lead to any functional changes.

14.8 Conclusion

The two complementary studies described in this thesis provide evidence of specific memory impairments, namely increased false recognition rates, associated with both delusions in AD and a distinct pattern of atrophy on structural MRI. These findings are in two cohorts with mild to moderate AD (MMSE scores > 16), and therefore relate to individuals who are most likely to be able to engage with simple cognitive interventions. Further research in this area could be of significant value, both to those experiencing these often distressing symptoms and their families.
References


259. Osterrieth PA. (1944) Le test de copie d’une figure complexe; contribution a l’etude de la perception et de la memoire. Archives de Psychologie. 30, 206-356.


Army Individual Test Battery. (1944) Manual of directions and scoring. DC, USA: War Department, Adjutant General’s Office.


293


434. IXI-Database [Available from: http://brain-development.org/].


**Appendix 1. Modified EPHPP Tool for Systematic Review 1: Neuroimaging Correlates of False Memory in Alzheimer’s Disease**

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### A. Selection bias

1. Are the individuals selected to participate likely to be representative of the target population? (yes = 1; no/unclear = 0)

**Total score (1 = strong; 0 = weak)**

### B. Study design

1. Was sample size ≥20, or derived from a power calculation? (yes = 1; no = 0)
2. Are inclusion and exclusion criteria clearly described? (yes = 1; no = 0)
3. Were details of sample characteristics included? (Age, MMSE and 1 other = 1; age and MMSE = 0.5; age and MMSE not included = 0)

**Total score (3 = strong; 1.5 – 2.5 = moderate; ≤1 = weak)**

### C. False memory task

1. Has the task been validated in any population? (yes/ not applicable = 1; no/unclear = 0)
2. Has the task been adapted for AD, or validated in an AD population? (yes = 1; no/unclear = 0)

**Total score (2 = strong; 1 = moderate; 0 = weak)**

### D. Neuroimaging methodology

1. Is a validated method of data acquisition used? (yes = 1; no/unclear = 0)
2. Is a validated method of data processing used? (yes = 1; no/unclear = 0)

**Total score (2 = strong; 1 = moderate; 0 = weak)**

### E. Analysis

1. Was analysis hypothesis driven? (yes = 1; no/unclear = 0)
2. Has the study controlled for potential confounders? (yes = 1; partially = 0.5; no = 0)
3. Has the study adjusted for multiple comparisons if appropriate? (yes/ not applicable = 1; no/unclear = 0)
4. Are statistical methods appropriate for study design/sample size? (yes = 1; no/unclear = 0)

**Total Score (≥3.5 = strong; 1.5 – 3 = moderate; <1.5 = weak)**

Global rating (Strong = 0 weak ratings; Moderate = 1 weak rating; Weak = ≥2 weak ratings)

**Total score (5-14)**
<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a discrepancy between the reviewers with respect to</td>
<td></td>
</tr>
<tr>
<td>component A-E ratings? (Yes/No)</td>
<td></td>
</tr>
<tr>
<td>If yes, indicate reason for discrepancy: 1 = Oversight; 2 =</td>
<td></td>
</tr>
<tr>
<td>Differences in interpretation of criteria; 3 = Differences in</td>
<td></td>
</tr>
<tr>
<td>interpretation of study</td>
<td></td>
</tr>
<tr>
<td>Final decision of both reviewers (global rating: Strong/Moderate/Weak)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 2. Modified EPHPP Tool for Systematic Review 2: Neuropsychological Correlates of Spatial Context Memory in Alzheimer's Disease

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

A. Selection bias
1. Are the individuals selected to participate likely to be representative of the target population? (yes = 1; no/unclear = 0)

**Total score** (1 = strong; 0 = weak)

B. Study design
1. Was sample size ≥20, or derived from a power calculation? (yes = 1; no = 0)
2. Are inclusion and exclusion criteria clearly described? (yes = 1; no = 0)
3. Were details of sample characteristics included? (Age, MMSE and 1 other = 1; age and MMSE = 0.5; age and MMSE not included = 0)

**Total score** (3 = strong; 1.5 – 2.5 = moderate; ≤1 = weak)

C. Spatial memory task
1. Has the task been validated in a healthy population? (yes/ not applicable = 1; no/unclear = 0)
2. Has the task been adapted for AD, or validated in an AD population? (yes = 1; no/unclear = 0)
3. Do AD participants show floor effects with the tool? (yes/ unclear = 0; no = 1)

**Total score** (≥2 = strong; 1 = moderate; 0 = weak)

D. Correlated neuropsychological task(s)
1. Has the task(s) been validated in a healthy population? (yes/ not applicable = 1; no/unclear = 0)
2. Has the task(s) been adapted for AD, or validated in an AD population? (yes = 1; no/unclear = 0)
3. Do AD participants show floor effects with the tool(s)? (yes/ unclear = 0; no = 1)

**Total score** (≥2 = strong; 1 = moderate; 0 = weak)

E. Analysis
1. Was correlation analysis hypothesis driven? (yes = 1; no/unclear = 0)
2. Has the study controlled for potential confounders? (yes = 1; partially = 0.5; no = 0)
3. Has the study adjusted for multiple comparisons if appropriate? (yes/ not applicable = 1; no/unclear = 0)

**Total Score** (≥2.5 = strong; 1.5 – 2 = moderate; <1.5 = weak)

Global rating (Strong = 0 weak ratings; Moderate = 1 weak rating; Weak = ≥2 weak ratings)

**Total score** (5-14)
<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Is there a discrepancy between the reviewers with respect to component A-E ratings? (Yes/No)</td>
<td></td>
</tr>
<tr>
<td>If yes, indicate reason for discrepancy: 1 = Oversight; 2 = Differences in interpretation of criteria; 3 = Differences in interpretation of study</td>
<td></td>
</tr>
<tr>
<td>Final decision of both reviewers (global rating: Strong/Moderate/Weak)</td>
<td></td>
</tr>
</tbody>
</table>
Appendix 3. Context Memory and Metamemory Study

Participant Information Sheet

THE WELLCOME TRUST CENTRE FOR NEUROIMAGING
(INC. LEOPOLD MULLER FUNCTIONAL IMAGING LABORATORY)
12 QUEEN SQUARE, LONDON WC1N 3BG

R&D / Sponsor Reference Number(s): 18/0038
IRAS Registration Number: 240572

Participant Information Sheet

Study title: COMET-AD

(Investigating Context Memory and Metamemory in the generation of delusions in Alzheimer's disease: a PhD student study)

Invitation
You are being invited to take part in a research study. This research study is part of an academic programme (PhD), and is being completed by student researcher Dr Emma McLachlan, supervised by Professor Robert Howard.

Before you make a decision, it is important for you to understand why the research is being carried out and what it will involve. Choosing not to take part will not disadvantage you in any way. Before you decide, please take time to read the following information carefully and discuss it with your relatives, friends or GP if you wish. Please ask us if there is anything that is not clear or if you would like more information.

Thank you for reading this.

What is the purpose of the research?
We are interested in how delusions form in Alzheimer's disease. Delusions are firmly held false beliefs, which are common in Alzheimer's disease. They can be distressing for patients and their relatives. At the moment, we treat delusions in Alzheimer's in the same way that we treat them in schizophrenia. However, this treatment has significant side effects and only works over a short period of time. We therefore think it is important to understand more about how delusions form, to find new ways to treat them.

This study aims to improve our understanding of how different types of memory error may lead to delusions. To study this, we will ask volunteers to perform simple 'pen and
paper’ and computer-based tasks. No prior experience of using a computer is required. We will ask participants to complete some of these computer-based tasks while their brain is being scanned.

This research is being completed as part of an academic programme (PhD) at University College London, by student researcher Dr Emma McLachlan.

**Why have I been invited?**

We are asking 50 people in the very early stages of Alzheimer's disease to take part in the study. You have been invited because you have expressed some interest in the study when it was briefly discussed with you by our research team or a member of your clinical care team.

**Do I have to take part?**

No, it is entirely up to you to decide whether or not you want to take part.

If you decide to take part you are free to withdraw at any time and without having to give a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive or your legal rights.

If you decide to take part then you will be asked to sign a consent form, which you will be given a copy of to keep.

**What will happen to me if I take part?**

The study will require up to three appointments for screening and completing questionnaires and memory tasks, and a fourth appointment for a brain scan.

1) **Screening (1 hr 30)**

It you agree to participate you will be contacted by Dr Emma McLachlan, the student researcher, to arrange an assessment session. This can take place at your home, or at the Wellcome Trust Centre for Neuroimaging in Queen Square in Central London, or at your local outpatient clinic: whichever is most convenient for you. We will pay for any transport costs to attend the session. You will be asked for your consent to take part in this assessment session, and it is possible that following this you or the researcher may decide that you are not suitable for the study.

You will first be asked to take part in an interview that will last about 90 minutes. This interview will follow a standardized format and you will be asked questions about yourself
including your age, level of education, current health and symptoms, past medical history and medication history. You will then be asked to complete questionnaires about your mood and mental state. You will also be asked to complete some standard memory tests and visual tasks (for example, counting the number of solid cube shapes on a board). These will be done to assess whether you are eligible to take part.

If you are eligible and wish to take part in the study, we will discuss the study further and you will have the opportunity to ask any questions you may have. You will be asked to sign a consent form to take part in the study. You will be free to withdraw from the study at any time, without having to give a reason. Deciding to withdraw from the study will not affect the standard of care you receive or your legal rights. We will then arrange a second appointment to complete further memory tests.

2) **Main Assessment Session (1 hr 30 – 2 hrs)**

You will not have to do any preparation before this appointment. Again, this can take place at your home, or at the Wellcome Trust Centre for Neuroimaging in Queen Square in Central London, or at your local outpatient clinic: whichever is most convenient for you. We will pay for any transport costs to attend the session.

You will meet with student researcher Dr Emma McLachlan and be asked to complete some computer-based memory tasks. No previous experience of using computers is needed. Dr McLachlan will explain what to do during the session, and you will be given time to practise to ensure you are comfortable with the tasks. The session will last up to 2 hours.

You will be asked to complete some computer-based tests of your verbal reasoning and attention (for example – you will be asked to press one button when a number appears on the computer screen). You will also be asked to complete three computer-based tests of context memory (for example *where* or *when* you saw something before) and one test of ‘metamemory’, which is the ability to self-evaluate your memory. Again, no computer experience is required. Before beginning each task, you will have the opportunity to practice the task so that you feel comfortable and familiar with what to do.

A quarter of participants will be asked for their feedback on these last three memory tasks, and we may make changes based on this. If large changes are made, those participants who are asked to provide feedback may be contacted to organise a further 45 minute appointment, during which the four computed-based memory tasks will be repeated.
3) **Brain scan (1 hr)**

You will then be asked to have a brain scan. A technique called functional magnetic resonance imaging (fMRI) will be used. Functional MRI is a painless and safe imaging technique for taking pictures of the brain. It does not involve being exposed to radiation (as it works through the use of a magnet) and has no known side-effects. Before being invited to the centre we will check that you are suitable for scanning using MRI. Dr McLachlan, the student researcher, will be running the scanning sessions, alongside radiology professionals. Once you have removed any metal you are wearing or carrying, you will be asked to lie still on a table inside the scanner for about 60 minutes (including preparation time) while the images of your brain are taken. Dr McLachlan will be in constant contact with you via an intercom in the scanner, and if you feel uncomfortable or anxious in any way the scan can be stopped at any time.

During the scan, you will be asked to complete up to two memory tasks that will be presented on a screen inside the scanner, giving your responses using a simple button box. You will have an opportunity to practice these tasks and ask any questions prior to the scan. While you complete these tasks we will collect pictures of your brain using functional MRI. This allows us to learn about how the brain works by looking at the blood flow to different parts of the brain whilst the brain performs the different tasks. We will also take more detailed anatomical images of your brain – a structural scan. While structural images of your brain are made, you will be asked to relax, and will not have to complete any tasks. Please see the separate 'your functional MRI scan' leaflet for further information on having a brain scan.

Your visit to the Centre (including practise, scanning and debriefing) will take about an hour. The brain scans will take place at the Wellcome Trust Centre for Neuroimaging, UCL. Transport for you and a friend or relative will be provided to travel to and from the scan.

**What are the possible disadvantages and risks of taking part?**

There are no known risks or side-effects associated with carrying out the questionnaires or computerized tasks. You may feel anxious before or tired after taking part in the tasks, but we will do everything we can to minimise or prevent this. You will be asked about your well-being often, and you will be given the opportunity to have either a short rest break or for the testing to be stopped.

There are no known risks involved in undergoing an fMRI scan. No radiation or injections are involved and you will not feel anything during the scan. However, you will be required
to lie still for approximately an hour, and some people find the enclosed space of the scanner uncomfortable. The MRI scans can be very noisy. To reduce the noise, you will wear headphones or earplugs that are designed to make the experience more comfortable. You will have access to a call button at all times and can press this to stop the scan and ask to be removed from the scanner. There is a possibility you may feel anxious or tired after taking part, but we will do everything we can to minimise this. All transport and refreshments on the day of the scan will be provided. You may find the appointments tiring or inconvenient, however the timing of these visits will be arranged when convenient for you and by providing all transport by taxis we aim to reduce any inconvenience as much as possible.

**Are there any benefits to taking part?**

There will be no direct benefit to you from taking part in this research. However, your data may contribute important theoretical information to our understanding of how the brain works, and in the future may aid attempts to manage delusional beliefs in patients with Alzheimer’s disease.

**What if I have a complaint?**

If you have any comments or concerns about any aspect of the study, e.g. the way you have been approached or treated during the course of the study, you should speak to Dr McLachlan (Tel: XXXX XXX XXXXX, XXXXXXXXXXX). If you remain unhappy and wish to make a formal complaint, please write to Professor R. Howard, Division of Psychiatry, 6th Floor, Maple House, 149 Tottenham Court Road, London, W1T 7NF. All correspondence will be addressed in strict confidence. The department is covered by UCL/UCLH liability insurance. If you are under the care of the Camden and Islington NHS Foundation Trust and wish to speak to an independent person, please contact the Advice and Complaints Service using the details below:

Camden and Islington NHS Foundation Trust, Advice and Complaints Service 1st Floor, East Wing, St Pancras Hospital, 4 St Pancras Way, London, NW1 0PE Phone number: 02033177102, email address: feedback@candi.nhs.uk

If you are under the care of Barnet, Enfield and Haringey Mental Health NHS Trust and wish to speak to an independent person, please contact the Advice and Complaints Service using the details below:

Patient Experience Advisor, Barnet Enfield and Haringey Mental Health NHS Trust Ivy House, Chase Farm Hospital, The Ridgeway, Enfield, EN2 8JL Phone number: 02087024700, email address: beh-tr.patient.experience@nhs.net
If you are under the care of East London NHS Foundation Trust and wish to speak to an independent person, please contact the PALS and Complaints Service using the details below:

Complaints and PALS Manager, FREEPOST RTXT-HJLG-XEBE, Trust Headquarters
The Green, 1 Roger Dowley Court, Russia Lane, London E2 9NJ9
Phone number: 08000858354, email address: elft.complaints@nhs.net

If you have joined the study via Join Dementia Research and wish to speak to an independent person, please contact Join Dementia Research using the details below:

Join Dementia Research, c/o NIHR Clinical Research Network Coordinating Centre
Minerva House, 5 Montague Close, London SE1 9BB
Phone number: 03001115111, email via:
https://www.joindementiaresearch.nihr.ac.uk/contactus

*Will my taking part be kept confidential?*

All information that is collected from you during the course of the research will be kept strictly confidential, anonymised, and will be collected and stored in accordance with the Data Protection Act 1998. Data will be kept in secured accommodation and on secured computers in the Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL and Division of Psychiatry, UCL. The data will be used only for the purpose of informing the research questions in this study, and only accessible by the relevant research teams at the Centre. The data will be retained securely, and may be accessed by the research teams for comparison with future data relating to the research question.

The research you are taking part in may be published, and as part of this process the anonymised results of the research may be published in scientific journals. You will never be identified and these data are always presented in an anonymised fashion. We may also share such anonymised data and results with other accredited researchers. Again, you or the data you provide will never be identified as all data will be anonymised.

*What will happen if I decide to withdraw from the study?*

If you decide to take part you are free to withdraw at any time and without having to give a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care you receive or your legal rights. If you decide to withdraw from the study we will ask for your permission to store any information we have already collected. Any information will be kept strictly confidential, anonymised and the anonymised data will be accessible only by the research team. Alternatively, if you prefer, any data we have collected prior to you withdrawing from the study will be destroyed.
Will my GP be informed?
If it is necessary to clarify whether you are safe to have an MRI scan (for example if you are unsure whether you have had surgery in the past) we will ask for your written permission to contact your GP to inform them you are interested in participating in the study and ask for relevant past medical information. We will not pass any other information to your doctor unless it is important for your health and you have agreed that we do so. Should the MRI scan unexpectedly reveal any clinically relevant abnormality, we would like your permission to notify your GP and usual clinical care team.

What will happen to the results of the research study?
Anonymous results will be published in peer-reviewed academic journals and presented in posters and talks at academic conferences. You will be asked whether you wish to receive a full copy of the report by the research team which will then be sent to you after the study had finished. You will not be identified personally in any publication.

Who funds this research?
Our work is funded by the National Institute for Health Research Biomedical Research Centre UCL.

Who has reviewed this research?
A Research Ethics Committee reviews all proposals for research using human participants before they can proceed. This research has been approved by the Health Research Authority and the National Research Ethics Service.

Contacts for further information
If you have any questions after reading this information sheet please ask Dr M![Lachlan: Dr Emma M![Lachlan, student researcher: XXXXXXXXXXX (e.mclachlan@ucl.ac.uk)

Other members of the research team include:
Prof Robert Howard, Chief Investigator, Division of Psychiatry (Robert.howard@ucl.ac.uk)

Thank you for taking the time to consider participating in our research.
Appendix 4. Wellcome Trust Centre for Neuroimaging MRI Exclusion Criteria

**NEUROIMAGING: VOLUNTEER SAFETY QUESTIONNAIRE**

The handling, processing, storage and destruction of data will be conducted in accordance with the Data Protection Act (2018).

I. The *absolute* contra-indications for MRI scanning are listed below. If you answer yes to any of the following we will *not* be able to scan you:

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>Do you have a cardiac pacemaker or active implant?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you have cerebral aneurysm clips?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you undergone permanent eye-lining as a cosmetic procedure?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you any cochlear implants (ear implants)?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

II. Before entering the MRI scan room please inform us if any of the following apply:

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you claustrophobic?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you had any surgery?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are any artificial devices implanted into your body (e.g. joint replacements, coils, implants or clips)?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you ever had a job in the metal-working industry or have you ever been exposed to metal dust or splinters?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you ever had an injury to your eyes involving metal at high speed?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you wear a hearing aid or have dentures, bridges, braces, or dental implants?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you have any shrapnel from a war injury or explosion?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you have an infusion pump or Hickman line?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you diabetic, epileptic or ever had a seizure?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you wearing Nicotine patches?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Do you have one or more tattoos?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Have you removed all loose metal objects, wallets, watches, and jewellery from your person?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Full Name (print)........................................................................................................
**

**Signature....................................................................................................................
**

**Date................................................................................................................................
**
For female volunteers only:

<table>
<thead>
<tr>
<th>Question</th>
<th>Yes</th>
<th>No</th>
<th>Unsure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Could you be pregnant?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you wearing a diaphragm / Intrauterine Device (Coil) or any other contraceptive device?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are you wearing any hormone replacement contraceptive patches?</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Signature.................................................................................. | Date.........................

Checked by:

_________________________  _____________  _____________
Radiographer/QU              Date                 Signature
Appendix 5. Context Memory and Metamemory Study Consent Form

INVESTIGATING CONTEXT MEMORY AND METAMEMORY IN THE GENERATION OF DELUSIONS IN ALZHEIMER’S DISEASE: A PhD STUDENT STUDY

Study Number: IRAS 240572
Participant Reference Number for study:

CONSENT FORM - Confidential

Name of Researcher:

Please initial box

1. I confirm that I have read and understood the information sheet dated 5th September 2018 (Version 2) for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.

2. I confirm that I have had sufficient time to consider whether or not I want to be included in the study.

3. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected.

4. I understand that a member of the research team may contact my GP or access my medical records to clarify aspects of my medical history if necessary. I understand my GP will only be informed if any of the results of the investigations carried out as a part of the research are important for my health and this has been discussed with me.

5. I understand that if I lose capacity to provide my consent to remain involved in the study, any data from my involvement up to that point will continue to be used for the study.

6. I agree to take part in the above study.

__________________________ ____________________________
Name of Participant Date Signature
<table>
<thead>
<tr>
<th>Name of Person taking consent</th>
<th>Date</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>(If different from Chief Investigator)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Researcher to be contacted | Email | Phone number |
| (If there are any problems) |       |             |
| Dr E Mclachlan | e.mclachlan@ucl.ac.uk | XXXXXXXXXX |
Appendix 6. Context Memory and Metamemory Study Case Report Form

CRF COVER PAGE

COntext memory and METamemory in the generation of delusions in Alzheimer’s Disease (COMET-AD):

a PhD student study

SCREENING NUMBER: ___ ___ ___

PARTICIPANT STUDY NUMBER: ___ ___ ___ (IF ASSIGNED)

PARTICIPANT INITIALS: ___ ___ ___
<table>
<thead>
<tr>
<th>Screening Number</th>
<th>Participant Study Number</th>
<th>Group Assigned:</th>
</tr>
</thead>
</table>

**VISIT 1 (SCREENING)**

**DEMOGRAPHIC DATA**

**Participant Informed Consent:**

Date consent form signed: __ __ / __ __ __ / __ __ __ __ (DD / MMM / YYYY)

**Demographic Data:**

Date of Birth: __ __ / __ __ __ / __ __ __ __ (DD / MMM / YYYY)

Age: __ __/ __ __

Origin:  
- White / Caucasian
- Black or African
- Asian
- Other, specify: ________________________________________

Sex:  
- Male
- Female

Diagnosis:

Diagnosis made by:  
- GP  
- Hospital team  
- Other  
- Unknown

Date of diagnosis:  

Age at diagnosis:  

As part of diagnosis:  
- Blood tests?  
- Brain scan?  
- Psychological tests?

Where was brain scan performed?

Date of Scan:  

Results:  

Years in Education:  

Education level:
<table>
<thead>
<tr>
<th>Screening Number</th>
<th>Participant Study Number</th>
<th>Group:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>VISIT 1 (SCREENING)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HISTORY OF PRESENTING COMPLAINT</td>
</tr>
</tbody>
</table>

DATE:
<table>
<thead>
<tr>
<th>Appearance and Behaviour:</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Speech:</td>
<td></td>
</tr>
<tr>
<td>Mood/Affect:</td>
<td>Subjective:</td>
</tr>
<tr>
<td>Thoughts:</td>
<td></td>
</tr>
<tr>
<td>Perceptions:</td>
<td></td>
</tr>
<tr>
<td>Cognition:</td>
<td></td>
</tr>
<tr>
<td>Insight:</td>
<td></td>
</tr>
<tr>
<td>From MSE:</td>
<td>□ Delusions □ No delusions</td>
</tr>
</tbody>
</table>
 Has the patient had any relevant medical or psychiatric history? □ No □ Yes, Complete below  

<table>
<thead>
<tr>
<th>Condition / illness /surgical procedure</th>
<th>Start date (DD/MMM/YYYY)</th>
<th>Stop date (DD/MMM/YYYY)</th>
<th>Or tick if ongoing at Screening Visit?</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Has the participant taken any concomitant medications at screening?

<table>
<thead>
<tr>
<th>Medication (Record &lt;specifyGeneric or Brand&gt; name)</th>
<th>Reason for use (Enter Medical History diagnosis or other reason for use, e.g. Prophylaxis)</th>
<th>Start Date (DD/MMM/YYYY)</th>
<th>Stop Date (DD/MMM/YYYY)</th>
<th>Or tick if ongoing at Screening Visit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**VISIT 1 (SCREENING)**

<table>
<thead>
<tr>
<th>Parkinson's Disease</th>
</tr>
</thead>
<tbody>
<tr>
<td>Have you ever been diagnosed with Parkinson's disease?</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Drugs/Alcohol</th>
</tr>
</thead>
<tbody>
<tr>
<td>Are you a smoker?</td>
</tr>
<tr>
<td>How much alcohol do you drink in an average week?</td>
</tr>
<tr>
<td>Have you ever been a heavy drinker?</td>
</tr>
<tr>
<td>Has drinking ever caused you any problems, such as losing jobs or with driving?</td>
</tr>
<tr>
<td>Have you ever used recreational drugs?</td>
</tr>
</tbody>
</table>

**Any other relevant info:**
<table>
<thead>
<tr>
<th>Screening Measures:</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Standardised Mini Mental State Examination (sMMSE)</td>
<td></td>
</tr>
<tr>
<td>2. Geriatric depression scale (GDS)</td>
<td></td>
</tr>
<tr>
<td>3. Neuropsychiatric Inventory (NPI)</td>
<td></td>
</tr>
<tr>
<td>4. Modified version of the Unified Parkinson’s Disease Rating scale (UPDRS) motor component</td>
<td></td>
</tr>
<tr>
<td>5. Bayer Activities of Daily Living (B-ADL)</td>
<td></td>
</tr>
</tbody>
</table>
### VISIT 1 (SCREENING)  
**INCLUSION CRITERIA**

The following criteria MUST be answered YES for participant to be included in the trial (except where NA is appropriate):

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>Yes</th>
<th>No</th>
<th>NA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Age 55 years or older</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.</td>
<td>Meet National Institute of Neurological and Communicative Disorders and Stroke/Alzheimer’s Disease and Related Disorders (NINCDS/ADRDA) criteria for AD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.</td>
<td>Scores ≥22 on sMMSE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.</td>
<td>Scores &lt;6 on the Geriatric depression scale (GDS-15)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.</td>
<td>Scores &lt; 8 on a modified version of the Unified Parkinson’s Disease Rating scale (UPDRS) motor component</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6.</td>
<td>Safe to have MRI</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If any of the above criteria is answered NO, the participant is NOT eligible for the trial and must not be included in the study. Please list reason(s) for ineligibility for screen failure on Participant Eligibility Review page.
The following criteria MUST be answered NO for the participant to be included in the trial:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Current or past psychiatric illness (such as schizophrenia), addiction (drug or alcohol), or neurological disorder (inc. traumatic brain injury or epilepsy)</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2. Medical conditions that might affect a person's ability to tolerate a brain scan, such as significant respiratory or cardiac disease or severe kyphosis, or claustrophobia</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3. Overt parkinsonian (facial masking, action tremor, resting tremor rigidity, bradykinesia) features (scores &gt;7 on the modified UPDRS) or other features suggestive of a Lewy Body Dementia including fluctuating conscious level, frequent falls, or visual hallucinations as a predominant feature</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

If any of the above criteria is answered YES, the participant is NOT eligible for the trial and must not be included in the study. Please list reason(s) for ineligibility for screen failure on Participant Eligibility Review page.
<table>
<thead>
<tr>
<th>Screening Number</th>
<th>Participant Study Number</th>
<th>Group:</th>
</tr>
</thead>
</table>

**VISIT 1 (SCREENING)**

**PARTICIPANT ELIGIBILITY REVIEW**

### End of Screening Visit Checklist:

<table>
<thead>
<tr>
<th></th>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Does the participant satisfy the inclusion and exclusion criteria to date?</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>2. Have all Screening Visit procedures been completed?</td>
<td>☐</td>
<td>☐</td>
</tr>
<tr>
<td>3. Has consent to participate been obtained?</td>
<td>☐</td>
<td>☐</td>
</tr>
</tbody>
</table>

### Participant's eligibility Investigator Sign-Off:

Is the participant eligible to take part in the study?

- ☐ Yes
- ☐ No, Please give reason for screen failure below

**Investigator's Signature:** __________  **Date:** ____/__/____ (DD / MMM / YYYY)

**Investigator's Name:** ________________

**Reason(s) for screen failure:**

1. 

2. 

3. 
Appendix 7. Temporal Context Confusion Task MATLAB Script

Below code written by Professor John King.

```matlab
function sch = schnidertask(subj, run);
    % Schneider task
    
    % This is just checking the parameters make sense
    assert(isnumeric(subj), 'Function must be called with an integer subject number and run(1-4), for example schnidertask(23,2)');
    assert(~mod(subj,1), 'Function must be called with an integer subject number and run(1-4), for example schnidertask(23,2)');
    assert(isnumeric(run), 'Function must be called with an integer subject number and run(1-4), for example schnidertask(23,2)');
    assert(~mod(run,1), 'Function must be called with an integer subject number and run(1-4), for example schnidertask(23,2)');
    assert(((run>0)&&(run<5)), 'Function must be called with an integer subject number and run(1-4), for example schnidertask(23,2)');
    
    % Create a filename for the subject
    filename = sprintf('sch%d02d_data.mat', subj);
    
    % Initialize Cogent.
    warning ('off', 'all');
    mon = 0; % 0 = sub window, 1 = full screen, 2 =
    2nd monitor
    scRes = 6; % 1 = 640x480, 2 = 800x600, 3 =
    1024x768, 4 = 1280x1024, 5 = 1600x1200
    bgCol = [0.4 0.4 0.4]; % Background colour - Grey
    fntCol = [1 1 1]; % Font colour - white
    fntType = 'Arial'; % Font type
    fntSz = 70; % Font size
    
    config_display(mon,scRes,bgCol,fntCol,fntType,fntSz,10);
    config_log;
    config_keyboard;
    config_sound (1, 8, 44100, 100);
    % Specify data file - it contains a list of 260 picture names
    config_data('schniderobj.dat');
    
    rng('default');
    rng(subj); % Initializing rng with subj gives a seed which we can use
to regenerate the same sequences next run.
    
    %replace with rng(12345); if you want every subject to have the same seq
    % (could be any number, needs to never change across subjects)
    
    %If we are on run=1, we will set up the item lists and save them.
    % later runs will load this list instead
    if(run==1)
        % specify the experimental sequence.
        
        % 1. select 52 random pictures from my set of 260
        objects = randperm(260,52);
        
        % 2. select 4 targets for each of the 4 runs (ie use the first 16 items)
```

336
all_targets = objects(1:16);
all_foils = objects(17:52);  % the remainder are going to be foils
run_targets = reshape(all_targets, [4,4]);  % each row of run_targets is going to be a run

runobj=zeros(4,80);  % a blank array to put the run data in 4 rows of 80 items
target_id=zeros(4,80);  % a blank array identify targets in each run
responses=zeros(4,80);  % a blank array to put the responses in
(Y=1)
resp_rt=zeros(4,80);  % a blank array to put the reaction times in

% 3. create the run sequences
for x=1:4
  % in each run, the foils include targets from the other runs - the following line concatenates all foils with those targets not used in the current run
  runfoils = cat(2, all_foils, all_targets([1:((x-1)*4 r*4+1:end)]));

  % now we have our set of foils for this run, we must shuffle them
  runfoils_shuff = runfoils(randperm(size(runfoils,2)));

  % there are 8 blocks of ten items. In each block there are the 4 targets for this run plus 6 randomly chosen foils
  for blk=1:8
    % the following line concatenates the 4 targets with 6 foils (which % it gets by stepping through the foils list in 6's)
    blk_items=cat(2,all_targets(((r-1)*4+1 : (r-1)*4+4) , runfoils_shuff((blk-1)*6+1 : (blk-1)*6+6)));

    % the block is shuffled
    blk_shuff=blk_items(randperm(size(blk_items,2)));

    % the block items are inserted into the big array that holds all the $run orders
    runobj(r,(blk-1)*10+1 : (blk-1)*10+10)=blk_shuff;
  end
end  % run definition loop
% save the subject-specific run orders
save(filename, 'runobj', 'run_targets', 'target_id');
else %if run>1 we load the run definitions
load(filename, '-mat');
end

aborted_at=[0 0]; % a place to store [run, item] in case we break cut using "q"
start_cogent;

% Prepare instruction screen
prearestring('You will see some objects',1,0,100);
preparestring('Say YES when you see a repeat',1);
preparestring('*Try to forget the objects you saw in the previous test*',1,0,-100);
preparepuretone(500,100,1); %make a beep to signal stimulus onset
playsound(1); % Show instructions
drawpic(1);

% Wait until space is pressed
waitkeyup(inf,71);

% Set up inter trial interval and stimulus duration (milliseconds)
i=1000; % Original Schindler paper had 700, but we are inserting the 100ms beep so this is 600 to give the correct ITI
scn=2000;

for pic = 1:80 % main presentation loop

% Find name of picture file
file = getdata(runobj(run, pic), 1 )
filenm=strcat('SVLO\', file);
clearpic( 1 );

% load picture into buffer 1
loadpic( filenm, 1 );
clearkeys % clear the keyboard buffer

drawpic( 1 );% Display buffer1 and wait sdur
playsound(1); % ***********to be resolved - currently the sound duration (100ms) is inserted after ITI so ITI is actually 700ms
prec_time-time;
drawpic( 1 );
wait(sdur);

% Clear screen and wait iti
drawpic( 2 );
wait( iti );
%read the keyboard buffer
readkeys;
[k, t, n] = getkeydown; %extract keyID, time and number of keypresses from keyboard buffer
if(~isempty(k))
   if(k(1)==17) % key 17 = q - this allows us the break out of the run

aborted_at = [run pic];
break
end
if |k(i)==71|
    responses(run,pic)=1; % log a response
    resp_rt(run,pic)=t(i)-pres_time; % log the RT
end
end % main presentation loop
% Save all the data
save(filename, 'runobj','run_targets', 'target_id', 'responses',
      'resp_rt', 'aborted_at');
stop_logent;
Appendix 8. Associative Inference Task MATLAB Script

Below code written by Dr Daniel Bush.

```matlab
function [exp] = Inf_Task(part_no)
% Paired associate and inference task
% Daniel Bush, UCL (2017) drdanielbush@gmail.com
%
% Inputs:
% part_no = participant number (data will be saved with this index)
%
% Assign some parameters
nFix = 1000; % Fixation period before each pair
nStim = 9000; % Length of encoding period (ms) - was 9000
nWait = 1000; % Length of post-stimulus wait period (ms)
retFix = 1000; % Fixation period before each retrieval trial (ms)
cueOnly1 = 1000; % Period that the cue is on the screen before recognition judgement (ms)
recogTime = 10000; % Length of recognition memory judgement (ms)
cueOnly2 = 1000; % Period that the cue is on the screen before associative judgement (ms)
assocTime = 10000; % Length of associative memory judgement (ms)
retWait = 1000; % Length of post-retrieval wait period (ms)
ncs = [-300 300]; % Co-ordinates to display the items on screen during encoding and retrieval (below)
retcs = [-520 -150; 0 -150; 520 -150; -520 -300; 0 -300; 520 -300];
nblocks = 2; % Number of blocks
oldEvents = 12; % Number of 'old' events
newEvents = 6; % Number of 'new' events
oldNewKey = [28 29]; % Cogent key codes for old / new response
assocKeys = [28 29 30 17 23 5]; % Cogent key codes for associative memory response
%
% Initialise necessary Cogent modules and start Cogent
config_display(0, 6, [0.4 0.4 0.4], [1 1 1], 'Arial', 70, 6, 0);
config_keyboard;
start_Cogent;

% Load the experimental stimuli and seed the random number generator
load('exp.mat')
rand('state',sum(clock)); % rng('shuffle')
ext.data = date;
ext.clock = clock;
%
% Randomly assign stimuli to 'events' (i.e. A-B-C triplets)
totalEvents = nblocks*(oldEvents+newEvents);
ext.events = [randperm(totalEvents) randperm(totalEvents) randperm(totalEvents)];
```
% Randomise encoding pair order for each event
pairOrd = nan(nblocks*oldEvents,3);
for trial = 1 : nblocks*oldEvents
    pairOrd(trial,:) = randperm(3);
end

clear trial

% Randomise event order for encoding
exp.encOrd = [];
exp.encPairOrd = [];
for block = 1 : nblocks
    exp.encOrd = [exp.encOrd [randperm(oldEvents)]]; 
    randperm([oldEvents,']'+(block-1)*oldEvents);
    exp.encPairOrd = [exp.encPairOrd 
        pairOrd(exp.encOrd(1:oldEvents,block),1) ;
        pairOrd(exp.encOrd(oldEvents+1:end,block),2)];
end
exp.encOrd(:,2) = exp.encOrd(:,2)+newEvents;
inferrdPairs = [pairOrd(1:oldEvents,3) : nan |newEvents,1) ;
pairOrd(oldEvents+1:end,3) ; nan(newEvents,1)]; clear pairOrd

% Randomise retrieval pair order for each event
pairOrd = nan(totalEvents,3);
for trial = 1 : totalEvents
    pairOrd(trial,:) = randperm(3);
end
clear trial

% Randomise event order for retrieval
exp.retOrd = [];
exp.retPairOrd = [];
for block = 1 : nblocks
    exp.retOrd = [exp.retOrd [randperm(oldEvents+newEvents)]]; 
    randperm([oldEvents+newEvents,']'+(block-1)*(oldEvents+newEvents));
    exp.retPairOrd = [exp.retPairOrd 
        pairOrd(exp.retOrd(1:oldEvents+newEvents,block),1) ;
        pairOrd(exp.retOrd(oldEvents+newEvents+1:2*(oldEvents+newEvents),block 
            ,2) : pairOrd(exp.retOrd(2*(oldEvents+newEvents)+1:end,block),3)];
end
clear block pairOrd

% Assign some memory
exp.encStim = nan(nblocks*oldEvents,3);
exp.encTimes = nan(nblocks*oldEvents,2);
exp.encTrialPos = nan(nblocks*oldEvents*2,2);
exp.retTimes = nan((oldEvents+newEvents)*3*nblocks,5);
exp.retCues = nan((oldEvents+newEvents)*3*nblocks,3);
exp.retTargs = nan((oldEvents+newEvents)*3*nblocks,3);
exp.retFoils = ceil((oldEvents+newEvents)*3*nblocks,3);
exp.retPos = ceil((oldEvents+newEvents)*3*nblocks);
exp.oldNew = nan((oldEvents+newEvents)*3,nblocks);
exp.oldNewPress = zeros((oldEvents+newEvents)*3,nblocks);
exp.oldNewRT = zeros((oldEvents+newEvents)*3,nblocks);
exp.oldNewCorr = zeros((oldEvents+newEvents)*3,nblocks);
exp.assocPress = zeros((oldEvents+newEvents)*3,nblocks);
exp.assocRT = zeros((oldEvents+newEvents)*3,nblocks);
exp.assocCorr = zeros((oldEvents+newEvents)*3,nblocks);
exp.infTrials = nan((oldEvents+newEvents)*3,nblocks);

% Begin the experiment
for block = 1 : nblocks

    % Prepare the start screen
    clearpict(1);
    preparestring([['Learning phase ' int2str(block) ' of ' int2str(nblocks)]], 1, 0, 100);
    preparestring(['Ready to start?', 1, 0, 0]);
    preparestring(['Press any key to begin', 1, 0, -100]);
    drawpict(1);
    waitkeydown(inf);
    clearpict(1);

    % Study phase
    for trial = 1 : oldEvents*2

        % Clear keyboard and prepare stimuli
        clearkeys
        clearpict(1); clearpict(2); clearpict(3);
        pos = randperm(2);
        ind = (block-1)*oldEvents*2+trial;
        stim = exp.events(exp.encOrd(trial,block),:);
        exp.encTrialPos(ind,:) = pos;

        preparestring('+',1);
        if exp.encPairOrd(trial,block) == 1          % Person / Location
            preparestring(exp.stim.person(stim(1)),2,encPos(pos(1)),0);
            preparestring(exp.stim.location(stim(2)),2,encPos(pos(2)),0);
            exp.encStim(ind,:) = [stim(1) stim(2) nan];
        elseif exp.encPairOrd(trial,block) == 2      % Location / Object
            preparestring(exp.stim.location(stim(2)),2,encPos(pos(1)),0);
            preparestring(exp.stim.object(stim(3)),2,encPos(pos(2)),0);
            exp.encStim(ind,:) = [nan stim(2) stim(3)];
        elseif exp.encPairOrd(trial,block) == 3      % Person / Object
            preparestring(exp.stim.person(stim(1)),2,encPos(pos(1)),0);
            preparestring(exp.stim.object(stim(3)),2,encPos(pos(2)),0);
            exp.encStim(ind,:) = [stim(1) nan stim(3)];
        end
        clear pos stim

    end

end
% Present stimuli
exp.encTimes(ind,1) = drawpic(1);
wait(encFix);
exp.encTimes(ind,2) = drawpic(2);
wait(encStim); % waitkeydown(inf);
drawpic(3);
[-,~,stop_n] = waitkeydown(encWait,71);
if stop_n > 0
    clearpic(1);
    preparestring('Task paused - press SPACE to continue',1);
    drawpic(1);
    waitkeydown(inf,71);
end

clear ind stop_n

% Save the data as we go along
save(['Part_' int2str(part_no) '_' Data.mat'], 'exp');
end
clear trial

% Get ready to start the retrieval period
clearpic(1);
preparestring('Learning phase complete', 1, 0, 100);
preparestring('Ready to start retrieval phase?', 1, 0, 0);
preparestring('Press any key to begin', 1, 0, -100);
drawpic(1);
waitkeydown(inf);
clearpic(1);

clearpic(1);

% Retrieval phase
for trial = 1 : (oldEvents+newEvents)*3

    % Identify stimuli
    ind = (block-1)*oldEvents+newEvents*3 + trial;
    stim = exp.events|exp.retOrd(trial,block);
    if exp.retOrd(trial,block) > (block-1)*oldEvents+newEvents
        exp.oldNew(trial,block) = 1;  % New trial
    else
        exp.oldNew(trial,block) = 0;  % Old trial
    end

cued = round(rand);
foils = (block-1)*oldEvents+newEvents+(1:oldEvents);
foils = foils==exp.retOrd(trial,block);
foils = zeros(foils);
foils = foils(randperm(length(foils)));
foils = foils(1:5);
exp.fos(trial,block) = randperm(6);
if exp.retPairOrd(trial,block) == inferredPairs(exp.retOrd(trial,block),1)
    exp.infTrials(trial,block) = 1;
else
    exp.infTrials(trial,block) = 0;
end
end
if exp.retPairOrd(trial,block) == 1
    if exp.oldNew(trial,block) == 1  % New / Cue person
        cue = exp.stim.person(stim(1));
        exp.retCues(ind,1| = stim(1);
    else
        if cued == 0  % Old / Cue person /
            retrieve location
                cue = exp.stim.person(stim(1));
                target = exp.stim.location(stim(2));
                foils = exp.events(foils,2);
                for f = 1 : 5
                    foil(f) = exp.stim.location(foils(f));
                end
                exp.retCues(ind,1| = stim(1);
                exp.retTargs(ind,2) = stim(2);
                exp.retFois(ind,2) = foils;
            else  % Old / Cue location /
                retrieve person
                    cue = exp.stim.location(stim(2));
                    target = exp.stim.person(stim(1));
                    foils = exp.events(foils,1);
                    for f = 1 : 5
                        foil(f) = exp.stim.person(foils(f));
                    end
                    exp.retCues(ind,2) = stim(2);
                    exp.retTargs(ind,1) = stim(1);
                    exp.retFois(ind,1) = foils;
                end
            end
        elseif exp.oldNew(trial,block) == 2  % New / Cue location
            if exp.retPairOrd(trial,block) == 1
                cue = exp.stim.location(stim(2));
                target = exp.stim.object(stim(3));
                foils = exp.events(foils,3);
                for f = 1 : 5
                    foil(f) = exp.stim.object(foils(f));
                end
                exp.retCues(ind,2) = stim(2);
                exp.retTargs(ind,3) = stim(3);
                exp.retFois(ind,3) = foils;
            else  % Old / Cue object /
                retrieve location
                    cue = exp.stim.object(stim(3));
                    target = exp.stim.location(stim(2));
                    foils = exp.events(foils,2);
                    for f = 1 : 5
                        foil(f) = exp.stim.location(foils(f));
                    end
                    exp.retCues(ind,3) = stim(3);
                    exp.retTargs(ind,2) = stim(2);
                    exp.retFois(ind,2) = foils;
                end
            end
        end
    end
end
elseif exp.retPairOrd(trial,block) == 3
if exp.oldNew(trial,block) == 1  % New / Cue object
    cue = exp.stim.object(stim(3));
    exp.retCues(ind,3) = stim(3);
else
    if cued == 0  % Old / Cue object / retrieve person
        cue = exp.stim.object(stim(3));
        target = exp.stim.person(stim(1));
        foils = exp.events(foils,1);
        for f = 1 : 5
            foil(f) = exp.stim.person(foils(f));
        end
        exp.retCues(ind,3) = stim(3);
        exp.retTargs(ind,1) = stim(1);
        exp.retFoils(ind,1) = foils;
    else  % Old / Cue person / retrieve object
        cue = exp.stim.person(stim(1));
        target = exp.stim.object(stim(3));
        foils = exp.events(foils,3);
        for f = 1 : 5
            foil(f) = exp.stim.object(foils(f));
        end
        exp.retCues(ind,1) = stim(1);
        exp.retTargs(ind,3) = stim(3);
        exp.retFoils(ind,3) = foils;
    end
end
clear stim f cued foils

% Prepare stimuli for display
for i = 1 : 5
    preparestring('+',i);
end
preparestring(cue,2,0,0);
preparestring(target,4,0,0);
preparestring('OLD or NEW',3,0,-100);
preparestring(cue,4,0,0);
if exp.cloldNew(trial,block) == 0
    preparestring(target,4,retPos(exp.retPos(trial,block)(i,1)),retPos(exp.retPos(trial,block)(i,2)));
    for f = 1 : 5
        preparestring(foil(f),4,retPos(exp.retPos(trial,block)(f+1,1)),retPos(exp.retPos(trial,block)(f+1,2)));
    end
    clear f
end
clear cue target foil
% Present trial
exp.retTimes(ind,1) = drawpict(1);
wait(retFix);
exp.retTimes(ind,2) = drawpict(2);
wait(cueOnly1);
exp.retTimes(ind,3) = drawpict(3);
[oldnew_kt,oldnew_krt,oldnew_n] = waitkeydown(recogTime); %
waitkeydown(inf);
if exp.oldNew(trial,block) == 0
    exp.retTimes(ind,4) = drawpict(2);
    wait(cueOnly2);
    exp.retTimes(ind,5) = drawpict(4);
    [assoc_kt,assoc_krt,assoc_n] = waitkeydown(assocTime); %
waitkeydown(inf);
else
    drawpict(5);
    [~,~,stop_n] = waitkeydown(retWait,71);
    if stop_n > 0
        clearpict(1);
        preparestring('Task paused - press SPACE to continue',1);
        drawpict(1);
        waitkeydown(inf,71);
    end
    clear stop_n
end

% Store data and compute accuracy
if oldnew_n > 0
    exp.oldNewPress(trial,block) = oldnew_kt(1);
    exp.oldNewRT(trial,block) = oldnew_krt(1) -
    exp.retTimes(ind,3);
    if exp.oldNew(trial,block) == 0 && oldnew_kt(1) ==
    oldNewKey(1)
        exp.oldNewCorr(trial,block) = 1;
    elseif exp.oldNew(trial,block) == 1 && oldnew_kt(1) ==
    oldNewKey(2)
        exp.oldNewCorr(trial,block) = 1;
    end
else
    if exp.oldNew(trial,block) == 0 && assoc_n > 0
        exp.assocPress(trial,block) = assoc_kt(1);
        exp.assocRT(trial,block) = assoc_krt(1) -
        exp.retTimes(ind,5);
        if assoc_kt(1) == assocKeys(exp.retPos(trial,block)(1))
            exp.assocCorr(trial,block) = 1;
        end
    elseif exp.oldNew(trial,block) == 1
        exp.assocPress(trial,block) = nan;
        exp.assocCorr(trial,block) = nan;
        exp.assocRT(trial,block) = nan;
    end
    clear ind oldnew_kt oldnew_krt oldnew_n assoc_kt assoc_krt assoc_n

346
% Save the data as we go along
save(['Part_' int2str(part_no) '_Data.mat'],'exp');
end

end
clear block trial

% Finish the experiment
clearpic(i)
preparestring('End of the experiment', 1, 0, 50);
drawpic(i);
waitkeydown(inf,71);
clearpic(i)
stop_cogent;
Appendix 9. DRM/Metamemory Task MATLAB Script

Below code written by Dr James Bisby.

function exp = DRM_encode(sub, direction, grp)
%% DRM Task encoding phase
% James A. Bisby 2017
% To start the task, type DRM_encode(enter a subject number, enter
group number)
% Data recorded includes the stimuli presented and the onset times for
each
% word shown. This is all recorded within a subject specific .mat
file. The
% test script requires this saved file.

% Create a filename for the subject
filename = sprintf('sub%d_data.mat',sub);

% Read the word lists from the .mat file
load 'ExpData.mat';

% Record which group the subject is in (default = 1 if no number
entered)
if exist('grp','var')
    exp.grp = grp; else exp.grp = 1;
end

% Record the direction you want to present stimuli
% 1 = forwards / 2 = backwards
if exist('direction','var')
    exp.direction = direction; else exp.direction = 1;
end

warning ('off','all');
mcn = 0; % 0 = sub window, 1 = full screen, 2 =
2nd monitor
scrRes = 6; % 1 = 640x480, 2 = 800x600, 3 =
1024x768, 4 = 1280x1024, 5 = 1600x1200
bgCol = [0.4 0.4 0.4]; % Background colour - Grey
fcntCol = [1 1 1]; % Font colour - white
fcntType = 'Arial'; % Font type
fcntSz = 100; % Font size

% Number of lists to be used in the experiment. There are 40 in total.
%  6 Lists - Roediger and McDermott, 1995
% 20 lists - Watson et al., 2002
% 40 lists - Chadwick et al., 2016
exp.nbLists = 15;

% Number of words taken from each category list. There are a total of
15
% words for each list. The lists for each category are in order of how
% likely a word will result in the associated lure being recognised.
% 15 words - Roediger and McDermott, 1995
% 4 words - Chadwick et al., 2016
exp.listLength = 4;
\% Number of words taken from each category to be used at test
exp.testListLength = 2;

\% Time that each word should be presented for.
\% Watson et al., 2002 uses 1500ms
\% Chadwick et al., 2016 uses 500ms
exp.word_t = 3000;

\% Time between each word within a list
\% 250ms - Watson et al., 2002
\% 000ms - Chadwick et al., 2016
exp.trial_iT_iT = 1000;

\% Time between each new list
\% 0000ms - Watson et al., 2002
\% 3000ms - Chadwick et al., 2016
exp.bl_iT_iT = 0000;

\% For presenting words in reverse order
seedNb = exp.listLength+1;

\% Prepare word lists for encoding

\% Seed the random number generator - so random order isn't always the same
\% rng('default');
\%rng(sub);
rand('state',sum(100*clock));

\% Create a random order for the number of lists
randOrd = randperm(length(exp.listItems));

\% Start experiment

\% configure the screen display for cogent
config_display(mon,scRes,bgCol,fontCol,fontType,fontSz,10,16);
config_keyboard;
config_sound (l, 5, 44100, 20);
\% Start cogent
start_cogent;

\% Set up variables to record stimulus shown and onset times
exp.encoding.stim = [];
exp.encoding.stim_t = [];
exp.encoding.lure = [];

\% Prepare instruction screen
preparestring('Say each word out loud',1);

\% Show instructions
drawplot(1);

\% Wait until space is pressed
waitkryup(inf,71);
% Start trials
for bl = 1:exp.nbLists % Start of block

    exp.encoding.lure{end+1,1} = exp.lures{randOrd(bl)};

    for trial = 1:exp.listLength % Trials within block

        % Blank screen
drawpict(2);

        % Prepare stimulus on buffer before presentation
clearpict(1);
        if exp.direction == 1
            exp.encoding.stim{end+1,1} = exp.listItems{randOrd(bl),trial};
            preparestring(exp.listItems{randOrd(bl),trial},1);
        elseif exp.direction == 2
            exp.encoding.stim{end+1,1} = exp.listItems{randOrd(bl),seedNb-trial};
            preparestring(exp.listItems{randOrd(bl),seedNb-trial},1);
        end

        % Wait for iti time
    wait(exp.trial_iti_t);

        % Show stimulus and record the onsets time
    exp.encoding.stim_{end+1,1} = drawpict(1);

        % Wait for stimulus time
    wait(exp.word_t);

        % Save trial info to subject .mat file
    save(filename, 'exp');

        % Listen for q key
        % read the keyboard buffer
    readkeys;
    [k, t, n] = getkeydown; %extract keyID, time and number of
    keypresses from keyboard buffer
    if(~isempty(k))
        if(k(1)==17) % key 17 = q - this allows us the break out
            of the run
            stop_cogent:
                return
            end
        end
    end

    % Blank screen
drawpict(2);
    % Wait for time between each word category
    wait(exp.bl_iti_t);
end

% Stop cogent
stop_cogent;
end
function exp = DRM_test(sub)

% DRM Test
% James A. Eisby 2017
% Memory test for DRM. Requires a subject number (sub) and encoding data
% file

filename = sprintf('sub%02d_data.mat',sub);

warning ('off','all');

mcr = 0; % 0 = sub window, 1 = full screen, 2 = 2nd monitor
scRes = 6; % 1 = 640x480, 2 = 800x600, 3 = 1024x768, 4 = 1280x1024, 5 = 1600x1200
bgCol = [0.4 0.4 0.4]; % Background colour - Grey
fntCol = [1 1 1]; % Font colour - white
fntType = 'Arial'; % Font type
fntSz = 100; % Font size

% Prepare word lists for encoding

lcad(sprintf('sub%02d_data.mat',sub));

exp.test.nbTrials = (exp.nbLists * exp.testListLength)*2 + exp.nbLists; % Total number of trials at test (OLD + NEW + LURES)
exp.test.newTrials = exp.nbLists * exp.testListLength;
exp.test.trial_int_t = 500; % ITI between each word
exp.test.OldRespKeys = [25 29 52]; % Response keys for the task (currently laptop KeyPad 1 2) and Esc key
exp.test.OldConfRespKeys = [25 29 30 52]; % Response keys for the task (currently laptop KeyPad 1 2) and Esc key

rand('state',sum(100*clock)); % set random number generator
rng('default')
rng(sub)

testStim = [];
testCons = [];

% Add the OLD and LURE items to the stimuli list
for i = 1:exp.nbLists
    testStim(end+1,1) = exp.encoding.lure(i,1);
    testCons(end+1,1) = 2;
end

for j = 1:exp.test.newTrials
    testStim(end+1,1) = exp.encoding.stim(j,1);
    testCons(end+1,1) = 1;
end

% Randomise the list of NEW items
a = randperm(length(exp.novelItems)); exp.novelItems = exp.novelItems(a);
% Add the NEW items to the stimull list
for i = 1:exp.test.newTrials; testStim(end+1,1) = exp.novelItems{i,1};
end

exp.test.randOrd = randperm(length(testStim)); % Create
a random order by the number of stimuli
exp.test.stimList = testStim(exp.test.randOrd,:); % Reorder
the stimuli by the random order
exp.test.condList = testCons(exp.test.randOrd,:); % Reorder
the conditions by the random order

% Create empty variables to record test data
exp.test.stim_t = []; % Time stimulus was present
exp.test.resp = []; % Response to the stimulus
exp.test.rt = []; % Reaction time
exp.test.conf_t = []; % Record the response in the results file
exp.test.conf_rt = [];

%% Start experiment

% Configure Cogent
config_display(mon,scRes,bgCol,fontCol,fontType,fontSz,10);
config_keyboard;
start_cogent;

preparestring('Which of these words',1,0,75);
preparestring('did you read aloud?',1,0,-75);
drawpict(1);
waitkeyup(inf,71);
clearpict(1);

%% Start task
for trial = 1:exp.test.nbTrials

drawpict(3); % Draw blank screen
clearpict(1); % Clear buffer 1
clearpict(2); % Clear buffer 2

preparestring(exp.test.stimList{trial,1},1,0,200);
% Place word on buffer 1
preparestring('Did you read this word aloud?',1,0,0);
preparestring('YES NO',1,0,-200);
% Place responses on buffer 1

wait(exp.test.trial_it1_t);
% Wait for iti time
% Old or New question
exp.test.stim_t(end+1,1) = drawpict(1);
% Present buffer 1 and record the onset time
[rsp_k resp_t resp_n] = waitkeyup(inf,exp.test.OldNewrespKeys);
% Wait for a response and record when pressed

if resp_k(1,1) == 52
    break
end

exp.test.resp(end+1,1) = resp_k(1,1); % Record the response in the results file
exp.test.rt(end+1,1) = resp_t(1,1) - exp.test.stim_t(end,1); % Calculate and record the reaction time

clearpict(1);

preparestring('How sure are you?',1,0,200); % Place word on buffer 1
preparestring('Complete Fairly 100\%
',1,0,0); % Place responses on buffer 1
preparestring(' guess confident confident',1,0,-100); % Place responses on buffer 1

exp.test.Conf_t(end+1,1) = drawpict(1);
[rsp_k resp_t resp_n] = waitkeyup(inf,exp.test.OldConfRespKeys);

if resp_k(1,1) == 52
    break
end

exp.test.confResp(end+1,1) = resp_k(1,1); % Record the response in the results file
exp.test.conf_rt(end+1,1) = resp_t(1,1) - exp.test.Conf_t(end,1); % Calculate and record the reaction time

save(filename, 'exp'); % Save trial data
end

stop_cogent;
end
Appendix 10. Analysis of the Association Between ‘State’ Delusion and False Recognition on ADAS-Cog-13 and RAVLT, Excluding Outliers Identified on Box Plots

Appendix 10 Table Analysis of the association between ‘state’ delusion and false recognition on ADAS-Cog-13 and RAVLT, excluding outliers

<table>
<thead>
<tr>
<th>PREDICTOR</th>
<th>EXP(β) (95% CI)</th>
<th>P VALUE</th>
<th>EXP(β) (95% CI)</th>
<th>P VALUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>False Recognition</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ADAS-Cog 13</td>
<td>1.105 (.983, 1.242)</td>
<td>.094</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RAVLT</td>
<td>-</td>
<td>-</td>
<td>1.046 (.900, 1.217)</td>
<td>.557</td>
</tr>
<tr>
<td>Age</td>
<td>1.023 (.979, 1.069)</td>
<td>.318</td>
<td>1.017 (.973, 1.063)</td>
<td>.460</td>
</tr>
<tr>
<td>Gender</td>
<td>.821 (.416, 1.619)</td>
<td>.821</td>
<td>.762 (.383, 1.517)</td>
<td>.439</td>
</tr>
<tr>
<td>Years of Education</td>
<td>1.060 (.937, 1.198)</td>
<td>.354</td>
<td>1.067 (.941, 1.210)</td>
<td>.309</td>
</tr>
<tr>
<td>ChEi</td>
<td>.996 (.431, 2.300)</td>
<td>.992</td>
<td>1.026 (.443, 2.378)</td>
<td>.952</td>
</tr>
<tr>
<td>MMSE</td>
<td>.861 (.754, .983)</td>
<td>.027</td>
<td>.862 (.754, .986)</td>
<td>.031</td>
</tr>
<tr>
<td>Category Fluency</td>
<td>.948 (.882, 1.018)</td>
<td>.143</td>
<td>.937 (.869, 1.009)</td>
<td>.087</td>
</tr>
</tbody>
</table>

Notes:
All measures at baseline.
Model one: ‘State’ delusion with false recognition on ADAS-Cog 13; Model two: ‘State’ delusion with false recognition on RAVLT. Both excluding outliers.
ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; ChEi = cholinesterase inhibitor prescription; MMSE = Mini Mental State Examination; RAVLT = Rey’s Auditory Verbal Learning Test. Category fluency for animals only.
Appendix 11. Edited CAT12 cat_defaults.m File

Changes to default parameters in code written by Professor Christian Gaser.

```matlab
function cat_defaults
% Sets the defaults for CAT
% FORMAT cat_defaults
% This file is intended to be customised for the site.
% Care must be taken when modifying this file
% $Id: cat_defaults.m 1715 2020-09-08 15:51:01Z gaser $

clear global cat;
global cat

% Options for initial SPM12 segmentation that is used as starting point
% for CAT12.
%============================================================================
cat.opts.tpm = fullfile(spm('dir'),'tpm','TPM.nii')):
% SPM12 default = [1 1 2 3 4 2]) - alternative: [3 3 2 3 4 2]
cat.opts.ngaus = [1 1 2 3 4 2]: % Gaussians per class
(SPM12 default = mni) - 'eastern':'mni':
% Affine regularisation
(SPM12 default = mni) - 'rigid':
cat.opts.warpreg = [0 0.001 0.5 0.05 0.2]: % Warping regularisation
(SPM12 default) - no useful modification found
% Preprocessing accuracy (CAT only!) - 1e-2 very low accuracy (fast); 1e-4 default; 1e-6
% very high accuracy (slow)
cat.opts.tol = 1e-4;
% Preprocessing accuracy (CAT only!) - 0 very low accuracy (fast) .. 1 very high accuracy
% (slow): default = 0.5
% Strength of the bias correction that controls the biasreg and biasfwhm parameter
% % 0 - use SPM parameter;
cat.opts.biasstr = 0.5;
cat.opts.biasreg = 0.5;
% Higher values for
% Bias FWHM
% Higher values for stronger bias fields
% Bias FWHM
% Lower values for strong bias fields, but check for overfitting of the thalamus (values <45 mm)
cat.opts.samp = 3;
% Sampling distance - alternative: 1.5
% Initial SPM segmentation resolution, whereas the AMAP runs on the full or specified resolution
```

355
% described by
cat.extopts.resctype and cat.extopts.resval. Higher resolution did not
improve the
% results in most results
(but increase calculation time were.
cat.opts.redspres = 0.0;  % limit image resolution
for internal SPM preprocessing output in mm (default: 1.0)

% Writing options
%==================================================================

% options:
% native 0/1  (none/yes)
% warped 0/1  (none/yes)
% mod    0/1/2/3 (none/affine+nonlinear/nonlinear only/both)
% dartel 0/1/2/3 (none/rigid/affine/both)

% save surface and thickness
cat.output.surface = 1;  % surface and thickness creation: 0 -
no (default), 1 - lh+rh, 2 - lh+rh+cerebellum,
% 3 - lh, 4 - rh, 5 - lh+rh (fast, no
registration, only for quick quality check and not for analysis),
% 6 - lh+rh+cerebellum [fast, no
registration, only for quick quality check and not for analysis]
% 9 - thickness only [for ROI analysis,
experimental!]
% +10 to estimate WM and CSF
width/depth/thickness (experimental!)

% save ROI values
cat.output.ROI = 1;  % write xml-file with ROI data (0 - no, 1
- yes (default))

% bias and noise corrected, global intensity normalized
cat.output.bias.native = 0;
cat.output.bias.warped = 1;
cat.output.bias.dartel = 0;

% bias and noise corrected, (locally - if LAS>0) intensity normalized
cat.output.las.native = 0;
cat.output.las.warped = 0;
cat.output.las.dartel = 0;

% GM tissue maps
cat.output.GM.native = 1; %changed from 0 to 1
cat.output.GM.warped = 0;
cat.output.GM.mod = 1;
cat.output.GM.dartel = 0;

% WM tissue maps
cat.output.WM.native = 0;
cat.output.WM.warped = 0;
cat.output.WM.mod = 1;
cat.output.WM.dartel = 0;

% CSF tissue maps
cat.output.CSF.native = 0;
cat.output.CSF.warped  = 0;
cat.output.CSF.mod    = 0;
cat.output.CSF.dartel = 0;

% WMH tissue maps (only for opt.extopts.WMHC==3) - in development
cat.output.WMH.native = 0;
cat.output.WMH.warped = 0;
cat.output.WMH.mod   = 0;
cat.output.WMH.dartel = 0;

% stroke lesion tissue maps (only for opt.extopts.SLC>0) - in development
cat.output.SL.native = 0;
cat.output.SL.warped = 0;
cat.output.SL.mod   = 0;
cat.output.SL.dartel = 0;

% label
% background=0, CSF=1, GM=2, WM=3, WMH=1 (if opt.extopts.WMHC==3), SL=1.5 (if opt.extopts.SLC>0)
cat.output.label.native = 1;
cat.output.label.warped = 0;
cat.output.label.dartel = 0;

% Tissue classes 4-6 to create own TPMs
cat.output.TPMC.native = 0;
cat.output.TPMC.warped = 0;
cat.output.TPMC.mod   = 0;
cat.output.TPMC.dartel = 0;

% cortical thickness (experimental)
cat.output.ct.native = 0;
cat.output.ct.warped = 0;
cat.output.ct.dartel = 0;

% percentage position (experimental)
cat.output.pp.native = 0;
cat.output.pp.warped = 0;
cat.output.pp.dartel = 0;

% jacobian determinant: 0/1 (none/yes)
cat.output.jacobian.warped = 0;

% deformations
% order is [forward inverse]
cat.output.warps = [0 0];

% transformations
% order is offline rigid (both forward and inverse)
cat.output.xmat = 0;

% Expert options
%----------------------------------------------------------------------------------------
% general GUI compatible definition of most *str parameter:
% 0  - no correction
% eps - ultralight correction
% 0.50 – medium correction
% 0.75 – strong correction
% 1 – heavy correction
% [inf – automatic correction]

% skull-stripping options
cat.extopts.gcutstr = 0.5;  % Strength of skull-stripping:
0 – SPM approach; eps to 1 – gcut; 2 – new AFRG approach; -1 – no skull-
stripping (already skull-stripped); default = 2

cat.extopts.cleanupstr = 0.5;  % Strength of the cleanup process:
0 to 1; default 0.5

% segmentation options
cat.extopts.spm_kmap = 0;  % Replace initial SPM by k-means ANAP
segm. 0 – Unified Segmentation, 2 – k-means ANAP

cat.extopts.NCstr = Inf;  % Strength of the noise correction:
0 to 1; 0 – no filter, -Inf – auto, 1 – full, 2 – ISARMLM (else SNLMM),
default -Inf

cat.extopts.LA5str = 0.5;  % Strength of the local adoption:
0 to 1; default 0.5

cat.extopts.BV2str = 0.5;  % Strength of the Blood Vessel
Correction: 0 to 1; default 0.5

cat.extopts.regstr = 0.5;  % Strength of Shooting registration:
0 – Darts, eps (fast), 0.5 (default) to 1 (accurate) optimized Shooting,
4 – default SPM Shooting

cat.extopts.WMHC = 1;  % Correction of WM hyperintensities:
0 – no correction, 1 – only for Darts/Shooting

2 – also correct segmentation to WM), 3 – handle as separate class;
default 1

cat.extopts.WMHCstr = 0.5;  % Strength of WM hyperintensity
Correction: 0 to 1; default 0.5

cat.extopts.SLC = 0;  % Stroke lesion correction (SLC):
0 – no correction, 1 – handling of manual lesion that have to be set to
zero!

2 – automatic lesion detection (in development)
cat.extopts.mrf = 1;  % MRF weighting:
0 to 1; <1 – weighting, 1 – auto; default 1

% resolution options
cat.extopts.retype = ’optimal’;  % resolution handling:
’native’, ’fixed’, ’best’, ’optimal’
cat.extopts.resval = [1.0 0.1];  % resolution value and its
tolerance range for the ’fixed’ and ’best’ retype

% check for multiple cores is different for octave
if sstmpsi(spm_check_version,’octave’)
   cat.extopts.nproc = nproc; % not changed
else
   cat.extopts.nproc = feature(’numcores’);
end

{%
 Preprocessing with native resolution.

358
In order to avoid interpolation artifacts in the Dartel output the lowest spatial resolution is always limited to the voxel size of the normalized images (default 1.5mm).

Examples:

<table>
<thead>
<tr>
<th>native resolution</th>
<th>internal resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.95 0.95 1.05</td>
<td>&gt; 0.95 0.95 1.05</td>
</tr>
<tr>
<td>0.45 0.45 1.70</td>
<td>&gt; 0.45 0.45 1.70</td>
</tr>
</tbody>
</table>

**Best:**

Preprocessing with the best (minimal) voxel dimension of the native image or at least 1.0 mm.

The first parameter defines the lowest spatial resolution for every dimension, while the second is used to avoid tiny interpolations for almost correct resolutions.

In order to avoid interpolation artifacts in the Dartel output the lowest spatial resolution is always limited to the voxel size of the normalized images (default 1.5mm).

Examples:

<table>
<thead>
<tr>
<th>Parameters</th>
<th>native resolution</th>
<th>internal resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1.00 0.10]</td>
<td>0.95 1.05 1.25</td>
<td>&gt; 0.95 1.00 1.00</td>
</tr>
<tr>
<td>[1.00 0.10]</td>
<td>0.45 0.45 1.50</td>
<td>&gt; 0.45 0.45 1.00</td>
</tr>
<tr>
<td>[0.75 0.10]</td>
<td>0.45 0.45 1.50</td>
<td>&gt; 0.45 0.45 0.75</td>
</tr>
<tr>
<td>[0.75 0.10]</td>
<td>0.45 0.45 0.80</td>
<td>&gt; 0.45 0.45 0.80</td>
</tr>
<tr>
<td>[0.50 0.10]</td>
<td>0.45 0.45 0.50</td>
<td>&gt; 0.45 0.45 0.50</td>
</tr>
<tr>
<td>[0.50 0.30]</td>
<td>0.50 0.50 1.50</td>
<td>&gt; 0.50 0.50 0.50</td>
</tr>
<tr>
<td>[0.50 0.30]</td>
<td>1.50 1.50 3.00</td>
<td>&gt; 1.00 0.00 1.00 % here the internal minimum of 1.0 mm is important.</td>
</tr>
<tr>
<td>[0.00 0.10]</td>
<td>0.45 0.45 1.50</td>
<td>&gt; 0.45 0.45 0.45</td>
</tr>
</tbody>
</table>

**Fixed:**

This option prefers an isotropic voxel size that is controlled by the first parameter.

The second parameter is used to avoid tiny interpolations for almost correct resolutions.

In order to avoid interpolation artifacts in the Dartel output the lowest spatial resolution is always limited to the voxel size of the normalized images (default 1.5mm).

There is no upper limit, but we recommend to avoid unnecessary interpolation.

Examples:

<table>
<thead>
<tr>
<th>Parameters</th>
<th>native resolution</th>
<th>internal resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1.00 0.10]</td>
<td>0.45 0.45 1.70</td>
<td>&gt; 1.00 1.00 1.00</td>
</tr>
<tr>
<td>[1.00 0.10]</td>
<td>0.95 1.05 1.25</td>
<td>&gt; 0.95 1.05 1.00</td>
</tr>
<tr>
<td>[1.00 0.02]</td>
<td>0.95 1.05 1.25</td>
<td>&gt; 1.00 1.00 1.00</td>
</tr>
<tr>
<td>[0.75 0.10]</td>
<td>0.75 0.95 1.25</td>
<td>&gt; 0.75 0.75 0.75</td>
</tr>
</tbody>
</table>

**Optimal:**

This option prefers an isotropic voxel size that is controlled by the median voxel size and a volume term that deals with highly anisotropic voxels.

The first parameter controls the lower resolution limit, while the second parameter is used to avoid tiny interpolations for almost correct resolutions.

Examples:
% registration and normalization options
% Subject species: -
  'human'; 'ape_greater'; 'ape_lesser'; 'monkey_oldworld'; 'monkey_newworld' (in development)
cat.exopts.species = 'human';
% Affine PreProcessing (ADP) with rough bias correction and brain
  extraction for special anatomies (nonhuman/nconates)
cat.exopts.ADP = 2; % 0 - none; 1070 - default; [1 - light; 2 - full; 1114 - update of 1070, 5 - animal (no affreg)] chanced from 1070 to 2
cat.exopts.vox = 1.5; % Voxel size for normalized data
(EXPERIMENTAL: inf - use template values)
cat.exopts.darteltpm =
  {fullfile(spm('dir'),'toolbox','cat12','templates_volumes','Template_1_IX ISSN_MNI152_sh.nii')}; % Indicate first Dartel template (Template 1)
cat.exopts.shootingtpm =
  {fullfile(spm('dir'),'toolbox','cat12','templates_volumes','Template_0_IX ISSN_MNI152_SH_nii')}; % Indicate first Shooting template (Template 0) - not working

cat.exopts.cat12atlas =
  {fullfile(spm('dir'),'toolbox','cat12','templates_volumes','cat.nii')}; % CAT atlas with major regions for VEM, SBM & ROIs
cat.exopts.brainmask =
  {fullfile(spm('Dir'),'toolbox','FieldMap','brainmask_nii')}; % Brainmask for affine registration

cat.exopts.tl =
  {fullfile(spm('Dir'),'toolbox','FieldMap','T1_nii')}; % T1 for affine registration

% surface options

cat.exopts.pbetres = 0.5; % internal resolution for thickness estimation in mm (default 0.5)
cat.exopts.collcorr = 0; % correction of surface collisions (experimental, not yet working properly!): 0 - none; 1 - Deformation; 20 - none; 21 - Deformation; 22 - CAT SI; 23 - PSI; 24 - PSI opt.
cat.exopts.reduce_mean = 1; % optimize surface sampling: 0 - PSI res. (slow); 1 - optimal res. (default); 2 - internal res.; 3 - SPM init; 4 - MATLAB init; 5 - SPM full; 6 - MATLAB full; 7 - MATLAB full ext.
cat.exopts.vdist = 4/3; % mesh resolution (experimental, do not change!)
cat.exopts.pbtlas = 0; % reduce myelination effects (experimental, not yet working properly!)
cat.exopts.thick_measure = 1; % distance method for estimating thickness: 1 - TLs: Freesurfer method using mean(Thcarea, Thcarea2) (default in 12.7+); 0 - Thlk: linked distance (used before 12.7)
cat.exopts.thick_limit = 5; % upper limit for TLs thickness measure similar to Freesurfer (only valid if cat.exopts.thick_measure is set to "i")
cat.exopts.close_parahipp = 0; % optionally apply closing inside mask for parahippocampal gyrus to get rid of the holes that lead to large
% cuts in gyri after topology
correction. However, this may also lead to poorer quality of topology
% correction for other data and should
be only used if large cuts in the parahippocampal areas occur

cat.extopts.scale_cortex = 0.7; % scale intensity values for cortex to
start with initial surface that is closer to GM/WM border to prevent that
gyri/sulci are glued
% if you still have glued gyri/sulci
(mainly in the occ. lobes) you can try to decrease this value (start with
0.6)
% please note that decreasing this
parameter also increases the risk of an interrupted parahippocampal gyrus

% if you still have glued gyri/sulci
parameter also increases the risk of an interrupted parahippocampal gyrus

% increases values in the
parahippocampal area to prevent large cuts in the parahippocampal gyrus
(initial surface in this area
% will be closer to GM/CSF border)
% if the parahippocampal gyrus is still
cut you can try to increase this value (start with 0.15)

% visualisation, print, developing, and debugging options

cat.extopts.colormap = 'BCCWHz'; % {'BCCWHz','BCCWHz'} and matlab
colormaps {'jct','gray','bone',...};
cat.extopts.verb = 2; % verbose output: 1 - default; 2 - details; 3 - write debugging files
cat.extopts.ignoreErrors = 1; % catch errors: 0 - stop with error (default): 1 - catch preprocessing errors and proceed with next
subject (requires MATLAB 2008 or higher);
% 2 - catch preprocessing errors and try backup function if this also fails then
proceed with the next subject (requires MATLAB 2008 or higher)
cat.extopts.expertgui = 0; % control of user GUI: 0 - common
user modus with simple GUI; 1 - expert modus with extended GUI; 2 -
developer modus with full GUI
cat.extopts.subfolders = 1; % use subfolders such as mri, surf,
report, and label to organize your data
cat.extopts.experimental = 0; % experimental functions: 0 - default,
1 - call experimental unsafe functions
cat.extopts.print = 2; % display and print out pdf-file of
results: 0 - off, 1 - volume only [use this to avoid problems on servers
that do not support openGL], 2 - volume and surface (default)
cat.extopts.fontsize = get(0,'defaultuicontrolFontSize'); % default
font size for GUI:
%cat.extopts.fontsize = spm('FontSize',7); % set default font size
for GUI manually; increase value for larger fonts or set it to
% nspmat.mat, send info = 1; % send Matlab and CAT12 version to SPM
server for internal statistics only. If you don't want to send this
% information set this flag to "0". See
% online help CAT12->CAT12 user statistics for more information

cat.extopts.gifti_dat = 1; % save gifti files after resampling
with external dat-file, which increases speed of gifti-processing and
keeps SPM.mat file small
% because the cdata field is not saved
with full data in SPM.mat.
% always use expert mode for standalone installations
if isdeployed, cat.extopts.expertgui = 1; end

% Expert options - ROIs
%============================================================================
% ROI maps from different sources mapped to Dartel CAT-space of IXI-template
% { filename , GUIlevel , tissue , use }  
% filename = ''  - path to the ROI-file 
% GUIlevel = [ 0 | 1 | 2 ]  - available in GUI level 
% tissue = [{'csf', 'gm', 'wm', 'brain', 'none'}]  - tissue classes for volume estimation 
% use = [ 0 | 1 ]  - default setting to use this atlas 

cat.extopts.atlas = { ...

fullfile(spm['dir'], 'toolbox', 'cat12', 'templates_volumes', 'neuromorphometrics.nii') 0  {'csf', 'gm'} 1: ... % atlas based on 35 subjects 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'lpba40.nii') 0  {'gm'} 0: ... % atlas based on 40 subjects 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'coobra.nii') 0  {'gm', 'wm'} 1: ... % hippocampus-amygda-cerebellum, 5 subjects, 0.6 mm voxel size 

fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'hammers.nii') 0  {'csf', 'gm', 'wm'} 0: ... % atlas based on 20 subjects 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'ibsr.nii') 1  {'csf', 'gm'} 0: ... % less regions, 18 subjects, low T1 image quality 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'aal3.nii') 1  {'gm'} 0: ... % many regions, but only labeled on one subject 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'mori.nii') 1  {'gm', 'wm'} 0: ... % only one subject, but with WM regions 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'anatomy.nii') 1  {'gm', 'wm'} 0: ... % ROIs requires further work >> use Anatomy toolbox 

fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'julichbrain.nii') 1  {'gm'} 0: ... %

fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'Schaefer2018_100Parcels_17Networks_order.nii') 1  {'gm'} 0: ... % atlas based on rsfMRI data from 1489 subjects 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'Schaefer2018_200Parcels_17Networks_order.nii') 1  {'gm'} 0: ... % atlas based on rsfMRI data from 1489 subjects 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'Schaefer2018_400Parcels_17Networks_order.nii') 1  {'gm'} 0: ... % atlas based on rsfMRI data from 1489 subjects 
fullfile(spm('dir'), 'toolbox', 'cat12', 'templates_volumes', 'Schaefer2018_600Parcels_17Networks_order.nii') 1  {'gm'} 0: ... % atlas based on rsfMRI data from 1489 subjects 

} 

% { fileid GUIlevel use }  - in development 
cat.extopts.satlas = { ...
Desikan

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.aparc_a2009s
 freesurfer.annot') 0 1;

Destrieux

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.aparc_DKI0.T
reesurfer.annot') 0 1;

HCP

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.aparc_HCP_MM
PI.freesurfer.annot') 0 0;

... Schaefer atlases ...

'Schaefer2016_100P_17N'

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.Schaefer2016
_100Parcels_17Networks_order.annot') 1 0;

'Schaefer2016_200P_17N'

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.Schaefer2016
_200Parcels_17Networks_order.annot') 0 0;

'Schaefer2016_400P_17N'

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.Schaefer2016
_400Parcels_17Networks_order.annot') 1 0;

'Schaefer2016_600P_17N'

fullfile(spm('dir'), 'toolbox', 'cat12', 'atlases_surfaces', 'lh.Schaefer2016
_600Parcels_17Networks_order.annot') 1 0;

};

%------------------------------------------------------------------------
% PRIVATE PARAMETERS (NOT FOR GENERAL USE)
%------------------------------------------------------------------------

% Additional maps
%------------------------------------------------------------------------
% atlas maps (for evaluation)
cat.output.atlas.native = 0;
cat.output.atlas.warped = 0;
cat.output.atlas.dartel = 0;

% IDs of the ROIs in the cat atlas map (cat.nii). Do not change this!
cat.extopts.LAB.NE = 0; % no brain
cat.extopts.LAB.CT = 1; % cortex
cat.extopts.LAB.CB = 3; % Cerebellum
cat.extopts.LAB.BG = 5; % BasalGanglia
cat.extopts.LAB.BV = 7; % Blood Vessels
cat.extopts.LAB.TH = 9; % Hypothalamus
cat.extopts.LAB.ON = 11; % Optical Nerve
cat.extopts.LAB.NE = 13; % MidBrain
cat.extopts.LAB.BS = 13; % BrainStem
cat.extopts.LAB.VT = 15; % Ventricle
cat.extopts.LAB.NV = 17; % no Ventricle
cat.extopts.LAB.HC = 19; % Hippocampus
cat.extopts.LAB.HD = 21; % Head
cat.extopts.LAB.HI = 23; % WM hyperintensities
cat.extopts.LAB.PH = 25; % Gyrus para hippocampalis
cat.extopts.LAB.LE = 27; % lesions
Appendix 12. Shell Script to Run Image Preprocessing

Below code written by Dr Ian Malone.

```
#S -l tmem=4G
#S -l h_vmem=4G
#S -l h_rt=2:0:0
#S -S /bin/bash
#S -N ADNI2TEST3
#S -j y
#S -pe smp 6
#S -R y

hostname
date
MATLAB=/share/apps/matlabR2018b
IMAGE=$(sed -ne"${SGE_TASK_ID}"p /SAN/neuroscience/Mcl_ADNI_Delusions/Scans/Baseline/list.txt)
DEFAULTS=/SAN/neuroscience/Mcl_ADNI_Delusions/Scripts/cat_defaults_ADNI.m
export PATH=$MATLAB/bin:$PATH
/home/emclachl/spm-12/toolbox/cat12/cat_batch_cat.sh --fg .p 6 -d $DEFAULTS $IMAGE
```
Appendix 13. ADNI False Memory Measures Region of Interest Analysis: Model Diagnostics (Pearson Residuals Plots)

Residual plots from Poisson regression modelling of predictors of false recognition on RAVLT, including covariates as described in Chapter 10 and ROIs as listed:

Left hippocampus

![Residual plots for left hippocampus](image1)

Left entorhinal cortex

![Residual plots for left entorhinal cortex](image2)
Left parahippocampal gyrus

Right parahippocampal gyrus
Left superior frontal gyrus

Left superior medial frontal gyrus
Right superior medial frontal gyrus

Residual plots from Poisson regression modelling of predictors of false recognition on ADAS-Cog 13, including covariates as described in Chapter 10 and ROIs as listed:

Left hippocampus
Right hippocampus

Left entorhinal cortex
Right entorhinal cortex

Left parahippocampal gyrus
Right parahippocampal gyrus

Left fusiform gyrus
Right fusiform gyrus

Left middle frontal gyrus
Right middle frontal gyrus

Right superior frontal gyrus
Left inferior frontal gyrus

Left superior medial frontal gyrus
Right superior medial frontal gyrus

Right medial frontal cerebrum
Left anterior cingulate gyrus

Right anterior cingulate gyrus
Appendix 14. ADNI False Memory Measures Region of Interest Analysis: Model Diagnostics (Cook’s Distance Bar Charts)

Cook’s distance plots from Poisson regression modelling of predictors of false recognition on RAVLT, including covariates as described in Chapter 10 and ROIs as listed. Horizontal lines show 4/(n – k – 1) threshold for outlier exclusion.

**Left hippocampus**

![Graph showing Mean Cook’s Distance for Left hippocampus](image)

**Left entorhinal cortex**

![Graph showing Mean Cook’s Distance for Left entorhinal cortex](image)
Left parahippocampal gyrus

Right parahippocampal gyrus
Left superior frontal gyrus

Left superior medial frontal gyrus
Right superior medial frontal gyrus

Cook’s distance plots from Poisson regression modelling of predictors of false recognition on ADAS-Cog 13, including covariates as described in Chapter 10 and ROIs as listed. Horizontal lines show $4/(n - k - 1)$ threshold for outlier exclusion.

Left hippocampus
Right entorhinal cortex

Left parahippocampal gyrus
Right parahippocampal gyrus

![Graph of Right parahippocampal gyrus]

Left fusiform gyrus

![Graph of Left fusiform gyrus]
Right fusiform gyrus

Left middle frontal gyrus
Right middle frontal gyrus

Right superior frontal gyrus
Left inferior frontal gyrus

Left superior medial frontal gyrus
Right superior medial frontal gyrus

Right medial frontal cerebrum
Left anterior cingulate gyrus

Right anterior cingulate gyrus
Appendix 15. Poisson Regression Models for False Recognition on RAVLT and ADAS-Cog 13 in the ADNI Cohort, With ROIs as Predictors: Excluding Outliers, Defined as Cook’s Distance > 4/(n – k – 1)

**Appendix 15 Table** Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest, excluding outliers

<table>
<thead>
<tr>
<th></th>
<th>RAVLT FALSE RECOGNITION MODELS</th>
<th>ADAS-COG 13 FALSE RECOGNITION MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL LIKELIHOOD RATIO (df = 8, All ps &lt; .001)</td>
<td>EXP(β)</td>
</tr>
<tr>
<td>LEFT HIPPOCAMPUS</td>
<td>$X^2 = 38.777$, n = 695</td>
<td>.965</td>
</tr>
<tr>
<td>RIGHT HIPPOCAMPUS</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>LEFT ENTORHINAL CORTEX</td>
<td>$X^2 = 41.341$, n = 694</td>
<td>.961</td>
</tr>
<tr>
<td>RIGHT ENTORHINAL CORTEX</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Left PHG</td>
<td>$X^2 = 38.852$, n = 697</td>
<td>.961</td>
</tr>
<tr>
<td>Right PHG</td>
<td>$X^2 = 39.619$, n = 694</td>
<td>.969</td>
</tr>
<tr>
<td>VENTRAL VISUAL STREAM</td>
<td>Left FFG</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Right FFG</td>
<td>-</td>
</tr>
<tr>
<td>Region</td>
<td>RAVLT False Recognition Models</td>
<td>ADAS-Cog 13 False Recognition Models</td>
</tr>
<tr>
<td>-------------</td>
<td>--------------------------------</td>
<td>-------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Model Likelihood Ratio (df = 8, All ps ≤ .002)</td>
<td>EXP(β)</td>
</tr>
<tr>
<td>Left MFG</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Right MFG</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Left SFG</td>
<td>$X^2 = 45.882, \ n = 696$</td>
<td>.987</td>
</tr>
<tr>
<td>Right SFG</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td>Vl-PFC</td>
<td>Left IFG</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>Left SMFG</td>
<td>$X^2 = 42.488, \ n = 694$</td>
</tr>
<tr>
<td></td>
<td>Right SMFG</td>
<td>$X^2 = 41.573, \ n = 695$</td>
</tr>
<tr>
<td>M-PFC</td>
<td>Right MFC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X^2 = 108.298, \ n = 687$</td>
<td>.949</td>
</tr>
<tr>
<td></td>
<td>LEFT ANTERIOR CINGULATE GYRUS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X^2 = 103.574, \ n = 687$</td>
<td>.974</td>
</tr>
<tr>
<td></td>
<td>RIGHT ANTERIOR CINGULATE GYRUS</td>
<td></td>
</tr>
<tr>
<td></td>
<td>$X^2 = 117.860, \ n = 685$</td>
<td>.966</td>
</tr>
</tbody>
</table>
### Appendix 15 Table Cont. Relationship between false recognition on RAVLT and ADAS-Cog 13 and volume of regions of interest, excluding outliers

Notes:
- Reduction is % reduction in false recognition per unit increase in ROI volume (one unit = 0.01% of TIV).
- ADAS-Cog 13 = Alzheimer’s Disease Assessment Scale–Cognitive Subscale, 13 item; DL-PFC = dorsolateral prefrontal cortex; FFG = fusiform gyrus; FR = false recognition; IFG = inferior frontal gyrus; MFC = medial frontal cerebrum; MFG = middle frontal gyrus; M-PFC = medial prefrontal cortex; PHG = parahippocampal gyrus; RAVLT = Rey’s Auditory Verbal Learning Test; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus; VL-PFC = ventrolateral prefrontal cortex.
- Poisson regression models include covariates: age, gender, years of education, MMSE score, cholinesterase inhibitor prescription, category fluency for animals and MRI field strength.
Appendix 16. Context Memory and Metamemory Study, False Memory Measures Region of Interest Analysis: Model Diagnostics (Pearson Residuals Plots)

Residual plots from Poisson regression modelling of ROI predictors of false recognition on TCC task, including ROIs as listed:

Left entorhinal cortex

Right entorhinal cortex
Left parahippocampal gyrus

Right parahippocampal gyrus
Left fusiform gyrus

Left superior frontal gyrus
Residual plots from Poisson regression modelling of ROI predictors of false recognition on DRM task, including ROIs as listed:

**Left medial frontal cerebrum**

**Right medial frontal cerebrum**
Appendix 17. Context Memory and Metamemory Study False Memory Measures Region of Interest Analysis: Model Diagnostics (Cook’s Distance Bar Charts)

Cook’s distance plots from Poisson regression modelling of ROI predictors of false recognition on TCC task, including ROIs as listed. Horizontal lines show \( 4/(n - k - 1) \) threshold for outlier exclusion.

**Left entorhinal cortex**

![](image1)

**Right entorhinal cortex**

![](image2)
Left parahippocampal gyrus

Right parahippocampal gyrus
Left fusiform gyrus

![Bar graph showing mean Cook's distance for different participant IDs.]

Left superior frontal gyrus

![Bar graph showing mean Cook's distance for different participant IDs.]
Cook's distance plots from Poisson regression modelling of ROI predictors of false recognition on TCC task, including ROIs as listed. Horizontal lines show $4/(n - k - 1)$ threshold for outlier exclusion.

**Left medial frontal cerebrum**

![Left medial frontal cerebrum graph]

**Right medial frontal cerebrum**

![Right medial frontal cerebrum graph]
Appendix 18. Poisson Regression Models for False Recognition on TCC and DRM Tasks in my Patient-based Study, With ROIs as Predictors: Excluding Outliers, Defined as Cook’s Distance $> 4/(n – k – 1)$

**Appendix 18 Table** Relationship between false recognition on the temporal context confusion and DRM/metamemory tasks and volume of regions of interest, excluding outliers

<table>
<thead>
<tr>
<th></th>
<th>TCC FALSE RECOGNITION MODELS</th>
<th></th>
<th></th>
<th></th>
<th>DRM FALSE RECOGNITION MODELS</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL (df = 1)</td>
<td>EXP(β)</td>
<td>95% CI</td>
<td>$P$ VALUE</td>
<td>MODEL (df = 1)</td>
<td>EXP(β)</td>
<td>95% CI</td>
</tr>
<tr>
<td>LEFT ENTORHINAL CORTEX</td>
<td>$X^2 (n = 7) = .493, p = .482$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>RIGHT ENTORHINAL CORTEX</td>
<td>$X^2 (n = 6) = 1.697, p = .193$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>VENTRAL VISUAL STREAM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left PHG</td>
<td>$X^2 (n = 8) = 9.281, p = .002$</td>
<td>.481</td>
<td>.276, .836</td>
<td>.010</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Right PHG</td>
<td>$X^2 (n = 8) = 10.693, p = .001$</td>
<td>.548</td>
<td>.351, .855</td>
<td>.008</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Left FFG</td>
<td>$X^2 (n = 8) = 14.667, p = .000$</td>
<td>.835</td>
<td>.756, .922</td>
<td>.000</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>DL-PFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left SFG</td>
<td>$X^2 (n = 7) = .018, p = .894$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Right SFG</td>
<td>$X^2 (n = 7) = 1.603, p = .205$</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>M-PFC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left MFC</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>$X^2 (n = 6) = 5.859, p = .015$</td>
<td>.249</td>
<td>.056, 1.105</td>
</tr>
</tbody>
</table>

402
**Appendix 18 Table Cont.** Relationship between false recognition on the temporal context confusion and DRM/metamemory tasks and volume of regions of interest, excluding outliers

<table>
<thead>
<tr>
<th></th>
<th>TCC FALSE RECOGNITION MODELS</th>
<th>DRM FALSE RECOGNITION MODELS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MODEL (df = 1)</td>
<td>EXP(β)</td>
</tr>
<tr>
<td>M-PFC</td>
<td>Right MFC</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RIGHT ANTERIOR CINGULATE GYRUS</td>
<td>X² (n = 7) = 1.246, p = .264</td>
<td>-</td>
</tr>
</tbody>
</table>

Notes:
DL-PFC = dorsolateral prefrontal cortex; DRM = Deese-Roediger-McDermott paradigm; FFG = fusiform gyrus; IFG = inferior frontal gyrus; MFC = medial frontal cerebrum; MFG = middle frontal gyrus; M-PFC = medial prefrontal cortex; PHG = parahippocampal gyrus; SFG = superior frontal gyrus; SMFG = superior medial frontal gyrus; TCC = temporal context confusion task; VL-PFC = ventrolateral prefrontal cortex.
Poisson regression models run without covariates.