



Trinity College Dublin

Coláiste na Tríonóide, Baile Átha Cliath

The University of Dublin

**A smartphone-based approach to understanding how
dementia risk factors relate to cognitive impairment**

Anna Marie Rosická

School of Psychology and Trinity College Institute of Neuroscience

Trinity College Dublin, University of Dublin

A thesis submitted to

Trinity College Dublin, University of Dublin

For the degree of

Doctor of Philosophy

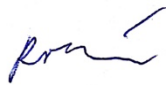
2025

Declaration

I declare that this thesis has not been submitted as an exercise for a degree at this or any other university and it is entirely my own work.

I agree to deposit this thesis in the University's open access institutional repository or allow the Library to do so on my behalf, subject to Irish Copyright Legislation and Trinity College Library conditions of use and acknowledgement.

I consent to the examiner retaining a copy of the thesis beyond the examining period, should they so wish (EU GDPR May 2018).



Signed: Anna Marie Rosická

Date: 30/09/2024

Summary

As populations age and dementia prevalence increases, promoting brain health has become more important than ever before. Non-modifiable factors such as age, sex/gender, or genetics (Riedel et al., 2016) affect the risk of dementia, but they are not fully deterministic. Estimates suggest that almost a half of dementia cases worldwide could be prevented or delayed by reducing a range of potentially modifiable risk factors throughout life (Livingston et al., 2024). Therefore, researchers are increasingly recognising the importance of understanding the factors that shape cognitive functioning pre-dementia, across the lifespan. This could enhance early detection of at-risk individuals and offer insight into the aetiology of dementia. The aims of this thesis were to explore novel, smartphone-based cognitive assessments as a tool for assessing and studying brain health, and to use data collected this way for improving our understanding of how dementia risk factors relate to cognitive impairment.

Chapter 2 of this thesis validated a new gamified task, Memory Match, as a measure of visual working memory that can be self-administered remotely using a smartphone app. This new task was based on an existing tool that holds promise for the early detection of Alzheimer's disease, the Visual Short-Term Memory Binding Task (Parra et al., 2010), but expands on it in several ways, as it aimed to be more engaging, short and easy to administer, and appropriate for use in healthy adults. Using both data from an in-lab validation study and a large dataset crowdsourced from the general public, we found that the new task showed acceptable internal consistency and test-retest reliability despite its reduced length, as well as substantially lower ceiling effects compared to the traditional task, supporting its use in online crowdsourced research in the healthy population. The new task had only moderate convergence with the traditional task ($r = .39$), suggesting partial construct overlap; this could be explained by the reduced ceiling effects and by small changes to the task paradigm. Despite these differences, the new task showed good construct validity as a measure of visual working memory, and we used it as such in the following studies.

Chapter 3 examined if objective and subjective cognitive impairments differ in their sensitivity to previously established dementia risk factors. Early detection of

both objective and subjective cognitive impairment is important. Subjective complaints in healthy individuals can even precede more consolidated, objective deficits. However, their differential associations to modifiable dementia risk factors have not been directly compared. We gathered a large cross-sectional sample ($N = 3,327$, age 18 to 84) via a smartphone app and quantified the associations of 13 previously established dementia risk factors with subjective memory problems and three gamified objective measures of executive function (visual working memory, cognitive flexibility, and model-based planning). We found that all our cognitive measures were sensitive to a range of self-reported health, lifestyle, and socioeconomic factors, but the associations differed by measure, mainly between subjective and objective cognition. Subjective memory problems had the strongest link to most factors and these differences persisted even after controlling for depression symptoms severity, a likely confound. Additionally, age did not moderate most of these associations. Together, these findings suggest that a single-item self-evaluation of cognition could be more sensitive to dementia risk factors than objective cognitive measures, highlighting the value of subjective assessments in the earliest stages of cognitive decline. This study also demonstrated the sensitivity of our smartphone-based measures to many modifiable dementia risk factors even in healthy adults.

Following this, **Chapter 4** aimed to evaluate our smartphone-based cognitive assessments in midlife as potential markers of increased genetic risk of AD. APOE- $\epsilon 4$ is the strongest known genetic risk factor for dementia, but the literature concerning its effects on cognition earlier in the lifespan has been inconclusive, with evidence showing that any effects are likely to be domain-specific (O'Donoghue et al., 2018). Here, we invited over 600 participants of an ongoing prospective cohort study to self-administer smartphone assessments, including measures of processing speed, cognitive flexibility, model-based planning, visual working memory, and subjective memory problems. In a final sample of $N = 87$ healthy middle-aged adults, we tested the associations between the effects of APOE status on cognition. We found that APOE- $\epsilon 4$ carriers outperformed non-carriers on visual working memory, yet this effect did not survive controlling for multiple comparisons. This trend could be due to antagonistic pleiotropy effects but needs to be replicated in larger samples. An AUC-ROC analysis tested if

APOE status could be predicted from norm-adjusted cognitive scores, i.e. adjusting observed visual working memory performance for what we would predict for given demographics based on an additional benchmark sample ($N = 3,327$). We found a modest effect ($AUC = .648$), implying low clinical utility of the task for detecting APOE status. APOE- $\epsilon 4$ was not linked to any other cognitive measures or to demographic characteristics, suggesting that if our smartphone measures are related to genetic risk in middle-aged cognitively unimpaired adults, the effects are smaller than we were powered to detect.

Finally, **Chapter 5** employed a machine-learning approach to investigate if the multivariate associations between modifiable dementia risk factors and cognition differ by sex. Previous studies reported considerable sex differences in the prevalence of some of the modifiable dementia risk factors, but it remains unclear if these factors confer similar magnitudes of risk for cognitive impairment in men and women. Using a large dataset ($N = 4,995$) collected through a smartphone app, we trained regularised regression models to predict cognition from dementia risk factors. We then evaluated these models in a held-out subset of the data. Dementia risk factors contributed significantly to cognition, explaining 14.1% of variance in objective cognition and predicting subjective memory problems with $AUC = .699$ in the hold-out set. Although we found sex differences in risk factor prevalence in our sample, the associations of risk factors with cognition were not moderated by sex. These findings give little support that risk factor interventions need to be tailored to sex, beyond accounting for the differences in prevalence of these risk factors.

In sum, this thesis demonstrates the potential of smartphone-based assessments for scaling up cognitive assessments and brain health research. Harnessing the power of citizen science and robust statistical methodologies, the findings bring novel insights into the associations of dementia risk factors and cognition across the lifespan.

List of Publications and Presentations

This thesis incorporates material already published or currently in press in the following manuscripts:

Rosická, A. M., Teckentrup, V., Fittipaldi, S., Ibanez, A., Pringle, A., Gallagher, E., ... & Gillan, C. M. (2024). Modifiable dementia risk factors associated with objective and subjective cognition. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*. (in press)

Whelan, R., Barbey, F. M., Cominetti, M. R., Gillan, C. M., & Rosická, A. M. (2022). Developments in scalable strategies for detecting early markers of cognitive decline. *Translational Psychiatry*, 12(1), 473.
<https://doi.org/10.1038/s41398-022-02237-w>

The following are presentations arising from this work:

Rosická, A. M., Fittipaldi, S., Gallagher, E., Hanlon, A. K., Ibáñez, A., Bridgeman, K., ... & Gillan, C. M. APOE-ε4 associations with smartphone-based cognitive markers. Poster presentation at Alzheimer's Association International Conference 2024, Philadelphia, Pennsylvania, USA.

Rosická, A. M., Teckentrup, V., Fittipaldi, S., Ibanez, A., Pringle, A., Gallagher, E., ... & Gillan, C.M. Modifiable dementia risk factors associated with objective and subjective cognition. Poster and oral presentation at the Neuroscience Research Day 2024, Dublin, Ireland.

Rosická, A. M., Teckentrup, V., Pringle, A., Gallagher, E., Hanlon, A. K., Lawlor, B., ... & Gillan, C.M. Modifiable dementia risk factors associated with distinct subjective and objective cognitive outcomes. Poster presentation at Alzheimer's Association International Conference 2023, Amsterdam, Netherlands.

Rosická, A. M., Teckentrup, V., Pringle, A., Gallagher, E., Hanlon, A. K., Lawlor, B., ... & Gillan, C.M. Modifiable risk factors associated with distinct cognitive outcomes: a smartphone-based approach. Poster presentation at the School of Psychology Research Day 2023, Dublin, Ireland.

Rosická, A. M., Teckentrup, V., Donegan, K., Pringle, A., Gallagher, E., Hanlon, A. K., & Gillan, C.M. Neureka: A smartphone-based approach to brain health assessment. Poster and oral presentation at the PREVENT General Assembly 2022, Cambridge, UK.

Rosická, A. M., Pringle, A., Gallagher, E., Hanlon, A. K., & Gillan, C.M.: Validating a gamified smartphone measure of visual short-term memory binding as a tool to detect early cognitive decline. Poster presentation at the British Association for Psychopharmacology Summer Meeting 2022, London, UK.

Authorship Contribution Statement

Chapter 2:

Anna M Rosická: Literature review, data preprocessing, data analyses, investigation, writing – original draft, review & editing, final draft. **Andrew Pringle:** Study conception & design, material preparation – task design, participant recruitment, data collection, initial data preprocessing. **Claire M Gillan:** Funding, study conception & design, review & editing, supervision. **Anna K Hanlon:** Participant recruitment, data collection, project administration. **Eoghan Gallagher:** Material preparation – app and task design, software development. **Nathalie Claus:** Participant recruitment, data collection. **Vanessa Teckentrup:** Data analyses (Star Racer validation).

Chapter 3:

Anna M Rosická: Literature review, study conception & design, data preprocessing, data analyses, investigation, writing – original draft, review & editing, final draft. **Claire M Gillan:** Funding, study conception & design, review & editing, supervision. **Vanessa Teckentrup:** Data analyses – supplement (Star Racer validation), review & editing. **Andrew Pringle:** Material preparation – task design. **Agustin Ibanez:** Study conception and design, review & editing. **Anna K Hanlon:** Participant recruitment, data collection, project administration. **Eoghan Gallagher:** Material preparation – app and task design, software development. **Sol Fittipaldi, Cathal McCrory, Brian Lawlor, Lorina Naci:** Review & editing.

Chapter 4:

Anna M Rosická: Literature review, study conception & design, project administration (data sharing agreement), material preparation – recruitment materials, data preprocessing, data analyses, investigation, writing – original draft, review & editing, final draft. **Claire M Gillan:** Funding (Neureka), study conception & design, review & editing, supervision. **Andrew Pringle:** Material preparation – task design (Neureka). **Anna K Hanlon:** Project administration (Neureka, data sharing agreement). **Katie Bridgeman:** Participant recruitment, data collection, project administration (PREVENT Dementia study, data sharing agreement). **Hannah Pullen, Amy Heneghan, Marianna Pope, Diana Benitez Jimenez,**

Sandra El-Sherbiny, Rebecca Smith, Hannah Darwin: Participant recruitment.
Karen Ritchie, Craig W. Ritchie, John O'Brien, Ivan Koychev: Supervision (PREVENT Dementia study). **Eoghan Gallagher:** Material preparation – app and task design, software development (Neureka). **Sol Fittipaldi, Graciela Muniz-Terrera, Brian Lawlor, Lorina Naci, Agustin Ibanez:** Study conception & design.

Chapter 5:

Anna M Rosická: Literature review, study conception & design, data preprocessing, data analyses, investigation, writing – original draft, review & editing, final draft. **Claire M Gillan:** Funding, study conception & design, review & editing, supervision. **Chi Tak Lee:** Data analyses. **Andrew Pringle:** Material preparation – task design. **Anna K Hanlon:** Participant recruitment, data collection, project administration. **Eoghan Gallagher:** Material preparation – app and task design, software development. **Agustin Ibanez, Brian Lawlor:** Study conception and design.

Acknowledgments

I would like to thank the many wonderful people who have supported me along my PhD journey. Without you all, this work would not have been possible.

First, I would like to thank my supervisor, Professor Claire Gillan. Claire, you have been the best PhD supervisor I could have wished for. I am truly grateful for your commitment to making the lab a fair and supportive environment and for providing us all with the best possible working conditions. Your passion for getting to the root of the problem, your superhuman work efficiency, and your way with words have always been an inspiration to me. Thank you for all your advice, guidance, and support, for your time and thorough feedback, and for pulling me out of numerous rabbit holes. Thank you for letting me do a PhD with you; it has been a pleasure.

I am also thankful to all Gillan Lab members, past and present, as well as to our neighbours from the Meta-LAB and the GBHI, for bringing joy to my three years in Dublin. Special thanks to Vanessa Teckentrup, Alice Cox, Anna Hanlon, Celine Fox, Eike Buabang, Kelly Donegan, Judith Mildner, Parnian Rafei, Sarah Moran, Sharon Chi Tak Lee, Sol Fittipaldi, and Steph Suddell. Thank you to our collaborators, in particular Agustín Ibañez, Brian Lawlor, Lorina Naci, and the PREVENT study team. Thanks to everyone involved in building *Neureka*, especially Eoghan Gallagher and Andy Pringle. Thanks to the Science Foundation Ireland for funding my PhD.

Thank you to all participants who kindly gave their time to brain health research.

A huge thank you goes to the Tramyard Climbing Squad for defying Gravity together and making Dublin worth returning to.

Thank you, Vít'o, for keeping me sane, cheering on me, making random jokes, listening to me, coming to rainy Dublin 18(!) times, and just existing. We made it!

Thank you, Mum and Dad, my dear brother, all of my family, for your support and care during my never-ending academic adventures. This has been the last one, I think. And anyway, I have learnt the most important things from all of you. /

Děkuju, mami a tati, bráško, celá rodino, za vaši podporu a péči během mých nekonečných akademických dobrodružství. Tohle bylo poslední, tedy aspoň myslím. A stejně jsem se to nejdůležitější naučila od vás všech.

Table of Contents

DECLARATION	1
SUMMARY	2
LIST OF PUBLICATIONS AND PRESENTATIONS	5
AUTHORSHIP CONTRIBUTION STATEMENT	7
ACKNOWLEDGMENTS	9
CHAPTER 1: GENERAL INTRODUCTION	13
1.1 EARLY DETECTION OF BRAIN HEALTH DECLINE	14
1.2 OBJECTIVE VS. SUBJECTIVE COGNITIVE DECLINE	16
1.3 RISK FACTORS FOR COGNITIVE DECLINE	19
1.3.1 MODIFIABLE RISK FACTORS FOR COGNITIVE DECLINE	20
1.3.2 THE LIFE-COURSE MODEL OF DEMENTIA RISK FACTORS	22
1.3.3 GENETIC RISK FACTORS FOR COGNITIVE DECLINE	23
1.3.4 APOE-E4 AND COGNITION ACROSS THE LIFESPAN	25
1.4 SEX/GENDER DIFFERENCES IN COGNITIVE IMPAIRMENT AND RISK	28
1.5 MULTIVARIATE APPROACHES TO STUDYING RISK FACTORS	31
1.6 COGNITIVE ASSESSMENTS USING SMARTPHONES	35
1.6.1 ADVANTAGES AND LIMITATIONS OF SMARTPHONE-BASED COGNITIVE ASSESSMENTS	35
1.6.2 A CROWDSOURCING APPROACH TO BRAIN HEALTH RESEARCH	39
1.7 STUDIES AND HYPOTHESES	42
CHAPTER 2: VALIDATING A GAMIFIED SMARTPHONE MEASURE OF VISUAL WORKING MEMORY AS A TOOL TO DETECT EARLY COGNITIVE DECLINE	45
2.1 INTRODUCTION	45
2.2 METHODS	47
2.2.1 PARTICIPANTS	47
2.2.2 PROCEDURE	48
2.2.3 MATERIALS	49
2.2.4 ANALYSES	54
2.3 RESULTS	56
2.4 DISCUSSION	58
CHAPTER 3: MODIFIABLE DEMENTIA RISK FACTORS ASSOCIATED WITH OBJECTIVE AND SUBJECTIVE COGNITION	62
3.1 INTRODUCTION	62
3.2 METHODS	64

3.2.1 PARTICIPANTS	64
3.2.2 PROCEDURE	68
3.2.3 MEASURES	69
3.2.4 DATA ANALYSIS	74
3.3 RESULTS	76
3.3.1 DESCRIPTIVE AND DEMOGRAPHIC ANALYSES	76
3.3.2 RISK FACTORS AND COGNITION	78
3.3.3 COMPARISON OF EFFECT MAGNITUDE FOR SUBJECTIVE VS. OBJECTIVE COGNITION	79
3.3.4 DEPRESSION-ADJUSTED ANALYSES	84
3.3.5 INTERACTIONS BETWEEN AGE AND RISK FACTORS	86
3.4 DISCUSSION	86

CHAPTER 4: APOE-E4 ASSOCIATIONS WITH SMARTPHONE-BASED COGNITIVE MARKERS **93**

4.1 INTRODUCTION	93
4.2 METHODS	98
4.2.1 PARTICIPANTS AND PROCEDURE	98
4.2.2 MEASURES	99
4.2.3 DATA ANALYSIS	101
4.3 RESULTS	103
4.3.1 SAMPLE DESCRIPTION	103
4.3.2 ASSOCIATIONS BETWEEN APOE-E4 STATUS AND COGNITION	104
4.3.3 CLASSIFYING GENETIC GROUPS BASED ON VISUAL WORKING MEMORY	108
4.4 DISCUSSION	109

CHAPTER 5: SEX DIFFERENCES IN MODIFIABLE RISK FACTORS ASSOCIATED WITH COGNITION: A MACHINE-LEARNING APPROACH **114**

5.1 INTRODUCTION	114
5.2 METHODS	117
5.2.1 PARTICIPANTS AND PROCEDURE	117
5.2.2 MEASURES	118
5.2.3 STATISTICAL ANALYSES	119
5.3 RESULTS	123
5.3.1 DESCRIPTIVE SEX DIFFERENCES	123
5.3.2 MODEL PERFORMANCE	126
5.3.3 PREDICTOR IMPORTANCE	128
5.3.4 SENSITIVITY ANALYSES	131
5.4 DISCUSSION	131

CHAPTER 6: GENERAL DISCUSSION **138**

6.1 SUMMARY	138
6.1.1 UTILITY OF SMARTPHONE-BASED COGNITIVE ASSESSMENTS	139
6.1.2 DEMENTIA RISK FACTORS AND COGNITION	141

6.2 IMPLICATIONS	146
6.2.1 THE VALUE OF SUBJECTIVE COGNITIVE ASSESSMENTS	146
6.2.2 GENETIC RISK: HOW SPECIFIC ARE 'DOMAIN-SPECIFIC' EFFECTS?	148
6.2.3 BRAIN HEALTH RESEARCH CAN INFORM TARGETED PREVENTION EFFORTS	149
6.2.4 PRACTICAL AND ETHICAL CONSIDERATIONS FOR SMARTPHONE ASSESSMENTS	152
6.3 LIMITATIONS AND FUTURE DIRECTIONS	155
6.4 CONCLUSION	159
 7 REFERENCES	 160
 8 APPENDICES	 208
 8.1 APPENDIX I: SUPPLEMENTAL INFORMATION FOR CHAPTER 3	 208
8.1.1 AGE DISTRIBUTION IN THE FINAL SAMPLE	208
8.1.2 ANALYSIS OF SAMPLE REPRESENTATIVENESS WITH REGARDS TO SES	208
8.1.3 SENSITIVITY ANALYSIS EXCLUDING NON-NATIVE SPEAKERS	211
8.1.4 NEUREKA RECRUITMENT STRATEGIES	212
8.1.5 VALIDATION OF GAMIFIED COGNITIVE MEASURES	218
8.1.6 ASSOCIATIONS BETWEEN RISK FACTOR MEASURES	227
8.1.7 RESULTS FOR NON-CISGENDER PARTICIPANTS	227
8.1.8 COMPARISON OF EFFECT MAGNITUDE FOR SUBJECTIVE VS. OBJECTIVE COGNITION	229
8.1.9 SMOKING HISTORY X AGE INTERACTION	231
8.1.10 INTERACTIONS BETWEEN AGE AND RISK FACTORS	231
8.2 APPENDIX II: SUPPLEMENTAL INFORMATION FOR CHAPTER 4	233
8.2.1 RECRUITMENT PROCESS	233
8.2.2 ASSOCIATIONS BETWEEN VARIABLES	242
8.2.3 SENSITIVITY ANALYSES	243
8.3 APPENDIX III: SUPPLEMENTAL INFORMATION FOR CHAPTER 5	245
8.3.1 DATA DISTRIBUTION BY SEX	245
8.3.2 ASSOCIATIONS BETWEEN PREDICTORS	246
8.3.3 SENSITIVITY ANALYSES IN A SEX-BALANCED SUBSET OF THE DATA	247

Chapter 1: General Introduction

“It is not enough for a great nation merely to have added new years to life — our objective must also be to add new life to those years.”

(J. F. Kennedy, 1963, p. 189)

The world population is ageing. In the past few decades, medicine and societies have made great advances to enable individuals to achieve longevity, globally increasing life expectancy by 10 to 20 years since the 1950s (World Health Organization, 2015). Increased expectancy together with population growth have unfortunately also been linked to an increase in the number of people suffering with dementia. Dementia is a debilitating syndrome that involves not only cognitive decline, but also changes to behaviour, mood, and personality. In 2019, there were around 57 million estimated cases of dementia worldwide and this number is expected to increase almost threefold by 2050 as populations get older (E. Nichols et al., 2022). This increase is particularly notable in lower- and middle-income countries (E. Nichols et al., 2022), which have seen faster increases in life expectancy and greater population growth in recent years. Adding years to life is not enough — we must also consider ways to add life to years.

Having a healthy brain is a crucial determinant of well-being in older age, but to facilitate that, we need to start much earlier in life. The term ‘brain health’ has recently received increasing emphasis, although its exact definition is still subject to debate (Gorelick & Sorond, 2023). Shared across definitions is the understanding that brain health encompasses both optimal brain structure and its optimal function, occurs across the entire lifespan, and is shaped by environmental, biological, psychological, and social determinants throughout life, many of which can be modified (Y. Chen et al., 2021; Gorelick & Sorond, 2023). Importantly, brain health is understood as more than just a lack of brain disease, just as overall health is more than simply a lack of any disease, and it is different from being ‘average’ or ‘normal’ because what is average or normal varies across populations (Y. Chen et al., 2021). Some authors have proposed to refer to brain health using the term ‘eumentia’ (literal translation: ‘good mindedness’) as the

opposite of dementia to better reflect this new perspective (Hachinski & Avan, 2022). Researchers and policy-makers are increasingly recognising that to improve quality of life, we need to focus on monitoring and promoting brain health across the lifespan rather than just reducing disease in older adulthood (Y. Chen et al., 2021; Hachinski & Avan, 2022; Lisko et al., 2021; Sabayan et al., 2023).

This thesis focuses on cognition, which is a key brain function and one of the core components of brain health (Gorelick & Sorond, 2023). Given its importance, cognition is increasingly becoming the target for early detection of individuals at an increased risk of brain health decline. As research and clinical practice alike strive to increase the precision of cognitive assessments, understanding the individual differences that contribute to cognitive performance is key. Beside implications for early detection, studying how modifiable and non-modifiable factors relate to cognition across the lifespan can help increase our understanding of the mechanisms that shape brain health and potentially develop more targeted preventative interventions.

1.1 Early Detection of Brain Health Decline

On a structural level, human brain development starts in the third gestational week and continues at least throughout the first few decades of life, possibly even longer (Stiles & Jernigan, 2010). This is followed by a period of plateau and then more progressive brain atrophy (i.e., volume reduction) around the sixth decade of life (Grydeland et al., 2019; Vinke et al., 2018). The developmental and ageing trajectories of different brain regions and measures peak at different points in life, as recently mapped in detail by Bethlehem et al. (2022). On a functional level, normal physical brain development and atrophy are paralleled by the development and decline of cognitive functions across the lifespan, many of which reach their peak level only later in life (Tucker-Drob, 2019; Turrini et al., 2023).

Besides brain atrophy, a variety of different structural neuropathologies often start to appear from mid to late adulthood. Accumulation of these neuropathologies is understood to cause dementia, but not everyone with some neuropathology will develop dementia (E. Nichols et al., 2023). This variability is attributed to individual

differences in the brain's ability to withstand neuropathology known as 'cognitive reserve' or 'cognitive resilience' (Stern et al., 2020). The most common neuropathological condition underlying dementia is Alzheimer's disease (AD; Beach & Malek-Ahmadi, 2021), which is marked by amyloid β protein accumulation outside neurons and abnormal tau protein inside neurons (Jack Jr. et al., 2024). Recently, Gustavsson et al. (2023) estimated that there might be as many as 416 million people on the AD neuropathology continuum globally, which means a staggering 22% of the 1.9 billion people aged 50 and older across the world have AD-related neuropathology. Besides AD, other neuropathologies are also common, including cerebrovascular disease (i.e., cerebral blood vessel damage and/or associated brain tissue damage), frontotemporal lobar degeneration (FTLD; i.e., neuronal damage to the frontal and temporal lobes marked by abnormal tau or TDP-43 protein), or Lewy body disease (i.e., accumulation of the alpha-synuclein protein in the cortex), and multimorbidity is highly prevalent in both mid and late adulthood (Beach & Malek-Ahmadi, 2021; E. Nichols et al., 2023).

Early detection of brain health decline and associated risk factors is crucial. One reason is that neuropathological conditions like AD have a long asymptomatic stage, whereby the underlying neuropathology such as amyloid β deposition is known to appear up to decades before the onset of dementia symptoms (Braak et al., 2011; Villemagne et al., 2013). Furthermore, the susceptibility to neuropathologies is influenced by a wide range of factors exerting their effects across the entire lifespan, with both genes (Corder et al., 1993) and modifiable risk factors (Livingston et al., 2024) playing a role. Finally, there is currently lack of satisfactory treatment options for dementia and the underlying neuropathologies. For instance, the recently developed anti-amyloid therapies for AD have limited effect, are costly, not yet globally available, and not suitable for everyone (Cummings et al., 2023; Self & Holtzman, 2023). The hope is to detect at-risk individuals at an earlier stage to be able to intervene before neuropathologies fully develop.

Cognitive assessments are a promising tool for early detection. Subtle objective cognitive changes and/or subjective cognitive decline without marked impact on daily functioning can appear early in the presence of neuropathology, sometimes

preceding full-blown dementia by up to a decade (Jack Jr. et al., 2024; Mortamais et al., 2017). Cognitive assessments can thus serve as proxy measures of healthy brain function, assisting early detection in the clinic and in research alike. This is even more important in randomised clinical trials, where cognition can serve as a more immediate outcome as compared to dementia onset or pronounced neuropathology (Y. Y. Lim et al., 2021; Ngandu et al., 2015).

1.2 Objective vs. Subjective Cognitive Decline

Cognitive impairments can be both objectively measured and subjectively perceived, and both types of insight are important. Cognitive assessments can reflect a wide range of brain health domains and as such are crucial for the differential diagnostics of different pathological processes. However, cognitive performance is also subject to ageing even in healthy adults (Tucker-Drob, 2019). To discern pathological cognitive decline from normal ageing-related cognitive changes, the selection of cognitive assessment methods needs to be informed by the patterns of decline typical to potential neuropathologies. While some aspects of cognitive impairments are shared across a variety of conditions including neurological disease (Markopoulou, 2010) or depression (Rock et al., 2014), impairments to some cognitive domains are more condition-specific. A trade-off between sensitivity and specificity is therefore a key consideration.

Let us take the example of AD-related cognitive decline. In AD, the first signs of neurodegeneration usually appear in subregions of the medial temporal lobe (Braak & Braak, 1991), and these changes seem to be mirrored by very early impairments to visuospatial processing and memory, i.e., a group of cognitive abilities needed to process visual and spatial relationships among objects (K. Ritchie et al., 2017). Specific impairments included visual working memory binding, which is the ability to temporarily remember combinations of multiple features such as objects and their colours (Parra et al., 2009) or objects and their locations (Liang et al., 2016). Visuospatial impairments were also found in paired associate recognition, measured as the ability to recognise previously presented pairs of objects from new pairings (Lowndes & Savage, 2007), and in wayfinding,

measured as the distance travelled to navigate to previously learnt landmarks in a virtual environment (Allison et al., 2016). Seemingly closer to dementia diagnosis, visuospatial deficits are followed by impairments to episodic memory (i.e., the memory for events), semantic memory (i.e., the memory for facts), and executive functioning (i.e., higher-level processes involved in behaviour control such as inhibition or cognitive flexibility) (Josefsson et al., 2019; Mortamais et al., 2017). Notably, these early cognitive changes can be subtle and may only progress over time into Mild Cognitive Impairment (MCI). MCI is a diagnostic category that describes a later stage of notable cognitive decline, both objective and subjective, where an individual may start having difficulties in activities of daily living but remains functionally independent (Bradfield, 2023; Jack Jr. et al., 2024). Impairments to episodic memory are understood as the hallmark of AD neuropathology (M. Gallagher & Koh, 2011). As such, they are commonly assessed in research and clinical practice through objective, standardised tests. However, episodic memory impairment also appears in normal ageing (Tromp et al., 2015) and conditions like FTLN (Pennington et al., 2011). As a result, some authors argue for prioritising other cognitive markers for AD to attain higher specificity, such as impairments to spatial navigation and orientation (Coughlan et al., 2018) or working memory binding deficits (Cecchini et al., 2023). For example, egocentric spatial navigation, measured as performance on a computer task that involves moving through a virtual supermarket and recalling one's starting position, differentiated people with AD from those with FTLN (Tu et al., 2015). Similarly, impaired working memory binding, measured as the ability to temporarily remember the conjunctions of multiple features such as shapes and their colours, differentiated people with AD from those with dementias of other type (Della Sala et al., 2012). Although more research is needed to assess the sensitivity of these cognitive markers in the earliest stages of AD (Cecchini et al., 2023; Coughlan et al., 2018), they are promising given their high specificity to AD neuropathology.

One issue with early cognitive decline is that often it is not detectable with standard objective assessments that are conducted at a single time-point and compared to a norm. In many cases, even though cognitive decline has occurred, a person may still fall within the normal range because they may have had a high pre-morbid level of functioning (Jessen, 2014). For this reason, assessing

cognition against one's own baseline is thought to be crucial. This has been recognised for example by the latest AD clinical staging criteria from the American Alzheimer's Association (Jack Jr. et al., 2024), intended for research and diagnosis of individuals with AD neuropathology, which place early cognitive changes as a transitional stage between the asymptomatic stage and the MCI stage. Comparisons against an individual's baseline can be achieved through repeated testing. However, thorough cognitive assessments are not routinely done in the clinic unless there is already evidence of cognitive decline, and so information on baseline cognitive functioning is usually limited or unavailable. Furthermore, the time and cost associated with repeated testing creates a barrier that prevents many individuals from accessing these services. In such cases, subjective cognitive assessments can provide much needed additional insights.

Subjective cognitive impairment can be assessed in a variety of ways, including longer standardised questionnaires and simple screening questions in various formats (Burmester et al., 2016). Typically, however, subjective cognitive impairment manifests in the form of memory complaints, which can motivate a more thorough cognitive examination. Individuals can become aware of a change in their cognitive abilities long before these become objectively detectable.

Subjective cognitive impairment impacts wellbeing (Braun et al., 2017; Yates et al., 2017) and is linked to an increased risk of dementia, whereby older adults without objectively detectable cognitive impairment are twice as likely to develop dementia if they have subjective memory decline than if they do not (Mitchell et al., 2014). Subjective cognitive impairment has therefore been recognised as an important early marker of dementia (Jessen, 2014).

However, subjective cognitive assessments are prone to bias. They become less reliable as an individual progresses in the disease and their awareness of cognitive decline decreases (Cacciamani et al., 2021; Pennington et al., 2021). Additionally, declines in meta-cognitive abilities (Rouault et al., 2018) or depression (Cherry et al., 2019) can negatively impact self-perceived cognition. Furthermore, changes with age are common — Ponds et al. (2010) reported that starting at the age of 50, subjective cognitive functioning worsens with age even in adults without dementia, and this applies to a range of cognitive domains including memory, attention, mental speed, planning, and decision-making. This suggests

that subjective cognitive impairment is a sensitive but non-specific measure of cognitive decline.

Objective and subjective cognitive assessments can complement each other in the earliest stages of dementia progression. This is even more important because 1) neither type of assessment is fully deterministic (Mitchell et al., 2014; Pandya et al., 2016), and 2) objective and subjective cognition have only small, albeit significant correlation across studies (Burmester et al., 2016). Some accounts also support a model in which subjective cognitive complaints precede more pronounced, objectively detectable cognitive deficits (Cacciamani et al., 2021; Jessen, 2014). Irrespective of the order of appearance, it seems that both objective and subjective cognition can be useful; the former might enable higher specificity and the latter higher sensitivity.

1.3 Risk Factors for Cognitive Decline

Brain health is inextricably connected to other aspects of general health. It is important to remember that describing each aspect as a separate entity is more of a useful simplification than an accurate representation of reality. For example, in diseases that primarily impact a single bodily system (e.g., the metabolic system in diabetes), there is evidence for associated biological ageing across other organs, including the brain (Tian et al., 2023). Perhaps best established is the link between brain and cardiovascular health, in that factors like hypertension, high cholesterol, or obesity can all contribute to small vessel disease in the brain, thus increasing the risk of stroke and dementia (Kivipelto et al., 2006; Lane et al., 2019; Tian et al., 2023). Accordingly, research has long recognised that general health, as well as various environment and lifestyle factors that affect it, play a vital role in keeping our brains healthy throughout life.

In the last decades of epidemiological research, much has been learnt about the factors that contribute to an increased risk of cognitive decline, many of which could be modified to help delay or prevent dementia of any cause (Livingston et al., 2024). These are described in the next section. Alongside potentially modifiable risk factors, a range of non-modifiable factors are also associated to

brain health. These include factors like age, genetic factors, and sex/gender differences (Riedel et al., 2016; Scarabino et al., 2016), which will be discussed in later sections.

1.3.1 Modifiable Risk Factors for Cognitive Decline

According to recent estimates, up to 45% of dementia cases worldwide could be attributed to potentially modifiable risk factors (Livingston et al., 2024). These include a range of factors related to lifestyle, health, and environment that are assumed to exert their effects across the lifespan. The Lancet Commission for Dementia periodically reviews existing high-quality evidence in this field (Livingston et al., 2017, 2020, 2024), most recently identifying 14 modifiable risk factors including early-life low education, mid-life hearing loss, high LDL-cholesterol, depression, physical inactivity, diabetes, smoking, hypertension, obesity, and excessive alcohol use, and late-life social isolation, air pollution, and visual loss (Livingston et al., 2024). Furthermore, many other potential risk factors have been identified beyond these 14, including socioeconomic status (SES) and other social determinants of health (Majoka & Schimming, 2021; A.-Y. Wang et al., 2023), sleep (S. Wang et al., 2024), diet (Nucci et al., 2024), infections (Sipilä et al., 2021), psychiatric symptoms (Miniawi et al., 2022), early menopause (Fu et al., 2022), tinnitus (Jafari et al., 2019), and cardiovascular diseases like stroke (Kuźma, Lourida, et al., 2018). Proposed mechanisms of action include direct effects on neuropathology, vascular damage, and inflammation in the brain, as well as effects on cognitive reserve or resilience, which enable individuals to withstand brain health insults (Livingston et al., 2024; Turrini et al., 2023). Although this research is based on evidence from all around the world, the importance of modifiable risk factors for dementia differs by region (Mukadam et al., 2019; Santamaria-Garcia et al., 2023). For example, one study estimated that hypertension accounted for a higher percentage of dementia cases in Latin America (9.3%), China (6.4%), and India (4%) than it did worldwide (2%), and the impact of low education or obesity was also estimated to be higher in these countries (Mukadam et al., 2019). The finding that age-specific dementia incidence is decreasing in some high-income countries such as the USA and Sweden

(Mukadam et al., 2024; Wolters et al., 2020) suggests that policy-level changes addressing modifiable risk factors are effective in reducing dementia rates.

A methodological challenge in risk factors research is finding evidence for causality. Epidemiological data is mostly observational, albeit often longitudinal, which means that reverse causation and confounding can threaten the ability to infer causality. For instance, research has found an association between alcohol abstinence and cognitive decline (Kesse-Guyot et al., 2014; M. C. Norton et al., 2012), which has been interpreted as evidence for health benefits of moderate alcohol consumption. However, alcohol abstinence can be motivated by confounding health factors or previous heavy drinking, which can be independently causing an increased risk of cognitive decline (Andrews et al., 2019; Sallis et al., 2023). Where possible, randomised control trials (RCTs) have provided causal evidence for the effects of various risk factors. For example, RCTs have supported the efficacy of hearing aids (Lin et al., 2023), antihypertensive treatments (Peters et al., 2022), and multidomain risk factor interventions (Salzman et al., 2022) in addressing cognitive impairment related to hearing loss, hypertension, and other risk factors. However, in other cases, RCT evidence does not align with findings from observational studies. A recent meta-analysis of RCTs by Ciria et al. (2023) found only a negligible protective effect of regular physical activity on cognitive functioning in healthy humans across the lifespan, contesting the strong evidence from observational studies (Iso-Markku et al., 2022). Another issue is that optimal experimental designs are not always possible: there is often a large lag between treatment and the expected outcomes of interest (i.e., changes in brain structure and function), the costs are high, validity in ‘real world’ scenarios limited, and allocating individuals to treatment groups can be ethically challenging or impossible, especially in situations where risks for other aspects of health have already been established, such as hypertension or smoking (Lindhout et al., 2023). For these reasons, observational evidence in the field is being complemented by creative quasi-experimental methods (Lindhout et al., 2023). Notably, natural experiments have allowed researchers to estimate the effects of education on dementia risk in countries that introduced population-wide changes to the schooling system (Courtin et al., 2019; Seblova et al., 2021). Furthermore, Mendelian randomisation studies have also enabled causal inferences about

dementia risk factors, relying on genes associated with the risk factors in question as proxies for random treatment assignment (Kuźma, Hannon, et al., 2018).

Where converging causal evidence is available, observational studies take on the assumption that the correlational evidence they report is driven by underlying causal mechanisms. Public health recommendations typically adopt a pragmatic approach: Even small effects from highly heterogeneous studies with limited causal evidence are often deemed sufficient evidence to act on a risk factor, especially if interventions could lead to general health benefits, as illustrated by the earlier example of exercise (Ciria et al., 2023; Livingston et al., 2024).

1.3.2 The Life-Course Model of Dementia Risk Factors

Age is one of the most established non-modifiable risk factors for dementia, in that in older age, prevalence of dementia is estimated to double about every five years (Cao et al., 2020; E. Nichols et al., 2022). Additionally, cognitive performance typically follows domain-specific trajectories across the lifespan, with so-called ‘fluid abilities’ peaking on average in early adulthood and ‘crystallised knowledge’ in early midlife, and both progressively declining after that (Tucker-Drob, 2019; Turrini et al., 2023). Subjective perceptions of cognition across domains also reflect a progressive decline from fifty years onwards (Ponds et al., 2010). Importantly, age interacts with other risk factors, as it not only reflects normal ageing processes, but also the time over which other risk factors exert their effects.

An important concept is that of a ‘life-course model’ of dementia risk (Livingston et al., 2017, 2020, 2024). The central assumptions of this model are that 1) risk factors cumulatively affect brain health throughout the lifespan, and 2) risk factors can have different ‘sensitive windows’ at which they are particularly — although not exclusively — relevant for brain health. These windows are partially based on data, e.g., type 2 diabetes and obesity seem to be risk factor for dementia in midlife but not in late life (Barbiellini Amidei et al., 2021; Singh-Manoux et al., 2018), and partially on the age in which different risk factor effects are commonly studied (Livingston et al., 2024). However, although many studies have

investigated risk factor effects on brain health outcomes in older adults, relatively few studies have explicitly considered the effects of previously established dementia risk factors on cognition across the lifespan.

Lifespan evidence supports the first assumption of the life-course model, showing cumulative effects of risk factors on cognition. LaPlume et al. (2022) studied over 22,000 individuals aged 18 to 89 and found additive effects of modifiable risk factors, in that on average, each additional risk factor was linked to a worsening in objective cognitive performance equal to 3 years of ageing. Similarly, a study by Bransby et al. (2023) focusing on 1,610 adults aged 40 to 70 found that the higher the number of modifiable risk factors, the worse was an individual's objective working memory performance and the greater their subjective cognitive complaints. There is also ample evidence that modifiable dementia risk factors do impact cognition at different stages of life outside of old age (Barbiellini Amidei et al., 2021; Biessels et al., 2008; Borgeest et al., 2020; Lövdén et al., 2020; E. S. Nichols et al., 2021), although less research has explicitly studied the differences in the relevance of risk factors at different ages (Anstey et al., 2019, 2021; H. Chen et al., 2023; LaPlume, McKetton, Levine, et al., 2022).

1.3.3 Genetic Risk Factors for Cognitive Decline

Genetic factors play a significant, albeit often elusive role in determining brain health. The term heritability is used to describe the proportion of the variation in a trait that can be attributed to genetic differences in a given population. Heritability is high for cognitive functioning (N. P. Friedman et al., 2008; McGue & Christensen, 2002) and the effects of genetic differences become even more pronounced with increasing age (Lindenberger et al., 2008). Additionally, brain diseases tend to run in families: family history of dementia conveys a risk for brain health, particularly for individuals with first-degree relatives and parents with dementia history (Cannon-Albright et al., 2019; Rohrer et al., 2009; Scarabino et al., 2016). Heritability estimates range from 30–50% for FTLD and 33–79% for late-onset AD, yet heritability is much lower for vascular dementia (>1%; Bergem et al., 1997; Ferencz & Gerritsen, 2015; Gatz et al., 2006; Rohrer et al., 2009).

Many genetic risk factors for known neuropathologies have been identified, explaining a part of these heritability estimates (K. Chen et al., 2024; Kia et al., 2021; Wightman et al., 2021). Besides several relatively rare genetic variants linked to early-onset autosomal dominant AD (S. Hsu et al., 2018), the apolipoprotein $\epsilon 4$ allele (APOE- $\epsilon 4$) is the strongest known and most studied genetic risk factor for dementia, best known for its association to AD but implicated in a range of brain diseases (Corder et al., 1993; Fernández-Calle et al., 2022; Wightman et al., 2021).

The human APOE gene has three forms, the alleles $\epsilon 2$, $\epsilon 3$, and $\epsilon 4$. Humans inherit one allele from each parent, resulting in different allele combinations. The APOE- $\epsilon 4$ allele occurs in about 15% of population (Eisenberg et al., 2010) and is linked to an increased risk of all-cause dementia and cognitive decline in a dose-dependent manner, whereby $\epsilon 4$ -heterozygotes (i.e., people with one $\epsilon 4$ allele) have a higher risk and earlier onset of dementia than non-carriers and $\epsilon 4$ -homozygotes (i.e., people with two $\epsilon 4$ alleles) have a higher risk and earlier onset than both non-carriers and $\epsilon 4$ -heterozygotes (Riaz et al., 2021). By the age of 85, about 50% of $\epsilon 4$ -homozygotes will develop AD, whereas only 10% of non-carriers will (Genin et al., 2011). Interestingly, the APOE- $\epsilon 2$ allele seems protective (Corder et al., 1993). The exact mechanisms of APOE's involvement in AD and other neuropathologies are still undergoing research, with prominent hypotheses suggesting a role in the generation of amyloid- β , tau pathology, and neuroinflammation (Fernández-Calle et al., 2022). According to the recently proposed multi-hit hypothesis (Steele et al., 2022), APOE- $\epsilon 4$ might exert its detrimental effects through multiple biological pathways simultaneously, independent of amyloid- β and tau pathologies, whereby insults across different systems probabilistically add up, resulting in an increased susceptibility to neuropathologies.

In spite of APOE- $\epsilon 4$ being the strongest known genetic risk factor for dementia, genetic screening for it is not common nor recommended and a prevailing practice in clinical practice has been that of non-disclosure (Galluzzi et al., 2022; Largent et al., 2021). This is based on the non-deterministic nature of the gene, which is not necessary nor sufficient to cause dementia and is estimated to only explain 27.3% of AD heritability (Lambert et al., 2013). Additionally, APOE- $\epsilon 4$ seems to interact in its effects on dementia risk with other factors such as sex (Altmann et al., 2014;

Polsinelli et al., 2023) or cardiovascular risk factors (Caselli et al., 2011; Sapkota et al., 2017). However, the detection and disclosure of genetic status might be beneficial for several reasons. First, identifying APOE- ϵ 4 carriers can be useful for clinical trial recruitment, which often aims to especially target at-risk individuals (e.g., Langbaum et al., 2019). Second, APOE status can be used in risk profiling and personalised risk reduction interventions, with initial strides in this direction being made in the US (Isaacson et al., 2018; Ranson et al., 2021). Third, having APOE- ϵ 4 is associated with an increased risk of side effects for the recently developed amyloid- β -targeting therapies for AD (Honig et al., 2023; VandeVrede et al., 2020), which makes the knowledge of APOE status potentially important in future clinical care and decision-making. Finally, many people, particularly first-degree relatives of dementia patients, are interested in and worried about their genetic risk (Kohler et al., 2017; Roberts, 2000) and might benefit from disclosure in a protocolled clinical or research context that incorporates genetic counselling as opposed to direct-to-consumer commercial testing, which is becoming increasingly available (Harkins et al., 2015; Largent et al., 2021; Messner, 2011).

1.3.4 APOE- ϵ 4 and Cognition across the Lifespan

Given the well-established detrimental effects of APOE- ϵ 4 on cognition in later life, many researchers have asked if APOE- ϵ 4 also impacts cognitive functioning in healthy adults across the lifespan. This is important because 1) understanding the associations of APOE- ϵ 4 and cognition could help disentangle the mechanisms underlying APOE- ϵ 4 risk-conferring effects, and 2) if sensitive, cognitive assessments could be used for the early detection of individuals at an increased genetic risk. However, as reviewed by several meta-analyses, the associations between APOE status and cognition across lifespan have been inconsistent in literature (Daly et al., 2024; Ihle et al., 2012; Lancaster, Tabet, et al., 2017; O'Donoghue et al., 2018; Small et al., 2004; Weissberger et al., 2018; Wisdom et al., 2011). Possible explanations for these incongruencies include competing mechanisms of action and methodological limitations.

Two main mechanisms of action have been suggested to explain APOE- ϵ 4 effects on cognitive decline, with possibly complementary effects (O'Donoghue et al., 2018). First, APOE- ϵ 4 could be linked to an increased likelihood of AD neuropathology, which could then manifest as cognitive impairments especially apparent in older adults (i.e., a 'prodromal' mechanism) (Greenwood et al., 2014; Knight et al., 2014). Indeed, meta-analytical evidence suggests older APOE- ϵ 4 carriers on average performed worse on some, but not all domains including episodic memory, executive function, and global cognitive measures (Small et al., 2004; Wisdom et al., 2011). Additionally, some — but not all — studies reported more pronounced cognitive decline in APOE- ϵ 4 carriers on these measures over time (O'Donoghue et al., 2018). However, studies on these topics were highly heterogeneous in their methodology, warranting caution when interpreting these findings (O'Donoghue et al., 2018; Small et al., 2004; Wisdom et al., 2011). Furthermore, evidence from animal studies suggests that even in the absence of AD-related neuropathology, APOE- ϵ 4 is linked to worse cognitive outcomes and to various forms of cellular damage including neurodegeneration and neuroinflammation, suggesting a prodromal mechanism cannot be the only explanation (Steele et al., 2022).

Second explanation is that of a 'cognitive phenotype' mechanism, whereby APOE- ϵ 4 could be associated with a life-long cognitive phenotype that is independent of future AD diagnosis (Greenwood et al., 2005, 2014; Greenwood & Parasuraman, 2003). Although meta-analyses found no consistent APOE- ϵ 4 effects on the performance across standard neuropsychological tests in early life (Ihle et al., 2012; Weissberger et al., 2018) or midlife (Lancaster, Tabet, et al., 2017), some studies employing more sensitive cognitive assays detected subtle differences in cognitive performance already in these periods, suggesting phenotype effects might be isolated to specific cognitive sub-domains (see Lancaster, Tabet, et al., 2017; O'Donoghue et al., 2018 for a review). A recent meta-analysis of APOE- ϵ 4 effects on visuospatial abilities found that across 54 studies and 9,587 individuals across the lifespan, APOE- ϵ 4 was linked to a small average decrement in tasks measuring spatial long-term memory, even though other subdomains of visuospatial abilities (spatial construction, spatial working memory, spatial reasoning, and navigation) were unaffected on average (Daly et al., 2024). Further

muddling the waters, the popular ‘antagonistic pleiotropy’ hypothesis posits that APOE- ϵ 4 might exert beneficial phenotype effects in earlier life despite its detrimental effects in later life, which would explain why the gene has remained so common (Han & Bondi, 2008). Supporting this, some studies in young or middle-aged adults found highly domain-specific protective effects of APOE- ϵ 4 on the ability to quickly switch attention at visuospatial tasks (Greenwood et al., 2014; Rusted et al., 2013; but cf. Lancaster, Forster, et al., 2017) or on visual working memory binding for objects and their location (Zokaei et al., 2017, 2020). However, meta-analyses and large-scale studies alike reported that APOE- ϵ 4 had null average effects on a range of cognitive domains in early life, including the broader domains of attention and working memory (Henson et al., 2020; Ihle et al., 2012; Weissberger et al., 2018), and middle-aged APOE- ϵ 4 carriers even showed impairments on some aspects of episodic memory and executive function (Lancaster, Tabet, et al., 2017; O’Donoghue et al., 2018). Some neuroimaging studies have suggested more efficient neural processing in early and mid-adulthood, which has been interpreted as another example of antagonistic pleiotropy (e.g., Koelewijn et al., 2019; Mondadori et al., 2007; but cf. Henson et al., 2020), but the degree to which this is reflected in cognitive function is unknown and there is little consensus as to how to interpret these APOE- ϵ 4-related differences (Trachtenberg et al., 2012).

A likely explanation of the incongruent findings lies in methodological drawbacks of the existing studies. Although heritability is high for cognitive abilities (N. P. Friedman et al., 2008; McGue & Christensen, 2002), the effects of specific genes on complex human traits are known to be quite small (Marian, 2012), suggesting highly sensitive tools, very large samples, and robust methodology would be necessary to detect any cognitive phenotype effects. Regarding tool sensitivity, some authors have argued APOE- ϵ 4 effects on cognition could be highly domain-specific, meaning very precise cognitive assays would be required to isolate cognitive sub-domains most affected by APOE- ϵ 4-related neurobiological mechanisms such as impaired synaptic plasticity (Daly et al., 2024; Lancaster, Tabet, et al., 2017; O’Donoghue et al., 2018). However, some apparently domain-specific findings could also be due to chance in studies that included many cognitive outcomes and did not correct for family-wise error, which is unfortunately

common (O'Donoghue et al., 2018). Regarding sample size, significant effects in either direction are more common in smaller studies in the field (Ihle et al., 2012) and there is evidence of publication bias in the literature (Daly et al., 2024), possibly reflecting high rates of Type I Error. Larger samples would be better powered to detect any effects. Regarding robust methodology, there has been large heterogeneity in the way cognition is measured and cognitive domains are defined across studies (Ihle et al., 2012; O'Donoghue et al., 2018; Weissberger et al., 2018). Focusing on specific cognitive measures rather than broad cognitive domains (e.g., distance to target on a wayfinding task rather than the broad term 'navigation ability') might help remedy this. Additionally, the effects of APOE- ϵ 4 might be obscured by the way carrier status is defined in research. Studies are not consistent in their definitions of APOE- ϵ 4 carriers and non-carriers with some but not all studies excluding carriers of the rarer APOE- ϵ 2 allele (see Weissberger et al., 2018, for a review), which seems to be protective against AD (Corder et al., 1993). In cases where only ϵ 3 ϵ 3 carriers were included in the non-carrier group, comparisons against ϵ 3 ϵ 4 and ϵ 4 ϵ 4 carriers could have yielded less pronounced differences. Overall, replications in large samples are needed, ideally using study designs that handle multiple of these methodological issues at once (O'Donoghue et al., 2018)

1.4 Sex/Gender Differences in Cognitive Impairment and Risk

An important source of individual differences in brain health are sex and gender. The term sex usually refers to a biological and physiological characteristic, whereas the term gender refers to a socially constructed aspect of one's identity and the norms, behaviours, and roles assigned to it (Mielke et al., 2022). Binary labels are commonly used to describe sex as male/female and gender as man/woman, although science is increasingly recognising that such categorical labels are insufficient to reflect the complex reality of intersex conditions and non-binary and gender non-conforming identities (Ainsworth, 2015). Biological definition of sex is further complicated by the fact that sex is usually assigned at birth based on anatomy but not based on gonadal or chromosome testing, which are neither routinely done nor necessarily conclusive (Kronk et al., 2021; Shteyler

et al., 2020). Sex and gender overlap and the terms are often used interchangeably in common speech and research, which makes their effects on brain health hard to disentangle (Mielke et al., 2022).

Women and men are differently impacted by dementia. Dementia prevalence (i.e., the total number of individuals with the disease at a given timepoint) and incidence (i.e., the number of new cases with the disease within a given time period) differ between men and women in some, but not all countries (Huque et al., 2023; Wolters et al., 2020). A recent meta-analysis found that the prevalence, but not incidence of any-cause dementia and AD worldwide was higher in women, which was associated to higher life expectancy and lower educational achievement of women at the country level (Huque et al., 2023). Additionally, in the oldest age bracket of adults over 90, women had higher incidence of any-cause dementia and AD than men (Huque et al., 2023). This study, along with an earlier meta-analysis, have shown a trend towards higher prevalence of vascular dementia in men (Cao et al., 2020; Huque et al., 2023), suggesting that patterns of disease and/or diagnosis differ by sex/gender.

Sex/gender differences have also been reported with regards to the cognitive functioning of healthy adults across the lifespan. On average, women tend to outperform men on objective tests of verbal episodic memory, which is memory for verbal information such as stories or word lists, and processing speed, which is the time needed to complete cognitive tasks; whereas men tend to outperform women on tests of visuospatial abilities (Avila et al., 2019; Daseking et al., 2017; de Frias et al., 2006; Ferreira et al., 2014; Herlitz et al., 1999; LaPlume, McKetton, Anderson, et al., 2022; McCarrey et al., 2016; Proust-Lima et al., 2008; Siedlecki et al., 2019). Subjective cognitive complaints were more prevalent in women in some (Gagnon et al., 1994; Pérès et al., 2011; Wen et al., 2021) but not all studies (Bassett & Folstein, 1993; Jonker et al., 1996). Moreover, women tend to subjectively overestimate the extent of their objective cognitive impairment more than men (Sundermann et al., 2018). However, women across the lifespan are also more likely to have depressive symptoms (Salk et al., 2017), which are a prominent correlate of subjective cognitive impairment (Hill et al., 2016) and might reduce the diagnostic utility of subjective cognitive assessments (Edmonds et al., 2014). To detect cognitive impairment and decline with increased precision, we

need to take these differences into account. One way to do so is by using sex/gender-adjusted norms for objective cognitive assessments, as the sex/gender-related advantage on some domains could lead to lowered sensitivity and delayed diagnosis if domain-specific tests are used (Sundermann, Biegon, et al., 2017; Sundermann et al., 2021). Additionally, subjective cognitive assessments might be particularly useful for early detection of cognitive decline in women where such complaints have a closer relationship to objective cognitive performance (Sundermann et al., 2018) and better predict eventual cognitive decline (Pérès et al., 2011), but comorbidities such as depression need to be considered.

The exact drivers of sex/gender differences in brain health are still debated. The possible explanations include 1) differences in exposure to individual risk factors in women and men, 2) differences in susceptibility to the risk factor effects in women and men, 3) differences in risk-factor-related survival, and 4) differences in diagnosis and treatment. First, prevalence of risk factors differs by sex and gender. Some risk factors are sex-specific, such as the hormonal changes linked to early menopause (Fu et al., 2022; Gurvich et al., 2018) or andropause (Sindi, Toopchiani, et al., 2021). Other influential factors are more common in one or the other gender, including lower educational achievement, less mentally stimulating jobs, and SES inequality (Ilinca & Suzuki, 2021; Rocca et al., 2014). Second, there is evidence that the susceptibility to the effects of some risk factors also differs by sex/gender. For example, women are more susceptible to the genetic risk conferred by the APOE- ϵ 4 allele (Altmann et al., 2014), although the exact mechanism driving this difference remains unknown. Third, some differences have been attributed to differences in survival rates. For instance, some studies have shown that women are more susceptible to the effects of midlife hypertension on late-life cognition (Anstey et al., 2021; Lyu & Kim, 2016; Singh-Manoux & Marmot, 2005; Gilsanz et al., 2017; Hébert et al., 2000). However, this finding is not universal, as other studies found no significant sex/gender difference (Geraets & Leist, 2023; Nianogo et al., 2022). The relatively stronger association of midlife hypertension and cognitive decline in women may be attributed to higher mortality rates among men due to cardiovascular risk factors in midlife, which can lead to selective survival of men into old age when cognitive decline typically occurs

(Chêne et al., 2015). Finally, men and women differ in the diagnosis and treatment of brain health risk factors (Klungel et al., 1997), in health symptom reporting (Barsky et al., 2001) and in the diagnosis of brain health declines (Sundermann, Biegon, et al., 2017; Sundermann et al., 2018), possibly misrepresenting the true prevalence of one or the other variable on either side of the equation.

1.5 Multivariate Approaches to Studying Risk Factors

With so many different factors at play, studying the complex interplay between brain health and its determinants is challenging. This is further complicated by the co-occurrence of risk factors, as well as their interactions with each other. Although research has identified many modifiable risk factors that independently predict cognitive decline, this needs to be complemented by a multivariate approach to disentangle the true effects and better understand the risk-conferring mechanisms.

Modifiable risk factors often co-occur. In a recent analysis of the Lancet commission, 54% of the total variance of 14 prominent modifiable risk factors for dementia was explained by five principal components, suggesting the risk factors co-occurred to a large extent (Livingston et al., 2024). It seems that socioeconomic factors underlie this co-occurrence. A systematic review conducted in the UK found that the co-occurrence of risky health behaviours is more likely in individuals with unfavourable social determinants of health such as lower SES and lower education (Meader et al., 2016). Another study of more than 6,000 participants of a German population-based cohort reported similar patterns, where socioeconomic disadvantage predicted stronger clustering of modifiable lifestyle and health risk factors for dementia (Röhr et al., 2022). As emphasised by the Health Lifestyle Theory, lifestyle behaviours are based not only on an individual's choices, but also on the options available to them dependent on the circumstances of their social class (Cockerham, 2005). Given the high co-occurrence of risk factors, it is important to understand the effects they have in combination, as opposed to just studying each risk factor in isolation.

One approach has been to compute a dementia risk index or composite risk score, which can help evaluate the additive effects of having multiple risk factors and identify individuals at an increased risk of dementia both in research and in the clinic (Anstey et al., 2013; Barnes et al., 2009; Dhana et al., 2022; Jessen et al., 2011; Kivipelto et al., 2006; Reitz et al., 2010; Vos et al., 2017). Well-established risk indices include a score based on the Cardiovascular Risk Factors, Aging, and Dementia study (CAIDE score), which predicted an increased risk of dementia up to four decades downstream with moderate accuracy (area under the ROC curve (AUC) = 0.747) (Exalto et al., 2014; Kivipelto et al., 2006), or the self-report-based Australian National University Alzheimer's Disease Risk Index (ANU-ADRI; Anstey et al., 2013), which predicted AD and dementia at a shorter follow-up than CAIDE but has better translated across populations (Anstey et al., 2014). Another prominent example is the Lifestyle for BRAin health (LIBRA) index, which is based on a broader range of weighted lifestyle risk factors (Vos et al., 2017) and predicts both cognitive impairment and dementia up to 16 years later (Schiepers et al., 2018). Some risk indices such as the CAIDE or LIBRA weight the relative importance of risk factors based on a multivariate analysis in a given cohort, whereas in other cases the risk ascertainment is based on an unweighted sum of lifestyle risk factors, i.e., the number of areas in which participants cross a set threshold (Dhana et al., 2022). Additive approaches like these have brought evidence that modifiable risk factors have cumulative effects of both objective and subjective cognition, as well as other brain health measures (Bransby, Rosenich, Buckley, et al., 2023; Jia et al., 2023; LaPlume, McKetton, Levine, et al., 2022; Reuben et al., 2022). A meta-analysis of studies using composite risk scores found that some modifiable risk factors can have a strong association with dementia and cognitive decline in combination, even though having just one of them was not associated with an independent significant effect (Peters et al., 2019). A caveat is that risk indices only provide a simplified, summative account of risk factors without fully considering the shared variance between risk factors or the relative importance of individual factors and risk factor clusters.

Latent variable approaches such as factor analysis or structural equation modelling (SEM) can shed more light into the latent structures that underlie risk factors. This is important because public health research tends to focus on

individual health-related behaviours while overlooking the latent causes of these behaviours, a phenomenon referred to as 'lifestyle drift' (Marmot & Allen, 2014). For example, Kesse-Guyot et al. (2014) used SEM to find that some lifestyle factors like reduced physical activity or diet were not individually linked to cognition but they contributed to an overall dose-dependent effect of unhealthy lifestyle in combination, represented by the latent variable of unhealthy lifestyle. In a more complex model, Borgeest et al. (2020) assessed the complementary contributions of modifiable lifestyle factor clusters towards different aspects of cognition across the lifespan, using exploratory structural equation modelling (Asparouhov & Muthén, 2009). They found that latent factors of socioeconomic status, physical health, and social and intellectual engagement, but not mental health, each made a unique, complementary contribution towards cognition, explaining 14% of fluid and 16% of crystallised abilities. Cluster analysis is another common approach, where frequently co-occurring risk factors are expressed statistically as risk factor profiles or dimensions to which individuals can belong. For example, Norton et al. (2012) identified distinct clusters of lifestyle behaviours (including diet, exercise, church attendance, social interaction, smoking, and alcohol intake) that were associated with incident dementia in a longitudinal cohort study. Xiong et al. (2024) recently employed latent class analysis to identify three clusters of risk factors — cardiometabolic, substance-use-related, and low risk, which differed in their associations to brain health measures including cognition and neuroimaging outcomes. Interestingly, the structure and correlates of these risk factor profiles differed significantly by sex/gender. For instance, the cardiometabolic risk profile was associated to cognition in men, but not in women. Some limitations of these latent variable approaches include their underlying statistical assumptions, which do not necessarily reflect the complexity of real-life data. SEM assumes linear relationships between variables, but some risk factors like hypertension can have U-shaped associations with cognitive decline (e.g., Gong et al., 2021; van Dalen et al., 2022). Latent class analysis assumes that variables within each class are conditionally independent, but risk factors are known to interact with each other (e.g., Röhr et al., 2022).

Finally, machine learning approaches have become increasingly popular in the past few decades. Machine learning is an extension of traditional statistics which

focuses on developing accurate multivariable statistical models that generalise well to new, unseen data (Yarkoni & Westfall, 2017). Beside from its utility for the prediction of brain health outcomes (e.g., Pankratz et al., 2015; You et al., 2022), machine learning can be used as a data-driven way to uncover generalisable patterns in multivariate data. Generalisability can be a problem in risk factors research, in part because the relative importance of risk factors for brain health outcomes varies between populations (Demnitz et al., 2023; Mukadam et al., 2019). Another known advantage of machine learning is its flexibility. Machine learning can incorporate the statistical approaches described above, but it can also involve other, robust model types that are better suited to handle complex data. Some multivariate machine learning approaches such as support vector regressions excel at handling non-linear patterns in the data. For example, Fittipaldi et al. (2024) used this method to study the relative importance of different factors for social cognition, showing that factors like education, general cognitive abilities, country income, or sex/gender were more relevant for social cognition than age or clinical diagnosis (i.e., AD, FTLT, MCI, subjective cognitive complaints vs. no diagnosis). Other machine learning models such as tree-based methods or regularised regression can effectively handle shared variance between risk factors. For example, random forest models have been employed in older adults to study the relative importance of different risk factors for the trajectory of cognitive decline (Casanova et al., 2020) or the relative importance of different risk factors for memory resilience in men vs. women (McDermott et al., 2017). As another example, Santamaria-Garcia et al. (2023) used a range of regularised regression models on both cross-sectional and longitudinal data to investigate how countries with different income levels differ in the importance of dementia risk factors for cognition and functional outcomes.

One limitation common to the multivariate approaches reviewed here is that they require large study samples with many observations to provide reliable and generalisable results. This becomes even more crucial with increasing complexity of the data (e.g., more risk factor variables) or the models fit to it (e.g., tree-based models compared to simple linear regression) (Yarkoni & Westfall, 2017). Inherent to any dataset is noise, but if we select our models only based on how well they fit our dataset, they might be biased by sample-specific noise, a phenomenon known

as overfitting. Although machine-learning procedures such as cross-validation (i.e. training and testing the models on subsets of the data to ensure generalisability) and regularisation (i.e. constraining the models so that simpler solutions are preferred) can help address overfitting, this can only be done provided the sample size is sufficient (Yarkoni & Westfall, 2017). The following section describes how large-scale data collection can be facilitated through smartphones.

1.6 Cognitive Assessments Using Smartphones

To study risk factors for cognitive health, large study samples are necessary. This is particularly true for genetic studies, where effect sizes are expected to be small, (O'Donoghue et al., 2018) and for multivariate analyses, especially if these involve machine learning (Yarkoni & Westfall, 2017). Online data collection, particularly using smartphones, is one way to address this (Gillan & Rutledge, 2021; Whelan et al., 2022). The increasing smartphone and tablet usage presents new opportunities for expanding the availability and reducing the cost of cognitive assessment and for improving the precision and reducing the burden of cognitive testing, thus facilitating large-scale data collection. The following sections first describe the benefits and drawbacks of smartphone-based cognitive assessments and then review some notable examples of brain health research that crowdsourced data online.

1.6.1 Advantages and Limitations of Smartphone-based Cognitive Assessments

Digital cognitive assessments have a range of advantages including accessibility, scalability, greater flexibility in designing stimuli, greater precision and standardization, and ease of scoring and data processing (Van Patten, 2021). The following paragraphs discuss each of these advantages in turn, while taking into account the associated limitations, including digital inclusivity, development cost, technical issues, and threats to data quality.

Improved accessibility is one of the key advantages of smartphone-based cognitive assessments. Standard in-person assessments are not always available in rural areas or to individuals with limited access to travel (Van Patten, 2021). Barriers to access are even more prominent in low- and middle-income countries (LMICs), presenting a particular opportunity for smartphone-based assessments and tele-health services in these regions (Chakrabarti, 2024; McCool et al., 2022; Muniz-Terrera et al., 2019). Smartphone-based assessments could accelerate recruitment in research and allow for larger sample diversity compared to traditional lab-based studies (Birnbaum, 2004; H. R. Brown et al., 2014; Doraiswamy et al., 2018; Gillan & Rutledge, 2021). In 2024, nine in ten U.S. adults owned a smartphone (Pew Research Center, 2024). Smartphone usage is increasing worldwide, including LMICs, although the rates are higher in high-income countries (Austin, 2024). Many people, particularly those from socioeconomically disadvantaged groups, are smartphone-dependent in their internet access (Pew Research Center, 2024), meaning that smartphones are their only means of getting online and thus participating in crowdsourced research. There is also evidence that touchscreen-based devices might be easier to use compared to a computer mouse for older adults, even those with a cognitive impairment (F. S. Lim et al., 2013; Valladares-Rodriguez et al., 2017).

Despite these substantial increases in smartphone accessibility, mobile devices are still not omnipresent. To understand what smartphone assessments can say about people, we need to also look at those who remain out of their reach. Access to smartphones (Pew Research Center, 2024) and participation in smartphone-based psychology research (Keusch et al., 2022) both decrease with increasing age. Other factors, such as low digital literacy, belonging to a racial minority, or low socioeconomic status, are all linked to lower participation in smartphone research (Keusch et al., 2022; Stone et al., 2024). This can reduce the representativeness of smartphone research, especially in older or otherwise vulnerable populations. However, the prevalence of mobile technologies is increasing even in older adults. In 2024, three quarters of U.S. adults over 65 owned a smartphone (Pew Research Center, 2024), an almost 6-fold increase since 2012 (Faverio, 2022). Smartphone apps have been successfully employed for crowdsourcing data from individuals across the lifespan including older adults

and to detect age effects on cognition (McNab et al., 2015; Rutledge et al., 2016; Smittenaar et al., 2015; Teki et al., 2016; Walkowiak et al., 2023). Among older participants, those with high digital literacy are most likely to self-select to smartphone-based studies. However, self-selection can pose problems also in lab-based research, which requires the participants to be physically able to arrive at the research site. The overlap and impact of each of these requirements is currently unknown.

Smartphone-based cognitive assessments are also praised for their relatively low cost compared to lab-based studies (Birnbaum, 2004; Feenstra et al., 2017), which makes them easier to scale (Van Patten, 2021). At the same time, it is important to consider the initial costs of developing an engaging smartphone-based test or a user-friendly interface, together with the more long-term costs of advertisement for recruitment purposes and app maintenance (H. R. Brown et al., 2014). Investments into the tool's usability can be substantial, as a good user experience requires iterative and potentially costly user-centred development. Nonetheless, higher usability can improve technology acceptance, which is a crucial factor that is often omitted when developing new digital assessment tools for older populations (Cubillos & Rienzo, 2023). Development costs need to be offset by tangible benefits and, ideally, a wide use of the resultant tool, requiring a cost-benefit analysis at the researcher side.

Smartphones and other digital interfaces are also appealing due to the new, flexible ways of task administration and stimulus presentation. First, they allow for greater precision (e.g., precise detection of on-screen position) and stimuli standardisation (e.g., pre-recorded task instructions) in comparison to pen-and-paper tasks, as well as for automated scoring and data processing (Germine et al., 2019; Van Patten, 2021). Together, this can help reduce experimenter bias (Birnbaum, 2004). Second, smartphone assessments facilitate repeated, longitudinal data collection, which enables to compare performance towards an individual's own baseline (Nicosia et al., 2023), to administer repeated memory assessments across varied retention intervals (Lancaster et al., 2020), as well as to capture cognitive fluctuations across different time intervals (Brewster et al., 2021; Sliwinski et al., 2018; Wilks et al., 2022). This type of variability would be very difficult to detect using laboratory-based protocols. However, the flexibility of

smartphones can also pose a technical challenge. There can be substantial differences across devices and operational systems with respect to screen size, speed, notification permissions, connection speed, and many other technical parameters (Germine et al., 2019; Passell et al., 2021), which need to be taken into account when designing a smartphone study.

Another way in which smartphone assessments facilitate flexible task administration is through gamification. Cognitive tasks can be repetitive, boring, too difficult, too easy (Germine et al., 2021) and may have practice effects that affect their validity (Scharfen et al., 2018). To solve these problems, there is growing interest in the use of gamification or ‘serious games’ in cognitive assessment (Gillan & Rutledge, 2021; Lumsden et al., 2016; Valladares-Rodríguez et al., 2016). Gamified assessments aim to reduce testing anxiety (Leduc-McNiven et al., 2018) and increase task engagement and enjoyment without affecting performance (Lumsden et al., 2016). This is even more important in a remote setting where engagement is harder to monitor and distractions and low motivation could detract from data quality (Germine et al., 2019; Van Patten, 2021). Many studies have successfully adapted gamified assessment methods in older adults with or without cognitive impairment (Isernia et al., 2021; T. Tong et al., 2016; Valladares-Rodríguez et al., 2018; Zygouris et al., 2015). Gamified assessment can provide better construct and ecological validity than simple laboratory-based tasks by presenting richer, more realistic contexts (Pedersen et al., 2023; Shute et al., 2016). Gamification is especially well suited for the assessment of navigation abilities or spatial memory where games can provide an immersive experience in 3D environments. An example of this is the Sea Hero Quest, a smartphone game designed to assess goal-oriented wayfinding in a high-quality 3D environment, in which participants navigate an on-screen boat towards check-points, using a map (Coughlan et al., 2019; see next section for study details). Finally, ‘adaptive testing’ is a common feature of gamified tasks, which is a procedure wherein tasks become more or less challenging to provide a tailored experience for the end-user. This allows cognitive tests to arrive at a reliable estimate of a person’s ability faster and more efficiently and avoids ceiling and floor effects that can harm test sensitivity. Adaptive testing also has the benefit that it promotes engagement on the part of the user and reduces frustration,

making people more likely to play for longer and more frequently (Mishra et al., 2016). A driving scenario game took this approach to assess attention and executive function via tablet (W.-Y. Hsu et al., 2021). The game is a closed-loop system that dynamically adjusts to keep the difficulty at 80% for all players and takes approximately 7 min to complete. In a recent proof-of-concept study, the game was able to distinguish those with cognitive impairment in a multiple sclerosis (MS) population from those without (W.-Y. Hsu et al., 2021).

1.6.2 A Crowdsourcing Approach to Brain Health Research

Leveraging the advantages listed above, citizen science is increasingly aiming to recruit smartphone users as a way to facilitate research participation. The term 'citizen science' refers to the organised involvement of the general public in scientific data collection or processing (Bonney et al., 2016). Apart from benefiting researchers, citizen science can also be rewarding for participants, as it can help increase public involvement in scientific research, potentially enhancing their understanding of scientific research, contributing to well-being, and increasing perceived sense of meaning in participants' hobbies (Bonney et al., 2016). Researchers can use citizen science platforms to give back to the community, provide science education, and communicate their findings.

As a prominent example of crowdsourcing for brain health research, the smartphone game Sea Hero Quest used a citizen-science approach to engage people from over 46 countries in measuring their spatial navigation abilities, receiving over 4.3 million downloads and over half a million usable datapoints (Coutrot et al., 2018; Walkowiak et al., 2023). This has helped researchers study the cultural determinants of navigation abilities (Walkowiak et al., 2023) or APOE-ε4 effects on navigation abilities (Coughlan et al., 2019), as described in more detail further below. Crowdsourced data collection via smartphone apps is also popular in, but not limited to, the field of computational psychiatry (Donegan et al., 2023; Fox et al., 2024; McNab et al., 2015; Rutledge et al., 2016; Smittenaar et al., 2015; Teki et al., 2016). Other research groups in the field of brain health research use web-based applications to engage the general public in large scale data

collection. The online platform TestMyBrain.org has been utilising a citizen-science approach to gather large-scale data on a battery of self-administered cognitive tests, which probe a range of cognitive abilities including memory, processing speed and executive function. Data from over 135,000 such citizen scientists was then used to develop the Test My Brain Digital Neuropsychology Toolkit (Singh et al., 2021) and make it rapidly available to meet a growing need for remote neuropsychological assessment during the COVID-19 pandemic.

Fewer studies have applied citizen science to studying risk factors for cognitive decline. As a particularly relevant example, the computerised online platform Cogniciti helped study a range of cognitive abilities in over 40,000 citizen scientists (LaPlume, Anderson, et al., 2022), but also the associations of dementia risk factors to cognition across the lifespan in a subset of these participants (LaPlume, McKetton, Anderson, et al., 2022; LaPlume, McKetton, Levine, et al., 2022). One study used Cogniciti data from over $N = 22,000$ participants aged 18–89 to study the prevalence of eight modifiable risk factors across the lifespan and to test their dose-response association with objective cognition (LaPlume, McKetton, Levine, et al., 2022). Not only did this cross-sectional study find evidence for cumulative effects of risk factors on cognition, as described in Section 1.3.2, but the large sample also enabled authors to study the interactions of such cumulative effects with age, finding that older participants had more risk factors and stronger dose-response effect of risk factors on cognition. Another study using Cogniciti data further tested how gender comes into play, reporting that the cumulative effects of modifiable risk factors on a paired associate recognition test were more pronounced in older women, although men had more risk factors and worse cognition on average (LaPlume, McKetton, Anderson, et al., 2022). These large-scale studies demonstrate the power of citizen science to study how risk factors for dementia interact with each other. However, both studies only investigated the cumulative effects of modifiable risk factors rather than the role of individual factors.

Although smartphones cannot be used to directly collect genetic data, smartphone-based cognitive assessments can be combined with other research methods to help study individuals at a genetic risk of dementia. For example, Coughlan et al. (2019) administered Sea Hero Quest to a lab-based genotyped

sample and found that APOE- ϵ 4 carriers ($n = 31$) had impaired performance on wayfinding relatively to non-carriers ($n = 29$). While these effects must be interpreted with caution given the small sample size, the authors additionally used crowdsourced, age-matched benchmark data ($n > 27,000$) as a normative comparison for the smaller lab-based dataset. Taking advantage of the fact that the majority of the general population — and so a majority of this benchmark sample — are not APOE- ϵ 4 carriers, the authors tested if navigational performance on the Sea Hero Quest can be used to differentiate genotyped APOE- ϵ 4 carriers from a dataset including mostly non-carriers (i.e. the crowdsourced benchmark data combined with APOE- ϵ 4 non-carriers of the lab-based sample), reaching better sensitivity than a classic visual episodic memory test (AUC = 0.714). This proof-of-concept study shows the potential of smartphone assessments for studying and detecting genetic risk, although replications in larger genotyped samples are needed to get clinically useful cognitive markers of such risk.

1.7 Studies and Hypotheses

Chapter 2 (Study 1): Validating a gamified smartphone measure of visual working memory as a tool to detect early cognitive decline

Chapter 2 of this thesis evaluated the psychometric qualities of a novel, gamified smartphone-based cognitive task as a measure of visual working memory, which can be self-administered through the smartphone app *Neureka*. This task, previously developed in the lab, was based on a traditional visual short-term memory binding task (VSMBT), which taps into working memory for conjunctions of shapes and colours and is promising for the detection of early AD-related decline. For the current validation study, we first used data from a small lab-based sample ($N = 33$) to evaluate the consistency of our new task with the traditional, computer-administered VSMBT and its different conditions. We further assessed the validity and split-half reliability of the new task in a large online sample crowdsourced via *Neureka* ($N = 6,398$). Finally, we assessed test-retest reliability of the task in a subset of the online participants who chose to repeat the task ($n = 294$).

Chapter 3 (Study 2): Modifiable dementia risk factors associated with objective and subjective cognition

Chapter 3 aimed to assess the differential associations of objective vs. subjective cognition with potentially modifiable risk factors for dementia. To this aim, we used a large cross-sectional sample of individuals without dementia, spanning the entire adult age range ($N = 3,327$), collected with the smartphone app *Neureka*. We ran a series of regressions to individually quantify the associations of 13 previously identified risk factors with subjective memory problems and three gamified objective measures of executive function, including cognitive flexibility, model-based planning, and the measure of visual working memory validated in Chapter 2. These analyses controlled for age and gender effects, but a secondary analysis also included depression symptoms as a covariate to test if depression explains the associations between risk factors and subjective cognition. Finally, we tested

an interaction effect of age to assess to what extent age modifies the associations between risk factors and cognitive outcomes.

Chapter 4 (Study 3): APOE- ϵ 4 associations with smartphone-based cognitive markers

In Chapter 4, we tested the associations between a set of previously validated smartphone-based cognitive markers and APOE- ϵ 4. We invited $N = 646$ APOE-genotyped participants of the PREVENT Dementia study, which is a multi-site prospective cohort study of cognitively healthy middle-aged adults, to complete a self-administered module in the smartphone app *Neureka*, including gamified cognitive tasks and memory self-evaluation. Our final sample involved $n = 87$ genotyped participants with sufficient completion rates. Controlling for age, gender, education, and parental history of dementia, we compared APOE- ϵ 4 carriers with non-carriers in their performance on five cognitive measures: model-based planning, visual working memory, processing speed (~Trails A), cognitive flexibility (~Trails B), and subjective memory problems. We conducted receiver operating characteristic (ROC) analysis to assess if APOE- ϵ 4 status can be predicted from norm-adjusted cognitive scores, i.e., the difference between observed cognitive performance and what we would predict for given demographics based on an additional benchmark sample, crowdsourced through the app *Neureka* ($N = 3,327$).

Chapter 5 (Study 4): Sex differences in modifiable risk factors associated with cognition: A machine-learning approach

Chapter 5 examined the sex differences in modifiable dementia risk factors and their effects on cognition in a large, cross-sectional lifespan sample collected through the smartphone app *Neureka* ($N = 4,995$). We trained a series of regularised models using a robust cross-validation approach and evaluated their performance in a 20% hold-out set, individually predicting subjective or objective cognition from risk factors, age, and sex vs. a model that additionally included risk factor interactions with sex. This way, we 1) tested the relative importance of

modifiable dementia risk factors for cognition when considered all together, and 2) tested if sex moderates the effects of modifiable dementia risk factors on cognitive impairments.

Chapter 2: Validating a gamified smartphone measure of visual working memory as a tool to detect early cognitive decline

2.1 Introduction

Executive functions are a broad group of higher-level cognitive abilities involved in controlling simpler cognitive processes, employed during effortful, non-automated tasks (Diamond, 2013). Although definitions vary, executive functions are often understood to include abilities such as working memory, cognitive flexibility, or inhibition (Diamond, 2013). Impairments to executive functions can precipitate dementia symptoms by up to a decade (Mortamais et al., 2017), making their assessment of interest in both research and clinical settings. The detection and monitoring of such early cognitive decline could be scaled up with the help of gamified smartphone assessments. One way to do this is through adapting and gamifying existing cognitive tasks that have shown sensitivity to incipient cognitive decline and dementia.

Particularly promising for early detection of cognitive decline is the Visual Short-Term Memory Binding Task (VSMBT), which is a type of a visual working memory task that measures the ability to temporarily hold in memory combinations of surface-level object features such as shape and colour (Cecchini et al., 2023; Logie et al., 2015). Many variations of the VSMBT exist, both computer-based and in physical form, employing both recognition and recall paradigms and tapping into memory for both abstract and everyday objects (Cecchini et al., 2023). Across these variations however, the VSMBT usually involves two conditions — a single-feature or ‘non-binding’ condition, where participants have to remember simple features (e.g., shape) alone, vs. a feature ‘binding’ condition, where participants have to remember combinations of such object features (e.g., shape together with colour or orientation). The binding condition or its relative impairment compared to the non-binding condition are assumed to reflect binding deficits (Schneegans & Bays, 2019). Traditionally, researchers have differentiated between *conjunctive memory binding*, which is the memory for conjunctions of features within the same object (e.g., shape-colour), and *relational memory binding*, which is the memory

for object pairs or for object locations (Cecchini et al., 2023). The VSMBT is an example of a task measuring the former.

There are several reasons to believe VSMBT could be useful for early detection of individuals at-risk of Alzheimer's disease (AD). First, unlike relational binding, conjunctive binding assessed with VSMBT was found to be specifically impaired in AD including its early stages, but not in normal ageing (Bastin, 2018; Brockmole et al., 2008; Killin et al., 2018; Parra et al., 2009, 2010; Yassuda et al., 2020). Further supporting this, a recent meta-analysis across $N = 1,933$ individuals found that individuals with AD, mild cognitive impairment, and preclinical AD performed on average worse on VSMBT than healthy controls, and this effect was twice as large or even larger for the binding condition compared to the single-feature condition of the task (Cecchini et al., 2023). Second, one study reported that deficits in conjunctive working memory binding differentiated AD patients from patients with frontotemporal dementia, Parkinson's disease, vascular dementia, or Lewy body disease (Della Sala et al., 2012). Finally, initial studies found that conjunctive working memory binding was also sensitive to amyloid- β deposition in the brain, which has been identified as an AD biomarker (Cecchini et al., 2021; Parra et al., 2024). However, another study found amyloid β was only associated to performance on the non-binding condition, not on the one that assesses conjunctive binding in particular (D. J. Norton et al., 2020), which suggests both conditions of the task might be of value. Together, this evidence implies the task could be sensitive to AD even in preclinical populations, albeit there is still a lack of longitudinal studies that would evaluate the task's ability to predict AD onset over time (Cecchini et al., 2023).

While the task seems promising for detection of early AD-related changes, several issues need to be considered when adapting it for administration in healthy individuals. First, ceiling effects might reduce the task's sensitivity. In studies that compare dementia patients to healthy controls on VSMBT, task difficulty is often titrated so that all compared groups have similar performance on the single feature condition which serves as a baseline (Cecchini et al., 2023; Parra et al., 2009, 2010). This is done by administering a different memory load to each group, i.e., varying how many objects are there in a study set for participants to remember. Even then, healthy younger individuals often exhibit ceiling effects on the task,

especially with smaller sets of objects to remember (e.g., Brockmole et al., 2008). Second, the task can be boring and unappealing to administer, given that its most common version comprises 32 to 100 trials (Cecchini et al., 2023). Finally, existing task versions are costly in terms of time and resources, as they require in-person administration, either with a computer using specialised software, or using a set of cards.

To address these limitations, we developed a gamified smartphone-based task called 'Memory Match' with graded difficulty levels, which was loosely based on the VSMBT. Compared to the traditional task, our newly developed game takes less time to complete, aims to be more engaging, and can be fully self-administered via smartphone while maintaining a standardised presentation of stimuli. In the current study, we validated the task by 1) comparing the new task with its traditional counterpart (i.e., convergent and discriminant validity) in a lab-based study, and 2) assessing the new task's validity and reliability in a large online sample crowdsourced via the Neureka app (i.e., external & construct validity, test-retest reliability, internal consistency).

2.2 Methods

2.2.1 Participants

First, we conducted an in-lab validation study, in which paid participants completed both Memory Match and its traditional computerised counterpart (VSMBT; Parra et al., 2010). As part of the study, invited participants also completed another new task (Star Racer) and its traditional version (the Trail-Making Test; see the Appendix of Chapter 3, Section 8.1.5.1, for the full validation analyses). We recruited a convenience sample at Trinity College Dublin, Ireland, through posters on campus and online advertisements and from the Greater Dublin area via online advertisements in the Dublin volunteer centre. Of the originally recruited $N = 45$ participants, $n = 35$ had complete data for both Memory Match and VSMBT. Further two participants were excluded — one due to colour-blindness and the other due to not passing a perceptual check (see Procedure, section Traditional

binding task below). The final $n = 33$ Memory Match/VSMBT participants were aged 18–46 ($M = 24.7 \pm 7.1$) and consisted of 63% female participants. Of these participants, $n = 31$ (age 18–46; $M = 24.7 \pm 7.1$; 64% female) also had available data for the Trail-Making Test (Bowie & Harvey, 2006) and its newly developed gamified smartphone counterpart called Star Racer (see Appendix of Chapter 3, Section 8.1.5.1 for a full description and validation). The Trail-Making Test and Star Racer both come in two versions, measuring processing speed (version A) and cognitive flexibility (version B), which enabled us to assess discriminant validity of Memory Match.

Second, we used data crowdsourced from unpaid ‘citizen scientists’ (i.e., volunteer members of the general public) who have completed Memory Match within the ‘Risk Factors Science Challenge’ module in the *Neureka* app since the app’s release in June 2020 until January 2023. After removing incomplete datasets and participants who left and re-entered the app during gameplay, the final sample consisted of $N = 6,398$ participants. Participants were aged 18–85 ($M = 45.2 \pm 14.7$). With regards to gender, 4,161 (65.0%) identified as cisgender female, 107 (1.7%) identified as non-cisgender, and 15 (0.2%) preferred not to state their gender. The participants came from 74 countries, with the United Kingdom, United States, Ireland, Canada, and Germany being the most prevalent. Additionally, a subset of $n = 294$ participants completed a part of the ‘Free Play’ module in the *Neureka* app that was comparable to the ‘Risk Factors’ version of the game, providing us with data for a test-retest reliability estimate.

2.2.2 Procedure

Study 1 (in-lab validation study). Participants were administered Memory Match and the traditional task in randomised order. Both Memory Match and Star Racer were administered on a Google Pixel 3a phone. After completing the tasks, participants provided their demographic data, then they were debriefed and reimbursed.

Study 2 (citizen scientists). Memory Match was included with two other gamified tasks in the 'Risk Factors Science Challenge' module in the *Neureka* app and administered remotely to volunteer citizen scientists alongside self-report lifestyle and health questionnaires. The Risk Factors challenge is further described in Chapter 3. The order of presentation of the tasks was randomised and completion of the entire challenge could be distributed across time. In the current analyses, we included each participant's first complete attempt on Memory Match in the Risk Factors challenge. Additionally, a modified version of Memory Match was available in the 'Free Play' section of the *Neureka* app as a standalone game, where the participants could choose which difficulty level of the game they complete and how many times. We used performance on completed Free Play takes and compared them to Risk Factors takes to estimate test-retest reliability of the task.

Both studies were approved by the Trinity College Dublin School of Psychology Ethics Committee and informed consent was obtained from all participants prior to participation.

2.2.3 Materials

Traditional binding task. A 'change-detection recognition' paradigm of the VSMBT (Brockmole et al., 2008; Parra et al., 2009) was administered using E-Prime 2.0 Run Time (Psychology Software Tools Inc., 2012). Like Brockmole et. al (2008), we used an array of black shapes (non-binding condition) or coloured shapes (binding condition) as stimuli. The shapes on each trial were randomly selected out of a set of eight six-sided polygons and presented on light grey background. The colours in the binding condition were randomly selected from a set of eight non-primary colours. No shape or colour was repeated within any given array. On each trial, a study array was presented followed by a test array, whereby the participants had to recognise whether the items presented in the test and study arrays were the 'same' or 'different' and to respond accordingly by a button press (see Figure 1 for a detailed overview of the task). On the 'different' trials, studied shapes were replaced by new ones (non-binding condition), or two shapes swapped colours (binding condition). Location of the stimuli changed

randomly between study and test arrays to prevent participants from using location as a memory cue.

The task in the current study consisted of four conditions, with two-item and three-item versions of both the non-binding and the binding condition. There were 16 ‘same’ trials and 16 ‘different’ trials in each condition, presented in random order, and so participants completed 128 trials in total. Conditions were blocked: all participants first completed a two-item block of the non-binding condition, followed by a two-item block of the binding condition. They were then shown two blocks of three-item displays, with half of participants receiving the non-binding block first. The order of conditions (non-binding or binding first) for the three-item blocks alternated between participants. To ensure that the participants understood the task, the examiner showed them sample arrays, gave an oral explanation of the task, and asked for their response (‘same’ or ‘different’), before each block started.

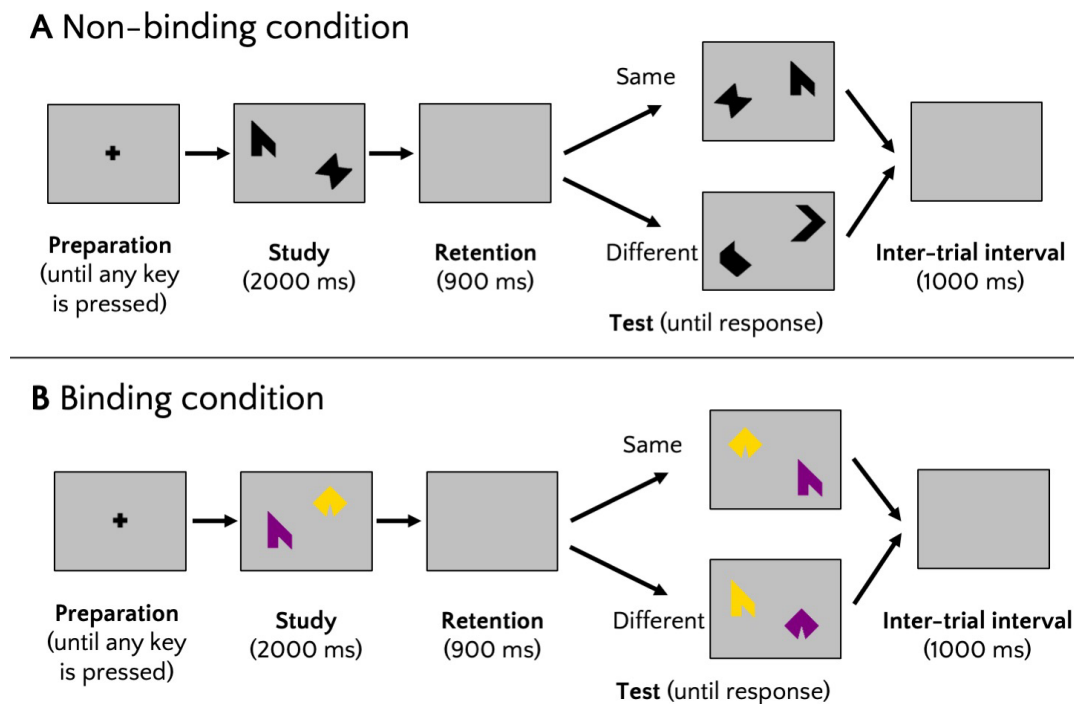


Figure 1. Traditional binding task (VSMBT), two-item version. **(A)** The non-binding condition, in which participants only need to memorise the shapes. **(B)** The binding condition, in which participants need to memorise shape-colour combinations. Note that in the ‘different’ trials, test set shapes are the same as in the study set but swap colours.

Between the two-item and three-item blocks of trials, participants additionally completed a brief perceptual task to check if their performance could be confounded by perceptual problems independent of memory. In each of the 10 trials of this perceptual task, participants were presented with two sets of three coloured shapes, which appeared simultaneously on-screen separated by a horizontal line. Participants were asked to indicate whether the two sets of shapes were the same or different, with five 'same' and five 'different' trials being ordered randomly across task. Scoring under 90% correct (less than 9 out of 10 trials) was considered indicative of perceptual binding difficulties, resulting in one participant being excluded from the study.

Memory Match. Memory Match is a novel gamified task with a visual search paradigm (Figure 2) that was developed in our lab and loosely based on the VSMBT. Memory Match begins after a set of self-paced instruction screens (Figure 2A). In each trial, participants need to memorise an array of symbols and then select the previously presented 'study symbols' from a 4x5 grid (Figure 2B). Like the traditional binding task, Memory Match incorporates 'non-binding' (symbol memory only) conditions as well as 'binding' conditions, where the colour and the symbol must both be remembered. Memory Match additionally includes two stimulus types (i.e., letters versus abstract shapes) and three difficulty levels with varying study set size (i.e., remembering 2, 3, or 4 symbols; see Figure 2C for an example of the stimuli used in each trial type). These manipulations served to increase the task's sensitivity at both the low and high end of performance and to improve gameplay experience through variety. In line with gamification principles, participants in Memory Match also receive immediate visual and sound feedback on their performance, collect points throughout the task, and are given a certain number of 'lives' per trial that they can lose in case of a mistake. The smallest study set size trials (i.e., 2 symbols to remember) are presented first, and the set size increases sequentially throughout the game. As more stimuli are to be remembered in the 3- and 4-set rounds, more lives are granted in these levels (N lives = N targets). Performance on this task is measured as the proportion of correctly memorised symbols, given the number of guesses the participant has made (i.e., correct responses/correct responses + errors). Given that the search

grid remains at 4x5 even when the study set size increases to 4, random responding leads to different accuracy scores on each level. Specifically, chance-level accuracy is .07, .12, and .17 for the 2-, 3- and 4-set size levels, respectively. No symbol or colour is repeated within any given study array, but they do repeat within the test displays of the binding condition, as 'lures' — i.e., symbols that match the target in one dimension (colour or shape) but not both. The trial ends when participants select all the correct symbols or lose all their 'lives' for the given trial, whichever comes first. There are 2 repeats of every trial type (2 binding conditions x 2 stimulus types x 3 study set sizes) totalling 24 trials per participant.

Besides the Risk Factors challenge, Memory Match was also available in a modified 'Free Play' version. In this version of the task within a separate section of the app, the participants could individually select the level of difficulty they want to play (Easy, Medium, or Hard). The task setup within each difficulty level was otherwise identical. Unlike in the Risk Factors version, the self-paced task instructions (Figure 2A) were not presented automatically at the start of each gameplay, but they were available for re-visiting individually within the 'Free Play' section.

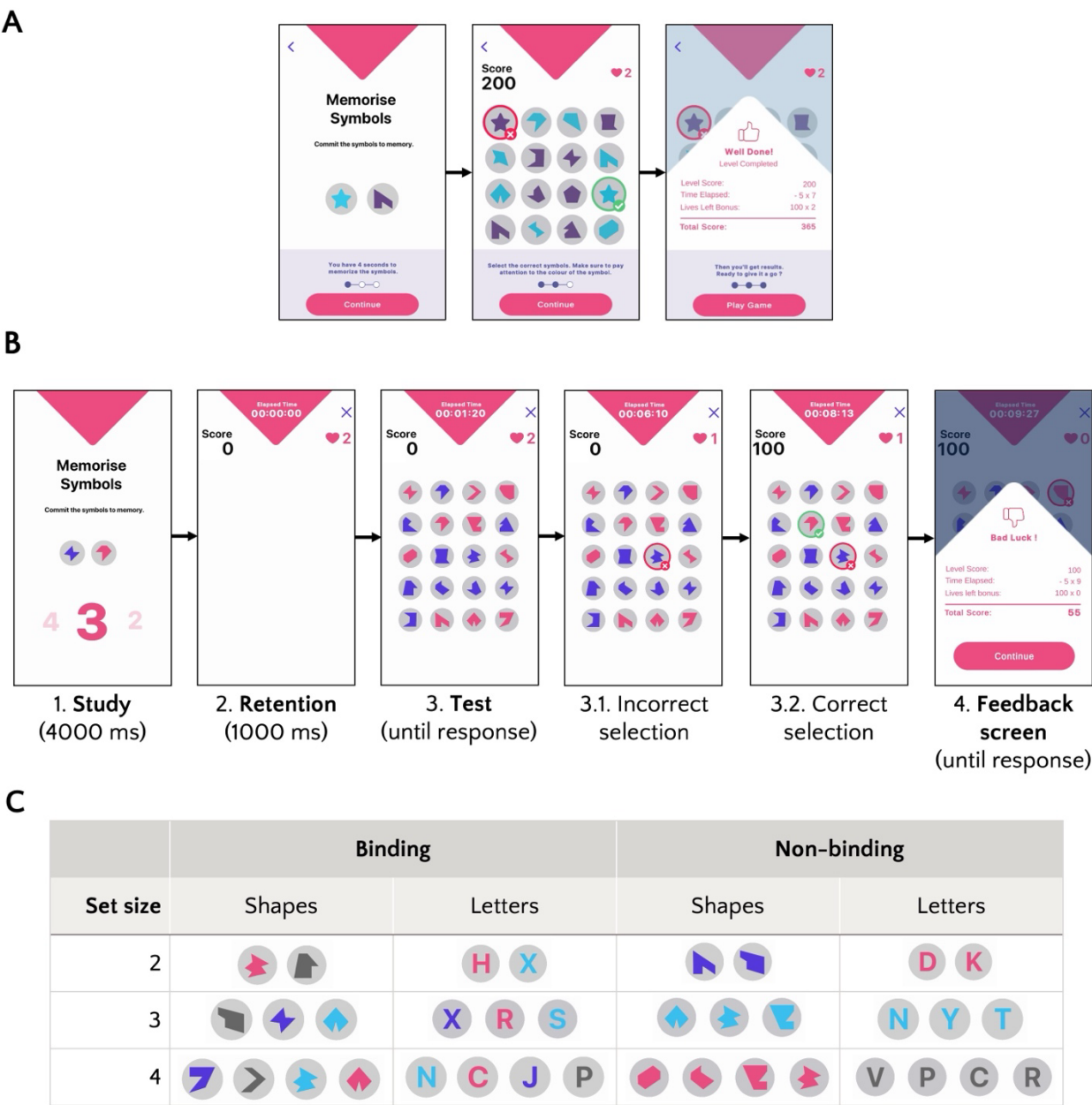


Figure 2. Memory Match overview. **(A)** Self-paced task instruction screens that appear at the start of Memory Match within the Risk Factors challenge. **(B)** An example trial of Memory Match. The presentation of the study array **(1.)** is followed by a retention interval **(2.)** until the test display **(3.)** fully loads. Participants can lose lives and points by making incorrect selections **(3.1.)** or earn points by making correct selections **(3.2.)**. Each trial is concluded by a feedback screen **(4.)**. **(C)** Examples of stimuli used in different trial types of Memory Match.

2.2.4 Analyses

All analyses were conducted using R Statistical Software v 4.3.1. (R Core Team, 2023a). Code to reproduce the findings and figures is available at <https://osf.io/vq6fa/>, alongside more details on how to request data.

Validity. First, we assessed the new task's *construct validity*. To do this, we ran an ANOVA with Tukey's post-hoc tests across both samples (i.e. in-lab and citizen scientist) to compare the task's conditions among themselves and established the extent to which the task behaves similarly to VSMBT. Second, in the in-lab validation study sample, we estimated *convergent validity* by calculating correlations of the new task and its different conditions with the traditional task equivalent. Here, the magnitude of Pearson product-moment correlations was understood as evidence for convergent validity, where correlations of $<.30$ were understood as low, $.31-.49$ as moderate, and $\geq .50$ as large in line with common effect size guidelines (Cohen, 1992). While there is no single cutoff for establishing convergent validity, some authors recognize $r = .30$ as a minimum acceptable coefficient at an early point of test development (Nunnally & Bernstein, 1994; Strauss et al., 2006). Complementing this analysis, we assessed *discriminant validity* by correlating the new task with other executive function measures, namely processing speed (Trails A) and cognitive flexibility (Trails B), and their gamified equivalents (Star Racer A and B). Finally, to assess the new task's ceiling effects, we used a paired t-test to compare mean performance and a Levene's test to compare variability of both the old and new task in the in-lab sample.

Internal consistency. Split-half reliability of Memory Match was calculated using the Pearson correlation across odd and even trials in the large citizen scientist sample. For comparison, we also calculated split-half reliability of the VSMBT from the in-lab validation study sample. Due to the VSMBT version we used, we did not have access to trial level data (only block-level), and so split-half reliability was calculated as the correlation of mean accuracy across the trials with 2 stimuli to remember vs. those with 3 stimuli to remember. To adjust for test length effects, Spearman-Brown formula was applied to the resultant correlations. The magnitude

of split-half reliability coefficients was evaluated against the guidelines by Strauss et al. (2006, p. 14), where reliability $<.59$ was understood as low, $.6-.69$ marginal, $.7-.79$ adequate, $.8-.89$ high, and $>.9$ very high.

Test-retest reliability. We calculated test-retest reliability as intraclass correlation between overall performance on Memory Match in the Risk Factors section and the equivalent subset of Free Play data. We only included participants who completed the Risk Factors and the Free Play versions of the game within 30 days from each other. The median distance between assessments was 1 day. As the Free Play version of the game had fewer trials compared to the full Risk Factors version and participants could complete the various difficulty levels in whatever order they preferred, we only used data from participants who completed trials in a number and difficulty that was equivalent to the Risk Factors version of the game, and only if they completed all these equivalent Free Play trials within the same day. Overall accuracy (i.e., mean proportion correct) of each participant was then calculated by averaging the proportion correct from their first attempt at each difficulty level in the Free Play section. We calculated the intraclass correlation coefficient (ICC) using a single-measurement, absolute-agreement, two-way mixed effects model (Koo & Li, 2016). In line with Koo & Li (2016), reliability $<.5$ was understood as poor, $.5-.75$ moderate, $.75-.9$ good, and $>.9$ excellent.

2.3 Results

The median time to complete the full task (i.e., all conditions, all 3 difficulty levels) was 7.6 minutes (M (SD) = 9.4 (21.5)). In citizen scientists, the overall accuracy on Memory Match decreased as memory load (i.e., number of memorised items) increased ($F(2) = 7729.17$, $p < .001$, $\eta^2 = .15$; Figure 3A). Performance on Memory Match also differed by condition ($F(1) = 1476.38$, $p < .001$, $\eta^2 = .02$; Figure 3B) and trial type ($F(1) = 52677.82$, $p < .001$, $\eta^2 = .38$), whereby it decreased as task conditions became more complex (i.e., binding < non-binding, abstract shapes < nameable letters; all $p < .001$). We also found a significant interaction of trial type (i.e., shapes/letters) and condition (i.e., binding/non-binding) in citizen scientists ($F(1) = 2936.44$, $p < .001$, $\eta^2 = .03$), whereby binding performance (i.e., remembering objects together with colours) was worse than non-binding performance (i.e., remembering objects alone) only on trials that assessed memory for letters, but not on trials that assessed memory for abstract shapes (Figure 3C), suggesting that the overall effect of condition was driven by the ‘letters’ trials. We found a comparable pattern of results in the smaller, in-lab sample (Figure 3A–C).

In the in-lab sample, the correlation between overall accuracy on Memory Match and VSMBT was $r(31)$ [95% CI] = .39 [.06; .65]; $p = .023$ (Figure 3D), and numerically lower for the binding condition — $r(31)$ [95% CI] = .38 [.05; .64]; $p = .028$ (Figure 3E). The correlation between the non-binding conditions of both tasks was not significant, $r(31)$ [95% CI] = .27 [-.08; .56]; $p = .13$ (Figure 3E). Overall accuracy on Memory Match was not significantly related to processing speed ($r(29)$ [95% CI] = -.09 [-.43; .27]; $p = .635$ for Trails A and $r(29)$ [95% CI] = -.29 [-.59; .07]; $p = .109$ for Star Racer A) or to cognitive flexibility ($r(29)$ [95% CI] = -.15 [-.48; .21]; $p = .412$ for Trails B and $r(29)$ [95% CI] = -.22 [-.53; .15]; $p = .241$ for Star Racer B). The mean overall performance was significantly lower ($t(32) = 7.77$; $p < .001$) and variability higher ($F(1, 64) = 16.38$; $p < .001$) on Memory Match ($M = .83 \pm 0.08$; skew = -0.28) than on VSMBT ($M = .94 \pm .04$; skew = -1.2), suggesting reduced ceiling effects (see also ranges of scores in Figure 3D, 3E).

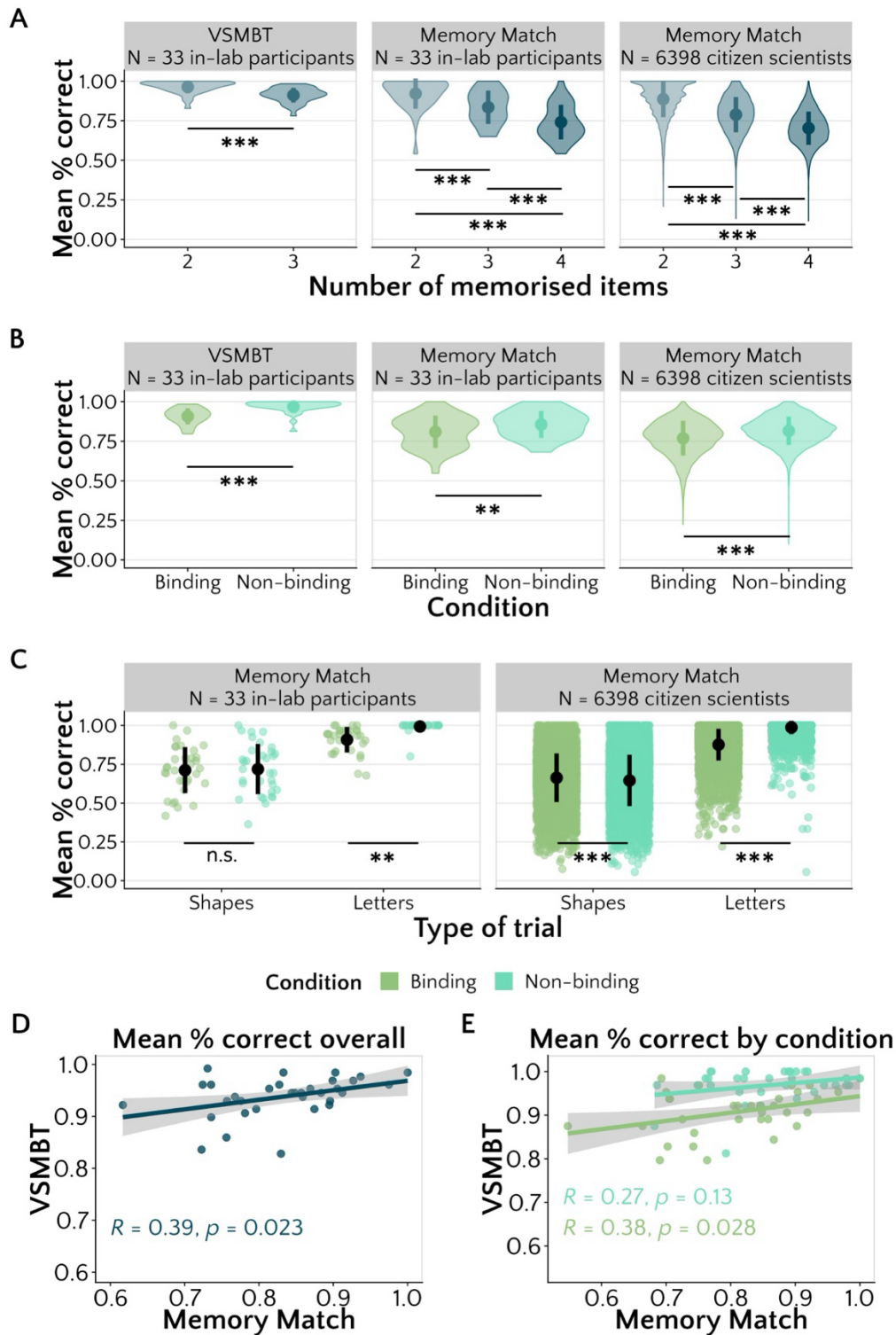


Figure 3. (A) Mean percentage correct in VSMBT and Memory Match, split by difficulty level (i.e., number of memorised items). (B) Mean percentage correct in VSMBT and Memory Match, split by condition (i.e., binding/non-binding). (C) Mean percentage correct in VSMBT and Memory Match, split by condition and trial type (i.e., shapes/letters). (D) Correlation of mean percentage correct (i.e., overall accuracy) in VSMBT and in Memory Match. (E) Correlation of mean percentage correct in VSMBT and in Memory Match, split by condition.

The split-half reliability of overall Memory Match accuracy was assessed in the larger sample who completed the game once, giving $r(6396) = .64$; $p < .001$ ($r = .78$ after adjusting for test-length effects using the Spearman-Brown formula), suggesting adequate internal consistency. These reliability estimates are comparable to the traditional version of the task where the correlation across the two levels of the task was $r(31) = .60$ ($r = .75$ after adjusting for test-length effects using the Spearman-Brown formula). The test-retest reliability of overall accuracy was assessed in those citizen scientists who played the game more than once and we found moderate reliability, $ICC(A,1) [95\% CI] = 0.59 [0.44; 0.70]$ ($p < .001$).

2.4 Discussion

In this study, we assessed the psychometric qualities of a novel gamified cognitive task called Memory Match, which aimed to assess visual working memory binding. We found that as expected, the overall accuracy on Memory Match decreased as memory load increased and as task conditions became more complex, i.e., participants were better at remembering abstract shapes than letters and at remembering symbols alone compared to coloured symbols, bringing evidence for construct validity of the task as a visual working memory measure. Despite its shorter length (24 as compared to 128 trials), Memory Match showed adequate split-half reliability comparable to that of the original task, and moderate test-retest reliability on repeated assessments. Importantly, we were successful in achieving lower ceiling effects compared to the traditional task, as suggested by higher score variability and lower average performance. Given this important enhancement to the task together with some changes to the task paradigm aimed to increase engagement, we did not expect perfect cross-task convergence. Indeed, the correlation between performance on Memory Match and VSMBT was low to moderate, although significant ($r = .39$), and even lower for individual task conditions. In line with existing recommendations for assessing task validity (Nunnally & Bernstein, 1994; Post, 2016; Strauss et al., 2006), a correlation of .30 and higher can be seen as initial but inconclusive evidence for convergent validity, calling for further task refinement, especially if the task was to be used in a diagnostic context. Importantly though, Memory Match performance was not

significantly associated to other executive function measures with a strong visual component (i.e., processing speed and cognitive flexibility), whether assessed through a pen-and-paper test or through a smartphone app, bringing evidence for discriminant validity.

One point of departure across the traditional and this new smartphone test concerned the ‘binding effect’. The binding condition is intended to be harder than the non-binding one even in healthy adults, as it requires individuals to remember more object features (Brockmole et al., 2008). While we replicated an overall binding effect (i.e., mean accuracy was lower on binding vs. non-binding trials) on aggregated performance, we noted that this was driven by the ‘letters’ trials in Memory Match. We suspect this is a feature of the use of a grid-search design instead of the traditional task’s forced-choice paradigm where participants have to categorise presented symbols as ‘same/different’. Because the target shapes are particularly difficult to identify within a grid of many other abstract shapes, on binding trials, it is possible participants utilised colour information to narrow down the number of shapes they had to consider and boost performance that way (see Figure 4 for an illustration). For these reasons, our new task cannot be used to assess the relative impairment on conjunctive feature binding compared to single-feature memory, unless we limit our assessment to the ‘letters’ condition, which, in turn, is likely not going to be sensitive to slight impairments due to its low difficulty. However, there is evidence that both conditions of the traditional task are valuable. A meta-analysis by Cecchini et al. (2023) reported that performance on both the binding *and* non-binding condition differed significantly between controls and individuals on the AD continuum, although this effect was stronger for the binding condition. Furthermore, the non-binding condition was sensitive to increased levels of amyloid β in one of three studies that assessed the task’s sensitivity to AD biomarkers (D. J. Norton et al., 2020; cf. Cecchini et al., 2021; Parra et al., 2024). Finally, another study using a *relational* binding paradigm found evidence that non-binding memory for objects was sensitive to amyloid β but object-location binding was not (Lu et al., 2021). The exact mechanisms underlying working memory binding are still debated, and as emphasised in the detailed review by Schneegans & Bays (2019), binding deficits might be driven by increased memory *load* on the binding condition (i.e., remembering more individual features) as

opposed to the binding aspect per se (i.e., the involvement of a separate mechanism for remembering integrated objects).

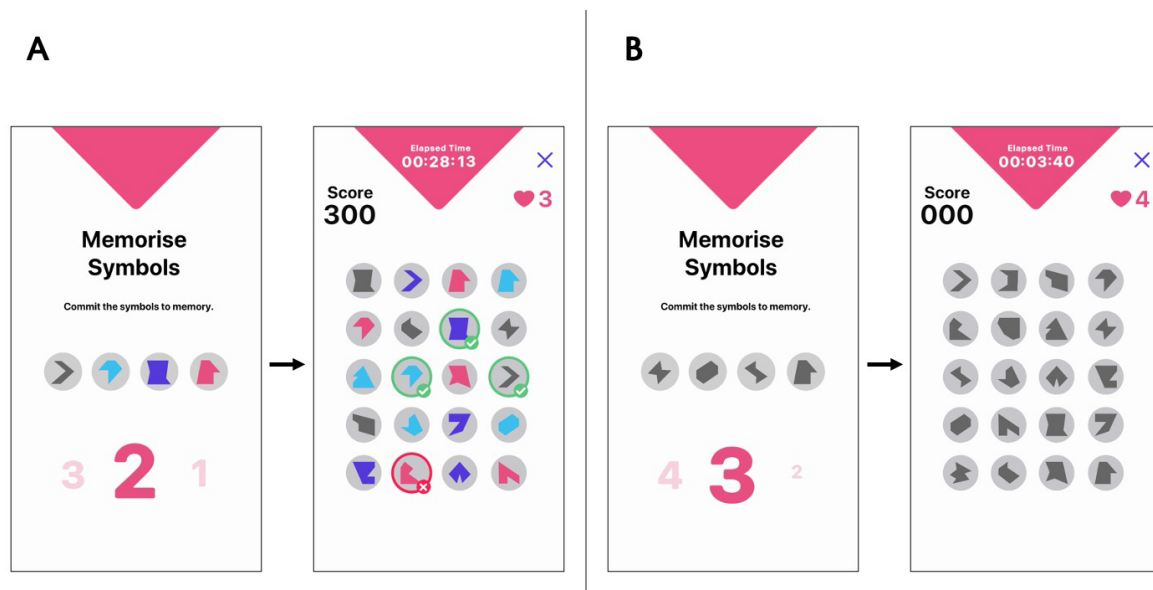


Figure 4. Comparison of the binding (A) and non-binding (B) condition of a shapes trial. The colour in the binding condition can serve as guidance for participants, whereby in comparison to the non-binding task, they have a reduced set of shapes of each colour to choose from.

Despite the limited binding effect in Memory Match, the task could be used as a measure of overall visual working memory performance. This is supported by a correlation between overall performance on this and the original working memory task irrespective of condition and by the task's psychometric qualities, which were adequate despite it being much shorter than its traditional counterpart. This, together with the task's gamified nature, could make it particularly valuable for large-scale, unsupervised remote studies where maintaining participant motivation can be a problem. However, further research needs to extend the current study and assess external validity of the task, e.g. by evaluating the effects of age on task performance, or further assess its convergent and discriminant validity by comparing it to other established working memory measures, such as the N-back task or the digit span task, as well as executive function measures more generally. Preferably, this should be done in a larger sample, which would allow to more precisely differentiate how strongly the task relates to various measures.

There were several limitations to the current study. First, although working memory binding tasks are promising for early detection, with the current study design we did not assess the task's ability to predict who will go on to develop AD or cognitive decline more generally, as we did not test cognitive changes over time and only included healthy individuals. Second, although the task aims to assess visual working memory, participants could have used various memorising strategies, with some of them involving verbal rehearsal. In line with this and as expected, we found an overall better performance on trials that required participants to remember letters compared to trials with abstract shapes. However, anecdotal evidence from some of our research participants suggested they were also naming the abstract shapes based on their similarity to real-life objects. This could possibly have resulted in improved performance in some, but not all participants. Although such effects cannot be precluded even in the case of the traditional task (Forsberg et al., 2019), future studies could explicitly test the contribution of verbal rehearsal to performance by explicitly instructing participants to use such strategies or asking them about their strategies at the end of the study. Third, the new task's test-retest reliability was assessed using an approximate equivalent of the full task version, i.e. by accessing data of individuals who selected to play all three task levels in the Free Play section of the app. This could have led to a different order of administration of the three difficulty levels for the retest and the retest data was also more likely to include participants who did not complete all three difficulty levels consecutively, although the median interval was 1 day. Ideally, this could be remedied in the future by controlled consecutive administration of the full task version, possibly providing more accurate task reliability estimates. Finally, like its traditional version, this task too relies on colour vision, as it requires participants to remember coloured shapes. In the current study, we excluded colour-blind participants, and the colour scheme used in the task is colour-blind friendly in the sense that the colours should seem as distinct for individuals with the most common form of colour-blindness, deuteranopia. However, future research should still exhibit caution if administering the task to potentially colour-blind participants. Despite these limitations, the task demonstrated promising psychometric qualities as a measure of visual working memory and can be used to facilitate and scale up research in the general population.

Chapter 3: Modifiable dementia risk factors associated with objective and subjective cognition

3.1 Introduction

Cognitive impairments are linked to important brain health and well-being outcomes, and thus their early detection and prevention have become a major public health concern. One key area of focus has been on the role of potentially modifiable social, health, and lifestyle factors in helping to maintain cognitive abilities as our populations age (Livingston et al., 2017, 2020). Cognitive impairments can be objectively measured through standardised cognitive tests, but also subjectively perceived, typically in the form of subjective memory complaints. Understanding how to preserve both objective and subjective cognition is important, as each have been shown to have a major impact on wellbeing (Yates et al., 2017), daily functioning (Farias et al., 2006), and also serious brain health outcomes like dementia (Jessen, 2014). For example, a decline in performance on objective tests of memory and executive function can precede dementia onset by up to a decade (Mortamais et al., 2017). Likewise, even before cognitive impairments can be objectively measured, subjective perceptions can reflect the earliest stages of cognitive decline (Jessen, 2014), and objectively unimpaired older adults who have subjective cognitive complaints are twice as likely to develop dementia than those who do not (Mitchell et al., 2014). However, neither are fully deterministic — not everyone with subjective cognitive complaints will eventually develop an objective form of cognitive decline (Mitchell et al., 2014) and more advanced objective cognitive impairment is typically accompanied by a lack of subjective insight into one's cognitive capacity, rendering subjective assessments less reliable (Cacciamani et al., 2021). Thus, although objective and subjective cognitive assessments overlap considerably in their functional significance, they have a surprisingly low correlation on the individual level (Burmester et al., 2016), which can additionally depend on age, gender, and culture (Walkowiak et al., 2023). This low correlation may be because they each reflect different stages of cognitive decline; subjective cognitive complaints could represent the earliest stages of cognitive decline that anticipate more consolidated, objective deficits (Cacciamani et al., 2021; Jessen, 2014). By

the time objective deficits emerge, a deterioration in metacognitive processes or other comorbidities like depression might affect the accuracy of subjective assessments (Cherry et al., 2019; Rouault et al., 2018). Another possibility is that cross-sectional objective assessments cannot distinguish low ‘premorbid’ ability from cognitive decline (Jessen, 2014), and so subjective cognitive complaints may be more sensitive because they reflect within-person cognitive changes that would go undetected by objective assessments administered at one moment in time (M. Martin & Zimprich, 2003).

With the importance of these different facets of cognition in mind, research has established robust associations between both objective and subjective cognition and a variety of potentially modifiable risk factors for dementia. Diabetes (Broadley et al., 2017), history of stroke (Lugtmeijer et al., 2021), hearing loss (Loughrey et al., 2018), tinnitus (Jafari et al., 2019), depression (D. Gallagher et al., 2016), social engagement (Boss et al., 2015), hypertension (Veldsman et al., 2020), physical inactivity (Colcombe & Kramer, 2003), smoking (Conti et al., 2019), socioeconomic status (SES) (Muhammad et al., 2022), and educational attainment (Brockmole & Logie, 2013) have all been linked to differences in performance on objective tests of executive function. Similarly, subjective cognitive complaints have also been linked to diabetes (Hill et al., 2021; Johar et al., 2023), history of stroke (van Rijsbergen et al., 2014), hearing loss (Borelli et al., 2022), tinnitus (Trevis et al., 2016), depression (Hill et al., 2016), and reduced social engagement (Cheng et al., 2023; P.-L. Lee et al., 2013). However, the evidence is more mixed for other risk factors like hypertension (Hill et al., 2021), physical inactivity (Borelli et al., 2022; P.-L. Lee et al., 2013; Mascherek et al., 2020; Wen et al., 2021), smoking (Mascherek et al., 2020; Paradise et al., 2011), SES (Simons et al., 2023; Zullo et al., 2021), and educational attainment (Borelli et al., 2022; Cheng et al., 2023; Zullo et al., 2021). Crucially, it is currently unclear how these risk factors differentially impact objective versus subjective cognition, if certain aspects of objective cognition (e.g., working memory, cognitive flexibility, or planning) are more or less affected by dementia risk factors than subjective cognition, and how these relationships change across the lifespan (Livingston et al., 2017, 2020).

To address this, we gathered a large multivariable, cross-sectional dataset from thousands of ‘citizen scientists’ participating in research via a smartphone app,

Neureka. Smartphone-based remote assessments such as these are increasingly popular tools for research and a promising way to detect both subjective and objective cognitive impairments, as they can easily combine short screening of subjective cognitive complaints with brief, gamified objective assessments of relevant aspects of cognition (Donegan et al., 2023; Ojakäär & Koychev, 2021; Walkowiak et al., 2023). A growing body of research shows good validity and acceptability of smartphone assessments in the field of cognitive decline (see Whelan et al., 2022 for a review). The benefits of smartphone-based assessment include low cost, low burden for examiners and examinees, scalability, accessibility, and ease of repeated testing (Van Patten, 2021). Here, we leveraged these features to gather rich data from a large, dementia-free sample spanning the entire adult age range. From each participant, we acquired information about previously established dementia risk factors, gamified objective assessments of cognition, and subjective cognitive complaints. Given the importance of objective executive function in daily functioning (Farias et al., 2006) and early detection of dementia (Mortamais et al., 2017), we focused on multiple aspects of executive function, namely working memory, cognitive flexibility, and model-based planning. Secondary analyses also tested if the associations between risk factors and cognition might be explained by individual differences in depression symptom severity (Johar et al., 2023), one of the best-established correlates of subjective complaints (Hill et al., 2016).

3.2 Methods

3.2.1 Participants

Between June 2020 and July 2023, $N = 9,918$ members of the public engaged with a module called ‘Risk Factors’ in the smartphone app *Neureka*, which contains a battery of self-report questionnaires and gamified cognitive assessments. 66.31% of these participants had missing data in at least one part of the battery. In the current study, we only used data from the first fully completed ‘Risk Factors’ attempt linked to each participant’s app account, resulting in a total sample of $N = 3,341$ full completers. The full completers did not significantly differ from the rest of

the participants either in age ($M_{\text{complete}} = 45.69$, $M_{\text{incomplete}} = 45.62$, $t(6845) = 0.23$, $p = .819$) or in the prevalence of self-reported memory problems ($X^2(1, N = 9,918) = 0.01$, $p = .998$). We further excluded participants with self-reported diagnosis of dementia ($n = 6$) or those who preferred not to state their gender ($n = 8$) as the sample size for those groups was too small to draw any meaningful conclusions, resulting in the final $N = 3,327$ participants.

The demographic characteristics of the final sample are summarised in Table 1. The age distribution split by gender is presented in Figure 5, and a histogram showing age distribution of the final sample is available in Appendix 8.1.1, Figures S1. Note that cisgender men were on average older ($M \pm SD = 47.83 \pm 14.71$) than cisgender women ($M \pm SD = 45.15 \pm 13.73$), with $t(2347.6) = -5.1424$ ($p < .001$). This was driven by the relatively higher proportion of women at the lower end of the age distribution, as seen in Figure 5. Transgender/non-binary participants were on average younger ($M \pm SD = 31.01 \pm 12.29$) in comparison to cisgender women, with $t(77.72) = 10.87$ ($p < .001$).

Participants came from 54 different countries, whereby the most frequent countries were United Kingdom ($n = 1,650$), United States ($n = 936$), Ireland ($n = 443$), Canada ($n = 57$), and Germany ($n = 46$). Notably, individuals with lower educational attainment appeared to be underrepresented in the current sample, based on comparisons with country-level OECD data for the five most frequent countries (see Appendix 8.1.2 for a full analysis of sample representativeness with respect to SES). Of the final sample, 2,986 (89.8%) participants reported having English as their first language (see Appendix 8.1.3 for sensitivity analyses with this reduced sample of native speakers).

The *Neureka* project was approved by the research ethics committee of the School of Psychology, Trinity College Dublin (approval number: SPREC072019-01). Data collected through *Neureka* is stored and processed in line with the EU General Data Protection Regulations.

Table 1. Demographic characteristics of study participants (N = 3,327).

Variable	Details	Descriptive statistics
Age	M ($\pm$ SD)	45.74 ($\pm$ 14.55)
	Range	18–84
Gender	Cisgender female	2160 (64.9%)
	Cisgender male	1099 (33.0%)
	Non-cisgender*	68 (2.0%)
Education	No formal education	135 (4.1%)
	Lower secondary	639 (19.2%)
	Upper secondary	1376 (41.4%)
	University/College degree or equivalent	894 (26.9%)
	Master's degree	243 (7.3%)
	PhD	40 (1.2%)

NOTE. *The 'non-cisgender' category includes participants who stated their gender is 'transgender male', 'transgender female', or 'non-binary'. Participants who preferred not to state their gender (i.e., 0.4% of all sign-ups) were not included in the current study.

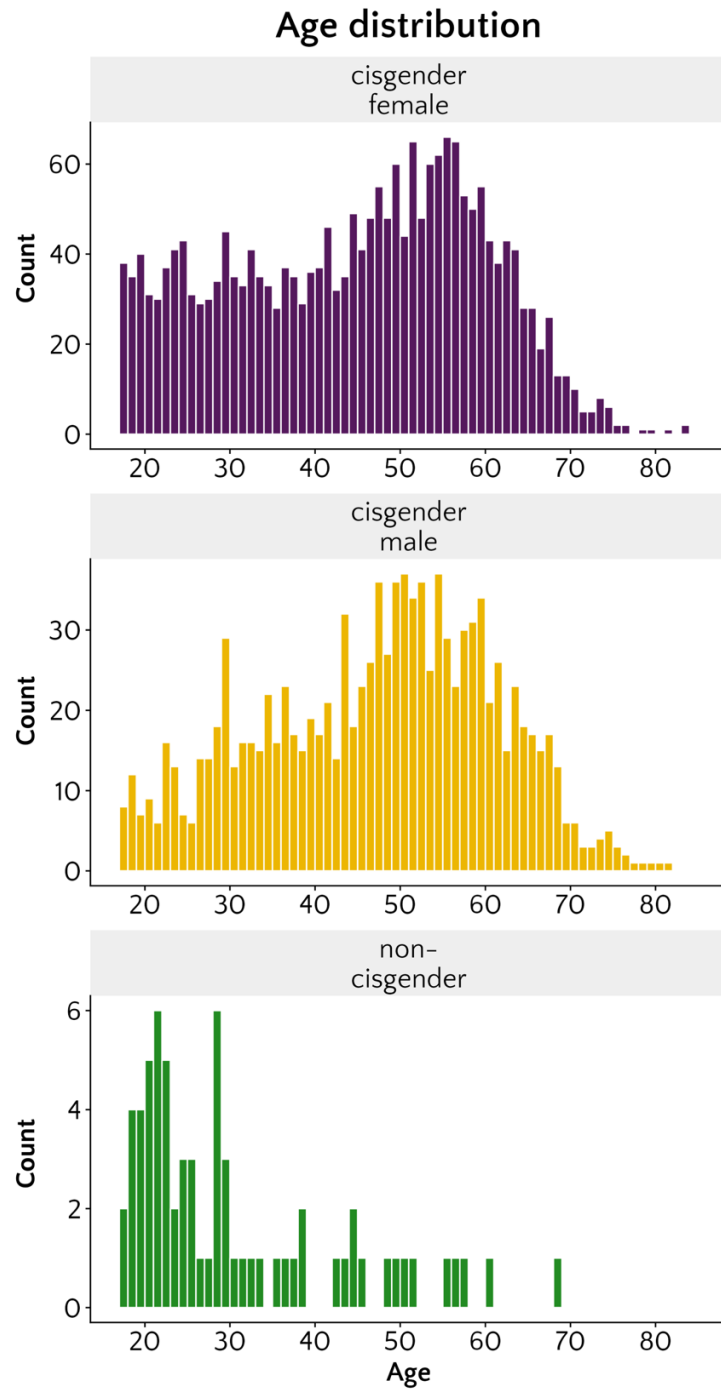


Figure 5. Age histogram showing the number of participants per years of age in the final sample, split by gender, where $n = 2160$ for cisgender women, $n = 1099$ for cisgender men and $n = 68$ for ‘non-cisgender’ participants. The ‘non-cisgender’ category includes participants who stated their gender is ‘transgender male’, ‘transgender female’, or ‘non-binary’.

3.2.2 Procedure

The *Neureka* app was developed in-house by our research team at Trinity College Dublin, Ireland, with the aim to facilitate large-scale online research of brain health. Prior to its public release, the app has been focus-grouped, piloted, validated, and iteratively improved in smaller paid research studies, and its validation continues as new modules are added to the app on an ongoing basis. The app is currently free to download from Google Play Store (Android users) and App Store (iOS users). Participants are recruited through a variety of methods, including radio and television publicity, ads on the Google Play Store, and through organised efforts such as the SciStarter project. Some participants are recruited as a part of smaller paid research studies at Trinity College Dublin, Ireland. For more detail on the recruitment methods, please see Appendix 8.1.4. All users of the *Neureka* app read a common information sheet and provide digital consent to participate in brain health and cognitive neuroscience research. During registration, participants select their age, gender, education, and country. Participants must be 18 years of age or older to participate in research on the app. After registration, participants can complete a variety of 'Science Challenge' modules in the app at their own leisure, without financial compensation. Data for this study come from the Risk Factors challenge, a module in the app that focuses on assessing dementia-related risk factors and cognition. It includes three gamified cognitive assessments that assess objective cognition (memory, cognitive flexibility, and model-based planning; see Chapter 2 of this thesis and Appendix 8.1.5 for more details on task validation) and two sets of questionnaires that assess subjective cognition together with a set of potentially modifiable risk factors for dementia (more detail provided in the following sections). The three gamified cognitive assessments and two sets of questionnaires are presented to participants in a semi-randomised order, whereby a game is always delivered first, followed by a set of questionnaires, followed by a second game, followed by another set of questionnaires, and ended with one final game. The module takes approximately 45 minutes to complete in full and can be completed in one sitting or with breaks, over an extended period of time.

3.2.3 Measures

Visual working memory. Visual working memory was assessed using a visual search game called 'Memory Match' (Figure 6A), which was loosely based on the Visual Short-Term Memory Binding Task (VSMBT; Parra et al., 2010). See Chapter 2 for a validation study comparing the gamified task to a traditional binding task. This game starts after a set of self-paced instruction screens (Figure 2A). In each trial, participants need to memorise an array of symbols and then select the previously presented 'study symbols' from a 4x5 grid (Figure 2B). Like the traditional binding task, this task design incorporates 'non-binding' (symbol memory only) conditions as well as 'binding' conditions, where the colour and the symbol must both be remembered. Memory Match additionally includes two stimulus types (i.e., letters versus abstract shapes) and three difficulty levels with varying study set size (i.e., remembering 2, 3, or 4 symbols; see Figure 2C for an example of the stimuli used in each trial type). These manipulations served to increase the task's sensitivity at both the low and high end of performance and to improve gameplay experience through variety. In line with gamification principles, participants in Memory Match also receive immediate visual and sound feedback on their performance, collect points throughout the task, and are given a certain number of 'lives' per trial that they can lose in case of a mistake. The smallest study set size trials (i.e., 2 symbols to remember) are presented first, and the set size increases sequentially throughout the game. As more stimuli are to be remembered in the 3- and 4-set rounds, more lives are granted in these levels (N lives = N targets). Performance on this task is measured as the proportion of correctly memorised symbols, given the number of guesses the participant has made (i.e., $\text{correct responses} / (\text{correct responses} + \text{errors})$). Given that the search grid remains at 4x5 even when the study set size increases to 4, random responding leads to different accuracy scores on each level. Specifically, chance-level accuracy is .07, .12, and .17 for the 2-, 3- and 4-set size levels, respectively. No symbol or colour is repeated within any given study array, but they do repeat within the test displays of the binding condition, as 'lures' — i.e., symbols that match the target in one dimension (colour or shape) but not both. The trial ends when participants select all the correct symbols or lose all their 'lives' for the given

trial, whichever comes first. There are 2 repeats of every trial type (2 binding conditions x 2 stimulus types x 3 study set sizes) totalling 24 trials per participant. Mean accuracy at the task was 80% ($SD = 8\%$) in the total sample, with a range of 25%-100%.

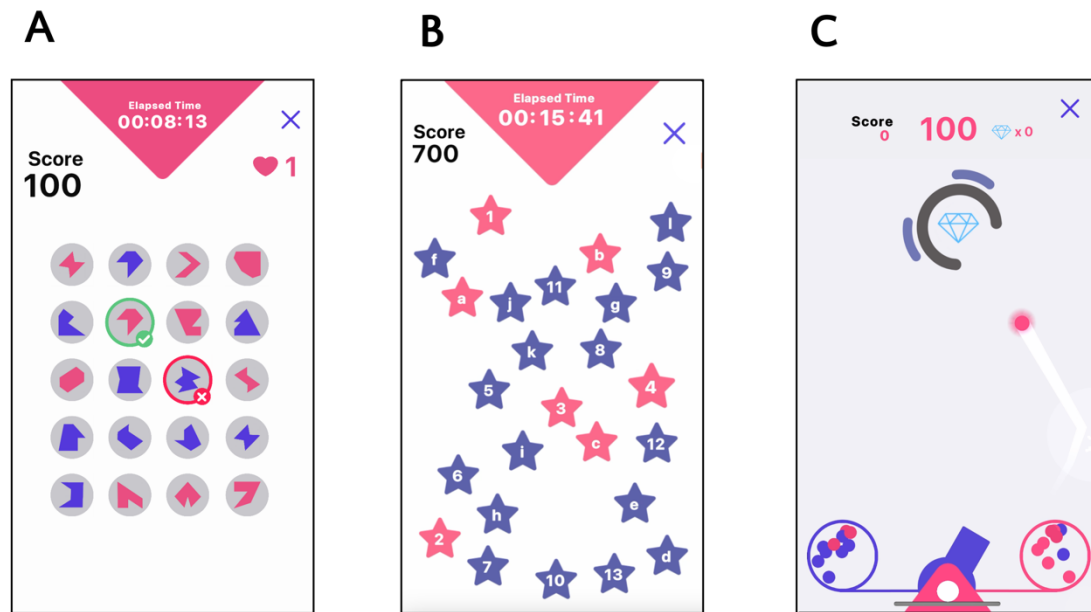


Figure 6. Screenshots from the gamified cognitive assessments used in the current study. **(A)** The game ‘Memory Match’ assesses visual working memory. Participants must recognise a set of previously presented shapes or letters and selected them from a grid. **(B)** The game ‘Star Racer’ assesses cognitive flexibility. Participants must tap stars in ascending order, alternating between numbers (1–13) and letters (a–l). **(C)** The game ‘Cannon Blast’ assesses model-based planning. Participants must pick one of two ball containers, aim a cannon, and shoot balls to collect diamonds, but their success relies on their ability to adjust to changing probabilities and picking the right container.

Cognitive flexibility. Cognitive flexibility was assessed through the game ‘Star Racer’ (Figure 6B), which is a gamified and extended version of the Trail-Making Test (TMT; Bowie & Harvey, 2006). See Appendix 8.1.5.1 for a validation study comparing it to the traditional TMT. Like the traditional TMT, Star Racer is divided into forms A and B, but contains multiple ‘runs’ of each form. The task begins with a set of self-paced tutorial screens and practice runs of each version (see Appendix 8.1.5.1, Figure S5A), which are then followed by 3 runs of each version (always ordered A-B-A-B-A-B; see Appendix 8.1.5.1, Figure S5B). In the A version of the game, participants are presented with a screen of 25 blue stars with

numbers and required to tap them in ascending order as quickly as possible. In the B version, the screen contains 25 stars labelled with either letters or numbers and participants must alternate between them as they ascend. Each run starts with a three-second countdown. Participants gather points throughout the task for each correct selection (indicated by the star turning pink) and lose points for elapsed time and an incorrect selection (indicated by the star turning red, shaking, and then reverting to blue again), with their score displayed on screen throughout and summarised at the end of each run. Participants cannot progress unless they correct any erroneous selections. Performance is calculated for versions A and B separately as mean time to complete all runs of the given type, resulting in two separate outcome variables (i.e., A: processing speed and B: cognitive flexibility). However, these two variables are strongly correlated (here, $r(3325) [95\%CI] = .77 [.75; .78], p < .001$). In the current study, we focus on cognitive flexibility as a measure of executive function. We did not use a B – A difference score to ensure adequate measure reliability, which tends to be problematic for difference scores (Kopp, 2011). Additionally, to identify inattentive players in this remotely administered task, we applied conservative time cut-offs: Individual runs with completions times of 100 or more seconds (for A version) or 300 or more seconds (for B version) were excluded; in such cases task performance was calculated as a mean time to complete the remaining trials. These cut-offs were based on approximately double the median completion times for the oldest and least educated group in a normative study of the traditional task (Tombaugh, 2004). The original mean time to complete Star Racer B was 64.31 seconds ($SD = 24.56$) with a range of 19.27–381.87 seconds, and after applying the cut-offs, this changed to the final 63.28 seconds ($SD = 21.73$) with a range of 19.27–204.37 seconds.

Model-based planning. Model-based (also known as ‘goal-directed’) planning was assessed through a game called ‘Cannon Blast’ (Figure 6C), which is a gamified version of the Two-Step Reinforcement Learning Task with a recent extensive validation (Daw et al., 2011a; Donegan et al., 2023). See Appendix 8.1.5.2 for a detailed description of the outcome measure and the task. In brief, the game requires participants aim and shoot balls from a cannon to hit diamonds, for which they earn reward points. Sometimes the balls are ‘bad’ and explode before

hitting the target and participants can use statistical features of the task to avoid receiving these balls. Specifically, they can keep a basic model-free policy that encourages them to repeat actions that lead to good balls and switch from actions that lead to bad ones. Alternatively, they can incorporate higher-order knowledge of the task structure to refine this as a model-based policy. The outcome measure of ‘Cannon Blast’ is the ‘model-based index’, which is an output of a hierarchical logistic regression analysis calculated across all 200 trials of the game available in Risk Factors (Donegan et al., 2023). In this analysis, we measure how much the choice to repeat an action of a given trial is influenced by (i) whether that choice resulted in a good ball on the last trial (model-free) and (ii) whether that is qualified by task structure. The mean model-based index value in the current sample was 0.27 ($SD = 0.33$) with a range of 0.46–2.22.

Subjective memory problems. Subjective memory problems were assessed with one questionnaire item within the Risk Factors challenge. Participants could respond ‘Yes’ or ‘No’ to the question ‘Do you currently experience memory problems?’. This single-item screener was the same as that used in the PREVENT study (K. Ritchie et al., 2017). In the current sample, 1,348 (40.5%) individuals reported subjective memory complaints.

Risk factors for dementia. A set of social, health and lifestyle factors, based on the risk factors for brain health identified previously (Jafari et al., 2019; Kuźma, Lourida, et al., 2018; Livingston et al., 2017; Scarabino et al., 2016; A.-Y. Wang et al., 2023), was assessed via questionnaires in the ‘Risk Factors’ challenge. Specifically, we measured educational attainment, SES, depression, loneliness, social network size, hearing handicap, tinnitus, stroke, diabetes, hypertension, smoking, exercise, and family history of dementia. Details on the risk factor measures are summarised in Table 2.

Table 2. *Description of risk factor measures.*

Measure	Description/question(s)	Descriptive statistics*	Reference
Education	0 = 'No formal education' 1 = 'Lower secondary education' 2 = 'Upper secondary education' 3 = 'University/College degree or equivalent' 4 = 'Master's degree or equivalent' 5 = 'PhD or equivalent'	M = 2.82 SD = 1.00 Range = 0–5	N/A
Socio-economic status (SES)	The Subjective SES scale: participants were asked to place themselves on a ladder representing where people stand (in the country they live) with respect to money, education, and respected jobs.	M = 5.95 SD = 1.77 Range = 1–10	(Adler et al., 2000)
Depression	Centre for Epidemiologic Studies Depression (CES-D) scale: a 20-item self-report scale for measuring depressive symptoms in the general population, which asks about the frequency of symptoms over the past week.	M = 17.96 SD = 13.59 Range = 0–60	(Radloff, 1977)
Loneliness	UCLA Loneliness Scale: a 20-item self-report scale for measuring general loneliness.	M = 21.10 SD = 15.54 Range = 0–60	(Russell et al., 1978)
Social network	Lubben Social Network Scale: a 6-item self-report measure of social engagement with family and friends	M = 15.96 SD = 6.46 Range = 0–30	(Lubben, 1988)
Hearing handicap	Hearing Handicap Inventory for the Elderly – Screening version (HHIE-S): a 10-item self-report scale measuring the emotional and social consequences of hearing impairment in adults.	M = 5.60 SD = 7.47 Range = 0–40	(Ventry & Weinstein, 1983)
Tinnitus	'In the past year have you had buzzing, ringing or noise (tinnitus) in your ears that lasts longer than 5 minutes?' (No/Yes, in the past week/Yes, not in the past week/Don't know)	Yes, in the past week/ Yes, not in the past week: 992 (29.8%) No/Don't know: 2335 (70.2%)	(Arnold et al., 2022)
History of stroke	'Have you ever had a stroke?' (Yes/No)	Yes: 45 (1.4%) No: 3282 (98.6%)	N/A
Diabetes	'Have you been diagnosed with diabetes by a doctor?' (Yes/No)	Yes: 174 (5.2%) No: 3153 (94.8%)	N/A
Hypertension	'Have you been diagnosed with high blood pressure (hypertension) by a doctor?' (Yes/No)	Yes: 511 (15.4%) No: 2816 (84.6%)	N/A

Smoking history	'Please indicate which of the following best describes you:' (1. Non-smoker/ 2. Ex-smoker/ 3. Current smoker/ 4. Unknown)	Current smoker/ Ex-smoker: 1134 (34.1%). Non-smoker/ Unknown: 2193 (65.9%)	Based on the PREVENT study (K. Ritchie et al., 2017)
Exercise	Godin-Shephard Leisure-Time Physical Activity Questionnaire: a self-report measure of physical activity, which asks about the frequency of strenuous, moderate, and mild exercise per week.	M = 32.70 SD = 25.48 Range = 0–119	(Godin, 2011)
Family history of dementia	'Was your biological mother ever diagnosed with dementia?' (Yes/No) 'Was your biological father ever diagnosed with dementia?' (Yes/No)	Yes (either parent or both): 560 (16.8%) No: 2780 (83.2%)	Based on the PREVENT study (K. Ritchie et al., 2017)

NOTE. **By design, there were no missing data for either measure in the total sample.*

3.2.4 Data Analysis

Data preparation. All analyses were conducted using R Statistical Software v 4.3.1 (R Core Team, 2023a). Code to reproduce the findings and figures is available at <https://osf.io/vq6fa/>, alongside more details on how to request data. Prior to analyses, variables were reverse-coded when necessary, so that in all cases higher values indicate worse cognition or higher risk. Furthermore, all continuous variables (both independent and dependent) were scaled by mean-centring and dividing by their standard deviation within each sample. To ensure the estimates for continuous and binary categorical predictors in our models are directly comparable to each other (see Gelman, 2008, for a discussion of the problem), we scaled binary variables (history of hypertension, smoking, stroke, diabetes, tinnitus, family history of dementia, or subjective memory problems) to have a standard deviation of 1 by coding them as 0/2. Gender was contrast-coded with 'cisgender male' as the reference category. We assessed the effects of age and gender on each objective cognitive measure using a series of linear regressions and on subjective memory problems using a logistic regression. For age, we included both a linear and a quadratic term (i.e., age squared) to capture potential nonlinear effects of chronological age across the wide age span in our sample. Effect sizes were reported as standardised betas and odds ratios with

associated confidence intervals. Associations between cognitive outcomes (i.e., visual working memory, cognitive flexibility, model-based planning, and subjective memory problems) were assessed using point-biserial correlation with a 2-step estimate for pairs of numeric and ordinal variables and Pearson product-moment correlation for continuous variables.

Primary analysis. To test the differential cognitive correlates of previously established risk factors for dementia, we ran a series of linear, respective logistic regressions, predicting objective, respective subjective cognition from each risk factor individually, controlling for gender and age. We calculated standardised betas, respective odds ratios with appropriate confidence intervals for each association. Because of their different interpretation (increase in y per SD change in x vs. increase in log odds of y per SD change in x), standardised betas in linear and logistic regression are not directly comparable. Therefore, to be also able to directly compare the magnitude of the risk factor effects on the continuous and categorical dependent variables, we additionally transformed the continuous dependent variables to dichotomous, resulting in new categorical measures of worse vs. better cognitive flexibility, visual working memory, and model-based planning. To ensure the group sizes (worse vs. better cognition) are comparable between all cognitive measures, we used the ratio of subjective memory problems as a quantile cut-off. That is, subjective memory problems were present in 40.5% of the sample, and so our binarization of objective cognitive scores returned 40.5% classified as having poor objective performance on visual working memory, cognitive flexibility, and model-based planning. We then ran an additional series of logistic regressions predicting each binarised objective cognitive outcome from each risk factor individually, controlling for gender and age, and calculated odds ratios for each association. For each cognitive outcome, Bonferroni correction was applied to correct for multiple comparisons, dividing $\alpha = 0.05$ by the number of risk factors (i.e., each independent variable was understood as one comparison).

Secondary analyses. First, we tested if the associations between risk factors and objective and subjective cognition change after controlling for depression. We ran

the same series of analyses as described in the Primary analysis section, this time additionally controlling for depression as measured by the CES-D scale. This was done for 12 of the 13 original risk factors, excluding depression itself.

Next, we tested if the associations between all 13 risk factors and cognition differed by age. This subgroup analysis was only tested on those combinations of independent and dependent variables whose associations were previously significant. We ran the same series of regressions as in Primary analysis, predicting objective, respective subjective cognition from each of this reduced set of risk factors, controlling for gender and age, this time with an added interaction term (i.e., age x risk factor). Bonferroni correction was applied to correct for multiple comparisons, dividing $\alpha = 0.05$ by the number of interaction terms of interest (i.e., each combination of independent and dependent variables was understood as one comparison).

3.3 Results

3.3.1 Descriptive and Demographic Analyses

As has been consistently reported in the literature, objective and subjective cognition were only weakly correlated ($r = .06-.16$; see Table 3 for correlations between all cognitive measures). The correlation matrix of all self-reported risk factors is presented in Appendix 8.1.6.

The older the participant, the worse was their objective cognitive performance across domains (standardised β [95% CI] = 0.20 [0.17, 0.24] for visual working memory, β [95% CI] = 0.34 [0.30, 0.37] for cognitive flexibility, and β [95% CI] = 0.09 [0.05, 0.12] for model-based planning, all $p < .001$; see Figures 5A–C). We also found a significant quadratic effect of age on all three objective cognitive measures: The older the participants, the stronger the association between age and visual working memory (β [95% CI] = 0.07 [0.04, 0.11], $p < .001$), cognitive flexibility (β [95% CI] = 0.13 [0.10, 0.17], $p < .001$), and model-based planning (β [95% CI] = 0.07 [0.03, 0.10], $p < .001$), indicating a steeper curve in older age groups. This quadratic effect was significant even though the raw cognitive scores

plateaued in the oldest age groups (i.e., around 70–75 years), as shown in Figures 5A–C, which could suggest evidence for selection bias in the oldest participants. Of note, very few participants reported being 75 and older ($n = 30$, i.e., 0.9% of the sample; see Appendix 8.1.1 for illustration). Subjective memory problems were significantly more likely in older participants (OR [95% CI] = 1.04 [1.01, 1.06], $p = .020$; see Figure 7D). However, unlike in the objective measures, there was no significant quadratic effect of age on subjective memory problems (OR [95% CI] = 1.00 [1.00, 1.00], $p = .121$). To ensure similarity between models of objective and subjective cognition, the main covariate-adjusted analyses (see next section) include age as a simple linear term.

Table 3. *Correlations between cognitive measures. Numbers represent Pearson product-moment correlation coefficients for pairs of continuous variables and polyserial correlations for pairs of continuous and categorical variables. 95% confidence intervals are provided in square brackets. All coefficients are significant at $p < .05$ or lower.*

	Subjective memory problems	Model-based planning	Cognitive flexibility
Visual working memory	.09 [.04; .13]	.25 [.22; .29]	.45 [.41; .47]
Cognitive flexibility	.16 [.11; .20]	.15 [.12; .19]	
Model-based planning	.06 [.02; .10]		

As for gender, we focused only on cisgender men and women for our main analyses due to sample size. We found that cisgender women did not significantly differ from cisgender men on visual working memory (β [95% CI] = 0.00 [-0.03, 0.04]; $p = .915$; Figure 7A), but they had significantly worse model-based planning (β [95% CI] = 0.05 [0.01, 0.08]; $p = .009$; Figure 7C) and were more likely to report subjective memory problems (OR [95% CI] = 1.24 [1.15, 1.34], $p < .001$; Figure 7D). Cisgender women showed higher cognitive flexibility than men, indexed by faster game completion times (β [95% CI] = -0.04 [-0.08, -0.01]; $p = .012$; Figure 7B). Comparisons of the cognition of transgender and non-binary participants are difficult to make due to the very small sample size of these groups ($n = 68$, i.e., 2% of the sample) but they are described the Appendix 8.1.7.

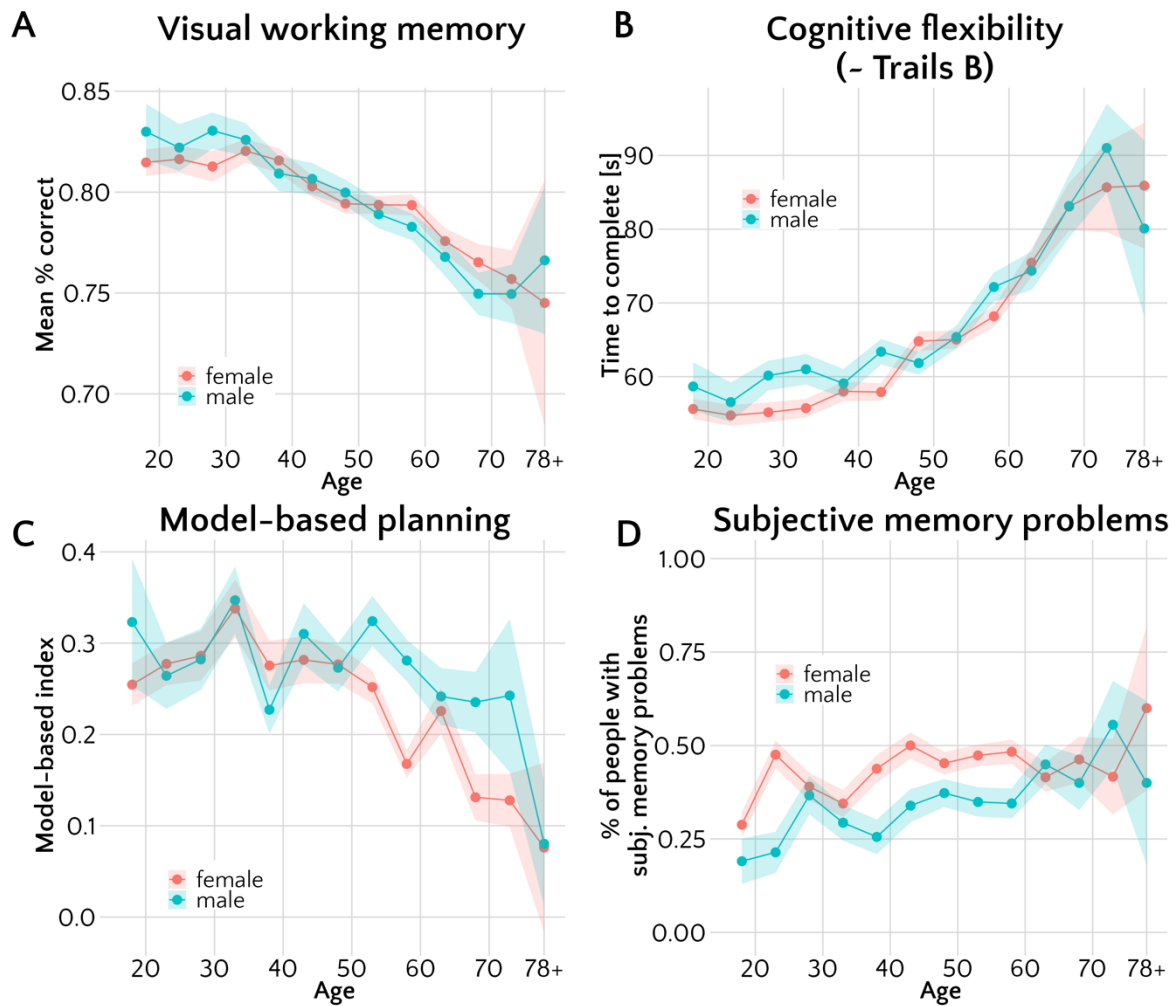


Figure 7. Associations of cognitive measures with age and gender. Points correspond to mean raw scores on **(A)** visual working memory, **(B)** cognitive flexibility, **(C)** model-based planning, and **(D)** mean proportion of participants with memory problems. The means were calculated for 5-year bins, split by gender. Error bars represent standard errors **(A–C)** or standard errors of proportion **(D)**. Only cisgender participants are displayed in this plot ($N = 3,276$); for plots including non-cisgender participants please see Appendix 8.1.7.

3.3.2 Risk Factors and Cognition

After applying Bonferroni correction per each dependent measure, we set our alpha level to $\alpha = .05/13 = .0038$. The associations between risk factors and visual working memory, cognitive flexibility, model-based planning, and subjective memory problems, controlling for age and gender, are visualised in Figure 8 and detailed results are presented in Table 4.

All three objective cognitive outcomes had significant associations with education, depression, and SES (ordered from largest to smallest effect size), in that worse cognition was associated with higher risk (Figure 8A). Model-based planning had no other significant associations with any other risk factor. Both visual working memory and cognitive flexibility, however, were significantly associated with history of stroke and hypertension, followed by loneliness, smoking history, social network size, and hypertension, in that worse cognition was associated with higher risk. Additionally, visual working memory was significantly linked to diabetes and cognitive flexibility to hearing handicap. None of the objective cognitive measures assessed here had significant associations with tinnitus, exercise, or family history of dementia. Overall, the effect sizes were comparable for visual working memory and cognitive flexibility.

Subjective memory complaints were significantly associated with 10/13 risk factors (Figure 8B), namely depression, SES, hearing handicap, loneliness, education, exercise, smoking history, tinnitus, social network size, and history of stroke (ordered from largest to smallest effect size). Subjective memory complaints were not significantly associated with family history of dementia, diabetes, or hypertension.

3.3.3 Comparison of Effect Magnitude for Subjective vs. Objective Cognition

Next, we compared the magnitude of the risk factor effects (expressed as odds ratios) on all cognitive outcomes, using a binarised version of the three objective cognitive measures as dependent variables in a series of logistic regressions. As illustrated in Figure 8C, objective and subjective cognitive measures differed in the strength of their associations with risk factors. Compared to the objective cognitive measures, subjective memory problems were more strongly associated with 8/13 of the risk factors studied, namely depression, SES, hearing handicap, loneliness, exercise, smoking, tinnitus, and social network. In case of higher risk on each of these eight factors, the odds of having subjective memory problems were increased by 27–82%, whereas the odds of having worse visual working memory or lower cognitive flexibility were increased only by 0–28% (Appendix 8.1.8).

Objective and subjective cognition had equivalent associations with education, history of stroke, family history of dementia, hypertension, and diabetes, as indicated by overlapping confidence intervals for the odds ratios expressing these effects. For example, in contrast to SES, which had about 50% higher OR for subjective vs. objective cognitive measures, objective and subjective cognitive outcomes were comparably sensitive to education, a highly related construct. Of note, although the confidence intervals of all measures were overlapping for hypertension, only the binarised objective measure of visual working memory was significantly associated to it, making it the only case where an objective measure outperformed subjective in sensitivity towards risk factors when put on the same binarised scale. Detailed results are presented in Appendix 8.1.8.

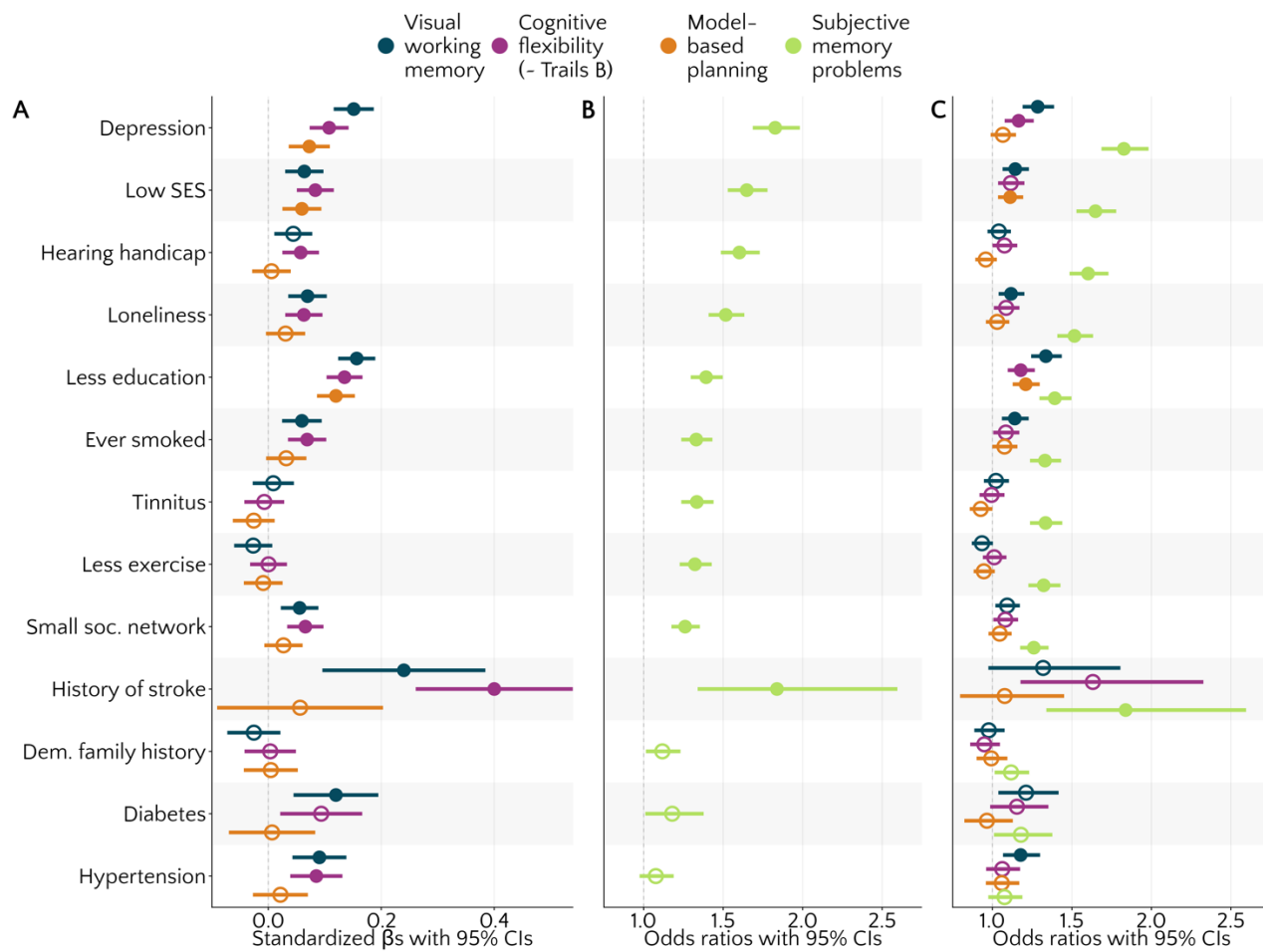


Figure 8. Associations between cognitive measures and risk factors, controlling for age and gender. Filled circles represent values significant at $p < .0038$, i.e., after applying Bonferroni correction per dependent variable. Higher estimates indicate worse cognitive outcomes. Circles and lines represent (A) standardised beta estimates with 95% confidence intervals or (B–C) odds ratios with 95% confidence intervals. In (C), objective cognitive scores were binarised to enable direct comparison of odds ratios with subjective memory problems.

Table 4. Associations of risk factors with cognitive outcomes ($n = 3,327$), controlling for gender and age or gender, age, and depression, expressed as estimates from logistic regressions (subj. memory problems) or linear regressions (all other DVs). Highlights indicate results significant after Bonferroni correction ($p < .0038$).

Risk factor	Covariates	Visual working memory			Cognitive flexibility			Model-based planning			Subjective memory problems			
		β (SE)	t	p	β (SE)	t	p	β (SE)	t	p	β (SE)	z	p	OR [CI _{95%}]
Depression	Age + gender	0.15 (0.02)	8.23	< .001	0.11 (0.02)	6.01	< .001	0.07 (0.02)	3.72	< .001	0.6 (0.04)	14.45	< .001	1.82 [1.68, 1.97]
Low SES	Age + gender	0.06 (0.02)	3.66	< .001	0.08 (0.02)	4.87	< .001	0.06 (0.02)	3.28	.001	0.5 (0.04)	12.89	< .001	1.65 [1.53, 1.78]
	Age + gender + depression	0.01 (0.02)	0.28	.778	0.05 (0.02)	2.62	.009	0.04 (0.02)	1.91	.056	0.33 (0.04)	7.86	< .001	1.39 [1.28, 1.51]
Hearing handicap	Age + gender	0.04 (0.02)	2.50	.012	0.06 (0.02)	3.40	.001	0 (0.02)	0.22	.823	0.46 (0.04)	11.89	< .001	1.59 [1.47, 1.71]
	Age + gender + depression	0.01 (0.02)	0.30	.767	0.03 (0.02)	1.85	.064	-0.01 (0.02)	-0.82	.413	0.35 (0.04)	8.79	< .001	1.43 [1.32, 1.54]
Loneliness	Age + gender	0.07 (0.02)	4.06	< .001	0.06 (0.02)	3.63	< .001	0.03 (0.02)	1.70	.089	0.42 (0.04)	11.08	< .001	1.52 [1.41, 1.64]
	Age + gender + depression	-0.04 (0.02)	-1.83	.068	-0.01 (0.02)	-0.46	.648	-0.02 (0.02)	-1.02	.307	0.09 (0.05)	1.87	.061	1.1 [1, 1.21]
Less education	Age + gender	0.16 (0.02)	9.33	< .001	0.14 (0.02)	8.34	< .001	0.12 (0.02)	6.93	< .001	0.33 (0.04)	9.01	< .001	1.39 [1.3, 1.5]
	Age + gender + depression	0.14 (0.02)	8.03	< .001	0.12 (0.02)	7.40	< .001	0.11 (0.02)	6.37	< .001	0.26 (0.04)	6.77	< .001	1.3 [1.2, 1.4]
Less exercise	Age + gender	0.02 (0.02)	1.37	.172	0.04 (0.02)	2.33	.020	0.03 (0.02)	1.77	.077	0.34 (0.04)	8.88	< .001	1.41 [1.31, 1.52]
	Age + gender + depression	0 (0.02)	-0.07	.941	0.02 (0.02)	1.30	.193	0.02 (0.02)	1.14	.256	0.26 (0.04)	6.63	< .001	1.3 [1.2, 1.41]
Ever smoked	Age + gender	0.06 (0.02)	3.23	.001	0.07 (0.02)	3.98	< .001	0.03 (0.02)	1.57	.118	0.29 (0.04)	7.66	< .001	1.33 [1.24, 1.44]
	Age + gender + depression	0.03 (0.02)	1.77	.077	0.05 (0.02)	2.94	.003	0.02 (0.02)	0.90	.371	0.2 (0.04)	5.16	< .001	1.22 [1.13, 1.32]
Tinnitus	Age + gender	0.01 (0.02)	0.52	.604	-0.01 (0.02)	-0.30	.763	-0.03 (0.02)	-1.40	.162	0.28 (0.04)	7.31	< .001	1.33 [1.23, 1.43]

Risk factor	Covariates	Visual working memory			Cognitive flexibility			Model-based planning			Subjective memory problems			
		β (SE)	<i>t</i>	<i>p</i>	β (SE)	<i>t</i>	<i>p</i>	β (SE)	<i>t</i>	<i>p</i>	β (SE)	<i>z</i>	<i>p</i>	OR [CI _{95%}]
	Age + gender + depression	-0.02 (0.02)	-0.82	.412	-0.02 (0.02)	-1.29	.196	-0.04 (0.02)	-2.03	.042	0.21 (0.04)	5.15	< .001	1.23 [1.14, 1.33]
Small soc. network	Age + gender	0.06 (0.02)	3.32	.001	0.07 (0.02)	3.96	< .001	0.03 (0.02)	1.61	.108	0.24 (0.04)	6.49	< .001	1.27 [1.18, 1.36]
	Age + gender + depression	0 (0.02)	0.19	.852	0.03 (0.02)	1.79	.073	0 (0.02)	0.19	.851	0.03 (0.04)	0.85	.397	1.03 [0.96, 1.12]
History of stroke	Age + gender	0.24 (0.07)	3.29	.001	0.4 (0.07)	5.63	< .001	0.06 (0.07)	0.74	.461	0.61 (0.17)	3.65	< .001	1.84 [1.34, 2.6]
	Age + gender + depression	0.19 (0.07)	2.64	.008	0.37 (0.07)	5.17	< .001	0.03 (0.08)	0.43	.666	0.46 (0.17)	2.65	.008	1.58 [1.14, 2.25]
Dem. family history	Age + gender	-0.02 (0.02)	-0.95	.340	0 (0.02)	0.13	.897	0 (0.02)	0.11	.909	0.11 (0.05)	2.25	.024	1.12 [1.01, 1.23]
	Age + gender + depression	-0.02 (0.02)	-0.93	.353	0 (0.02)	0.16	.876	0 (0.02)	0.13	.896	0.12 (0.05)	2.40	.016	1.13 [1.02, 1.25]
Diabetes	Age + gender	0.12 (0.04)	3.18	.002	0.09 (0.04)	2.52	.012	0.01 (0.04)	0.14	.889	0.17 (0.08)	2.11	.035	1.18 [1.01, 1.38]
	Age + gender + depression	0.1 (0.04)	2.72	.007	0.08 (0.04)	2.17	.030	0 (0.04)	-0.08	.934	0.1 (0.08)	1.25	.211	1.11 [0.94, 1.3]
Hyper-tension	Age + gender	0.09 (0.02)	3.65	< .001	0.09 (0.02)	3.62	< .001	0.02 (0.02)	0.82	.410	0.07 (0.05)	1.45	.146	1.08 [0.97, 1.19]
	Age + gender + depression	0.08 (0.02)	3.18	.001	0.08 (0.02)	3.27	.001	0.01 (0.02)	0.59	.553	0.03 (0.05)	0.53	.594	1.03 [0.93, 1.14]

3.3.4 Depression-adjusted Analyses

To test if depression symptoms might explain the stronger associations between subjective memory complaints and self-report risk factors, we repeated our primary analyses controlling for depression. This was done for 12 of the 13 original analyses, excluding the analysis of depression itself. As seen in Figure 9, in general, adding depression as a covariate nominally decreased the magnitude of associations between most risk factors and subjective memory problems.

However, it did not render them non-significant, in most cases — the total number of significant associations reduced from 9/12 to 6/12 risk factors. The associations with loneliness, social network, and history of stroke were no longer statistically significant at $p < .0038$ after adjusting for depression. The full results are presented in Table 4.

For objective cognitive outcomes, where the effects were more modest to begin with, the number of significant associations with risk factors reduced substantially when controlling for depression, here from 8/12 to 2/12 for visual working memory, from 8/12 to 4/12 for cognitive flexibility, and from 2/12 to 1/12 for model-based planning. Of note, adjusting for depression removed the associations of all objective cognitive outcomes with SES, whereas the associations between all objective cognitive outcomes and education held. See Table 4 for detailed results.

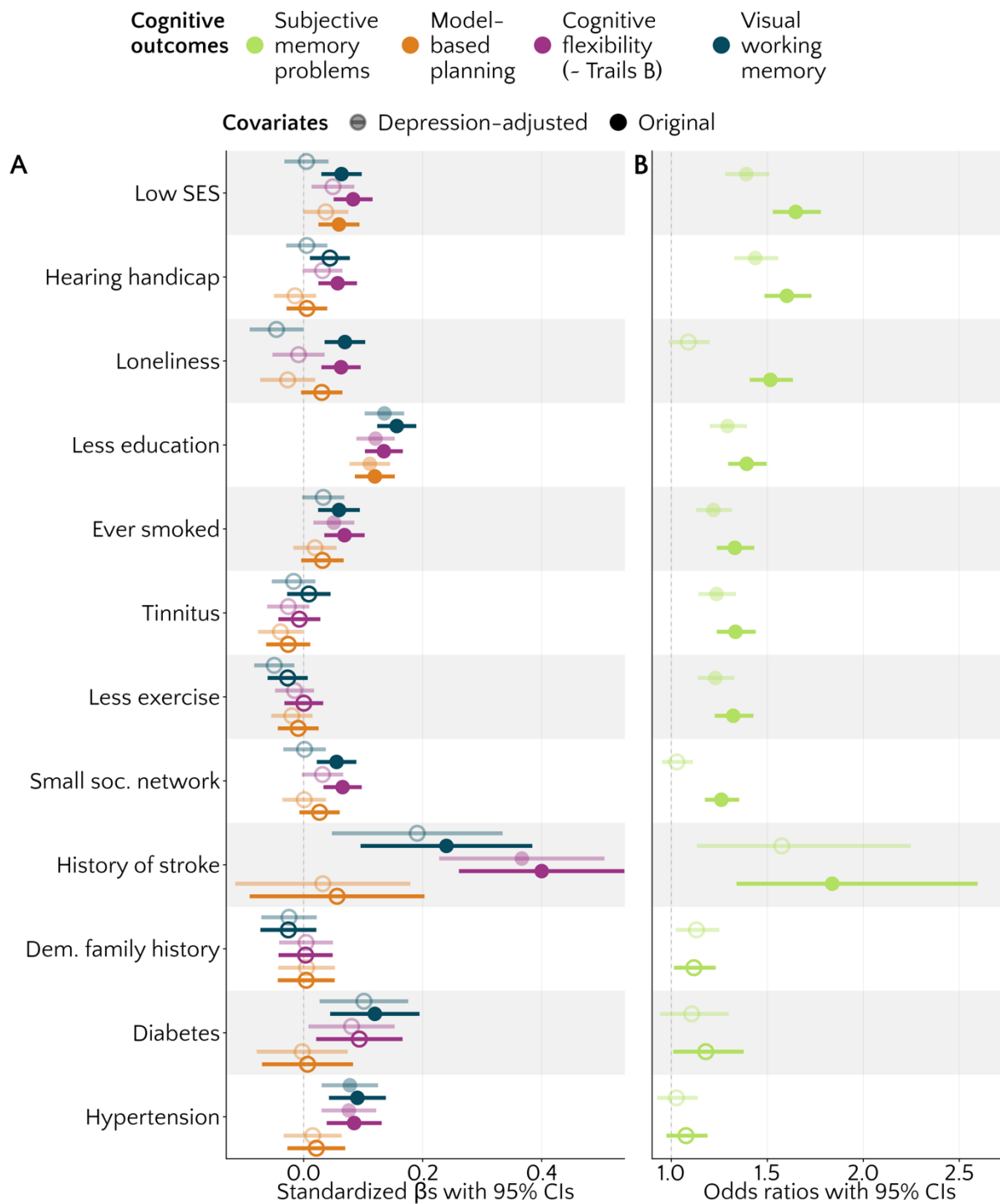


Figure 9. Associations between cognitive measures and risk factors, expressed as standardised beta estimates (**A**) or odds ratios (**B**) with 95% confidence intervals. Lighter shades represent depression-adjusted analyses, darker shades represent original analyses. All analyses additionally controlled for gender and age. Filled circles represent values significant at $p < .0038$, i.e., after applying Bonferroni correction per dependent variable. Higher estimates indicate worse cognitive outcomes.

3.3.5 Interactions between Age and Risk Factors

In the regression models described in the Primary analysis, 31 risk factors showed a significant association with cognitive outcome. For these 31 combinations of risk factors and cognitive outcomes, we ran an additional series of models with the added term of the risk factor's interaction with age, setting a new alpha level to $\alpha = .05/31 = .0016$ to control for multiple comparisons. Broadly, we observed that almost none of these effects were moderated by age. Only one factor showed a significant interaction with age: Smoking history differed in its association with subjective memory problems by age (OR [95% CI] = 0.87 [0.81, 0.94], $p < .001$), showing stronger associations in younger adults than older adults (Appendix 8.1.9, Figure S10). No other risk factor showed a significant interaction with age (see Appendix 8.1.10, Table S2, for a complete list of interaction terms of these 31 models).

3.4 Discussion

Objective and subjective cognitive impairments are independent predictors of dementia, yet their differential sensitivity to previously established risk factors is little understood. Here, we estimated the pattern and magnitude of associations between potentially modifiable dementia risk factors and subjective memory problems compared to multiple aspects of objective cognition using large-scale cross-sectional data ($N = 3,327$) gathered across the adult lifespan via smartphone. Consistent with previous findings (Burmester et al., 2016), objective cognitive measures were only weakly correlated with subjective memory problems. Both were more frequent in older participants, but women exhibited a higher frequency of subjective, but not objective cognitive impairments, in line with prior research (Wen et al., 2021). Both objective and subjective cognition were associated with depression, SES, and education, and with the exception of the objective measure of model-based planning, all measures were additionally associated with loneliness, social network size, history of smoking and stroke, and hypertension. Interestingly, several risk factors were selectively linked with objective or subjective cognition; only working memory or cognitive flexibility were

associated with hypertension and diabetes, whereas only subjective impairments were associated with less exercise and tinnitus. Crucially, our findings suggest that across the adult lifespan, subjective memory complaints are considerably more related to self-reported risk factors for dementia than objective differences in cognitive ability assessed via smartphone. When put on the same binarised scale, subjective memory was more strongly associated than any objective test with 8/13 factors, namely depression, SES, hearing handicap, loneliness, tinnitus, exercise, and social network, and the magnitude of associations was up to 50% stronger for subjective memory. It was outperformed by an objective cognitive measure (visual working memory) only in the case of hypertension. This implies high sensitivity of subjective cognitive assessment to dementia risk factors in healthy individuals, possibly reflecting relatively higher usefulness of subjective assessments in the earliest stage of cognitive decline before impairments can be objectively detected, a finding further supported by prior evidence (Jessen, 2014; Mitchell et al., 2014).

One possible explanation for the current findings is that the discrepancy between objective and subjective cognition could be driven by negative interpretative bias. That is, those who view the world and themselves more negatively might be more likely to respond negatively on subjective scales, biasing the measurement of risk factors and cognitive impairments alike. We repeated our analyses controlling for depression and found that magnitudes of effects decreased across the board when depression was included in the model. However, despite this reduction in magnitude, associations between most risk factors and subjective memory remained significant, whereas most associations with objective cognitive measures were rendered non-significant. The two factors that changed most substantially when depression was included were loneliness and social network size, which no longer showed associations with either objective or subjective cognition after controlling for depression. This suggests that the link between these measures of social engagement and both objective and subjective brain health may be mediated by depression, consistent with prior research (Cacioppo et al., 2010; Franco-O'Byrne et al., 2023; Uchida et al., 2019). Overall, the impact of depression on these models was marked, but not complete, suggesting that depression may be an important element of the causal path between self-report

risk factors and objective cognitive impairments, but longitudinal or quasi-experimental studies are needed to understand how these are precisely related.

We did not observe associations between most risk factors and model-based planning, which is to our knowledge the first time this question has been posed. We were interested to address this as previous research has suggested that model-based planning declines in older age (Eppinger et al., 2013), but also critically relies on the healthy functioning of the hippocampus (Vikbladh et al., 2019). However, against our prediction, model-based planning was only associated with education, depression, and SES. It is unlikely these null effects can be attributed to a lack of sensitivity of our smartphone task, as a recent paper showed it is sensitive to individual differences in compulsivity (Donegan et al., 2023), replicating work with the classic task (Gillan et al., 2016). Nonetheless, it must be acknowledged that all four cognitive assessments used in this study differed in not just their degree of objectivity, but also their duration and psychometric properties and direct comparisons are subject to those caveats. Despite this, it is striking that the largest correlates were observed for subjective memory, which was assessed using a single item self-report assessment. All three cognitive tests, in contrast, were validated against traditional benchmarks, aggregated across many trials, and included several features to improve their performance over traditional versions (e.g., reducing ceiling effects, increasing reaction time measurement precision).

This study raises the general question of whether objective assessments of cognition or its subjective perceptions should be prioritised as outcomes in research and public interventions. Objective measures gathered at a single timepoint cannot identify cognitive decline because of premorbid differences in ability; this might explain the overall reduced signal observed here for objective cognitive tests. However, it is notable that objective assessments of cognitive flexibility and working memory were nonetheless both sensitive to hypertension and only the latter to diabetes, while subjective cognition was not. This could be because cardiovascular factors and associated structural changes to brain health affect cognitive abilities more directly (Veldsman et al., 2020), while causal paths for lifestyle, sociodemographic and social factors might be more complex. Although cross-sectional, these findings may provide a useful starting point for

thinking about such paths. For example, our findings with respect to depression support the idea that the association between low social involvement and both objective and subjective cognition could be explained by depression (Cardona & Andrés, 2023).

The association of cognition with most risk factors in the current study did not differ by age, except for subjective memory problems and smoking. This finding contrasts with the life-course model of risk factors for dementia (Livingston et al., 2017, 2020), which posits that different periods of life might provide 'sensitive windows' during which various factors might particularly come into play.

Furthermore, the only interaction effect observed in the current study was opposite to those posited by the prevailing life-course model: We found that the association between smoking and subjective memory problems was strongest in the youngest participants, whereas previous meta-analytical work (Livingston et al., 2017, 2020) placed smoking as a risk factor for later life. However, another recent study using large cross-sectional data ($N > 20,000$) collected across the lifespan found that risk factor prevalence or associations with cognition do not match the prevailing life-course model, suggesting the mapping of 'sensitive windows' might be inaccurate (LaPlume, McKetton, Levine, et al., 2022). A caveat of this interpretation is the fact that both ours and the LaPlume et al. (2022) study included a large number of participants under 30, where cognitive decline due to neurological causes is unlikely. Even if certain risk factors at a younger age contribute to dementia risk in older age, this does not by definition imply cognitive decline would appear already at a younger age. Additionally, some of the risk factors in the current study are less likely to apply to young people in the first place (e.g., hearing impairment, diabetes). As a result, insufficient variability in lower age brackets could have attenuated any interaction effect, even if present. To verify the current null finding, researchers should strive to replicate it in longitudinal studies using dementia incidence as an outcome, preferably ensuring individuals with high risk factor prevalence are sufficiently represented amongst younger participants.

A limitation of the current study is that the methods of recruitment (i.e., online) and data collection (i.e., smartphone assessment) were likely to introduce selection bias, especially in older participants. Both these methods require digital literacy and are probably more appealing to individuals of higher socioeconomic status

(Stone et al., 2024). Indeed, our sample had overall higher levels of education than what we would expect based on representative national data. This could have diminished or seemingly reverted the effects of age on cognition, as reported in another smartphone-based study (Coutrot et al., 2018), where cognitive performance was found to markedly rise in adults over the age of 75. However, the present study did not find such biologically implausible inflection point: We showed an average decrease in cognitive performance even in older participants up until a plateau around 70–75 years of age, and very few Neureka participants were older than 70. Besides age, other risk factor effects could also have been obscured by higher socioeconomic status, as risk factor co-occurrence and effects on cognition seem to be partially driven by social disparities (Meader et al., 2016; Röhr et al., 2022). In our analytical approach, we chose not to control for SES or education, as we wanted to directly compare both variables with other risk factors while keeping covariates constant, but a future study could address this. Selection bias poses important issues for the representativeness of smartphone research, particularly — but not only — in older adults. However, the remote, self-paced nature of smartphone-based research can make it more representative in other respects, compared to in-person lab-based studies (Muniz-Terrera et al., 2019). Therefore, convergence across a variety of research methods is key.

Several other limitations are worth highlighting. First, due to the cross-sectional nature of this study, the risk factor variables described here could be understood as ‘diagnostic factors’ (Huitfeldt, 2016) or ‘risk markers’ (Capron, 2003) rather than true risk factors. This precludes us from drawing causal inferences about the etiology of cognitive impairment from this dataset alone. That said, prior work using a suite of methodologies has converged on a likely causal role for many of these risk factors in brain health and dementia (Courtin et al., 2019; Kuźma, Hannon, et al., 2018; Livingston et al., 2020; but cf. Desai et al., 2023; Seblova et al., 2021). Second, the current assessment of subjective memory problems could be improved. We relied on a common single-item, binary measure of subjective memory problems without any anchor, whereas previous research has shown that age-anchored comparisons and self-comparisons (i.e., comparing own memory to that of individuals of the same age, respective to own memory 5 years ago) differ in their associations with depression and age (Mogle et al., 2017; Ponds et al.,

2010). Third, the use of convenience sampling at a large scale was likely to introduce unknown biases. Neureka study participants are recruited through a wide range of methods including app store ads, citizen science projects, and a portal of clinical trials, using materials that mention both brain and mental health conditions. Information on participant motivation is not being collected, and participant data is not tied to recruitment information, making it impossible to fully understand participant background and motives. In light of this uncertainty, the current results need to be interpreted with caution and triangulated with other research methodologies. Although citizen science is a great way to scale up research, participants may be highly motivated and active also in other aspects of their lifestyle, making them possibly less representative. Despite this, we had sufficient variability in our data to detect associations between modifiable risk factors and cognition. Fourth, we cannot exclude the possibility that domain differences contributed to the differences between subjective and objective assessment. We focused on tasks with a strong executive function component based on prior work linking executive function to healthy brain ageing (Mortamais et al., 2017; Park et al., 2002), but it is possible that an objective test of episodic verbal memory could have shown even stronger effects (Mortamais et al., 2017). Episodic memory might have also been more salient for the current participants' self-evaluations, possibly contributing to the discrepancy. Finally, we assessed risk factors using self-report due to the remote nature of the study. Past research has shown different associations of objectively and subjectively measured factors with brain health outcomes. For example, informant-based subjective hearing impairment, but not objectively measured hearing loss predicted incident dementia (Marinelli et al., 2022).

Despite these limitations, the current study shows the respective value of objective and subjective cognitive assessment and is the first to our knowledge to directly compare these types of assessments on associations with previously established risk factors for dementia. This study also highlights the potential of self-administered, smartphone-based assessments to detect early cognitive impairment and its meaningful associations with established risk factors for dementia. This may be particularly valuable for studying how risk factors impact on cognition in low- and middle-income countries where access to cognitive

assessment is more difficult (Muniz-Terrera et al., 2019) and the relative importance of risk factors for dementia differs (Santamaria-Garcia et al., 2023). Future randomised controlled trials of prevention interventions should consider including brief subjective memory assessments as cognitive outcomes, alongside objective measures.

Chapter 4: APOE- ϵ 4 associations with smartphone-based cognitive markers

4.1 Introduction

Midlife may be a critical window for preventing and treating Alzheimer's disease (AD; C. W. Ritchie et al., 2016). To allow health systems to direct these efforts to those most likely to benefit, highly sensitive markers that can identify individuals at an increased risk of AD are required. Cognitive tests could serve this purpose, constituting an affordable and accessible alternative to traditional biomarkers of early AD-related pathology (Glymour et al., 2018). This is supported by the finding that preclinical AD-related neurodegeneration is mirrored by a decline in specific cognitive domains that can be observed up to decades before dementia diagnosis (Mortamais et al., 2017). Impacted domains include navigation abilities (Allison et al., 2016), visuospatial processing (K. Ritchie et al., 2017), and executive functions (Mortamais et al., 2017), with some of these impairments possibly appearing even before a decline in episodic memory, a hallmark feature of AD (Coughlan et al., 2018). Beyond objective cognitive tests, subjectively perceived cognitive decline also has potential as an important risk marker. For example, it has been shown to be associated with brain atrophy similar to that of preclinical AD (Saykin et al., 2006) and unlike cognitive tests, often incorporates a baseline comparison to prior subjective functioning (Jessen, 2014). Both objective and subjective cognitive markers are suitable for remote assessment via smartphones, which can offer even more cost-effectiveness and scalability. We have previously shown that such markers are sensitive to a wide range of potentially modifiable risk factors for dementia across the healthy adult lifespan, with subjective cognition showing larger effects compared to objective tests (Chapter 3). However, it remains unclear if such cognitive markers are also sensitive to genetic risk for dementia, particularly during the critical midlife period.

The strongest known genetic risk factor for late-onset AD is the APOE- ϵ 4 allele (Corder et al., 1993; Wightman et al., 2021), which appears in around 15% of the world population (Eisenberg et al., 2010). Although not deterministic, APOE- ϵ 4 is associated with a higher risk and earlier onset of AD and with steeper cognitive decline in older adults in a dose-dependent manner (Riaz et al., 2021). About a

half of APOE- ϵ 4-homozygotes (i.e., individuals carrying two APOE- ϵ 4 alleles) will develop AD by age 85, compared to only 10% of non-carriers (Genin et al., 2011). The exact risk-conferring mechanisms of APOE- ϵ 4 are not yet understood, with research suggesting links to neuroinflammation, amyloid- β load, and tau pathology (Fernández-Calle et al., 2022). Despite the evident link to AD, the effects of APOE- ϵ 4 on cognition in healthy individuals have been inconsistent across studies, age groups, and cognitive domains (O'Donoghue et al., 2018). Although two meta-analyses (Small et al., 2004; Wisdom et al., 2011) found an average negative effect on episodic memory, executive function, and global cognitive measures in older adults, these effects do not consistently extend to midlife. A meta-analysis including 23 studies with a mean sample age between 35–60 years showed no average effects of APOE- ϵ 4 across domains (Lancaster, Tabet, et al., 2017). However, some of the included studies found an accelerated cognitive decline over time in APOE- ϵ 4 carriers and subtle impairments on behavioural measures designed to maximise sensitivity (Lancaster, Tabet, et al., 2017), suggesting better targeted studies could reveal a different picture.

One reason for the lack of robust effects in midlife could lie in methodological issues such as low sensitivity of cognitive tasks used in previous studies. Any genetic effects on complex phenotypes like cognition are likely to be small (Marian, 2012) and they tend to amplify in older age (Lindenberger et al., 2008; McGue & Christensen, 2002), meaning in midlife they might only be detectable using highly sensitive and targeted neurocognitive tasks, as opposed to the crude cognitive outcomes typically utilised in AD research (Lancaster, Tabet, et al., 2017; O'Donoghue et al., 2018). Another reason could be that APOE- ϵ 4 affects cognition through multiple pathways, with each manifesting most markedly at a different point in the lifespan (Greenwood et al., 2014; O'Donoghue et al., 2018), but all possibly converging in midlife. The proposed pathways include a 'prodromal' mechanism, whereby APOE- ϵ 4 might contribute to AD pathology and thus affect cognition particularly in older adults at the prodromal stage of AD (Greenwood et al., 2014; Knight et al., 2014), and a 'cognitive phenotype' mechanism, whereby APOE- ϵ 4 might give rise to specific life-long patterns in cognitive abilities that are independent of future AD diagnosis (Greenwood et al., 2005, 2014; Greenwood & Parasuraman, 2003). It is also possible that some of the effects of APOE- ϵ 4 on

cognition are positive at certain periods of the lifespan. The ‘antagonistic pleiotropy’ theory posits that APOE- ϵ 4 might have some evolutionary value in youth despite its deleterious effects in higher age (Han & Bondi, 2008). This theory has been supported by evidence from neuroimaging (Koelewijn et al., 2019; Mondadori et al., 2007; cf. Henson et al., 2020), but there is no evidence of antagonistic pleiotropy from studies using classic neuropsychological tests across the lifespan (Henson et al., 2020; Weissberger et al., 2018). Finally, midlife-specific effects are hard to evaluate as midlife has been inconsistently defined across studies (Lancaster, Tabet, et al., 2017), with some including participants as young as 24 (Flory et al., 2000) and other participants as old as 85 (Greenwood et al., 2005). Taken together, the effects of APOE- ϵ 4 on cognition in midlife could differ based on cognitive domain, task sensitivity, age, and possibly other factors, which need to be considered.

To determine if cognitive markers could be used to detect genetic risk in midlife, studies need to investigate APOE- ϵ 4 effects in larger samples and within narrower age windows, focusing on cognitive tasks that have been previously identified as particularly promising for AD risk detection. Promising candidates are tasks that recruit the hippocampus, which is one of the first brain areas impacted by early AD-related neurodegeneration (Braak & Braak, 1991). Hippocampus-reliant tasks include measures of episodic memory (Burgess et al., 2002), navigation abilities (Coughlan et al., 2018), and other cognitive assays. Episodic memory is one of the most consistently impaired domains in older APOE- ϵ 4 carriers (O’Donoghue et al., 2018; Small et al., 2004; Wisdom et al., 2011), but a meta-analysis reported that in midlife the effect of APOE- ϵ 4 on episodic memory was null on average (Lancaster, Tabet, et al., 2017). Navigation abilities can be affected in AD even before episodic memory and show a higher specificity to AD (Coughlan et al., 2018). Initial work also suggested that tasks assessing navigational abilities and allocentric spatial memory might be particularly sensitive to APOE- ϵ 4 in young to middle-aged carriers (Coughlan et al., 2019, 2020; Kunz et al., 2015). Evidence against this hypothesis comes from a more recent-meta analysis of APOE- ϵ 4 effects on spatial processing, which found no effect on navigational abilities averaging across 15 studies and $N = 1,384$ individuals throughout the lifespan (Daly et al., 2024). However, true effects could have been obscured by the meta-

analysis averaging across a variety of navigation measures, including both egocentric navigation (i.e., finding location of objects relative to the self) and allocentric navigation measures (i.e., finding location of objects relative to other objects, relying on mental maps), which are underpinned by distinct cognitive and neural processes (Tu et al., 2015). A potentially more sensitive assay of hippocampal function might be model-based planning, which is the ability to create more general, non-spatial mental models of the world based on experience, typically measured through the Two-step Reinforcement Learning Task (Daw et al., 2011b). The neural substrate of model-based planning overlaps with allocentric navigation, as individuals with hippocampal lesions had impairments in both (Vikbladh et al., 2019). Additionally, model-based planning relates to real-world functional impairments (Donegan et al., 2023) and declines in older age (Eppinger et al., 2013). While the associations between model-based planning and APOE status have not been investigated yet, Coughlan et al. (2019) found that in $N = 60$ adults aged 53–73, worse performance on an allocentric navigation task differentiated APOE- $\epsilon 4$ carriers from non-carriers. Hence, it would be instructive to assess if similar decrements are observed for model-based planning in middle-aged APOE- $\epsilon 4$ carriers.

Other potentially useful early markers of incipient AD include visual working memory binding tasks, executive function measures, and subjective memory problems, but evidence for their sensitivity to APOE- $\epsilon 4$ in midlife is even more limited. First, a recent series of studies across the lifespan has shown that APOE- $\epsilon 4$ carriers outperformed non-carriers on visual working memory for abstract shapes and their locations (Lu et al., 2021; Zokaei et al., 2017, 2020), which is a measure sensitive to amyloid β pathology (Cecchini et al., 2021; Lu et al., 2021). However, studies using other working memory paradigms reported no consistent APOE- $\epsilon 4$ effects both across the lifespan (Daly et al., 2024; O'Donoghue et al., 2018) or in midlife (Lancaster, Tabet, et al., 2017). It is unclear if the protective effects extend to more closely related visual working memory measures such as object-colour binding tasks, which are also promising as AD markers (D. J. Norton et al., 2020; Parra et al., 2024). Another useful early marker is the Trail-Making Test (TMT), which taps into executive functions like processing speed (version A) or cognitive flexibility (version B) (Bowie & Harvey, 2006), is sensitive to

impairments in frontoparietal network integrity (Varjacic et al., 2018), and its B version has also shown sensitivity to abnormal levels of the AD biomarker A β ₁₋₄₂ (J. K. Ho et al., 2018; Schott et al., 2010). APOE- ϵ 4 was linked to impaired TMT B performance in older adults with prodromal AD (Knight et al., 2014), but a large population-based study found no effects of APOE- ϵ 4 on TMT B performance in healthy middle-aged adults ($N = 7,526$, age 40–79) (Rodriguez et al., 2017). Lastly, subjective cognitive assessments provide a unique opportunity to tap into very early cognitive changes that cannot yet be detected by objective tests (Jessen et al., 2020). While subjective memory decline predicts objective cognitive decline and dementia (Mitchell et al., 2014), the evidence of its sensitivity to genetic risk factors across the lifespan is contradictory (Ali et al., 2018; Huang et al., 2020; T. Zhang et al., 2017). In summary, a range of well-targeted cognitive measures could serve as early markers of AD-related changes, but there is little or mixed evidence for their sensitivity to APOE- ϵ 4 in midlife.

The aim of the current study was to assess the associations of APOE- ϵ 4 with a set of previously validated cognitive markers that can be assessed remotely via the smartphone app *Neureka*, focusing on middle-aged individuals. We previously found these markers were sensitive to age and modifiable dementia risk factors in healthy adults across the lifespan (Chapter 3). Here, we wanted to further evaluate their sensitivity to genetic risk of AD. Specifically, we hypothesised that APOE- ϵ 4 carriers will have impaired model-based planning compared to non-carriers. Furthermore, given the literature does not provide strong support for directional hypotheses, we tested the associations of APOE- ϵ 4 status with visual working memory, processing speed (Trails A), cognitive flexibility (Trails B), and subjective memory problems in a more exploratory way. To assess the clinical utility of these cognitive markers, we tested if we can use these smartphone-based cognitive assessments, adjusted for normative performance using benchmark cohort data ($N = 3,327$), to classify individuals with vs. without an increased genetic risk of AD (i.e., having at least one APOE- ϵ 4 allele) with a clinically meaningful level of precision.

4.2 Methods

4.2.1 Participants and Procedure

APOE genotyped sample. Between January and December 2022, all current participants of the PREVENT study ($N = 646$) were invited via email by PREVENT study coordinators to take part in the *Neureka* x PREVENT add-on study (see Appendix 8.2.1 for more detail on the recruitment process and materials). The PREVENT study protocol has been described in detail elsewhere (C. W. Ritchie & Ritchie, 2012), but in brief, this is a UK- and Ireland-based multi-site prospective cohort study focused on biomarkers and risk factors for dementia in middle-aged adults. All PREVENT study participants provided their informed consent as specified elsewhere (C. W. Ritchie & Ritchie, 2012).

The current add-on study asked PREVENT participants to download the *Neureka* app and complete a module within it called 'Risk Factors'. PREVENT study site coordinators provided invitation links to individuals who responded to an initial email. $N = 147$ invitation links were opened (i.e., 23% of all invited participants) and $n = 134$ participants (i.e., 21% of all invited participants) signed an informed consent form online to have their data connected and shared bidirectionally between the PREVENT and *Neureka* projects. Participants also provided consent for data processing on the *Neureka* app itself. $N = 41$ participants who signed the online consent form did not sign up to *Neureka* or their *Neureka* data was not connected correctly; $n = 4$ participants did not interact with the app after signing up; and further $n = 2$ participants did not have APOE data available in PREVENT. This left us with a sample of $n = 87$ participants (i.e., 13% of all invited participants), who interacted with the required module called 'Risk Factors' in the smartphone app *Neureka* and had APOE data available.

The 'Risk Factors' module takes about 45 minutes to complete and contains 1) two blocks of questionnaires that together assess potentially modifiable biopsychosocial factors and subjective memory problems, and 2) three gamified cognitive tasks that together measure four aspects of cognition (model-based planning, visual working memory, processing speed, and cognitive flexibility). The gamified cognitive tasks and questionnaire blocks were presented to participants

in a semi-randomised order (task — questionnaires — task — questionnaires — task). Participants were free to complete the module in one sitting or over time. $N = 77$ (88.5%), $n = 72$ (82.8%), respective $n = 70$ (80.5%) participants passed exclusion criteria and had data available for at least one of each of the gamified cognitive tasks ('Star Racer', 'Cannon Blast', and 'Memory Match' respectively, see description of the tasks and their measures below).

Benchmark sample. For benchmark comparisons, we used the data from $N = 3,327$ members of the public, from whom we had no genetic information, but who completed the entire 'Risk Factors' module between June 2020 and April 2023. This sample spans individuals aged 18–84 (65% cisgender female) and is described in detail in Chapter 3. Of this benchmark sample, we further selected $n = 1,805$ participants age-matched to the APOE genotyped sample (i.e., same age brackets). Members of the Benchmark sample were individuals who downloaded the *Neureka* app from the app store and completed the required sections at a time and date of their choosing. These participants gave informed consent to participate in neuroscience research within the app during sign-up.

4.2.2 Measures

Model-based planning. Model-based (also 'goal-directed') planning was assessed in *Neureka* via the game 'Cannon Blast' (Donegan et al., 2023), a gamified adaptation of the Two-Step Reinforcement Learning Task (Daw et al., 2011b). In this game, participants aim and fire balls from a cannon to hit diamonds, thus earning reward points. Some balls explode before hitting the desired target in a probabilistic manner, which participants can avoid by employing a basic model-free policy (i.e., repeating actions as long as the outcomes are good and switching to another action after a ball explodes) or a model-based policy (i.e., incorporating more advanced understanding of the probabilistic task structure to refine one's actions) The outcome measure of 'Cannon Blast' is the 'model-based index', which is derived from a hierarchical logistic regression analysis across all

200 trials. A detailed task description and validation are available elsewhere (Donegan et al., 2023).

Visual working memory. Visual working memory was assessed in *Neureka* via the game 'Memory Match', where participants memorise a set of symbols and then search for these in a 4x5 grid. This task is loosely based on the Visual Short-Term Memory Binding Task (VSMBT) (Parra et al., 2010) with a 'binding' and 'non-binding' condition (i.e., memory for symbols only, vs. memory for symbols and colours). 'Memory Match' additionally has varying stimuli (letters vs. abstract shapes) and three graded difficulty levels (remembering two, three, or four symbols), with each combination of conditions, stimuli, and levels occurring twice. In the current study, visual working memory was calculated as proportion of correct responses, averaged across all 24 trials. A detailed task description and validation are available in Chapter 2.

Processing speed and cognitive flexibility. Processing speed and cognitive flexibility were assessed in *Neureka* via the game 'Star Racer', where participants must tap labelled stars onscreen in ascending order, as quickly as possible. This task is a gamified version of the Trail-Making Test (TMT) (Bowie & Harvey, 2006). Similar to the traditional TMT, 'Star Racer' consists of trials in two versions (A and B), but Star Racer has three trials of each version with varying star position. In the current study, processing speed, respective cognitive flexibility, were calculated as mean time to complete three trials of version A of the task (i.e., numbers only), respective three trials of version B (i.e., switching between numbers and letters). Individual trials with completions times of 100 or more seconds (for A version) or 300 or more seconds (for B version) were excluded from the calculation; in such cases task performance was calculated as a mean of the remaining trials. A detailed task description and validation are available in Chapter 3 and its Appendix 8.1.5.1.

Subjective memory problems. We assessed subjective memory through a single-item question taken over from the PREVENT study ('Do you currently experience memory problems?' Yes/No). Here, it was presented as part of a questionnaire on modifiable risk factors for dementia within the 'Risk Factors' module.

APOE-ε4 status. In the PREVENT study, APOE genotyping was done based on blood samples, as described in detail in Ritchie et al. (K. Ritchie et al., 2017). Similarly to the PREVENT study (C. W. Ritchie et al., 2013), here we also classify participants with at least one APOE-ε4 allele as APOE-ε4 carriers. Excluding participants with the protective APOE-ε2 allele (Corder et al., 1994) did not change our results, and so we included APOE-ε2 carriers in our study.

Covariates. Family history of dementia (FHD) was defined as having one or both parents with dementia, based on what participants report via *Neureka* or PREVENT study, whichever data was available and/or more recent. Self-reported dementia in participant siblings or more distant relatives was not included as FHD, similarly to the main PREVENT study (C. W. Ritchie et al., 2013). For age and gender, we used data collected at sign-up to *Neureka*, whereas for education we used years of education as collected in PREVENT, which could be a more precise measure compared to that collected in *Neureka* (i.e., educational level).

4.2.3 Data Analysis

Data preparation. Analyses were conducted using R Statistical Software v4.3.2. (R Core Team, 2023b). Analysis scripts were uploaded to a private GitHub repository (<https://github.com/amrosicka/PREVENTxNeureka.git>) and can be made available upon request. Prior to analyses, the dependent variables were reverse-coded so that in all cases higher values indicate worse cognition. Furthermore, all continuous variables were scaled by mean-centring and dividing by their standard deviation. We scaled all binary variables by coding them as 0/2

to ensure a direct comparability of the estimates for continuous and binary independent variables in our models (see Gelman, 2008). Gender was coded with 'male' as the reference category. Missing data within each analysis were treated by case-wise deletion. All bootstrapping analyses were done with $n = 10,000$ bootstrap replicates.

Analyses. We tested if the performance on each of our five cognitive measures (model-based planning, visual working memory, processing speed, cognitive flexibility, and subjective memory problems) is sensitive to APOE- $\epsilon 4$ status. This was done in a series of regression analyses in the APOE genotyped sample, predicting each cognitive measure from APOE- $\epsilon 4$ status and relevant covariates (age, gender, education, and family history of dementia). We used linear regressions for all gamified cognitive tests and logistic regression for subjective memory problems, which was the only cognitive measure expressed as a binary categorical variable. We calculated bootstrapped 95% confidence intervals for all predictors. In univariate post-hoc analyses, we compared the cognitive performance of APOE- $\epsilon 4$ carriers and non-carriers separately with that of an age-matched subset of the benchmark sample ($n = 1,805$), using t-tests for gamified cognitive tests and chi-square tests for subjective memory problems.

For the significant model, we additionally conducted receiver operating characteristic (ROC) analyses to assess if APOE- $\epsilon 4$ status can be predicted from the given cognitive measure. We predicted APOE- $\epsilon 4$ status either from the scaled cognitive performance alone or from norm-adjusted scaled cognitive performance and calculated the total area under the curve (AUC) for each with bootstrapped 95% confidence intervals. Norm-adjusted cognitive performance was calculated as the difference between the observed performance and the performance predicted by a linear regression model fit in the full benchmark sample ($N = 3,327$), where cognitive performance was predicted from age, gender, education level, and family history of dementia. This enabled to more reliably account for the effects of these covariates on task performance based on a large normative dataset, as opposed to just controlling for the covariates in the smaller genotyped sample.

4.3 Results

4.3.1 Sample Description

Table 5 summarises the demographic characteristics of the total APOE genotyped sample ($N = 87$) as well as those of APOE- $\epsilon 4$ carriers and non-carriers separately. The genotyped sample was highly educated, and a large proportion of participants had parental history of dementia ($n = 38$; i.e., 43.7%). This is typical for the PREVENT study due to its recruitment strategy that targets the children of people with dementia (C. W. Ritchie & Ritchie, 2012). Of note, a large proportion of this sample was APOE- $\epsilon 4$ positive ($n = 35$; i.e., 40.2%), which is substantially higher than what is common in the general population (i.e., 15%; Eisenberg et al., 2010). However, participants of the PREVENT study are not informed about their own APOE status, which precludes them from conscious self-selection bias to the current add-on study. The demographic characteristics of the genotyped sample did not differ significantly from those of the full PREVENT sample (as presented in C. W. Ritchie et al., 2024). APOE- $\epsilon 4$ carriers did not differ significantly from non-carriers in any of their demographic characteristics (see Table 5). Correlations between demographics and cognitive measures are presented in Appendix 8.2.2. The demographic characteristics of the Benchmark sample have been described elsewhere (Chapter 3, Table 1).

Table 5: Demographic characteristics of the APOE genotyped sample.

Variable	Details	APOE genotyped sample total (N = 87)	APOE-ε4+ (n = 35)	APOE-ε4- (n = 52)	p
Age	M (±SD) Range	56.39 (±5.72) 44–67	56.34 (±5.98) 45–67	56.42 (±5.72) 44–66	0.934†
Gender	Female Male	50 (57.5%) 37 (42.5%)	18 (51.4%) 17 (48.6%)	32 (61.5%) 20 (38.5%)	0.475‡
Education (years)	M (±SD) Range	16.54 (±3.01) 9–23	16.06 (±3.24) 9–23	16.87 (±2.84) 10–22	0.170†
Education (level)	No formal education Lower secondary Upper secondary University/College Master's degree PhD	0 (0%) 5 (5.7%) 25 (28.7%) 35 (40.2%) 20 (23.0%) 2 (2.3%)	0 (0%) 3 (8.6%) 12 (34.3%) 13 (37.1%) 6 (17.1%) 1 (2.9%)	0 (0%) 2 (3.8%) 13 (25.0%) 22 (42.3%) 14 (26.9%) 1 (1.9%)	0.639‡
Dementia family history	No Yes	38 (43.7%) 49 (56.3%)	15 (42.9%) 20 (57.1%)	23 (44.2%) 29 (55.8%)	1.00†
APOE allele	ε2ε3 ε3ε3 ε2ε4* ε3ε4* ε4ε4*	8 (9.2%) 44 (50.6%) 2 (2.3%) 27 (31.0%) 6 (6.9%)			

Note: * APOE carriers; † Mann-Whitney U test; ‡ χ² test.

4.3.2 Associations between APOE-ε4 Status and Cognition

Descriptive statistics for cognitive outcomes, together with univariate associations between APOE-ε4 status and cognitive measures, are presented in Table 6. We ran covariate-adjusted analyses, predicting each cognitive measure from APOE-ε4 status in a regression model controlling for age, gender, years of education, and family history of dementia. As shown in the list of models in Table 7, having APOE-ε4 was linked to significantly better visual working memory (β [bootstrapped 95% CI] = -0.29 [-0.53, -0.06]; $t = -2.30$; $p = .025$) and no other effects were observed. After Bonferroni correction for five comparisons, this effect was no longer significant (under $\alpha = .05/5 = .01$). A sensitivity analysis suggested that this effect was not driven by participants incorrectly binding shapes with colours in their memory, but rather by their general misremembering of the stimuli (see Appendix

8.2.3 for more detail). Additionally, univariate post-hoc comparisons revealed that neither carriers nor non-carriers differed significantly in their cognition from the age-matched benchmark sample (see Figure 10, panels A–E).

Table 6: *Dependent cognitive measures in total sample and split by APOE-ε4 status.*

Variable	Genotyped sample total (n = 87)		APOE-ε4+ (n = 35)	APOE-ε4- (n = 52)	p
	M (± SD) / Count (%)	Complete cases (%)	M (± SD) / Count (%)	M (± SD) / Count (%)	
Model-based planning	0.24 (±0.28)	72 (82.8%)	0.23 (±0.28)	0.24 (±0.29)	0.891†
Processing speed (~ Trails A)	31.64 (±7.01)	76 (87.4%)	31.58 (±5.43)	31.68 (±7.89)	0.950†
Cognitive flexibility (~ Trails B)	69 (±18.4)	76 (87.4%)	66.64 (±15.18)	70.46 (±20.16)	0.352†
Visual working memory	0.79 (±0.08)	70 (80.5%)	0.81 (±0.07)	0.77 (±0.08)	0.028†*
Subjective memory problems		86 (98.9%)			1.00‡
No:	51 (59.3%)		21 (60.0%)	30 (58.8%)	
Yes:	36 (40.7%)		14 (40.0%)	21 (41.2%)	

Note: * $p < .05$; † *t*-test; ‡ χ^2 test.

Table 7: Summary of linear (respective logistic) regression models predicting cognitive performance (respective subjective memory problems) from APOE-ε4 status and covariates. APOE-ε4 status estimates highlighted in bold.

Cognitive outcome	Term	β / OR [bootstrapped 95% CI]	t / F	p
Model-based planning (n = 72)	(Intercept)	0.29 [-0.2, 0.88]	1.13	.263
	APOE-ε4+	-0.02 [-0.27, 0.19]	-0.16	.870
	Age	-0.02 [-0.26, 0.36]	-0.14	.893
	Gender	-0.22 [-0.5, 0.01]	-1.83	.072
	Education	-0.04 [-0.3, 0.19]	-0.32	.749
	DFH	-0.06 [-0.69, 0.41]	-0.25	.804
Processing speed (~ Trails A; n = 76)	(Intercept)	0.09 [-0.39, 0.8]	0.37	.715
	APOE-ε4+	0 [-0.22, 0.21]	-0.01	.993
	Age	0.28 [0.06, 0.52]	2.43	.018*
	Gender	-0.11 [-0.37, 0.11]	-0.96	.340
	Education	0.08 [-0.13, 0.3]	0.64	.522
	DFH	0.05 [-0.46, 0.5]	0.22	.830
Cognitive flexibility (~ Trails B; n = 76)	(Intercept)	0 [-0.48, 0.65]	-0.01	.988
	APOE-ε4+	-0.09 [-0.31, 0.11]	-0.76	.451
	Age	0.23 [0, 0.44]	1.93	.058
	Gender	-0.03 [-0.28, 0.2]	-0.29	.773
	Education	0.03 [-0.18, 0.29]	0.24	.811
	DFH	0.19 [-0.34, 0.66]	0.78	.439
Visual working memory (n = 70)	(Intercept)	0.39 [-0.09, 0.88]	1.50	.140
	APOE-ε4+	-0.29 [-0.53, -0.06]	-2.30	.025*
	Age	0.04 [-0.17, 0.27]	0.33	.741
	Gender	-0.14 [-0.37, 0.1]	-1.13	.264
	Education	0.03 [-0.23, 0.32]	0.23	.817
	DFH	0.03 [-0.43, 0.53]	0.12	.902
Subjective memory problems (n = 86) †	(Intercept)	0.96 [0.35, 2.91]	-0.09	.929
	APOE-ε4+	0.93 [0.55, 1.59]	-0.30	.766
	Age	0.9 [0.52, 1.55]	-0.43	.670
	Gender	0.74 [0.45, 1.21]	-1.31	.191
	Education	0.89 [0.54, 1.52]	-0.51	.610
	DFH	1.08 [0.36, 3.2]	0.17	.867

Note: Higher numbers indicate worse cognition. Gender was coded with male as reference point. * $p < .05$. † Model estimates from a logistic regression, i.e., odds ratios. DFH = Family history of dementia.

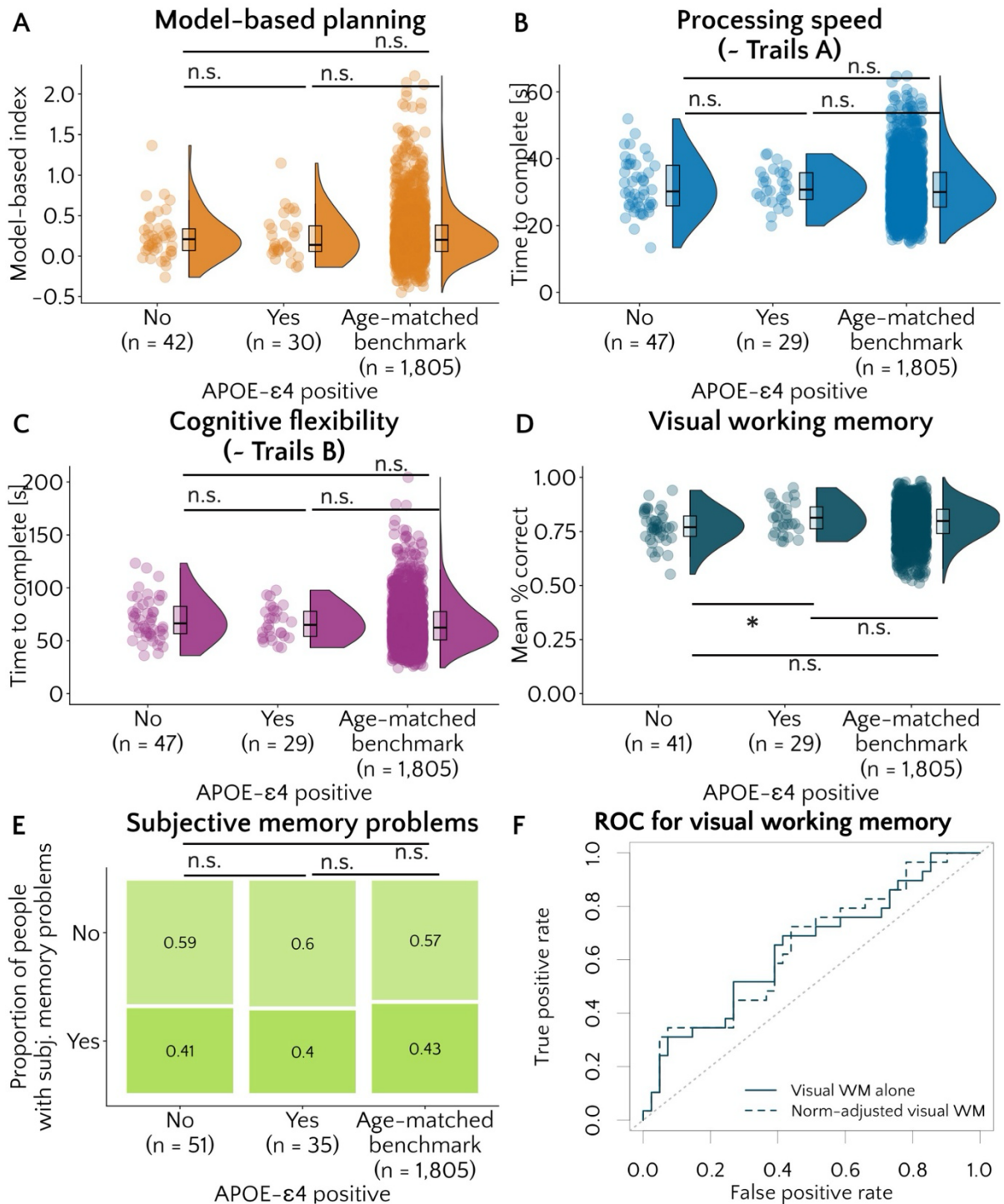


Figure 10: Raw scores on **(A)** model-based planning, **(B)** processing speed, **(C)** cognitive flexibility, **(D)** visual working memory, and **(E)** subjective memory problems, split by APOE-ε4 status in the APOE genotyped sample, with age-matched benchmark sample (n = 1,805) displayed for comparison. **(F)**: ROC curve for visual working memory predicting APOE-ε4 status. Visual working memory alone: AUC = 0.642; SE = 0.068. Norm-adjusted visual working memory (adjusting for age, gender, educational level, and family history of dementia based on normative benchmark data, N = 3,327): AUC = 0.648; SE = 0.067.

4.3.3 Classifying Genetic Groups Based on Visual Working Memory

We did ROC analysis to quantify the potential clinical utility of visual working memory, which was the only cognitive marker assessed here that showed a trend towards an association with APOE- ϵ 4 status. We found that visual working memory performance significantly predicted APOE- ϵ 4 status, either alone or when adjusted for norms (i.e., observed performance compared to that we would predict based on each individual's age, gender, education, and family history of dementia, see Analyses for details on calculation). The ROC curve is displayed in Figure 10, panel F. For visual working memory alone, AUC = .642 (SE = 0.068; bootstrapped 95% CIs = .507–.768); whereas for norm-adjusted visual working memory performance, AUC = .648 (SE = 0.067; bootstrapped 95% CIs = .515–.774). The confidence intervals of both AUCs overlapped, suggesting there was little value added by comparing visual working memory performance to normative data when predicting APOE status. A power analysis (using the formulas 2 and 3 in Obuchowski et al., 2004, p. 1123) showed that the present sample size ($n_{\text{cases}}/n_{\text{controls}} = 29/41$, with $\alpha = 0.05$, $1-\beta = 0.8$) was sufficient to detect what would constitute a clinically useful effect (i.e., AUC > .8)(Çorbacioğlu & Aksel, 2023), but that the smallest reliably detectable effect in this sample would constitute AUC = .689.

4.4 Discussion

Well-targeted, smartphone-administered cognitive assessments can be sensitive to a range of dementia risk factors in otherwise cognitively healthy adults, promising cheap and convenient early detection of AD risk at a large scale. We have previously established a set of such cognitive markers sensitive to ageing and lifestyle risk across the lifespan (Chapter 3), which can be self-administered remotely through the smartphone app *Neureka*. Here, we probed the associations between these markers and APOE- ϵ 4 in cognitively unimpaired middle-aged adults ($N = 87$). APOE- ϵ 4 carriers in our genotyped sample outperformed non-carriers in a visual working memory task, but this effect did not survive controlling for multiple comparisons. Additionally, performance on this visual working memory task was not able to classify APOE status with a clinically meaningful level of precision (AUC = .648), even when comparing individual performance against a large normative benchmark dataset ($N = 3,327$) collected from the general public. We found that APOE- ϵ 4 carriers did not perform significantly worse in either model-based planning (a hippocampus-reliant cognitive domain, Vikbladh et al., 2019) or in a gamified version of Trail-Making Test A or B, and were also not significantly more likely to report subjective memory problems. Together, these findings suggest limited clinical utility of the present set of smartphone-based measures in detecting genetic risk for AD.

The trend of APOE- ϵ 4 carriers in our study to perform better on visual working memory could be explained by multiple factors. First, the finding could be due to chance and the small sample size, supported by the fact that this result did not survive the Bonferroni correction. However, this type of correction is a rather conservative way of controlling for multiple comparisons, possibly obscuring smaller, albeit true effects. Genes constitute only one of the many factors shaping cognitive abilities, suggesting small effect sizes are likely (Marian, 2012; O'Donoghue et al., 2018). The second plausible explanation is that the current findings reflect a true positive effect of APOE- ϵ 4. Supporting this, previous studies by Lu et al. (2021) and Zokaei et al. (2020) in larger samples ($N = 398$, respective $N = 1277$) similarly found that APOE- ϵ 4 carriers outperformed non-carriers on a task that involved working memory for objects and their locations, reporting small

but significant effects ($OR = 0.86\text{--}0.93$, $p < .05$, respective partial $\eta^2 = 0.004$, $p = .05$). A replication of our current findings in a larger sample would be needed to provide more reliable evidence whether this previously reported positive effect truly extends to our measure of working memory for objects and their colours. If this is the case, these effects could be an example of antagonistic pleiotropy (Han & Bondi, 2008), whereby APOE- $\epsilon 4$ might exert positive effects on visual working memory in middle-aged carriers before they have developed a dementia prodrome, albeit such a positive effect on cognition is not confirmed and probably only limited to highly specific domains (Henson et al., 2020; Weissberger et al., 2018). Finally, sample specifics might be driving the current findings. Corroborating this, two previous studies in larger subsets of the PREVENT cohort also found better performance in APOE- $\epsilon 4$ carriers, although this was on a form-perception measure without a recall component in the first study ($N = 210$) (K. Ritchie et al., 2017), respective on a narrative recall task in the second study ($N = 162$) (McKiernan et al., 2020). Participants of the PREVENT study might represent atypical populations that are highly socially and cognitively active, which may both drive their participation in research and have a protective effect on their cognition. As such, this may outweigh the detrimental effects of APOE status, even if temporarily. Additionally, PREVENT participants have a higher-than-average rate of APOE- $\epsilon 4$ prevalence (C. W. Ritchie et al., 2024), as the study intentionally targets individuals with family history of dementia (K. Ritchie et al., 2017). However, PREVENT participants are not informed of their APOE status and in the current study, we found no relationship between family history of dementia and cognition or family history of dementia and APOE status, which suggests there is little evidence that atypically active APOE- $\epsilon 4$ carriers self-selected into the study more often than atypically active non-carriers.

The lack of an effect for model-based planning was unexpected. One possible explanation is that the current study was underpowered to detect a true effect or the lack thereof. This is the first study to our knowledge to directly test the associations between APOE status and model-based planning, albeit previous studies have repeatedly shown an association between APOE status and impaired performance on the related domain of navigation abilities in middle-aged to older adults (Berteau-Pavy et al., 2007; Coughlan et al., 2019, 2020; Laczó et al., 2014).

However, a very recent meta-analysis by Daly et al. (2024) put into question the replicability of previous findings and their generalizability across the lifespan, as they did not find a significant effect of APOE status on navigation when averaging across a range of wayfinding tasks, which included both egocentric and allocentric navigation. Our estimates of the sample size required for the current analyses were based on a small study ($N = 60$) that reported a large effect for allocentric navigation abilities (i.e., $d = 0.5$, based on data provided by Coughlan et al., 2019), a finding possibly due to chance or specific to allocentric navigation tasks. Whatever the true effect is, our current findings suggest the effects of APOE- $\epsilon 4$ on model-based planning are not compelling enough to attempt to detect APOE status using model-based planning tasks as a cognitive marker. Another possible explanation for the lack of an effect in our study could be that putative impairments to allocentric navigation ability in APOE- $\epsilon 4$ carriers are driven by impairments to neural substrate other than that implicated in the present model-based planning task. Although areas of the hippocampus are important for both allocentric navigation and model-based planning (Vikbladh et al., 2019), the neural underpinnings of navigation are more extensive than that, including the entorhinal cortex and particularly grid cells (Coughlan et al., 2018). Finally, the explanation could lie in small task differences between our model-based planning task (Donegan et al., 2023) and the exact model-based planning task used in the existing imaging evidence (Vikbladh et al., 2019), which could be linked to differences in what brain regions were implicated. While a validation study (Donegan et al., 2023) has shown a modest correlation ($r = .4$) between model-based planning index measured with the current smartphone-based task and that measured with the more traditional, computer-based task version, a validation of our newer, smartphone-based measure against neuroimaging data is yet to be done.

While the current study found little evidence for the usefulness of our measures as cognitive screening for APOE- $\epsilon 4$, it contributes more generally to the understanding that genetic effects on cognition are complex and non-specific, leading to unreliable outcomes, some of which might even be positive. This has several important implications. First of these is related to the disclosure of an individual's APOE status, which is currently not encouraged or practiced (Galluzzi

et al., 2022; Largent et al., 2021), although this is likely to change as advances in research and treatment make disclosure more useful (Honig et al., 2023; Solomon et al., 2018). Existing research suggests that disclosure is not necessarily an issue and can even help motivate lifestyle changes (Largent et al., 2021), however, it is crucial to accompany genetic status disclosure with appropriate education on the non-deterministic nature of genetic risk factors (Harkins et al., 2015). Second, researchers need to keep in mind the complex and still poorly understood mechanisms behind the link between APOE- ϵ 4 and cognitive decline. APOE- ϵ 4-reducing genetic therapies have been suggested as therapeutic strategies for AD, showing success in animal models and promise in humans (Y. Li et al., 2022). However, until the mechanisms are better understood, efforts to reduce APOE- ϵ 4 in humans are not sufficiently justified as a potential global health intervention, as there is still too much we do not know.

The current study has several limitations. First, as mentioned previously, the sample size in the current study was not sufficient to detect small effects, which are most probable in the case of genetic effects on cognition (Marian, 2012). While the current data were sufficient to test the *clinical* usefulness of using our existing cognitive tasks to detect an increased genetic risk of AD, the effects we found were unreliable due to the small sample size and need to be replicated in future larger studies. Second, interactions of APOE status with other factors such as gender, education, or lifestyle are common (Altmann et al., 2014; Caselli et al., 2011; X. Li et al., 2019; Polsinelli et al., 2023; Sapkota et al., 2017; X. Wang et al., 2019) and could have obscured the current findings, concealing a possibly true effect. Although we controlled for covariate effects, future studies should investigate gene interactions with such factors. Finally, current data do not provide much insight into the mechanistic explanations of APOE- ϵ 4 and its effects on cognitive decline. To study APOE- ϵ 4 effects on cognition, the ideal study design should be longitudinal, differentiating between prodromal and phenotype effects by including participants across the lifespan and excluding those who develop dementia on follow-up (O'Donoghue et al., 2018). Additionally, it is possible that while APOE- ϵ 4 was not reliably linked to cognitive outcomes cross-sectionally, it may still predict cognitive change over time, a hypothesis supported by mixed but perhaps even more consistent previous findings (O'Donoghue et al., 2018). Such

a hypothesis could be even more easily tested longitudinally and at a large scale with a smartphone approach as compared to in-person assessments.

To conclude, in middle-aged adults without dementia, we did not find a clinically significant relationship between APOE- ϵ 4 and smartphone-based cognitive assessments, spanning multiple domains including model-based planning, visual working memory, executive functions, and subjective memory. Overall, this proof-of-concept study gives little evidence that can be used to detect genetic risk for AD, although they are sensitive to lifestyle risk factors, as shown in Chapter 3. To more reliably identify at-risk individuals, smartphone assessments of modifiable risk factors and cognition could be complemented with genetic testing or other biomarker assessment.

Chapter 5: Sex differences in modifiable risk factors associated with cognition: A machine-learning approach

5.1 Introduction

Cognitive decline in ageing is common but not inevitable, with evidence suggesting that almost a half of dementia cases could be prevented or delayed by reducing modifiable risk factors across the lifespan (Livingston et al., 2024). While modifiable risk factors are increasingly becoming the targets of dementia prevention programmes (e.g., Baker et al., 2024; Ngandu et al., 2015), these programmes are typically not tailored to account for sex differences. This is potentially important because men and women differ in their cognitive functioning over the lifespan (Avila et al., 2019; Daseking et al., 2017; de Frias et al., 2006; Ferreira et al., 2014; Herlitz et al., 1999; LaPlume, McKetton, Anderson, et al., 2022; McCarrey et al., 2016; Proust-Lima et al., 2008; Siedlecki et al., 2019), in their risk of dementia (Huque et al., 2023; Wolters et al., 2020), and in the prevalence of different dementia types (Cao et al., 2020; Huque et al., 2023). At the same time, men and women have different prevalence and number of potentially modifiable risk factors (Anstey et al., 2021; Geraets & Leist, 2023; Gong et al., 2021; LaPlume, McKetton, Anderson, et al., 2022; Liu et al., 2020), suggesting that the mechanisms linking risk factors with cognition could differ between men and women. These could include both gender-related explanations like socio-economic inequality (Ilinca & Suzuki, 2021; Rocca et al., 2014) or differences in the diagnosis and treatment of cardiovascular risk factors (Chêne et al., 2015), and sex-related explanations such as genetic differences (Altmann et al., 2014), the effects of sex hormones (Gurvich et al., 2018) or survival bias (Chêne et al., 2015). Better understanding the differences in how risk factors affect cognition in men vs. women could help improve the tailoring and success of public health interventions.

Although controlling for sex is common in cognitive research, to truly understand the interplay of sex with risk factors and cognition, studies need to directly examine the interaction effects or stratify by sex (Gurvich et al., 2018). Existing evidence of this type is far from uniform. Some studies in older adults report that compared to men, cognition in women is significantly associated with an overall

higher number of risk factors (Liu et al., 2020; McDermott et al., 2017). In contrast, a recent US study using nationally representative survey data on risk factor prevalence from $N = 378,615$ individuals across the lifespan estimated that most of the modifiable risk factors alone did not differ in their importance for dementia risk in men and women (Nianogo et al., 2022). However, the authors also found that the same factors in combination accounted for a larger dementia risk in men, as shown by greater combined population-attributable risk. Importantly, these differences were estimated by combining risk factor prevalence data with meta-analytical data on risk factors — dementia associations rather than by recording true dementia cases in the dataset. Other studies that directly assessed cognitive outcomes suggest that cumulative effects of risk factors do not differ by sex, because 1) the associations between the number of risk factors a person has and cognitive performance were similar in men and women (Deckers et al., 2019; but cf. LaPlume, McKetton, Anderson, et al., 2022), and 2) lifestyle interventions targeting multiple risk factors simultaneously did not differently affect cognition in men and women (Sindi, Kåreholt, et al., 2021). However, without modelling the effect of sex on each risk factor individually, there is a risk that opposing effects of different risk factors might obscure effects.

Recent work has started to examine this and found early evidence for differences in the particular risk factors that are important for women vs. men. For example, women seem to be more susceptible to the detrimental effect of midlife hypertension on cognitive impairment (Anstey et al., 2021; Lyu & Kim, 2016; Singh-Manoux & Marmot, 2005) and dementia risk (Gilsanz et al., 2017; Hébert et al., 2000; but cf. Geraets & Leist, 2023; Nianogo et al., 2022), despite the fact that midlife hypertension has higher prevalence in men (Nianogo et al., 2022). Other possibly sex-specific risk factors include hearing (Lyu & Kim, 2016) and lower education (Sindi, Kåreholt, et al., 2021) in women, as well as low social engagement in men (Anstey et al., 2021; Maccora et al., 2021). The evidence is mixed for other risk factors, like in the example of physical inactivity, (Barha et al., 2017; Lyu & Kim, 2016; Sindi, Kåreholt, et al., 2021; cf. Fallah et al., 2009; Tolppanen et al., 2015; Anstey et al., 2021). Furthermore, some authors have argued that depression might be a stronger risk factor in women (Sindi, Toopchiani, et al., 2021), as depression symptoms and diagnoses are more

common in women than men across the lifespan (Salk et al., 2017). However, the evidence for sex-specific effects of depression on cognitive decline is mixed, too, with some studies in older adults finding a detrimental effect only in women (Kim et al., 2015; Nianogo et al., 2022; Sachdev et al., 2012; Sundermann, Katz, et al., 2017), other studies in men only (Dal Forno et al., 2005; Maccora et al., 2021; McDermott et al., 2017) and other not finding a sex-specific effect at all (Lyu & Kim, 2016; Sindi, Kåreholt, et al., 2021). Finally, yet other risk factors do not seem to differ in their effects on cognition between men and women, for instance smoking (Gong et al., 2021).

One reason for the inconclusiveness of this body of literature could lie in limitations of the statistical approaches adopted. It is increasingly recognised that we need to study the effects of modifiable risk factors on cognition using multivariate analyses (Borgeest et al., 2020; Demnitz et al., 2023) as some modifiable risk factors can have a strong association with dementia and cognitive decline in combination even when they do not all have an independent significant effect (Kesse-Guyot et al., 2014; M. C. Norton et al., 2012; Peters et al., 2019). Perhaps more importantly, many of the known risk factors often co-occur, which requires a thoughtful approach to the handling of multicollinearity. A second problem may simply be that the impact of dementia risk factors on cognition varies considerably across different samples and this produces low generalisability of findings (Demnitz et al., 2023; Mukadam et al., 2019). One way we could remedy both these issues is by using machine-learning methods such as regularised regression, which is a robust statistical approach that controls for multicollinearity and facilitates model generalisability, while enabling to rank the importance of predictors for the given outcome. Initial work has shown the potential of machine-learning for studying the effects of dementia risk factors on cognitive decline (Casanova et al., 2020; Fittipaldi et al., 2024; McDermott et al., 2017; You et al., 2022).

In this study, we investigated which modifiable risk factors were specifically important for cognitive impairment in men vs. women, across the lifespan. We leveraged a large dataset collected via a smartphone app to 1) learn about the relative importance of these previously established risk factors when considered all together, 2) learn how much variance they together explain, and 3) test the interactions of these risk factors with sex. We assessed cognition using a variety

of previously validated cognitive smartphone measures sensitive to dementia risk factors. We assessed subjective and objective cognition separately, as we have previously found that subjective cognition was more closely tied to modifiable risk factors (Chapter 3), although there is little evidence for sex-specific risk factors for subjective cognition (Schliep et al., 2022).

5.2 Methods

5.2.1 Participants and Procedure

Participants were recruited from the general public via the citizen-science research platform *Neureka*, which is a smartphone app freely available from app stores. All *Neureka* participants provide informed consent at sign-up, i.e. before participating in any research. Between June 2020 and March 2024, $N = 10,660$ *Neureka* users that were over 18 years old and without a self-reported diagnosis of dementia engaged with a module in the app called the ‘Risk Factors’ challenge. This module takes about 45 minutes to complete in full but can be completed over time, at participants’ leisure. It contains three gamified cognitive tasks, which are interspersed with two blocks of questionnaires that assess subjective cognition and potentially modifiable risk for dementia. The tasks and questionnaire blocks are presented in a semi-randomised order.

Of the initial sample, we excluded participants with incomplete cognitive data ($n = 5,560$, i.e. 52.2%). Furthermore, we only retained data of participants who identified as cisgender male ($n = 1,653$, i.e. 32.4%) or cisgender female ($n = 3,342$, i.e. 65.5%). This decision was motivated by the relatively small size of the transgender and non-binary gender subgroup ($n = 105$, 2%), which would have resulted in a low-variance predictor after categorical encoding. Following standard procedures for data preprocessing used in machine learning (Kuhn & Johnson, 2013), our pipeline excluded categorical predictors that were endorsed by fewer than 10% participants, as these contribute little to model performance and may introduce instability (see Figure 11 for more detail on our preprocessing pipeline).

As a result, the final sample included $N = 4,995$ individuals, and the gender variable became de facto equivalent to binary sex.

5.2.2 Measures

Objective cognition. We assessed objective cognition through two gamified cognitive tasks in the Risk Factors module, which together give the three measures of visual working memory, processing speed, and cognitive flexibility. Another task in the Risk Factors module assesses model-based planning and was not used in the current study as it previously has not shown much sensitivity to risk factors for dementia (Chapter 3). For details on the task structure, task validation, and exclusion criteria, please see Chapters 2 and 3.

In brief, visual working memory was assessed as mean proportion of correct selections across 24 trials of a task with graded difficulty levels where participants memorise an array of symbols of varying type (letters vs. abstract shapes, single colour vs. coloured) and then find and select these symbols from a 4x5 grid. Processing speed, respective cognitive flexibility was assessed as mean time to complete three trials of a gamified trail-making task test in which participants tap labelled stars in ascending order as fast as they can, selecting numbers (version A), respective switching between numbers and letters (version B). Objective cognition was calculated from these measures as an average of the z-scores for visual working memory, processing speed, and cognitive flexibility.

Subjective memory problems. Subjective cognition was assessed through a single-item questionnaire in the Risk Factors module, in which participants respond Yes/No to the question ‘Do you currently experience memory problems?’ This screening of subjective memory, while not comprehensive (Burmester et al., 2016), has previously been shown to be more sensitive to risk factors in our citizen scientist sample than any objective gamified cognitive measure per se (Chapter 3).

Risk factors. We assessed a set of previously established risk factors for dementia (Jafari et al., 2019; Kuźma, Lourida, et al., 2018; Livingston et al., 2020; Scarabino et al., 2016; A.-Y. Wang et al., 2023) through a set of questionnaires in the Risk Factors module. After excluding predictors with low variance (i.e., diabetes and history of stroke, which were endorsed in < 10% cases), the final set of risk factors included education (achieved level), socioeconomic status (Subjective SES Scale; Adler et al., 2000), depression (CES-D scale; Radloff, 1977), loneliness (UCLA Loneliness Scale; Russell et al., 1978), social network size (Lubben Social Network Scale; 1988), hearing handicap (HHIE-S; Ventry & Weinstein, 1983), exercise (Godin-Shephard Leisure-Time Physical Activity Questionnaire; Godin, 2011), and bespoke questions about tinnitus, hypertension, smoking, and parental history of dementia. Further details on the questionnaires are summarised in Table 8 of the Results section.

5.2.3 Statistical Analyses

Data preparation. Analyses were conducted using R Statistical Software v4.4.1 (R Core Team, 2024). Analysis scripts were uploaded to a private GitHub repository (https://github.com/amrosicka/multivariate_analyses.git) and can be made available upon request. Prior to analyses, some variables were reverse-coded so that in all cases higher values indicate worse cognition or higher risk. Sex was coded with 'male' as the reference category.

The full data processing and data analysis pipeline is shown in Figure 11. We split the total sample into 80/20 train/hold-out set stratified by age and sex. In the train set ($N = 4001$), we imputed missing data of the predictors using median/mode imputation, and then carried over the same median/mode values to impute missing data in the hold-out set ($N = 994$). Next, we mean-centred and z-transformed continuous variables in the train set, and then carried over the mean and standard deviation values to mean-centre and z-transform the hold-out set. Finally, using the resultant z-scores, we computed the composite scores for objective cognition, separately in each set, as described above.

Model training and cross-validation. We trained regression and classification elastic net models in the train set predicting objective and subjective cognition, respectively. We built three levels of models using different sets of predictors: 1) 'benchmark models' with sex and age only, 2) 'risk factor models' with sex, age, and dementia risk factors, and 3) 'interaction models' with all the aforementioned predictors, plus interactions between sex and each predictor.

We employed elastic net models, a type of machine learning models that combine lasso and ridge regularisation to penalise (i.e., reduce or nullify) regression coefficients (J. H. Friedman et al., 2010). The advantage of this approach is that it facilitates parsimony and model generalisation while effectively handling multicollinearity. For model training, we used repeated nested cross-validation with 10 repeats and 10 inner x 10 outer folds. Repeated nested cross-validation is an advanced model cross-validation technique whereby hyperparameter tuning and model selection are fully separated from model performance evaluation by iteratively partitioning data into inner and outer folds (used for hyperparameter tuning and performance evaluation, respectively) (Lewis et al., 2023). This method leads to less biased estimates and reduced overfitting. Model selection was done using RMSE for regression and logarithmic loss (logLoss) for classification models. A final round of cross-validation was then performed on the non-partitioned data to fit a final model that could then be used for predictions on external data. We extracted performance metrics for each model: R^2 , RMSE, and MAE for the regression models, and AUC, Accuracy, Balanced Accuracy, and logLoss for the classification models, with R^2 and AUC being of primary interest.

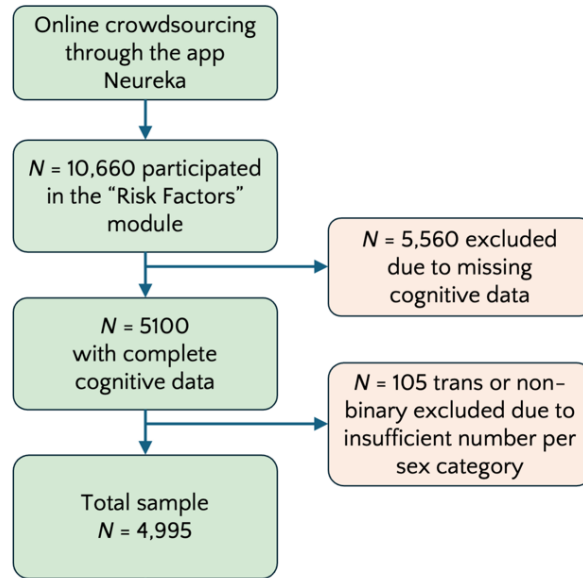
Model performance evaluation and comparison. Models were individually evaluated based on their cross-validated performance metrics, before being subjected to validation on fully held-out unseen data using the hold-out set ($N = 994$, i.e., 20% of the total sample). In both the training and hold-out set, we compared the performance (R^2 for regression and AUC for classification) of models with three different sets of predictors — 1) benchmark, 2) sex, age, and risk factors, and 3) sex, age, and risk factors with sex interactions. We used a permutation analysis for significance testing, whereby we compared the true

difference in the performance (R^2 for linear objective cognition models and AUC for binary subjective memory problems) of two models against a null distribution. To generate this null distribution of differences, we generated shuffled versions of the models, each containing 50% of the predictions from model A and 50% of the predictions of model B. We computed the R^2 /AUC for each model's predictions using the observed (true) data and calculated a difference between both shuffled models. We repeated this 1000 times to generate a distribution and derived a p -value based on the number of times the true difference was larger than the shuffled data. Differences were significant if the true difference outperformed the shuffled data 95% of the time.

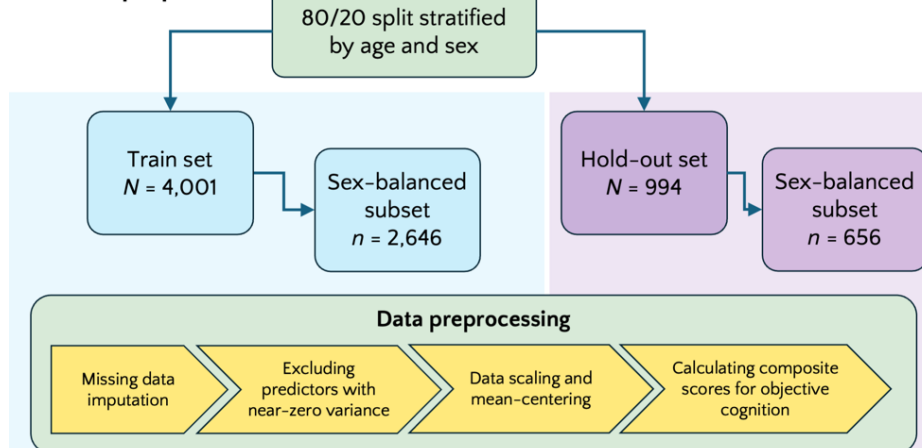
Predictor importance. We calculated mean Shapley additive explanation (SHAP) values to assess variable importance for the averaged final models (Lundberg & Lee, 2017). Mean SHAP values are a model-agnostic metric that can be calculated for each predictor, which represents the average marginal contribution of the given predictor, i.e., how much model performance would change if it was eliminated.

Sensitivity analyses. 66.9% of the sample were women. To test whether the sample size imbalance biased the results, we re-ran the analyses in a subset of the sample, whereby the size of the women sample was reduced to the size of the men sample, matching on age.

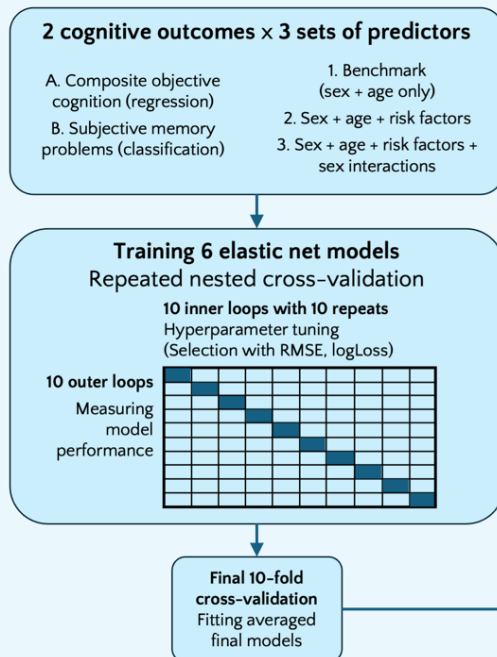
A. Data collection



B. Data preparation



C. Model training



D. Model testing

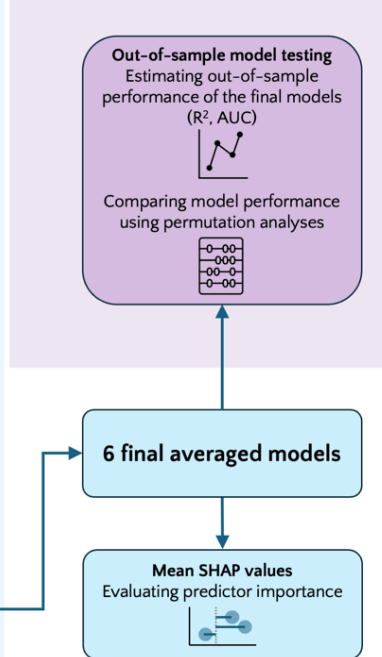


Figure 11: Data collection, processing, and analysis pipeline. **(A)** Data was collected online through the smartphone app Neureka, where $N = 10,660$ participants initiated the 'Risk Factors' module. After excluding participants with missing cognitive data ($N = 5,560$) and trans or non-binary participants due to insufficient sample size ($N = 105$), the final sample included $N = 4,995$ participants. **(B)** The dataset was split 80/20 into training ($N = 4,001$) and hold-out ($N = 994$) sets, used in the main analyses. For sensitivity analyses, these were then further reduced to sex-balanced subsets of $n = 2,646$ in the training set and $n = 656$ in the hold-out set. Data preprocessing included missing data imputation with median/mode, exclusion of predictors with near-zero variance, data scaling and mean-centering, and calculation of composite scores for objective cognition. Both missing data imputation and data scaling were done carrying over values from train to hold-out set to prevent leakage. **(C)** Composite objective cognition, respective subjective memory problems were predicted using elastic net models (regression, respective classification) with three sets of predictors: (1) sex and age only, (2) sex, age, and risk factors, and (3) sex, age, risk factors, and sex interactions, totalling 6 models of interest. Model training involved repeated nested cross-validation with 10 inner loops for hyperparameter tuning using RMSE or log-loss and 10 outer loops for performance evaluation. Final models were fitted using 10-fold cross-validation on the full training data. **(D)** Out-of-sample performance was assessed in the hold-out set using R^2 for regression and AUC for classification models, with permutation analyses to compare model performance. Predictor importance was evaluated in the training set using mean SHAP (SHapley Additive exPlanations) values, providing insights into the contribution of individual risk factors to cognitive outcomes.

5.3 Results

5.3.1 Descriptive Sex Differences

Table 8 summarises the descriptive statistics for the total sample ($N = 4,995$) and for men and women separately. Participants in the total sample were aged 18–84, of which 66.9% were women. Men in the total sample were on average older than women ($t(3428.6) = 6.0627$; $p < .001$) and reported having smaller social networks ($t(2992.2) = 3.2671$; $p = .001$). They were also more likely to have tinnitus ($\chi^2(1, N = 4,995) = 4.2566$; $p < .0391$) and hypertension ($\chi^2(1, N = 4,995) = 42.885$; $p < .001$). Women reported lower socio-economic status (SES; $t(3134.9) = -7.4944$; $p < .001$); more depression symptoms ($t(3581.8) = -10.669$; $p < .001$), more loneliness ($t(3181.6) = -4.7677$; $p < .001$), and less exercise ($t(2926.7) = -6.8002$; $p < .001$). There were no significant differences in education, family history of dementia, hearing handicap, or history of smoking.

Data distributions of each risk factor and cognitive measure by sex are plotted in Figure S13 in Appendix 8.3.1. For univariate associations between variables, please see Figure S14 in Appendix 8.3.2.

Table 8: *Descriptive statistics of the total sample and women vs. men separately.*

Variable	Details	Total sample (N = 4,995)	Women (n = 3344)	Men (n = 1651)	p*
Age	M (±SD) Range	45.68 (±14.49) 18–84	44.82 (±14.66) 18–84	47.41 (±13.99) 18–82	< .001
Education (level)	1 = 'No formal education'	70 (1.4%)	51 (1.5%)	19 (1.2%)	.09776
	2 = 'Lower secondary'	385 (7.7%)	243 (7.3%)	142 (8.6%)	
	3 = 'Upper secondary'	1336 (26.7%)	918 (27.5%)	418 (25.3%)	
	4 = 'University/ College degree or equivalent'	2047 (41.0%)	1296 (41.7%)	651 (39.4%)	
	5 = 'Master's degree'	955 (19.1%)	612 (18.3%)	343 (20.8%)	
	6 = 'PhD'	202 (4.0%)	124 (3.7%)	78 (4.7%)	
Dementia family history	'Was your biological mother ever diagnosed with dementia?' (Yes/No) 'Was your biological father ever diagnosed with dementia?' (Yes/No)	Yes (either parent or both): 835 (16.7%) No: 4160 (83.3%)	Yes (either parent or both): 582 (17.4%) No: 2762 (82.6%)	Yes (either parent or both): 253 (15.3%) No: 1298 (84.7%)	.0638
Socio- economic status (SES)	The Subjective SES scale (1 item) (Adler et al., 2000) — participants place themselves on a ladder representing where they stand in their country with respect to money, education, and respected jobs	M = 5.96 SD = 1.77 Range = 1–10 Missing: 183 (3.7%)	M = 5.83 SD = 1.75 Range = 1–10 Missing: 120 (3.6%)	M = 6.24 SD = 1.77 Range = 1–10 Missing: 63 (3.8%)	< .001

Depression	Centre for Epidemiologic Studies Depression (CES-D) scale (20 items) (Radloff, 1977) — measuring frequency of symptoms over the past week	M = 18.00 SD = 13.55 Range = 0–60 Missing: 41 (0.8%)	M = 19.38 SD = 13.86 Range = 0–60 Missing: 25 (0.7%)	M = 15.21 SD = 12.45 Range = 0–60 Missing: 16 (1.0%)	< .001
Loneliness	UCLA Loneliness Scale (20 items) (Russell et al., 1978) —measuring general loneliness	M = 20.89 SD = 15.64 Range = 0–60 Missing: 194 (3.7%)	M = 21.63 SD = 15.69 Range = 0–60 Missing: 123 (3.7%)	M = 19.36 SD = 15.47 Range = 0–60 Missing: 71 (4.3%)	< .001
Social network	Lubben Social Network Scale (6 items) (Lubben, 1988) — measuring social engagement with family and friends	M = 15.89 SD = 6.50 Range = 0–30 Missing: 209 (4.2%)	M = 16.11 SD = 6.38 Range = 0–30 Missing: 135 (4.0%)	M = 15.45 SD = 6.72 Range = 0–30 Missing: 74 (4.5%)	< .001
Hearing handicap	Hearing Handicap Inventory for the Elderly — Screening version (HHIE-S) (10 items) (Ventry & Weinstein, 1983) — measuring the emotional and social consequences of hearing impairment in adults	M = 5.52 SD = 7.48 Range = 0–40 Missing: 29 (0.6%)	M = 5.46 SD = 7.56 Range = 0–40 Missing: 18 (0.5%)	M = 5.62 SD = 7.31 Range = 0–40 Missing: 11 (0.7%)	.4778
Exercise	Godin-Shephard Leisure-Time Physical Activity Questionnaire (3 items) (Godin, 2011) — measuring frequency of strenuous, moderate, and mild exercise per week	M = 32.63 SD = 25.75 Range = 0–119 Missing: 172 (3.4%)	M = 30.82 SD = 24.88 Range = 0–119 Missing: 108 (3.2%)	M = 36.31 SD = 27.07 Range = 0–119 Missing: 64 (3.9%)	< .001
Tinnitus	‘In the past year have you had buzzing, ringing or noise (tinnitus) in your ears that lasts longer than 5 minutes?’ (No / Yes, in the past week / Yes, not in the past week / Don’t know)	Yes, in the past week/ Yes, not in the past week: 1424 (28.8%) No/Don’t know: 3520 (71.2%) Missing: 51 (1.0%)	Yes, in the past week/ Yes, not in the past week: 921 (27.9%) No/Don’t know: 2384 (72.1%) Missing: 39 (1.2%)	Yes, in the past week/ Yes, not in the past week: 503 (30.7%) No/Don’t know: 1136 (69.3%) Missing: 12 (0.7%)	.0391

Hypertension	'Have you been diagnosed with high blood pressure (hypertension) by a doctor?' (Yes / No)	Yes: 765 (15.5%) No: 4179 (84.5%) Missing: 51 (1.0%)	Yes: 433 (13.1%) No: 2872 (86.9%) Missing: 39 (1.2%)	Yes: 332 (20.3%) No: 1307 (79.7%) Missing: 12 (0.7%)	< .001
Smoking history	'Please indicate which of the following best describes you:' (1. Non-smoker/ 2. Ex-smoker/ 3. Current smoker/ 4. Unknown)	Current smoker/ Ex-smoker: 1717 (35.0%). Non-smoker/ Unknown: 3183 (65.0%) Missing: 95 (1.9%)	Current smoker/ Ex-smoker: 1123 (34.3%). Non-smoker/ Unknown: 2151 (65.7%) Missing: 70 (2.1%)	Current smoker/ Ex-smoker: 594 (36.5%). Non-smoker/ Unknown: 1032 (63.5%) Missing: 25 (1.5%)	.1233

Note: * Women and men were compared on continuous variables using *t*-tests and on categorical variables using the chi-square test.

5.3.2 Model Performance

Performance of all models fitted in the training sample is shown in Table 9. Benchmark model with sex and age alone explained 12.7% of variance in objective cognition and sex and age alone predicted subjective memory problems with an AUC = .570. In comparison, models including sex, age, and risk factors explained 16.1% of variance in objective cognition and predicted subjective memory problems with an AUC = .723. Against our expectation, adding sex interactions to the model did not increase model performance. The interaction models were nominally (though trivially) worse, explaining 16.0% of variance in objective cognition and predicted subjective memory problems with an AUC = .721. Overall, and as expected, the models performed slightly worse in the hold-out sample (Table 9). The benchmark model with sex and age alone explained 10.1% of variance in objective cognition and sex and age alone predicted subjective memory problems with an AUC = .552. The model that additionally included risk factors explained 14.1% of variance in objective cognition and an equivalent model predicted subjective memory problems with an AUC = .699. A permutation test confirmed this improvement in performance was significant (both $p < .001$; see Figure 12, panels A and B). The interaction model explained 13.6% of variance in objective cognition, which was again nominally

worse than the model without interactions (but not statistically different, $p = .657$; see Figure 12, panel C). Adding sex interactions to the equivalent model for subjective memory problems also did not significantly increase performance compared to the model with interactions (AUC = .701; probability that the true effect was larger than null distribution $p = .472$; see Figure 12, panel D).

Table 9: Model performance results from nested repeated cross-validation predicting objective vs. subjective cognition in the training sample ($N = 4,001$) and from applying the final models in the hold-out sample ($N = 994$).

Training set							
Model	Objective cognition			Subjective memory problems			
	Regression model performance			Classification model performance			
	MAE	R ²	RMSE	AUC	Accuracy	Balanced accuracy	Log-Loss
Benchmark (sex + age)	0.608	.127	0.774	.570	.599	.504	0.666
Sex + age + risk factors	0.597	.161	0.759	.723	.677	.638	0.595
Sex + age + risk factors + sex interactions	0.597	.160	0.760	.721	.677	.637	0.597
Hold-out set							
Model	Objective cognition			Subjective memory problems			
	Regression model performance			Classification model performance			
	MAE	R ²	RMSE	AUC	Accuracy	Balanced accuracy	Log-Loss
Benchmark (sex + age)	0.596	.107	0.769	.552	.593	.498	0.670
Sex + age + risk factors	0.581	.141	0.754	.699	.652	.610	0.609
Sex + age + risk factors + sex interactions	0.581	.136	0.755	.701	.651	.610	0.609

Note: R^2 = R squared; RMSE = Root mean square error; MAE = Mean absolute error; AUC = Area under the classification curve, LogLoss = Logarithmic loss.

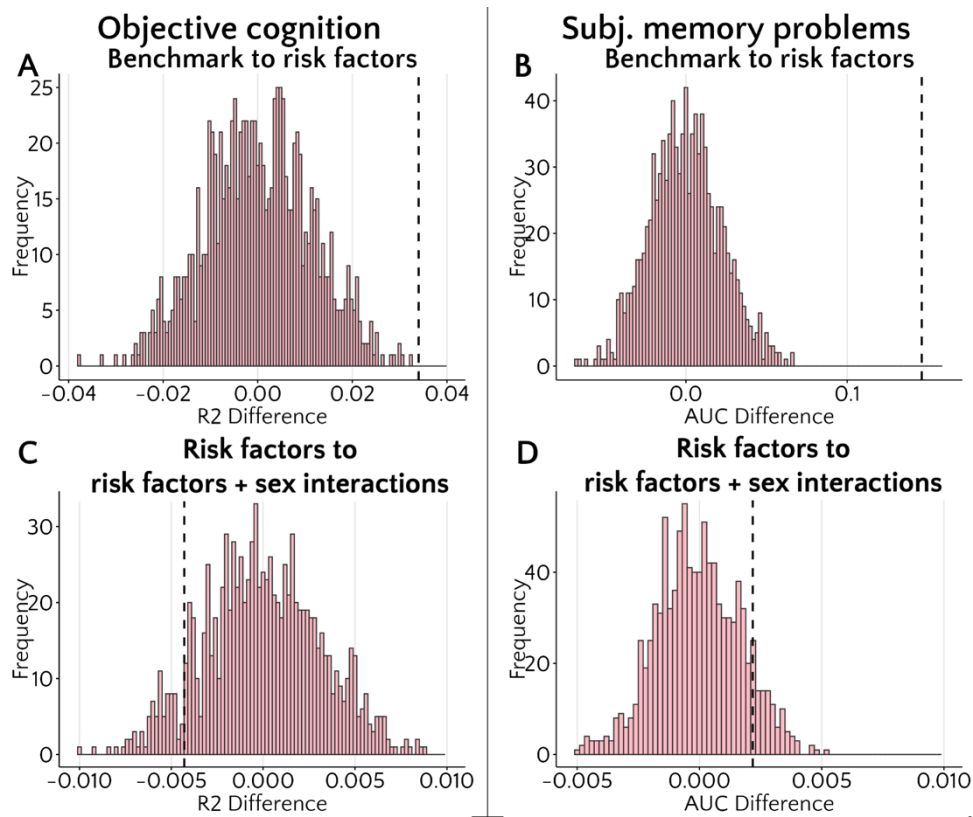


Figure 12: Permutation analysis of between-model differences in performance, based on the hold-out sample ($n = 994$). Bars represent the null distribution of 1000 permuted performance differences between models. Dashed lines represent true performance differences between models, specifically: **(A)**: R^2 difference between the benchmark model and the risk factors model; **(B)**: AUC difference between the benchmark model and the risk factors model; **(C)**: R^2 difference between the risk factors model and the interaction model; **(D)**: AUC difference between the risk factors model and the interaction model.

5.3.3 Predictor Importance

In the benchmark models, both sex and age were retained as predictors, with final hyperparameters being $\alpha = .0527$ and $\lambda = 0.0001$ for objective cognition and $\alpha = .0001$ and $\lambda = 0.0001$ for subjective memory problems, which means a prevalently ridge, respective fully ridge model with the lowest possible penalty parameter, i.e., a model very close to simple linear/logistic regression without regularisation. The mean SHAP values for the benchmark models (Figure 13, panel A) show that being a woman was linked to better cognition, but higher likelihood of subjective memory problems, and higher age was linked to both subjectively and objectively worse cognition. Age was more important for objective cognition than sex. This

pattern for age and sex held when brought forward to the other models as covariates.

In the risk factors models, all 13 predictors were retained, with final hyperparameters being $\alpha = .0527$ and $\lambda = 0.0001$ for objective cognition and $\alpha = .0001$ and $\lambda = 0.0001$ for subjective memory problems, which means a prevalently ridge, respective fully ridge model with the lowest possible penalty parameter. The mean SHAP values of the risk factors models are presented in Figure 13, panel B. The most important predictors of worse objective cognition were age, depression, and lower education. *Better* objective cognition in this model was predicted by being a woman, but also by tinnitus, less exercise, higher loneliness levels, and lower SES. The most important predictors of having subjective memory problems were depression, hearing handicap, closely followed by age, low SES, less exercise, and then other predictors. The low absolute mean SHAP value for hypertension suggests low importance of this predictor for the model. Notably, smaller social network was the only predictor linked to a *lower* likelihood of subjective memory problems.

In the risk factors model with interactions, four predictors were omitted for objective cognition and all predictors were retained for subjective memory problems. Final hyperparameters again suggest a prevalently ridge ($\alpha = .0527$), prevalently low penalty ($\lambda = 0.0001$) model for objective cognition and a fully ridge ($\alpha = .0001$), very low penalty ($\lambda = 0.0001$) model for subjective memory problems. The mean SHAP values of the interaction models are presented in Figure 13, panel C. The pattern of variable importance notably resembles that of models without interactions, with most interaction terms ranking rather low in importance for subjective memory problems and a more varied importance for objective cognition. The most important predictors of objective cognition were age, less education, and the interaction of depression and sex, with depression alone moving down the ranks.

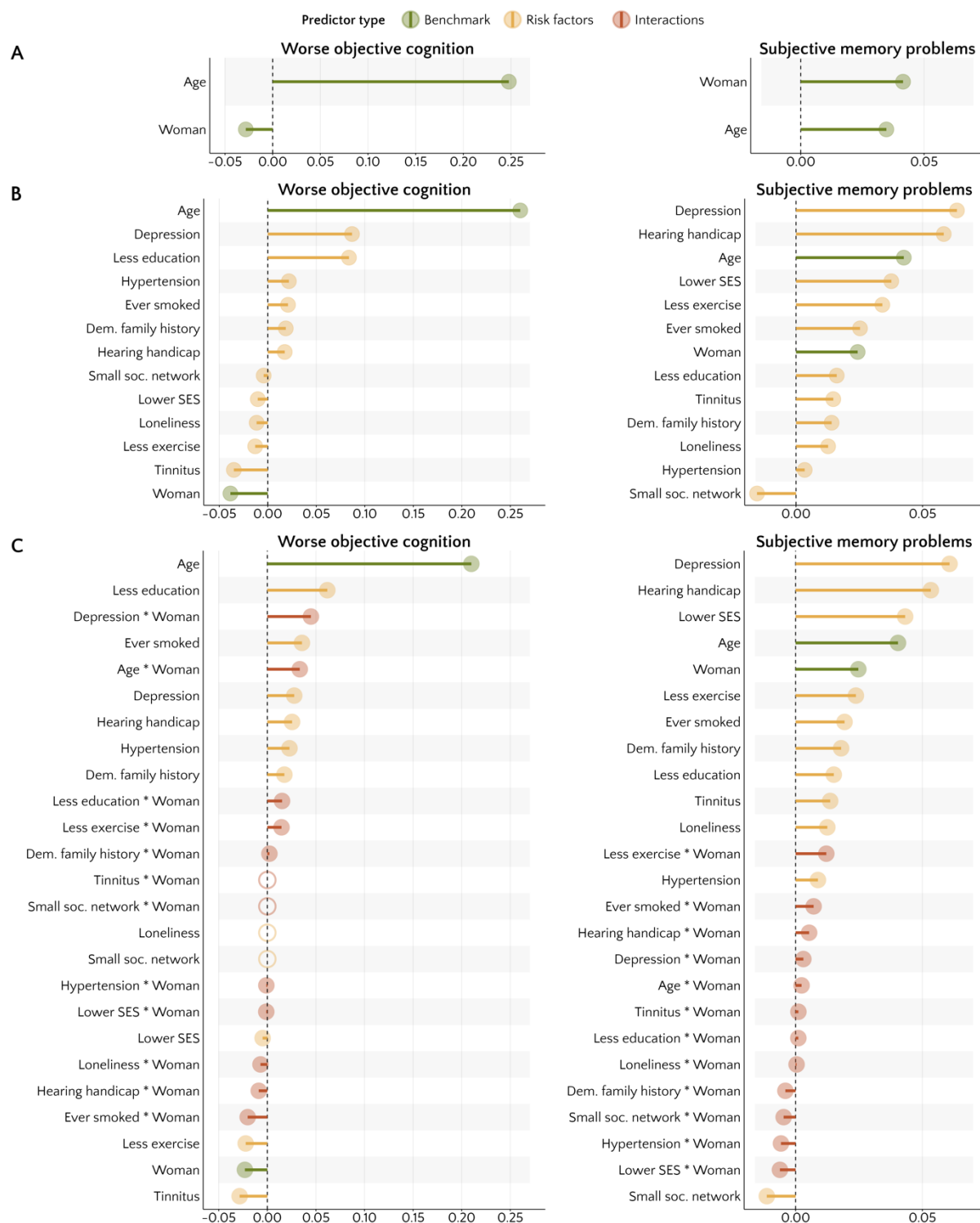


Figure 13: Ranked mean SHAP values for (A) benchmark models; (B) risk factors models without interactions; (C) risk factors models with sex interaction, shown separately for the objective cognition regression models (left) and subjective memory problems classification models (right). Higher absolute values indicate higher predictor importance for the model. Positive SHAP values indicate the predictor is associated with worse cognitive outcomes (e.g., higher depressive symptoms are linked with lower objective cognition and with a higher likelihood of subjective memory problems). Predictors not retained in the final model are depicted as empty circles.

5.3.4 Sensitivity Analyses

Our sample did not have an equal sex split (66.9%/33.1% women/men ratio), and it is possible this contributed to the lack of performance improvement associated with adding sex interactions. To rule this out, we additionally reran the same series of analyses in a subset of the data, whereby we down-sampled women in the training set to reach sex parity ($n = 1,323$ for both women and men) using a selection procedure that produced the same mean age ($M (\pm SD) = 47.43 (\pm 14)$) as that of men ($M (\pm SD) = 47.45 (\pm 14.04)$). In this sex-balanced sample, nested cross-validation resulted in better performing benchmark models compared to those fit in the full training sample: Sex and age alone explained 14.5% of variance in objective cognition and predicted subjective memory problems with an $AUC = .588$. Including risk factors in the models again significantly increased model performance to $R^2 = 16.6\%$ and $AUC = .720$ (both $p < .001$). Importantly, adding sex interactions to the models did not further increase model performance for either objective cognition ($R^2 = 16.7\%$; $p = .442$) or for subjective memory problems ($AUC = .716$; $p = .635$). The results were similar for the hold-out set. This suggests that the lack of sex interaction effects was not driven by class imbalance. Detailed results are presented in Appendix 8.3.3.

5.4 Discussion

Previous research has repeatedly shown that modifiable risk factors for dementia are also linked to deficits in cognitive performance across the entire lifespan (Borgeest et al., 2020; LaPlume, McKetton, Levine, et al., 2022; Chapter 3). Here, we applied machine learning in a large dataset of healthy adults ($N = 4,995$) crowdsourced through a smartphone app to assess if sex moderates these effects. In our sample, we found sex differences in the prevalence of some of the factors, in that smaller social networks, tinnitus, and hypertension were more common in men, whereas lower SES, depression, loneliness, and physical inactivity were more common in women. Educational achievement, family history of dementia, hearing handicap, or smoking did not differ by sex. The higher prevalence of physical inactivity (Azevedo et al., 2007) and depression symptoms (Salk et al.,

2017) in women are in line with previous studies. We found that compared to a baseline model with age and sex alone, modifiable dementia risk factors explained 3.5% additional variance in objective cognition and predicted subjective memory problems with 15% higher relative accuracy, totalling to 16.1% variance explained in objective performance and an AUC = .723 for predicting subjective memory problems. Importantly though, adding risk factor interactions with sex did not improve model performance, suggesting women and men did not differ significantly in what risk factors were relevant for them in relation to cognition. The results held even after matching the sample size for women and men in our sample, suggesting they were not driven by sex imbalance, and replicated in a fully held-out unseen dataset (20% of the total sample). This suggests the effect of dementia risk factors on cognition is not moderated by sex, although as reflected by the low performance improvement, most of the variability in cognition seems to be due to factors not assessed in our study.

The present study is one of a few that took a lifespan approach to directly test the moderating effects of sex on how modifiable risk factors relate to cognition. A previous study in a life-spanning sample of adults found that sex and dementia risk factors interacted with age in their effects on episodic memory (LaPlume, McKetton, Anderson, et al., 2022); whereby having cumulatively fewer risk factors was linked to a less pronounced effect of age on episodic memory ageing in women but not in men. Similarly, Anstey et al. (2021) studied a life-span sample stratified by age group and found a range of sex differences in the risk factors associated to episodic memory decline specific to middle-aged or older adults, suggesting age interactions might be relevant. These findings may align with prior sex-stratified studies that focused on middle-aged to older (McDermott et al., 2017) and on older adults (Liu et al., 2020) and found sex differences in individual risk factor effects on cognitive decline. Thus, it is possible that the sex differences in risk factor relevance are specific to older age groups or individuals with preclinical dementia.

Another reason for the discrepancy with previous literature could be that sex differences in risk factor relevance might be specific to certain cognitive domains and not others. LaPlume et al. (2022) have previously reported only negligible sex differences in risk factor associations to visual working memory, processing speed,

and cognitive flexibility, as compared to the more pronounced effects for episodic memory. Similarly, Anstey et al. (2021) found sex differences for episodic memory decline over time. Possibly, episodic memory might be more affected by sex-specific risk factor effects than the domains studied in the present study, i.e. subjective memory and executive function measures. This could even be an artifact of women having better episodic memory in the first place, as previous studies found an episodic memory advantage in women (Ferreira et al., 2014; LaPlume, McKetton, Anderson, et al., 2022; Sundermann, Biegon, et al., 2017), and therefore having more potential for decline. However, both subjective memory problems and executive function impairments are important cognitive markers in the preclinical stages of dementia that need to complement episodic memory assessments (Jessen et al., 2020; Mortamais et al., 2017). As further shown here, these non-episodic domains were also sensitive to the simple effects of many previously established risk factors, although their combined effects were modest to good, together explaining 3.5% additional variance in objective cognition and having 15% stronger association to subjective cognition compared to sex and age alone. It is unclear why sex differences in risk factor susceptibility should be limited to episodic memory. Further machine-learning studies should incorporate a wider range of cognitive domains to compare the sex differences in each.

Besides testing for sex differences, we assessed the relative importance of different dementia risk factors through a statistically robust, multivariate approach, addressing shared variance and improving on generalisability through machine learning. Several of the differences between predictors relevant for subjective and objective cognition deserve a mention here. First, age was the most important predictor of objective cognition, whereas for subjective memory problems it played a less pronounced role as compared to depression or perceived hearing handicap. This could reflect a relatively higher susceptibility of self-evaluated cognition to these other factors as compared to age-related cognitive decline. Second, being a woman was consistently linked to better objective cognition and worse subjective cognition, regardless of whether other dementia risk factors were included or not. This is in line with previous reports that besides episodic memory, middle-aged to older women performed better on global cognition measures and executive function (Casanova et al., 2020; Levine et al., 2021), but less in line with the

general position that visuospatial abilities are better in men across the lifespan (B. H. Lee et al., 2023; but cf. Hyde, 2016), given that our measures of objective cognition had a strong visuospatial component. Previous evidence on sex differences in subjective cognition is mixed (M. J. Brown & Patterson, 2020; Schliep et al., 2022; Wen et al., 2021) and a further analysis of third variables potentially contributing to these incongruencies is needed. Third, depression ranked as one of the most important predictors of both subjective and objective cognition, overtaken only by age in the case of objective cognition. This highlights the potential impact of depressive symptoms on cognition in adults without dementia. Fourth, there was an interesting discrepancy between subjective and objective cognition with regards to the direction of risk factor effects. We have previously shown in univariate analyses in Chapter 3 that subjective cognition was more sensitive than objective cognition to risk factors across the board, when studying each factor's effects individually and controlling only for age and sex. Congruent with this, in the current multivariate analysis where risk factor effects were considered all together, all but one risk factor studied here were associated with worse subjective cognition. The only exception was small social network, which had the opposite, 'protective' effect. This finding cannot be interpreted in isolation as it indicates smaller social network after accounting for other factors like depression and loneliness, which might be mediating the effects of social network (Cacioppo et al., 2010; Franco-O'Byrne et al., 2023). Additionally, our previous univariate analyses in Chapter 3 have shown a negative effect of social network size alone on both subjective and objective cognition, but a substantial reduction after controlling for depression alone. In contrast to these more straightforward effects on subjective cognition, objective cognition was negatively associated to fewer risk factors, which included depression, less education, hypertension, smoking history, family history of dementia, and hearing handicap. However, multiple risk factors were linked to better objective cognition, most prominently tinnitus and to a smaller extent also less exercise, loneliness, lower SES, or small social network. Even though in previous univariate analyses in Chapter 3 our objective cognitive outcomes were sensitive to loneliness, lower SES, and small social network, it seems that after accounting for all other factors in the current study these effects became negligible or even negative. This can be due to smaller sensitivity of objective cognitive assessments in healthy adults across the

lifespan, whereby many of the relationships to risk factors are weaker in the first place and therefore the risk factor effects more easily outweigh each other. Additionally, interactions between risk factors are likely. For example, tinnitus was not significantly associated to objective cognition in a previous univariate analysis in Chapter 3 but after accounting for all other factors here, including perceived hearing handicap, it became predictive of better objective cognition, possibly indicating a complex interaction effect. We also previously found that associations of risk factors to objective cognition were more often mediated by depression than those to subjective cognition, hinting at possibly stronger between-factor interactions for objective cognition. Overall, objective and subjective cognition are complementary but distinct constructs with surprisingly little overlap (Burmester et al., 2016), which might explain the differences in their sensitivity to dementia risk factors and in the underlying mechanisms.

Despite its large sample and robust statistical approach, this study was not without limitations. First, the current observational design precludes us from drawing causal explanations, with a high risk of reverse causation or third variables driving the current effects and masking potential sex differences. Future quasi-experimental studies should try to replicate our findings to test causal effects of risk factors on cognition separately by sex. Second, longitudinal data is needed to shed more light on whether there are sex differences in cognitive decline over time as compared to the current null findings for cross-sectional data. Third, we only addressed the interactions between sex and risk factors, while many of the other dementia risk factors are also likely to interact, as supported by the opposing effects in our data. Further studies investigating a larger number of interaction effects, most prominently the interactions with age, could enable to better describe the true interplay of risk factors, sex, and cognition. This is even more important as women in the current sample were on average younger than men, meaning that any interactions of gender with risk factors could be confounded with possible interactions of age with risk factors. Although regularized regression enables to deal with multicollinearity more efficiently than traditional statistical modelling, more complex interaction models could be useful in this context. Fourth, our study excluded transgender and non-binary participants, who made up 2% of the initial sample. Although this was motivated by preprocessing thresholds that filter out

low-variance categorical predictors — standard in machine learning pipelines — we acknowledge that this population is often underrepresented in research. As transgender and non-binary individuals constitute a small proportion of the population (e.g., 1.6% in the U.S.; A. Brown, 2022), they are unlikely to be adequately represented in machine-learning studies unless specific sampling strategies are employed. Future work should therefore consider targeted recruitment or alternative modelling approaches to allow for meaningful representation of individuals whose gender is different from their sex assigned at birth. Fifth, and on a related note, we based our exploration of sex differences on the participants' self-reported gender, as only cisgender participants were included in the final sample. Although this made the gender variable de facto equivalent to binary sex, the match between biological sex and gender identity could not be verified in a remote setting. Additionally, our study was unable to disentangle the effects of sex from the effects of gender, although this is the case in most brain health research in human subjects (Mielke et al., 2022). Future studies should explicitly ask about both gender and sex assigned at birth to better capture the experiences of intersex, trans, and non-binary people, for example using the question format recommended by Kronk et al. (2021). Sixth, we only employed a single modelling approach, i.e., elastic nets, which are easy to interpret and handle multicollinearity well (J. H. Friedman et al., 2010). However, our models were optimised with minimal regularisation, suggesting that models similar to simple linear/logistic regressions describe the data well and shared variance between risk factors is not fully accounted for. A more complex modelling approach (e.g., tree-based methods such as XGBoost; T. Chen & Guestrin, 2016) might outperform the current one while simultaneously better handling the underlying complexity of the relationships and their interactions, albeit at the cost of lower interpretability. Seventh, although this study covered many previously established modifiable dementia risk factors, new ones are being discovered on an ongoing basis and are now thought to collectively account for up to 45% dementia cases worldwide (Livingston et al., 2024). Future research of potential sex differences should reflect these additions and aim for a more comprehensive assessment of known dementia risk factors. Comprehensiveness could be increased by employing a data-driven approach to large, rich data, like in the example of You and colleagues (2022), who applied machine learning in a large dataset ($N > 400,000$) with

hundreds of candidate variables to predict incident dementia (irrespective of sex). Finally, the current sample was crowdsourced via a smartphone app to gather large numbers of participants, thus possibly attracting a non-representative subset of the population that might be more uniform, better educated, and higher socio-economically positioned than the general population, especially in the older age brackets (Stone et al., 2024). A resultant lack of variability might be driving the lack of sex interactions. That being said, it is unclear how the representativeness of our sample compares to traditional lab-based research, which is not without its disadvantages, either (Muniz-Terrera et al., 2019). Future studies specifically targeting underrepresented populations could help clarify the contribution of these factors to the current results.

In conclusion, we found that in a healthy adult sample spanning the entire adulthood, sex differences did not moderate the effects of dementia risk factor on objective or subjective cognition. These findings replicated in a hold-out set and held after accounting for sample imbalance between women and men. The current findings give little evidence that we need sex-specific dementia prevention strategies to target the risk factors studied here, beyond addressing the differences in prevalence. Replication in other, more diverse samples is needed to fully exclude this, comparing different cognitive outcomes, and investigating age interactions, together with more longitudinal or (quasi-)experimental evidence.

Chapter 6: General Discussion

6.1 Summary

As populations worldwide get to live longer and dementia prevalence is on the rise, the promotion of brain health is becoming a burning issue. Studying the antecedents of brain health is complicated by the sheer number of potentially relevant factors, the complexity of their interactions, as well as the delay between exposures and late-life outcomes. Despite this, converging research spanning a suite of methodologies has identified a range of modifiable and non-modifiable factors that contribute to an increased risk of dementia (Livingston et al., 2024; Riedel et al., 2016; Scarabino et al., 2016). Importantly, researchers and communities alike are increasingly recognising that promoting brain health is not a problem for the old, but rather a life-long endeavour (Gorelick & Sorond, 2023; Livingston et al., 2024). Accordingly, studies have repeatedly shown that modifiable dementia risk factors are linked to cognition also in adults without dementia, across the lifespan (Barbiellini Amidei et al., 2021; Biessels et al., 2008; Borgeest et al., 2020; Lövdén et al., 2020; E. S. Nichols et al., 2021). This makes life-long cognitive functioning and its determinants a key topic of interest in brain health research. Understanding the individual differences that shape cognitive performance throughout the healthy human lifespan could eventually 1) facilitate early detection of individuals at an increased risk of dementia using cognitive tests, and 2) give mechanistic insight into how risk factors influence cognitive decline.

Although links between various dementia risk factors and cognitive outcomes have been established, many aspects of these associations remain unclear. For example, both objective (Mortamais et al., 2017) and subjective (Mitchell et al., 2014) cognitive impairments are known to be associated with an increased risk of dementia, but their relative malleability by different dementia risk factors has received little attention. Genetic risk factors such as APOE- ϵ 4 contribute to dementia risk, but evidence for their effects on cognitive performance in healthy adults across the lifespan has been inconclusive (O'Donoghue et al., 2018). Additionally, modifiable dementia risk factors can interact with non-modifiable factors such as sex. Prior work found substantial sex differences in the prevalence

of some of the modifiable dementia risk factors (e.g., Salk et al., 2017; Azevedo et al., 2007). However, few studies have tested if these risk factors confer similar magnitudes of risk for cognitive impairment in men and women, beyond assessing additive effects of risk factors (LaPlume, McKetton, Anderson, et al., 2022).

Overall, many individual differences are likely to contribute to cognitive performance, often with small effect size, requiring large samples and robust study designs to enable sufficient power for studying these relationships.

This thesis investigated these topics in a series of studies using data collected with a smartphone app using both gamified tasks and questionnaires. Cognition across the human lifespan and its associations to dementia risk factors have been the central focus of this thesis, with two broad aims. The first aim was to explore the utility of novel, smartphone-based and remotely administered cognitive assessments as a platform for understanding brain health in adults across the lifespan. The second aim was to evaluate evidence gathered using these methods to answer some of the outstanding questions in the field. The following sections summarise the findings with respect to these two aims, then consider the implications and limitations of the research presented here, and finally, suggest future research directions.

6.1.1 Utility of Smartphone-Based Cognitive Assessments

Remote, smartphone-based assessments are a promising new platform for crowdsourcing brain health research and for assessing cognition in geographical locations where access to in-person assessments is limited. Before implementing such assessments, however, it is crucial to evaluate their psychometric qualities, i.e., validity and reliability. Chapter 2 validated a novel gamified cognitive task, Memory Match, as a measure of visual working memory for the use in brain health research. This task was loosely based on the visual short-term memory binding task (VSMBT), which promising for the early detection of AD (Cecchini et al., 2023; Della Sala et al., 2012; Parra et al., 2009). Our version of the task differed from the original version in several ways that aimed to make it more suitable for remote data collection by adding gamification elements and decreasing ceiling effects.

The validation study in Chapter 2 showed that we were successful at reducing the ceiling effects common to the original task. Despite the task's brevity (i.e., 24 trials completed in just under 8 minutes), it showed good psychometric qualities and correlated with the traditional VSMBT ($r = .39$). The relatively low cross-task convergence implies partial construct overlap and could be explained by a lower ceiling effect of Memory Match and by small changes to the task setup, i.e., the visual search paradigm in Memory Match as opposed to the forced-choice paradigm in the VSMBT. A detailed analysis of the task's conditions revealed that Memory Match did not show a consistent binding effect (i.e., the binding condition was not always more difficult than the non-binding condition), which led us to use overall visual working memory performance as an outcome measure of interest in subsequent analyses rather than focusing on any specific condition in Memory Match.

Extending this validation study, further evidence on the task's validity can be found in Chapter 3, where we analysed the effects of age, gender, and dementia risk factors on cognitive performance across the lifespan in a large sample of participants ($N = 3,327$). Speaking to the task's external validity, we found that Memory Match performance was lower for older participants, and this decline became steeper with increasing age. Such age effects on overall visual working memory performance are in line with previous evidence from a cross-sectional online study in almost 56,000 participants that assessed lifespan performance on a version of the VSMBT and reported similar patterns (Brockmole & Logie, 2013). There was no significant difference between cisgender men and women on Memory Match performance, which suggests that unlike age, sex/gender may not need to be considered if task norms are developed in the future. Furthermore, Chapter 3 also found that Memory Match was sensitive to a range of modifiable dementia risk factors including depression, SES, loneliness, education, history of smoking or stroke, social network size, diabetes, and hypertension, making it the most sensitive of our objective cognitive tasks. Regarding convergent and discriminant validity, Chapter 3 showed that overall accuracy on Memory Match was significantly correlated to performance on a measure of cognitive flexibility (i.e., the gamified Trail-Making Test B) with $r = .45$. This was not unexpected, considering that both tasks assess executive function, have a strong visual

component, require visual search, and were self-administered via smartphone. Somewhat unexpectedly, the correlation with TMT B was nominally stronger than the correlation of Memory Match with its traditional predecessor ($r = .39$). However, this latter estimate, derived from the smaller in-lab validation study in Chapter 2 ($N = 33$) may lack precision due to limited statistical power, as reflected in its wide confidence intervals ($r = .39$, 95% CI [.06; .65]). The overlap in confidence intervals between the two correlations ($r = .45$, 95% CI [.41; .47]) makes it hard to directly compare them. Furthermore, the discriminant validity of Memory Match is supported by 1) the lack of a significant association with cognitive flexibility in the in-lab sample (Chapter 2), and 2) its comparatively lower correlation with model-based planning ($r = .25$, 95% CI [.22; .29]) in Chapter 3, a less related construct. That being said, future work should gather more evidence on the validity of Memory Match, in particular with regards to other working memory measures.

Together, Chapters 2 and 3 brought evidence that Memory Match can be used as a measure of visual working memory in online crowdsourced research in the healthy population. The task shows substantially reduced ceiling effects compared to the traditional measure. It is short, sufficiently reliable, and sensitive to both age and a range of brain health correlates, and we found evidence for its construct validity as a visual working memory measure. However, due to the ways in which it diverges from the VSMBT, it is unlikely for our novel task to serve as an AD-specific measure of binding deficits. At the same time, Memory Match shares some of the limitations of the traditional binding task, such as its dependence on colour vision or not being a purely visual measure due to the possibility for participants to employ verbal rehearsal strategies. These aspects need to be considered in future research employing the task.

6.1.2 Dementia Risk Factors and Cognition

As summarised in the previous section, Chapter 2 validated a novel measure of visual working memory, Memory Match. Building on this work, Chapters 3 to 5 employed Memory Match together with two other gamified measures and a series

of self-report questionnaires administered through the smartphone app *Neureka* to study the associations of dementia risk factors with cognition in adults without dementia. Chapter 3 assessed the differential associations of objective and subjective cognition with modifiable dementia risk factors. Chapter 4 assessed the sensitivity of the smartphone-based cognitive outcomes to the genetic risk for AD poised by APOE- ϵ 4. Chapter 5 explored sex differences in how dementia risk factors relate to objective and subjective cognition. This section provides a summary of each of these studies in turn, focusing on the links between dementia risk factors and cognition.

Although many potentially modifiable dementia risk factors have been identified, there is a lack of comprehensive studies comparing their associations with different aspects of objective versus subjective cognition in healthy adults. This is important because both objective and subjective cognitive assessments are promising for early detection, but they might differ in their sensitivity. Chapter 3 used a large cross-sectional sample ($N = 3,327$, age 18 to 84) crowdsourced using the *Neureka* app to quantify the associations of 13 risk factors with subjective memory problems vs. three objective, gamified measures of executive function (visual working memory, cognitive flexibility, model-based planning). We found that all four smartphone-based cognitive outcomes were associated with a range of established dementia risk factors: Depression, socioeconomic status, hearing handicap, loneliness, education, smoking, tinnitus, low physical activity, small social network, stroke, diabetes, and hypertension were all associated with impairments in at least one cognitive measure. The single-item measure of subjective memory had the strongest link to most factors, suggesting that it was more sensitive to most risk factors than our objective tests. A recent study in middle-aged adults lent further support to this conclusion, as it also reported that subjective cognition was significantly associated with more types of modifiable dementia risk factors than two objective composite measures (Bransby, Rosenich, Buckley, et al., 2023). Because of this, we can hypothesise that cognitive self-evaluations might bring greater value than objective cognitive assessments at the earliest stages of cognitive impairment.

Although depressive symptoms are known to contribute to subjective cognitive complaints (Hill et al., 2016), most of the associations between risk factors and

subjective cognition in this study persisted after controlling for depression, suggesting depression was not fully driving the difference between objective and subjective cognition. Importantly, the pattern of associations was largely consistent across the lifespan as we did not find almost any significant interactions between risk factors and age. This means that the risk factor effects did not reflect the sensitive windows proposed by the life-course model of risk factors for dementia (Livingston et al., 2017, 2020, 2024). On the one hand, this discrepancy could suggest a bias in the literature, in that the proposed sensitive windows could be distorted simply by what is the most common age of research participants in studies of dementia risk factors. On the other hand, it is also possible that we did not observe significant effects because the age effects on cognition are non-linear or even more granular than what we tested here. Methodological issues—particularly the oversampling of younger female participants and those of higher socioeconomic status—could also have reduced the current study’s sensitivity to age interactions. Additionally, we only assessed participants aged 18 and older, meaning any effects specific to childhood and adolescence could not be detected with the current design.

Irrespective of when risk factors exert their effects most strongly, the period of midlife seems critical for the formation of brain pathology, as this is when first neurodegenerative changes start to appear in individuals who eventually go on to develop dementia (K. Ritchie et al., 2015). This makes midlife of particular interest both for interventions and for early detection, and we also turned our attention to midlife in the Chapter 4 study. After we had established our smartphone-administered cognitive measures as valid and sensitive to modifiable risk factors, we decided to test whether these measures can also be used to detect genetic risk (APOE- ϵ 4) in otherwise healthy middle-aged adults, comparing across different cognitive domains. To this end, we invited APOE-genotyped participants of a large cohort study (aged 44–67) to self-administer a module in *Neureka*. Due to low response rates, cognitive data on our smartphone measures was only available for $n = 87$ out of 646 invited participants (i.e. 13%), a number further lowered by incomplete data for some participants (completion rates 80.5%–98.9%), leading to a loss of power. Nonetheless, we assessed the associations of APOE status with five cognitive domains, controlling for relevant covariates. We found that APOE- ϵ 4

carriers outperformed non-carriers in visual working memory (β [bootstrapped 95% CI] = -0.29 [-0.53, -0.06]; $t = -2.30$; $p = .025$); however, this effect did not survive controlling for family-wise error. The positive effect of APOE- $\epsilon 4$ on visual working memory could be due to antagonistic pleiotropy effects (i.e., the risk gene conferring an advantage earlier in life; Han & Bondi, 2008), but this needs to be replicated in larger samples. Importantly, in a ROC analysis, norm-adjusted visual working memory scores only modestly differentiated between carriers and non-carriers (AUC (SE) = .648 (0.067); bootstrapped 95% CIs = 0.515–0.774). This effect size would not be clinically useful for detecting APOE status as the AUC was less than 0.8, which is a cutoff we would have been sufficiently powered to detect even with the limited sample size. There were no differences between non-carriers and carriers in their demographic characteristics or on any other cognitive outcome (i.e., model-based planning, Trail-Making Test A or B, and subjective memory problems), suggesting that if these smartphone measures are related to genetic risk in middle-aged cognitively unimpaired adults, the effects are smaller than we were powered to detect with the given sample size.

Chapter 5 turned to another (mostly) non-modifiable factor linked to dementia risk — sex. Building on prior work that found substantial sex differences in the prevalence of modifiable dementia risk factors (Anstey et al., 2021; Geraets & Leist, 2023; Gong et al., 2021; LaPlume, McKetton, Anderson, et al., 2022; Liu et al., 2020), Chapter 5 tested if dementia risk factors are linked to similar magnitudes of risk for cognitive impairment in men and women. Knowing this is important as women and men are differently impacted by cognitive decline and dementia and preventative strategies could be personalised to address this. We again used data crowdsourced through the smartphone app *Neureka*, this time spanning $N = 4,995$ participants aged 18–84, of whom 66.9% were women. We found sex differences in the prevalence of some, but not all, of the factors measured: Tinnitus, hypertension, and smaller social networks were more common in men and lower SES, higher depression, higher loneliness, and reduced physical activity were more common in women. Educational achievement, family history of dementia, hearing handicap, or smoking did not differ by sex. Going beyond differences in prevalence, we applied a multivariate machine-learning approach to develop models that predicted composite objective cognition

vs. subjective memory problems from a range of modifiable risk factors. We then compared these to models that additionally included sex as a moderator of these effects. We found that dementia risk factors contributed significantly to cognition, explaining 14.1% of variance in objective cognition and predicting subjective memory problems with AUC = .699 in the hold-out set. A model that included sex as a moderator did not improve performance for either objective or subjective cognition, thus not giving any evidence that sex moderates the effect of dementia risk factors on cognition. Overall, these findings gave little support that risk factor interventions need to be tailored to sex, beyond accounting for the differences in prevalence of these risk factors.

Chapter 5 also provided additional evidence on the relative importance of dementia risk factors for subjective and objective cognition, after accounting for shared variance, thus extending on Chapter 3 findings from univariate analyses. When we considered all the risk factors combined, they explained only little additional variance in cognition over and above sex and age (i.e., 2.5% increase in R^2 for objective cognition and 15% relative increase in AUC for subjective cognition), which is far from the estimates of risk factor contributions to dementia risk in existing literature. For example, based on the most recent review by Livingston and colleagues (Livingston et al., 2024), we could estimate that low early-life education (5%), together with midlife hearing loss (7%), depression (3%), physical inactivity (2%), diabetes (2%), smoking (2%), and hypertension (2%), and late-life social isolation (5%) together contribute to 27% of dementia cases. It is possible that modifiable risk factors convey much smaller risk for executive function and subjective memory in healthy adults as compared to the size of their effects on dementia risk. Another explanation for the difference could be that the Livingston estimates are based on high-quality epidemiological evidence, often quasi-experimental or at least longitudinal, thus possibly yielding more power. Apart from information on the total contribution of risk factors to cognition, Chapter 5 also showed the relative importance of individual risk factors when considering them in a multivariate analysis. Here, the most important predictors of objective cognition were age, followed by depression and education, whereas for subjective memory problems, depression was most important, followed by hearing handicap and age. The changes to risk factor relevance for cognition after accounting for all

other factors as compared to univariate analyses, especially in the case of objective cognition, hint at complex interaction patterns between risk factors. This lends further support to the need for employing more complex modelling approaches to describe these relationships in future research.

6.2 Implications

6.2.1 The Value of Subjective Cognitive Assessments

Subjective assessments are often understood as less trustworthy than objective assessments, as they are more likely to be affected by depression-related biases (Cherry et al., 2019; Rouault et al., 2018), disease-related lack of insight (Cacciamani et al., 2021; Pennington et al., 2021), cultural differences in response style (Wu, 2016), or even personality traits like agreeableness (Costa et al., 2023). However, the field of brain health is increasingly recognising the importance of early-stage cognitive complaints as a marker of cognitive decline (Jessen, 2014) and acknowledging the limitations of standard neuropsychological tests, especially in highly educated and cognitively active individuals who manifest high cognitive resilience to brain neuropathology or those who employ efficient compensatory strategies that enable them to score within the normative range (Stern et al., 2020; Tucker-Drob, 2019). Blind reliance on objective assessments and normative data could delay the early detection of cognitive decline unless subjective cognitive complaints are taken into consideration.

Chapters 3 and 5 of this thesis contribute to this debate, as they show subjective cognition was sensitive to more self-reported dementia risk factors compared to our gamified objective cognitive measures, both when considering each risk factor on its own (Chapter 3) or in combination, accounting for shared variance among risk factors (Chapter 5), and the associations of risk factors to subjective cognition were relatively stronger (Chapter 3). One way to interpret these findings is in the light of possibly higher utility of subjective cognitive assessments in otherwise healthy adults across the lifespan. This is even more striking given the simplicity of the subjective cognitive assessment we used: a single, yes-or-no questionnaire

item ('Do you currently experience memory problems?') was more closely related to self-reported health and lifestyle factors than a set of valid, objective cognitive tests targeting executive functions. Although this question is far from ideal as an assessment of subjective cognition — current evidence suggests an age-anchored question would be best practice, whereby participants should be asked to compare their current cognitive functioning with respect to their previous levels of cognitive functioning (Chapman et al., 2021) — it is similar to what people might encounter in clinical settings and probably sufficient to capture cases of subjective cognitive decline that would prompt someone to seek professional assessment.

Perhaps smartphone-based cognitive screening for cognitive decline could start with just this: a simple question. Without discounting the value of objective, well-targeted smartphone-based tests, a brief questionnaire would certainly be cheaper to both develop and administer than gamified cognitive assessments.

Nonetheless, we should be careful not to throw out the baby with the bathwater. One of the greatest promises of smartphone-based assessments is the ease of repeated test administration, which was not included in the studies presented here. Repeated assessment with smartphones would enable to more easily obtain a baseline score from people and thus incorporate decline into the assessment. Additionally, changes and fluctuations to cognition, measured at a fine-grained, day-to-day or even within-day level, could be even more sensitive to dementia risk (Brewster et al., 2021; Sliwinski et al., 2018; Wilks et al., 2022). Additionally, episodic memory tests, which remain the gold standard in the diagnosis of cognitive decline especially for AD, can be facilitated through smartphone-based administration, as they enable assessing recall across varied time-intervals without requiring the participants to travel to the clinic for administration (Lancaster et al., 2020). We did not compare subjective cognition against any of these repeated measures and so cannot say much about their relative value. Taken together, the utility of subjective vs. objective cognitive assessments will largely depend on the overlap between format and intended purpose.

6.2.2 Genetic Risk: How Specific Are ‘Domain-Specific’ Effects?

In Chapter 4, we tested the sensitivity of *Neureka*'s cognitive measures to genetic risk of AD, namely the presence of the APOE- ϵ 4 allele. Although our study was sufficiently powered to detect an effect that would be clinically significant (i.e., AUC ≥ 0.8) or an effect reported in a prior study that assessed navigation abilities via smartphone in healthy middle-aged adults ($d = 0.5$; Coughlan et al., 2019), the final sample size in Chapter 4 was likely too small to gain reliable insights into any cognitive phenotype effects related to APOE status. Our findings are therefore sufficient to draw practical conclusions (i.e., our smartphone-based cognitive measures cannot be used to detect individuals at genetic risk) but preclude us from making meaningful contributions to the literature on APOE- ϵ 4 effects, above and beyond possible inclusion in future meta-analyses. Nonetheless, smartphone-based assessments could still facilitate scaling up the research of genetic effects on cognition, employing the model we presented here. Large-scale studies with existing genetic data could invite participants to complete brief online cognitive assessments. Another path forward would be to combine remote cognitive assessments with at-home genetic testing, similar to what has been introduced in the Brain Health Registry study (Fockler et al., 2021).

Before opting for this, however, researchers need to carefully consider their methodological approach. The literature on APOE- ϵ 4 effects on cognition across the healthy lifespan is heterogeneous, especially in younger to middle-aged adults (Ihle et al., 2012; Lancaster, Tabet, et al., 2017; O'Donoghue et al., 2018; Weissberger et al., 2018). These dissatisfying inconsistencies have been attributed to the domain-specificity and small size of APOE- ϵ 4 effects (Lancaster, Tabet, et al., 2017; O'Donoghue et al., 2018), yet even a meta-analysis that investigated five narrow subdomains of visuospatial cognitive abilities across the lifespan found no effects across domains with the exception of spatial long-term memory (Daly et al., 2024). The most straightforward conclusion would be the possibility that APOE- ϵ 4 simply does not exert reliable effects on cognition across the lifespan, not even in middle-aged adults, and the significant differences found in some studies could have been driven by publication bias. Based on the current evidence, this explanation cannot be fully excluded. Alternatively, the

inconsistencies could imply that studies need to aim for even higher granularity with the cognitive assessments they use. For example, the subdomain of spatial navigation abilities assessed in the recent meta-analysis (Daly et al., 2024) averaged across both egocentric (Coughlan et al., 2020) and allocentric (Coughlan et al., 2019) navigation tasks, which are understood to measure distinct concepts (K. Ritchie et al., 2018), and involved a range of cognitive outcomes including both the accuracy (e.g., Coughlan et al., 2019) and speed (e.g., Acevedo et al., 2010; Yassen et al., 2015) of wayfinding. Prodromal effects of APOE- ϵ 4 (i.e., effects mediated by prodromal neuropathology) are likely to be more global and diffuse, making it easier to detect them in older adults across a range of domains. In contrast, phenotypical effects in midlife, if they exist, might be even more domain-specific than the broad, vague categories recognised by neuropsychology. For example, we might see a consistent effect on distance travelled on an egocentric wayfinding task but not when averaging across all the different tasks that aim to assess navigation abilities. Detecting phenotype effects in earlier life would therefore require a highly targeted approach with precise operationalisation of the measures used and replicating the exact task effects across large different samples with sufficient statistical power. Until we address these methodological issues, it is unlikely that we can get a good understanding of any potential effects of APOE- ϵ 4 on cognitive phenotypes across the lifespan, let alone that smartphone-based assessments could be used for the indirect detection of genetic risk.

6.2.3 Brain Health Research Can Inform Targeted Prevention Efforts

With so many modifiable risk factors for dementia and cognitive decline already identified, substantial efforts have been made to address these factors. Public health recommendations have been published, including reports based on scientific consensus (Livingston et al., 2017, 2020, 2024), dementia prevention guidelines from the World Health Organization (World Health Organization, 2019), and both national and international manuals for healthcare settings (Altomare et al., 2021; Frisoni et al., 2023). However, their wide application is hindered by the fact that many of these recommendations do not consider the feasibility and level

(individual vs. societal) at which they best should be applied (Bransby, Rosenich, Maruff, et al., 2023), so that interventions can reach those who most need them. Next to recommendations, over a hundred RCTs have been conducted to test the efficiency of specific interventions for dementia prevention (Coley et al., 2022), however, with surprisingly few positive findings. Successful dementia prevention interventions include multicomponent interventions (Ngandu et al., 2015), antihypertensive treatment (Forette et al., 1998), hearing aids (Lin et al., 2023), and amyloid-reducing therapies (Cummings et al., 2023), but effects were often specific to select high-risk populations. Overall, this suggests that there is still more to be learnt about the best ways to implement dementia prevention strategies.

Testing if these interventions work is not enough — we also need to test for whom. There is a growing appeal for personalised intervention strategies that would target high-risk individuals (Altomare et al., 2021; Coley et al., 2022; Frisoni et al., 2023; Livingston et al., 2024). Although multi-domain interventions seem most effective in dementia prevention (Coley et al., 2022; Hafdi et al., 2021), multi-domain does not equal universal, blanket recommendations, and both policymakers and researchers need to consider which groups could benefit the most. Information on risk factor prevalence can inform such considerations. The prevalence of modifiable dementia risk factors differs by geographic region (Mukadam et al., 2019; Santamaria-Garcia et al., 2023), age (H. Chen et al., 2023; LaPlume, McKetton, Levine, et al., 2022), and sex (Anstey et al., 2021; Geraets & Leist, 2023; Gong et al., 2021; LaPlume, McKetton, Anderson, et al., 2022; Liu et al., 2020). Chapter 5 of this thesis brought further evidence for the latter — we found risk factor profiles differed between men and women in our sample, with men reporting relatively fewer social connections, more tinnitus, and more hypertension, but also higher SES, less depression and loneliness, and more physical activity than women. While we assessed a wide range of risk factors, the list was not comprehensive and many new risk factors have been identified since the time this study was designed, which precludes us from saying that either sex had a higher overall prevalence of risk factors. Although we did not find evidence that sex would modulate risk factor effects on cognition, risk factor prevalence alone should still inform intervention strategies as differences in the prevalence of

any particular risk factor likely contribute to differences in the prevalence of resultant cognitive decline. This was illustrated for example by Nianogo et al. (2022), who used risk factor prevalence data together with meta-analytical information on the effects of each risk factor to calculate the relative contribution of each risk factor to dementia risk in men vs. women. Because of small sex differences in risk factor prevalence, men in their study had higher total dementia risk associated with modifiable risk factors than women. Applying the same logic to prevention, we can conclude that targeting risk factors that are most prevalent in each sex can lead to more successful dementia reduction. Based on converging evidence from Chapter 3 and other studies (e.g. Azevedo et al., 2007; Nianogo et al., 2022; Salk et al., 2017), this would include depression-reduction strategies and exercise interventions in women, as well as antihypertensive treatment in men.

Additionally, smartphones provide opportunities not only for risk assessment and early detection, but also for directly providing personalised lifestyle recommendations and interventions. Public knowledge of dementia risk factors and prevention remains limited (Parial et al., 2021), which makes education in this respect even more important. Smartphone-based interventions have been put into practice for example by the Five Lives app, which aims to both assess and intervene on lifestyle risk factors in the general population and has so far demonstrated high usability and reliability among older adults (Reid et al., 2024). Other research groups have focused on adapting a multi-domain lifestyle intervention from the successful FINGER study (Ngandu et al., 2015), including a smartphone-administered, 90-day intervention (Marquard, 2023) and an eight-week 'motivational enhancement' programme via an app (Roh et al., 2022). RCTs evaluating other digital lifestyle interventions for dementia prevention are underway (Eggink et al., 2021; Gray et al., 2022; Heffernan et al., 2019). A meta-analysis of 31 studies found that personalised mobile interventions were moderately effective for lifestyle behaviour change, and higher personalisation of the interventions was associated with better outcomes (H. L. Tong et al., 2021). However, more research is still needed to evaluate smartphone apps as a platform for lifestyle changes that can improve cognitive outcomes.

6.2.4 Practical and Ethical Considerations for Smartphone Assessments

Before implementing smartphone assessments at a large scale, either in brain health research or as a routine method of cognitive screening, several practical and ethical points need to be considered. One of the potential benefits of smartphone-based assessment lies in utilising the devices that people already possess. However, the low cost, high scalability and accessibility of such an approach must be weighed against issues that are outside of researchers' control. Data can be lost simply due to internet connection problems. Additionally, data privacy is a major concern (Öhman et al., 2021). Worryingly, not all smartphone applications intended for neuropsychiatric conditions seem to have a privacy policy, and if this is available at all, it is often inappropriately complex for lay people (Minen et al., 2021). Given these concerns, it is important to inform participants about what happens to their data in an accessible, transparent way.

Ideally, metrics of cognitive functioning should be interoperable (i.e., easily exchanged and interpreted across systems), although this can be difficult to achieve in practice. For example, differences in the technical parameters of smartphones affect the accuracy of task presentation and measurement particularly for timed tasks that involve very rapid presentation of stimuli or on tasks that become more or less difficult based on screen size (Germine et al., 2019). These differences are not random: cognitive task performance varies systematically across device types and operating systems, which is partially explained by demographic characteristics like education, age and sex, suggesting that different people access different devices (Passell et al., 2021). Interoperability is also hindered because researchers working in this area preferentially develop new assessment tools rather than validating and generalising existing tools (Van Patten, 2021) (cf. the Open Digital Health initiative, <https://opendigitalhealth.org>).

There is less experimental control with regards to distractions and identity verification for remotely collected data, but this is counteracted by greater precision and standardisation of the stimuli vs. pen-and-paper tasks (Germine et al., 2019; Van Patten, 2021). For example, a digital version of the Trail-Making Test can capture detailed information about the participant's trajectory on-screen and additional time data that a human examiner with a stopwatch could not (Linari

et al., 2022). Moreover, assessment in real-life conditions can be more ecologically valid than laboratory-based single assessments in carefully controlled conditions (e.g., for tasks that examine memory retention over several days). A further concern is that self-administered tasks need to be short to remain engaging, which means fewer trials and therefore less reliable measures. In citizen science, this can be partially offset by large samples (H. R. Brown et al., 2014).

Validation of smartphone-based cognitive measures against lab-based methods, as well as their thorough psychometric evaluation and norm creation are vital prerequisites to an informed use of such tasks in both brain health research and in clinical settings. For early detection of cognitive decline, this needs to be extended to assessing the predictive validity of the tasks, i.e., their sensitivity and specificity to relevant biomarkers and ideally also their ability to predict who will eventually develop cognitive decline. Chapter 2 evaluated the sensitivity of our smartphone-based measures against APOE- $\epsilon 4$; but other biomarkers might be better suited for this aim given their more consistent links to cognition. As an example of such an approach, a German research group developed a mnemonic discrimination task, later adapted for smartphone administration (Berron et al., 2022), and validated it against tau pathology in the cerebrospinal fluid and task-based fMRI measures of hippocampal activity (Berron et al., 2019). Not only high task sensitivity, but also high specificity (i.e., the test's ability to correctly classify negative/low-risk individuals as negative/low-risk) is crucial for any diagnostic tool because incorrectly classifying someone as high-risk may give rise to the 'worried well', causing unnecessary distress to the individual and a burden to the health system (Fornazzari & Fischer, 2022).

As we develop new and even more sensitive tools for the early detection of cognitive decline, the question is how — and if at all — they should best be implemented. A very important question is what should happen next after the risk assessment. Smartphone-based cognitive markers do not evade the ethical challenges associated with biomarker-based diagnostics in general. This includes risk disclosure: Individuals have a right to be informed about their own health, but many may wish not to be informed; and there is far too much uncertainty of when an 'at-risk' person develops further symptoms of cognitive decline and whether they will do so at all, which means the value of risk disclosure is not universal

(Vanderschaeghe et al., 2018). It is also important to give more space to the perspective of the potential end-users of such early detection: Although population-level screening for dementia can be perceived as acceptable by the public, patient and public involvement research has identified a range of important considerations, including low understanding of the dementia syndrome, questions about the cost and affordability of such services, the need for social support post-screening, and raised suspicion about economic (for-profit) motivations for the screening on the side of those implementing it (S. Martin et al., 2015).

Smartphone-based early detection tools would therefore need to carefully consider the type of feedback they give to participants, as well as their general messaging to the public. This is even more important given that self-administered smartphone assessments mostly rely on one-way communication, making it difficult for users to ask questions and for administrators to handle any potential user distress or dissatisfaction. This can be partially offset by thorough usability evaluation before implementation (Weichbroth, 2020) or by making it easy for users to remotely access support, as is the case in some remote smartphone-based studies (e.g., Thompson et al., 2022; Verhagen et al., 2019).

Finally, anyone considering the development of a smartphone app for the early detection of cognitive decline or dementia risk should be wary of falling for the 'law of the instrument'. New technologies are appealing, and it can be tempting to use them as universal instruments for a wide range of problems, even if they are not always the best solution. Implementation of any new service needs to be preceded by considering the problem at hand as well as the person who is having it and deciding whether new technologies are the right path forward (IJsselstein et al., 2020; Norman, 2019). Although smartphone usage is increasing worldwide and this includes older adults (Austin, 2024; Faverio, 2022; Pew Research Center, 2024), a large proportion of individuals still struggle with digital literacy (Kang et al., 2023) or can have negative attitudes towards digital technologies (M. Zhang, 2023). Smartphone-based assessments would not reach these individuals. In some cases, novelty outweighs practical applicability, resulting in interesting solutions that do not gain wide acceptance. Reflecting this, the field of smartphone assessments has seen considerable growth in the recent years, with many groups developing new smartphone-based tools and publishing proof-of-concept studies

(Van Patten, 2021). However, many such tools are abandoned soon after their development and few have made it to large-scale implementation in clinical studies (Mittermaier et al., 2023; Öhman et al., 2021), with the notable exception of several tablet adaptations of established computerised cognitive batteries. These include the Cambridge Neuropsychological Test Automated Battery (Ashford et al., 2024), National Institutes of Health Toolbox Cognitive Battery (Weintraub et al., 2013), and the Cogstate Brief Battery (Edgar et al., 2021). A systematic search of trials registered at [ClinicalTrials.gov](https://clinicaltrials.gov) found that 33 out of 1197 RCTs related to Alzheimer's disease between 2010–2021 used phones, tablets, or mobile applications to assess clinical trial outcomes (Mittermaier et al., 2023). The potential of smartphones as a tool for early detection of cognitive decline is undeniable, yet smartphones need to be seen as that: a tool, one of many, which needs to be considered and applied within a larger, person-centred system.

6.3 Limitations and Future Directions

The studies in this thesis applied a range of robust statistical approaches, leveraged smartphone-based data collection to gather large datasets, and covered a wide variety of dementia risk factors and other individual differences potentially contributing to brain health. Nonetheless, when interpreting the current findings, several limitations need to be taken into account.

A central caveat common to the studies in this thesis lies in their cross-sectional design. First, like any other observational study, we cannot exclude the risk of confounding or reverse causation (Lindhout et al., 2023). For example, depression can be both the cause and the consequence of cognitive impairment (Livingston et al., 2024). Ideally, future studies could assess if the relationships we found here also hold longitudinally, e.g. by testing whether baseline risk factors predict cognitive outcomes several years downstream or whether the reverse is true, which could give insight into the temporal dynamics at play. Even better, studies should aim to confirm the current findings using RCTs and quasi-experimental designs. Second, although cognitive impairment is linked to an increased risk of dementia (Mitchell et al., 2014; Mortamais et al., 2017), this relationship is not

deterministic. It is important to keep in mind that many factors throughout life, such as depression (Rock et al., 2014) or diabetes (van Duinkerken & Ryan, 2020), can impair cognition without necessarily leading to dementia. Functional cognitive disorders are a common presentation at memory clinics — these involve a range of cognitive problems that are often reversible and not associated to any obvious brain disease or general medical condition (McWhirter et al., 2020). Hence, cross-sectional lifespan data on the associations between dementia risk factors and cognition cannot give reliable mechanistic explanations of the aetiology of dementia. Nonetheless, observational evidence can still be used to formulate hypotheses which can then be tested with (quasi-)experimental methods. Furthermore, studying the antecedents of cognition alone is meaningful because cognition is linked to well-being and daily functioning (Farias et al., 2006; Yates et al., 2017) and its optimal functioning is understood as a key component of brain health (Gorelick & Sorond, 2023). Additionally, this type of research can inform the evaluation of an individual's cognitive performance; for example, knowing that women or individuals with a hearing impairment are more likely to report subjective memory problems but not to perform worse on objective tests can help set expectations about the normal performance of people who fall within these categories.

Another limitation common to the studies in this thesis involves selection bias. Online recruitment and smartphone-based data collection are likely to attract participants of higher digital literacy and socioeconomic status (Stone et al., 2023, 2024). This was the also the case with *Neureka* participants, as shown by the unusually high representation of participants with tertiary education (Chapter 3). Additionally, the reliance on convenience sampling puts into question the generalisability of the current findings, as no systematic efforts were made to ensure the study sample is representative of the broader population, particularly underserved or 'hard-to-reach' groups. Although remote data collection could in theory bridge some accessibility barriers that traditional in-person research cannot, for example by reaching individuals in remote rural areas or removing the need for travel to participate (Birnbaum, 2004; H. R. Brown et al., 2014; Doraiswamy et al., 2018; Gillan & Rutledge, 2021), the current approach did not systematically include individuals of lower socioeconomic status or participants from non-English

speaking countries. The risk of selection bias is even more pronounced for older research participants, as corroborated for example by another smartphone study, which found that individuals after the age of 75 performed better on a navigational task than their few years younger counterparts (Coutrot et al., 2018). However, in our sample, average cognitive performance decreased with age up until about 70 years of age, after which it stagnated (Chapter 3, Figure 7), suggesting that that age-related decline was still observable despite this bias. Furthermore, a smartphone-based assessment of brain health is more likely to attract not only unusually cognitively active individuals, but possibly also the ‘worried well’, i.e. individuals who worry about their potential cognitive decline, often because of family history of dementia. Although family history of dementia did not have a significant association to objective cognition in Chapter 3, a negative effect on cognition emerged after controlling for all other risk factors in Chapter 5 (see Figure 13B). Finally, language remains an issue for representativeness, as the app was only available in English, favouring participants from English-speaking countries or well-educated non-native speakers. Future smartphone studies could therefore improve generalisability through targeted sampling strategies, such as partnering with community organisations in underserved areas and LMICs, localising smartphone apps in terms of language and content, and providing technical support to participants with lower digital literacy. Of note, it is currently However, this question could be empirically tested. One possible approach would be to conduct a meta-analysis comparing the representativeness (e.g., in terms of demographic characteristics) of published brain health studies that used either design. Alternatively, an experimental design could be used, in which individuals are randomly sampled from a population (e.g., city address points) and assigned to different study formats. This approach has been employed, for example, to examine how the duration of online studies influences self-selection bias (Stone et al., 2023).

Chapter 3 of this thesis highlighted that subjective and objective cognitive assessments are distinct. Subjective and objective cognitive measures showed only low correlations, in line with previous research (Burmester et al., 2016), and differed in their sensitivity to risk factors. However, despite our best efforts, a contributing effect of self-report styles on the differences between subjective and

objective cognition cannot be fully excluded. Whoever is more likely to report problems with their memory could also be expected to report worse health and lifestyle. While secondary analyses controlling for depression in Chapter 3 showed that subjective cognition was more sensitive to dementia risk factors even after controlling for depressive symptoms, other aspects could be contributing to self-report patterns, such as worry or careless responding. It is notable that most closely associated with subjective cognition were those risk factors that might be seen as relatively more subjective — e.g., scores on a depression or loneliness scale, a subjective SES evaluation, and self-perceived handicap related to hearing impairment as compared to the more ‘objective’ risk factors such as a prior diagnosis of diabetes or hypertension. Future research could address this by replicating the current findings with more objective risk factor measures such as health record data or objective in-person assessments.

Chapter 4 attempted to utilise smartphones to facilitate data collection and expand an existing genotyped and deeply characterised dataset, the PREVENT cohort study (K. Ritchie et al., 2017). Unfortunately, participant uptake was low, and the resultant sample size was not sufficiently powered to examine the potentially small effects of APOE status on cognition, which is the main limitation of the current study. Nonetheless, our study is still informative in the sense that we did not observe a clinically meaningful effect of APOE status on any of our smartphone-based cognitive outcomes ($AUC < .800$), suggesting that further validation of these measures as tools for detecting *genetic* risk is not substantiated and would likely lead to null results. Future research could use the Chapter 4 study as a model for the validation of smartphone-based assessments against biomarkers through joining existing longitudinal cohort studies, but researchers need to keep in mind that such add-on studies can have high non-response rates and are unlikely to result in large-scale datasets.

Finally, Chapter 5 employed a robust, multivariate machine-learning approach to gain knowledge of risk factor associations to cognition and their interactions with sex. Testing models in a hold-out set is best practice in the field, which can help gain better insights into the generalisability of the models. However, in an ideal case this would be done in a separate dataset (Hastie et al., 2001; Yarkoni & Westfall, 2017). Additionally, more complex models than elastic nets, such as

XGboost (T. Chen & Guestrin, 2016) or support-vector machines (Suthaharan, 2016) could allow to better characterise the complex interactions and non-linear relationships between risk factors and cognition. A caveat of such models is their higher computational cost. However, for lower numbers of variables such as those included in the current dataset, this would not be unfeasible. Nonetheless, regularised regression models such as elastic nets are easier to interpret (J. H. Friedman et al., 2010), and the aim of the Chapter 5 study was to inform theory rather than to maximise prediction, making the current modelling approach sufficient. Future studies could consider extending the set of predictors to include the most recent comprehensive list of risk factors for dementia (Livingston et al., 2024) to gain a more comprehensive understanding of the question whether men and women differ in how important are dementia risk factors for their cognition.

6.4 Conclusion

In summary, this thesis showcased novel, smartphone-based cognitive assessments as a tool for assessing and studying brain health. The first study of this thesis validated a gamified cognitive task as a measure of visual working memory, and the following studies then used this task together with other smartphone-based cognitive measures to understand cognitive performance and its antecedents in healthy adults across the lifespan. Using data collected with a smartphone application, both in-person and remotely, we found that subjective cognition was more closely linked to a range of previously established modifiable dementia risk factors compared to objective cognition, but neither subjective nor objective cognition was affected by a genetic risk factor for AD in middle-aged adults. Finally, we employed a multivariate machine-learning approach in a large lifespan sample and found that the individual differences contributing to cognition did not differ between men and women. The final chapters then reviewed these findings and their implications in the context of the early detection of cognitive decline. Together, these results highlight smartphones as a promising platform for scaling up cognitive assessments both in research and in clinical settings. However, methodological, practical, and ethical limitations of this approach need to be considered.

7 References

- Acevedo, S. F., Piper, B. J., Craytor, M. J., Benice, T. S., & Raber, J. (2010). Apolipoprotein E4 and Sex Affect Neurobehavioral Performance in Primary School Children. *Pediatric Research*, 67(3), 293–299. <https://doi.org/10.1203/PDR.0b013e3181cb8e68>
- Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: Preliminary data in healthy, White women. *Health Psychology*, 19(6), 586. <https://doi.org/10.1037/0278-6133.19.6.586>
- Ainsworth, C. (2015). Sex redefined. *Nature*, 518(7539), 288–291. <https://doi.org/10.1038/518288a>
- Ali, J. I., Smart, C. M., & Gawryluk, J. R. (2018). Subjective Cognitive Decline and APOE ϵ 4: A Systematic Review. *Journal of Alzheimer's Disease*, 65(1), 303–320. <https://doi.org/10.3233/JAD-180248>
- Allison, S. L., Fagan, A. M., Morris, J. C., & Head, D. (2016). Spatial Navigation in Preclinical Alzheimer's Disease. *Journal of Alzheimer's Disease*, 52(1), 77–90. <https://doi.org/10.3233/JAD-150855>
- Altmann, A., Tian, L., Henderson, V. W., Greicius, M. D., & Investigators, A. D. N. I. (2014). Sex modifies the APOE-related risk of developing Alzheimer disease. *Annals of Neurology*, 75(4), 563–573. <https://doi.org/10.1002/ana.24135>
- Altomare, D., Molinuevo, J. L., Ritchie, C., Ribaldi, F., Carrera, E., Dubois, B., Jessen, F., McWhirter, L., Scheltens, P., van der Flier, W. M., Vellas, B., Démonet, J.-F., Frisoni, G. B., Abramowicz, M., Altomare, D., Barkhof, F., Berthier, M., Bieler, M., Blennow, K., ... on behalf of the European Task Force for Brain Health Services. (2021). Brain Health Services: Organization, structure, and challenges for implementation. A user manual for Brain Health Services—part 1 of 6. *Alzheimer's Research & Therapy*, 13(1), 168. <https://doi.org/10.1186/s13195-021-00827-2>
- Andrews, S. J., Goate, A., & Anstey, K. J. (2019). Association between alcohol consumption and Alzheimer's disease: A Mendelian randomization study. *Alzheimer's & Dementia*. <https://doi.org/10.1016/j.jalz.2019.09.086>
- Anstey, K. J., Cherbuin, N., & Herath, P. M. (2013). Development of a New Method for Assessing Global Risk of Alzheimer's Disease for Use in Population Health Approaches to Prevention. *Prevention Science*, 14(4), 411–421. <https://doi.org/10.1007/s11121-012-0313-2>
- Anstey, K. J., Cherbuin, N., Herath, P. M., Qiu, C., Kuller, L. H., Lopez, O. L., Wilson, R. S., & Fratiglioni, L. (2014). A Self-Report Risk Index to Predict Occurrence of Dementia in Three Independent Cohorts of Older Adults: The

ANU-ADRI. *PLOS ONE*, 9(1), e86141.
<https://doi.org/10.1371/journal.pone.0086141>

- Anstey, K. J., Ee, N., Eramudugolla, R., Jagger, C., & Peters, R. (2019). A Systematic Review of Meta-Analyses that Evaluate Risk Factors for Dementia to Evaluate the Quantity, Quality, and Global Representativeness of Evidence. *Journal of Alzheimer's Disease*, 70(s1), S165–S186.
<https://doi.org/10.3233/JAD-190181>
- Anstey, K. J., Peters, R., Mortby, M. E., Kiely, K. M., Eramudugolla, R., Cherbuin, N., Huque, M. H., & Dixon, R. A. (2021). Association of sex differences in dementia risk factors with sex differences in memory decline in a population-based cohort spanning 20–76 years. *Scientific Reports*, 11(1), 7710. <https://doi.org/10.1038/s41598-021-86397-7>
- Arnold, M. L., Dhar, S., Lee, D. J., Perreira, K. M., Pupo, D., Tsalatsanis, A., & Sanchez, V. A. (2022). Sex-Specific Prevalence, Demographic Characteristics, and Risk Factors of Tinnitus in the Hispanic Community Health Study/Study of Latinos. *American Journal of Audiology*, 31(1), 78–90. https://doi.org/10.1044/2021_AJA-21-00138
- Ashford, M. T., Aaronson, A., Kwang, W., Eichenbaum, J., Gummadi, S., Jin, C., Cashdollar, N., Thorp, E., Wragg, E., Zavitz, K. H., Cormack, F., Banh, T., Neuhaus, J. M., Ulbricht, A., Camacho, M. R., Fockler, J., Flenniken, D., Truran, D., Mackin, R. S., ... Nosheny, R. L. (2024). Unsupervised Online Paired Associates Learning Task from the Cambridge Neuropsychological Test Automated Battery (CANTAB®) in the Brain Health Registry. *The Journal of Prevention of Alzheimer's Disease*, 11(2), 514–524.
<https://doi.org/10.14283/jpad.2023.117>
- Asparouhov, T., & Muthén, B. (2009). Exploratory Structural Equation Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(3), 397–438.
<https://doi.org/10.1080/10705510903008204>
- Austin, J. P., Sneha Gubbala and Sarah. (2024, February 5). *8 charts on technology use around the world*. Pew Research Center.
<https://www.pewresearch.org/short-reads/2024/02/05/8-charts-on-technology-use-around-the-world/>
- Avila, J. F., Vonk, J. M. J., Verney, S. P., Witkiewitz, K., Arce Rentería, M., Schupf, N., Mayeux, R., & Manly, J. J. (2019). Sex/gender differences in cognitive trajectories vary as a function of race/ethnicity. *Alzheimer's & Dementia*, 15(12), 1516–1523. <https://doi.org/10.1016/j.jalz.2019.04.006>
- Azevedo, M. R., Araújo, C. L. P., Reichert, F. F., Siqueira, F. V., da Silva, M. C., & Hallal, P. C. (2007). Gender differences in leisure-time physical activity. *International Journal of Public Health*, 52(1), 8.
<https://doi.org/10.1007/s00038-006-5062-1>

- Baker, L. D., Snyder, H. M., Espeland, M. A., Whitmer, R. A., Kivipelto, M., Woolard, N., Katula, J., Papp, K. V., Ventrelle, J., Graef, S., Hill, M. A., Rushing, S., Spell, J., Lovato, L., Felton, D., Williams, B. J., Ghadimi Nouran, M., Raman, R., Ngandu, T., ... Group, for the U. S. P. S. (2024). Study design and methods: U.S. study to protect brain health through lifestyle intervention to reduce risk (U.S. POINTER). *Alzheimer's & Dementia*, 20(2), 769–782. <https://doi.org/10.1002/alz.13365>
- Barbiellini Amidei, C., Fayosse, A., Dumurgier, J., Machado-Fragua, M. D., Tabak, A. G., van Sloten, T., Kivimäki, M., Dugravot, A., Sabia, S., & Singh-Manoux, A. (2021). Association Between Age at Diabetes Onset and Subsequent Risk of Dementia. *JAMA*, 325(16), 1640–1649. <https://doi.org/10.1001/jama.2021.4001>
- Barha, C. K., Davis, J. C., Falck, R. S., Nagamatsu, L. S., & Liu-Ambrose, T. (2017). Sex differences in exercise efficacy to improve cognition: A systematic review and meta-analysis of randomized controlled trials in older humans. *Frontiers in Neuroendocrinology*, 46, 71–85. <https://doi.org/10.1016/j.yfrne.2017.04.002>
- Barnes, D. E., Covinsky, K. E., Whitmer, R. A., Kuller, L. H., Lopez, O. L., & Yaffe, K. (2009). Predicting risk of dementia in older adults: The late-life dementia risk index. *Neurology*, 73(3), 173–179. <https://doi.org/10.1212/WNL.0b013e3181a81636>
- Barsky, A. J., Peekna, H. M., & Borus, J. F. (2001). Somatic Symptom Reporting in Women and Men. *Journal of General Internal Medicine*, 16(4), 266–275. <https://doi.org/10.1046/j.1525-1497.2001.016004266.x>
- Bassett, S. S., & Folstein, M. F. (1993). Memory Complaint, Memory Performance, and Psychiatric Diagnosis: A Community Study. *Journal of Geriatric Psychiatry and Neurology*, 6(2), 105–111. <https://doi.org/10.1177/089198879300600207>
- Bastin, C. (2018). Differential age-related effects on conjunctive and relational visual short-term memory binding. *Memory*, 26(9), 1181–1190. <https://doi.org/10.1080/09658211.2017.1421228>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Beach, T. G., & Malek-Ahmadi, M. (2021). Alzheimer's Disease Neuropathological Comorbidities are Common in the Younger-Old. *Journal of Alzheimer's Disease*, 79(1), 389–400. <https://doi.org/10.3233/JAD-201213>
- Bergem, A. L., Engedal, K., & Kringlen, E. (1997). The role of heredity in late-onset Alzheimer disease and vascular dementia. A twin study. *Archives of General Psychiatry*, 54(3), 264–270. <https://doi.org/10.1001/archpsyc.1997.01830150090013>

- Berron, D., Cardenas-Blanco, A., Bittner, D., Metzger, C. D., Spottke, A., Heneka, M. T., Fließbach, K., Schneider, A., Teipel, S. J., Wagner, M., Speck, O., Jessen, F., & Düzel, E. (2019). Higher CSF Tau Levels Are Related to Hippocampal Hyperactivity and Object Mnemonic Discrimination in Older Adults. *Journal of Neuroscience*, 39(44), 8788–8797. <https://doi.org/10.1523/JNEUROSCI.1279-19.2019>
- Berron, D., Ziegler, G., Vieweg, P., Billette, O., Güsten, J., Grande, X., Heneka, M. T., Schneider, A., Teipel, S., Jessen, F., Wagner, M., & Düzel, E. (2022). Feasibility of Digital Memory Assessments in an Unsupervised and Remote Study Setting. *Frontiers in Digital Health*, 4. <https://www.frontiersin.org/article/10.3389/fdgth.2022.892997>
- Berteau-Pavy, F., Park, B., & Raber, J. (2007). Effects of sex and APOE ϵ 4 on object recognition and spatial navigation in the elderly. *Neuroscience*, 147(1), 6–17. <https://doi.org/10.1016/j.neuroscience.2007.03.005>
- Bethlehem, R. a. I., Seidlitz, J., White, S. R., Vogel, J. W., Anderson, K. M., Adamson, C., Adler, S., Alexopoulos, G. S., Anagnostou, E., Areces-Gonzalez, A., Astle, D. E., Auyeung, B., Ayub, M., Bae, J., Ball, G., Baron-Cohen, S., Beare, R., Bedford, S. A., Benegal, V., ... Alexander-Bloch, A. F. (2022). Brain charts for the human lifespan. *Nature*, 604(7906), Article 7906. <https://doi.org/10.1038/s41586-022-04554-y>
- Biessels, G. J., Deary, I. J., & Ryan, C. M. (2008). Cognition and diabetes: A lifespan perspective. *The Lancet Neurology*, 7(2), 184–190. [https://doi.org/10.1016/S1474-4422\(08\)70021-8](https://doi.org/10.1016/S1474-4422(08)70021-8)
- Birnbaum, M. H. (2004). Human Research and Data Collection via the Internet. *Annual Review of Psychology*, 55(1), 803–832. <https://doi.org/10.1146/annurev.psych.55.090902.141601>
- Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Understanding of Science*, 25(1), 2–16. <https://doi.org/10.1177/0963662515607406>
- Borelli, W. V., Zimmer, E. R., Bieger, A., Coelho, B., Pascoal, T. A., Chaves, M. L. F., Amariglio, R., & Castilhos, R. M. (2022). Subjective cognitive decline in Brazil: Prevalence and association with dementia modifiable risk factors in a population-based study. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1), e12368. <https://doi.org/10.1002/dad2.12368>
- Borgeest, G. S., Henson, R. N., Shafto, M., Samu, D., Cam-CAN, & Kievit, R. A. (2020). Greater lifestyle engagement is associated with better age-adjusted cognitive abilities. *PLOS ONE*, 15(5), e0230077. <https://doi.org/10.1371/journal.pone.0230077>
- Boss, L., Kang, D.-H., & Branson, S. (2015). Loneliness and cognitive function in the older adult: A systematic review. *International Psychogeriatrics*, 27(4), 541–553. <https://doi.org/10.1017/S1041610214002749>

- Bowie, C. R., & Harvey, P. D. (2006). Administration and interpretation of the Trail Making Test. *Nature Protocols*, 1(5), Article 5. <https://doi.org/10.1038/nprot.2006.390>
- Braak, H., & Braak, E. (1991). Neuropathological staging of Alzheimer-related changes. *Acta Neuropathologica*, 82(4), 239–259. <https://doi.org/10.1007/BF00308809>
- Braak, H., Thal, D. R., Ghebremedhin, E., & Del Tredici, K. (2011). Stages of the pathologic process in Alzheimer disease: Age categories from 1 to 100 years. *Journal of Neuropathology and Experimental Neurology*, 70(11), 960–969. <https://doi.org/10.1097/NEN.0b013e318232a379>
- Bradfield, N. I. (2023). Mild Cognitive Impairment: Diagnosis and Subtypes. *Clinical EEG and Neuroscience*, 54(1), 4–11. <https://doi.org/10.1177/15500594211042708>
- Bransby, L., Rosenich, E., Buckley, R. F., Yassi, N., Pase, M. P., Maruff, P., & Lim, Y. Y. (2023). Multidomain modifiable dementia risk factors are associated with poorer cognition in midlife. *Neuropsychology*, 37(5), 582–594. <https://doi.org/10.1037/neu0000900>
- Bransby, L., Rosenich, E., Maruff, P., & Lim, Y. Y. (2023). How Modifiable Are Modifiable Dementia Risk Factors? A Framework for Considering the Modifiability of Dementia Risk Factors. *The Journal of Prevention of Alzheimer's Disease*. <https://doi.org/10.14283/jpad.2023.119>
- Braun, T., Schmukle, S. C., & Kunzmann, U. (2017). Stability and change in subjective well-being: The role of performance-based and self-rated cognition. *Psychology and Aging*, 32(2), 105. <https://doi.org/10.1037/pag0000153>
- Brewster, P. W. H., Rush, J., Ozen, L., Vendittelli, R., & Hofer, S. M. (2021). Feasibility and Psychometric Integrity of Mobile Phone-Based Intensive Measurement of Cognition in Older Adults. *Experimental Aging Research*, 47(4), 303–321. <https://doi.org/10.1080/0361073X.2021.1894072>
- Broadley, M. M., White, M. J., & Andrew, B. (2017). A Systematic Review and Meta-analysis of Executive Function Performance in Type 1 Diabetes Mellitus. *Psychosomatic Medicine*, 79(6), 684. <https://doi.org/10.1097/PSY.0000000000000460>
- Brockmole, J. R., & Logie, R. (2013). Age-Related Change in Visual Working Memory: A Study of 55,753 Participants Aged 8–75. *Frontiers in Psychology*, 4(12). <https://doi.org/10.3389/fpsyg.2013.00012>
- Brockmole, J. R., Parra, M. A., Sala, S. D., & Logie, R. H. (2008). Do binding deficits account for age-related decline in visual working memory? *Psychonomic Bulletin & Review*, 15(3), 543–547. <https://doi.org/10.3758/PBR.15.3.543>

- Brown, A. (2022, June 7). *About 5% of young adults in the U.S. say their gender is different from their sex assigned at birth*. Pew Research Center. <https://www.pewresearch.org/short-reads/2022/06/07/about-5-of-young-adults-in-the-u-s-say-their-gender-is-different-from-their-sex-assigned-at-birth/>
- Brown, H. R., Zeidman, P., Smittenaar, P., Adams, R. A., McNab, F., Rutledge, R. B., & Dolan, R. J. (2014). Crowdsourcing for Cognitive Science – The Utility of Smartphones. *PLOS ONE*, 9(7), e100662. <https://doi.org/10.1371/journal.pone.0100662>
- Brown, M. J., & Patterson, R. (2020). Subjective Cognitive Decline Among Sexual and Gender Minorities: Results from a U.S. Population-Based Sample. *Journal of Alzheimer's Disease*, 73(2), 477–487. <https://doi.org/10.3233/JAD-190869>
- Burgess, N., Maguire, E. A., & O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, 35(4), 625–641. [https://doi.org/10.1016/s0896-6273\(02\)00830-9](https://doi.org/10.1016/s0896-6273(02)00830-9)
- Burmester, B., Leathem, J., & Merrick, P. (2016). Subjective Cognitive Complaints and Objective Cognitive Function in Aging: A Systematic Review and Meta-Analysis of Recent Cross-Sectional Findings. *Neuropsychology Review*, 26(4), 376–393. <https://doi.org/10.1007/s11065-016-9332-2>
- Cacciamani, F., Houot, M., Gagliardi, G., Dubois, B., Sikkes, S., Sánchez-Benavides, G., Denicolò, E., Molinuevo, J. L., Vannini, P., & Epelbaum, S. (2021). Awareness of Cognitive Decline in Patients With Alzheimer's Disease: A Systematic Review and Meta-Analysis. *Frontiers in Aging Neuroscience*, 13, 697234. <https://doi.org/doi:10.3389/fnagi.2021.697234>
- Cacioppo, J. T., Hawkley, L. C., & Thisted, R. A. (2010). Perceived social isolation makes me sad: 5-year cross-lagged analyses of loneliness and depressive symptomatology in the Chicago Health, Aging, and Social Relations Study. *Psychology and Aging*, 25(2), 453–463. <https://doi.org/10.1037/a0017216>
- Cannon-Albright, L. A., Foster, N. L., Schliep, K., Farnham, J. M., Teerlink, C. C., Kaddas, H., Tschanz, J., Corcoran, C., & Kauwe, J. S. K. (2019). Relative risk for Alzheimer disease based on complete family history. *Neurology*, 92(15), e1745–e1753. <https://doi.org/10.1212/WNL.00000000000007231>
- Cao, Q., Tan, C.-C., Xu, W., Hu, H., Cao, X.-P., Dong, Q., Tan, L., & Yu, J.-T. (2020). The Prevalence of Dementia: A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease*, 73(3), 1157–1166. <https://doi.org/10.3233/JAD-191092>
- Capron, L. (2003). What's in a risk factor? "He who strikes the ball". *Diabetes & Metabolism*, 29(1), 6–13. [https://doi.org/10.1016/S1262-3636\(07\)70001-0](https://doi.org/10.1016/S1262-3636(07)70001-0)

- Cardona, M., & Andrés, P. (2023). Are social isolation and loneliness associated with cognitive decline in ageing? *Frontiers in Aging Neuroscience*, 15, 1075563. <https://doi.org/10.3389/fnagi.2023.1075563>
- Casanova, R., Saldana, S., Lutz, M. W., Plassman, B. L., Kuchibhatla, M., & Hayden, K. M. (2020). Investigating Predictors of Cognitive Decline Using Machine Learning. *The Journals of Gerontology: Series B*, 75(4), 733–742. <https://doi.org/10.1093/geronb/gby054>
- Caselli, R. J., Dueck, A. C., Locke, D. E. C., Sabbagh, M. N., Ahern, G. L., Rapcsak, S. Z., Baxter, L. C., Yaari, R., Woodruff, B. K., Hoffman-Snyder, C., Rademakers, R., Findley, S., & Reiman, E. M. (2011). Cerebrovascular risk factors and preclinical memory decline in healthy APOE ϵ 4 homozygotes. *Neurology*, 76(12), 1078–1084. <https://doi.org/10.1212/WNL.0b013e318211c3ae>
- Cecchini, M. A., Parra, M. A., Brazzelli, M., Logie, R. H., & Della Sala, S. (2023). Short-term memory conjunctive binding in Alzheimer's disease: A systematic review and meta-analysis. *Neuropsychology*, 37(7), 769–789. <https://doi.org/10.1037/neu0000825>
- Cecchini, M. A., Yassuda, M. S., Squarzoni, P., Coutinho, A. M., de Paula Faria, D., Duran, F. L. de S., Costa, N. A. da, Porto, F. H. de G., Nitrini, R., Forlenza, O. V., Brucki, S. M. D., Buchpiguel, C. A., Parra, M. A., & Busatto, G. F. (2021). Deficits in short-term memory binding are detectable in individuals with brain amyloid deposition in the absence of overt neurodegeneration in the Alzheimer's disease continuum. *Brain and Cognition*, 152, 105749. <https://doi.org/10.1016/j.bandc.2021.105749>
- Chakrabarti, S. (2024). Digital psychiatry in low-and-middle-income countries: New developments and the way forward. *World Journal of Psychiatry*, 14(3), 350–361. <https://doi.org/10.5498/wjp.v14.i3.350>
- Chapman, S., Sunderaraman, P., Joyce, J. L., Azar, M., Colvin, L. E., Barker, M. S., McKeague, I., Kreisl, W. C., & Cosentino, S. (2021). Optimizing Subjective Cognitive Decline to Detect Early Cognitive Dysfunction. *Journal of Alzheimer's Disease*, 80(3), 1185–1196. <https://doi.org/10.3233/JAD-201322>
- Chen, H., Cao, Y., Ma, Y., Xu, W., Zong, G., & Yuan, C. (2023). Age- and sex-specific modifiable risk factor profiles of dementia: Evidence from the UK Biobank. *European Journal of Epidemiology*, 38(1), 83–93. <https://doi.org/10.1007/s10654-022-00952-8>
- Chen, K., Gao, T., Liu, Y., Zhu, K., Wang, T., & Zeng, P. (2024). Identifying risk loci for FTD and shared genetic component with ALS: A large-scale multitrait association analysis. *Neurobiology of Aging*, 134, 28–39. <https://doi.org/10.1016/j.neurobiolaging.2023.09.017>

- Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Chen, Y., Demnitz, N., Yamamoto, S., Yaffe, K., Lawlor, B., & Leroi, I. (2021). Defining brain health: A concept analysis. *International Journal of Geriatric Psychiatry*, 37(1). <https://doi.org/10.1002/gps.5564>
- Chêne, G., Beiser, A., Au, R., Preis, S. R., Wolf, P. A., Dufouil, C., & Seshadri, S. (2015). Gender and incidence of dementia in the Framingham Heart Study from mid-adult life. *Alzheimer's & Dementia*, 11(3), 310–320. <https://doi.org/10.1016/j.jalz.2013.10.005>
- Cheng, G.-R., Liu, D., Huang, L.-Y., Han, G.-B., Hu, F.-F., Wu, Z.-X., He, X.-M., Huang, Y.-W., Yu, Y.-F., Xu, L., Li, J.-Q., Chen, Y.-S., Wei, Z., Wu, Q., Mei, Y.-F., Chen, X.-X., Ou, Y.-M., Zhang, J.-J., Yang, M.-L., ... Zeng, Y. (2023). Prevalence and risk factors for subjective cognitive decline and the correlation with objective cognition among community-dwelling older adults in China: Results from the Hubei memory and aging cohort study. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 19(11), 5074–5085. <https://doi.org/10.1002/alz.13047>
- Cherry, K. E., Lyon, B. A., Boudreaux, E. O., Blanchard, A. B., Hicks, J. L., Elliott, E. M., Myers, L., Kim, S., & Jazwinski, S. M. (2019). Memory Self-Efficacy and Beliefs about Memory and Aging in Oldest-Old Adults in the Louisiana Healthy Aging Study (LHAS). *Experimental Aging Research*, 45(1), 28–40. <https://doi.org/10.1080/0361073X.2018.1560107>
- Ciria, L. F., Román-Caballero, R., Vadillo, M. A., Holgado, D., Luque-Casado, A., Perakakis, P., & Sanabria, D. (2023). An umbrella review of randomized control trials on the effects of physical exercise on cognition. *Nature Human Behaviour*, 7(6), 928–941. <https://doi.org/10.1038/s41562-023-01554-4>
- Cockerham, W. C. (2005). Health Lifestyle Theory and the Convergence of Agency and Structure. *Journal of Health and Social Behavior*, 46(1), 51–67. <https://doi.org/10.1177/002214650504600105>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037//0033-2909.112.1.155>
- Colcombe, S., & Kramer, A. F. (2003). Fitness Effects on the Cognitive Function of Older Adults: A Meta-Analytic Study. *Psychological Science*, 14(2), 125–130. <https://doi.org/10.1111/1467-9280.t01-1-01430>
- Coley, N., Giulioli, C., Aisen, P. S., Vellas, B., & Andrieu, S. (2022). Randomised controlled trials for the prevention of cognitive decline or dementia: A systematic review. *Ageing Research Reviews*, 82, 101777. <https://doi.org/10.1016/j.arr.2022.101777>

- Conti, A. A., McLean, L., Tolomeo, S., Steele, J. D., & Baldacchino, A. (2019). Chronic tobacco smoking and neuropsychological impairments: A systematic review and meta-analysis. *Neuroscience & Biobehavioral Reviews*, 96, 143–154. <https://doi.org/10.1016/j.neubiorev.2018.11.017>
- Çorbacioğlu, Ş. K., & Aksel, G. (2023). Receiver operating characteristic curve analysis in diagnostic accuracy studies: A guide to interpreting the area under the curve value. *Turkish Journal of Emergency Medicine*, 23(4), 195–198. https://doi.org/10.4103/tjem.tjem_182_23
- Corder, E. H., Saunders, A. M., Risch, N. J., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Rimmler, J. B., Locke, P. A., Conneally, P. M., Schmechel, K. E., Small, G. W., Roses, A. D., Haines, J. L., & Pericak-Vance, M. A. (1994). Protective effect of apolipoprotein E type 2 allele for late onset Alzheimer disease. *Nature Genetics*, 7(2), Article 2. <https://doi.org/10.1038/ng0694-180>
- Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G. W., Roses, A. D., Haines, J. L., & Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science (New York, N.Y.)*, 261(5123), 921–923. <https://doi.org/10.1126/science.8346443>
- Costa, A. N., Nowakowski, L. M., McCrae, C. S., Cowan, N., & Curtis, A. F. (2023). Discrepancies in Objective and Subjective Cognition in Middle-Aged and Older Adults: Does Personality Matter? *Gerontology and Geriatric Medicine*, 9, 23337214221146663. <https://doi.org/10.1177/23337214221146663>
- Coughlan, G., Coutrot, A., Khondoker, M., Minihane, A.-M., Spiers, H., & Hornberger, M. (2019). Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. *Proceedings of the National Academy of Sciences*, 116(19), 9285–9292. <https://doi.org/10.1073/pnas.1901600116>
- Coughlan, G., Laczó, J., Hort, J., Minihane, A.-M., & Hornberger, M. (2018). Spatial navigation deficits—Overlooked cognitive marker for preclinical Alzheimer disease? *Nature Reviews Neurology*, 14(8), Article 8. <https://doi.org/10.1038/s41582-018-0031-x>
- Coughlan, G., Zhukovsky, P., Puthusseryppady, V., Gillings, R., Minihane, A.-M., Cameron, D., & Hornberger, M. (2020). Functional connectivity between the entorhinal and posterior cingulate cortices underpins navigation discrepancies in at-risk Alzheimer's disease. *Neurobiology of Aging*, 90, 110–118. <https://doi.org/10.1016/j.neurobiolaging.2020.02.007>
- Courtin, E., Nafilyan, V., Glymour, M., Goldberg, M., Berr, C., Berkman, L. F., Zins, M., & Avendano, M. (2019). Long-term effects of compulsory schooling on physical, mental and cognitive ageing: A natural experiment. *Journal of Epidemiology and Community Health*, 73(4), 370–376. <https://doi.org/10.1136/jech-2018-211746>

- Coutrot, A., Silva, R., Manley, E., de Cothi, W., Sami, S., Bohbot, V. D., Wiener, J. M., Hölscher, C., Dalton, R. C., Hornberger, M., & Spiers, H. J. (2018). Global Determinants of Navigation Ability. *Current Biology*, 28(17), 2861–2866.e4. <https://doi.org/10.1016/j.cub.2018.06.009>
- Cubillos, C., & Rienzo, A. (2023). Digital Cognitive Assessment Tests for Older Adults: Systematic Literature Review. *JMIR Mental Health*, 10(1), e47487. <https://doi.org/10.2196/47487>
- Cummings, J., Apostolova, L., Rabinovici, G. D., Atri, A., Aisen, P., Greenberg, S., Hendrix, S., Selkoe, D., Weiner, M., Petersen, R. C., Salloway, S., & For the Alzheimer's Disease and Related Disorders Therapeutics Work Group. (2023). Lecanemab: Appropriate Use Recommendations. *The Journal of Prevention of Alzheimer's Disease*, 10(3), 362–377. <https://doi.org/10.14283/jpad.2023.30>
- Dal Forno, G., Palermo, M. T., Donohue, J. E., Karagiozis, H., Zonderman, A. B., & Kawas, C. H. (2005). Depressive symptoms, sex, and risk for Alzheimer's disease. *Annals of Neurology*, 57(3), 381–387. <https://doi.org/10.1002/ana.20405>
- Daly, J., De Luca, F., Berens, S. C., Field, A. P., Rusted, J. M., & Bird, C. M. (2024). The effect of apolipoprotein E genotype on spatial processing in humans: A meta-analysis and systematic review. *Cortex*, 177, 268–284. <https://doi.org/10.1016/j.cortex.2024.05.006>
- Daseking, M., Petermann, F., & Waldmann, H.-C. (2017). Sex differences in cognitive abilities: Analyses for the German WAIS-IV. *Personality and Individual Differences*, 114, 145–150. <https://doi.org/10.1016/j.paid.2017.04.003>
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011a). Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron*, 69(6), 1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>
- Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011b). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, 69(6), 1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>
- de Frias, C. M., Nilsson, L.-G., & Herlitz, A. (2006). Sex Differences in Cognition are Stable Over a 10-Year Period in Adulthood and Old Age. *Aging, Neuropsychology, and Cognition*, 13(3–4), 574–587. <https://doi.org/10.1080/13825580600678418>
- Deckers, K., Nooyens, A., van Boxtel, M., Verhey, F., Verschuren, M., & Köhler, S. (2019). Gender and Educational Differences in the Association between Lifestyle and Cognitive Decline over 10 Years: The Doetinchem Cohort Study. *Journal of Alzheimer's Disease*, 70(s1), S31–S41. <https://doi.org/10.3233/JAD-180492>

- Della Sala, S., Parra, M. A., Fabi, K., Luzzi, S., & Abrahams, S. (2012). Short-term memory binding is impaired in AD but not in non-AD dementias. *Neuropsychologia*, 50(5), 833–840. <https://doi.org/10.1016/j.neuropsychologia.2012.01.018>
- Demnitz, N., Hulme, O. J., Siebner, H. R., Kjaer, M., Ebmeier, K. P., Boraxbekk, C.-J., & Gillan, C. M. (2023). Characterising the covariance pattern between lifestyle factors and structural brain measures: A multivariable replication study of two independent ageing cohorts. *Neurobiology of Aging*, 131, 115–123. <https://doi.org/10.1016/j.neurobiolaging.2023.07.023>
- Desai, R., John, A., Saunders, R., Marchant, N. L., Buckman, J. E. J., Charlesworth, G., Zuber, V., & Stott, J. (2023). Examining the Lancet Commission risk factors for dementia using Mendelian randomisation. *BMJ Ment Health*, 26(1). <https://doi.org/10.1136/bmjment-2022-300555>
- Dhana, K., Franco, O. H., Ritz, E. M., Ford, C. N., Desai, P., Krueger, K. R., Holland, T. M., Dhana, A., Liu, X., Aggarwal, N. T., Evans, D. A., & Rajan, K. B. (2022). Healthy lifestyle and life expectancy with and without Alzheimer's dementia: Population based cohort study. *BMJ*, 377, e068390. <https://doi.org/10.1136/bmj-2021-068390>
- Diamond, A. (2013). Executive Functions. *Annual Review of Psychology*, 64, 135–168. <https://doi.org/10.1146/annurev-psych-113011-143750>
- Donegan, K. R., Brown, V. M., Price, R. B., Gallagher, E., Pringle, A., Hanlon, A. K., & Gillan, C. M. (2023). Using smartphones to optimise and scale-up the assessment of model-based planning. *Communications Psychology*, 1(1), Article 1. <https://doi.org/10.1038/s44271-023-00031-y>
- Doraiswamy, P. M., Narayan, V. A., & Manji, H. K. (2018). Mobile and pervasive computing technologies and the future of Alzheimer's clinical trials. *Npj Digital Medicine*, 1(1), 1–4. <https://doi.org/10.1038/s41746-017-0008-y>
- Edgar, C. J., Siemers, E., Maruff, P., Petersen, R. C., Aisen, P. S., Weiner, M. W., Albala, B., & Initiative, for the A. D. N. (2021). Pilot Evaluation of the Unsupervised, At-Home Cogstate Brief Battery in ADNI-2. *Journal of Alzheimer's Disease*, 83(2), 915–925. <https://doi.org/10.3233/JAD-210201>
- Edmonds, E. C., Delano-Wood, L., Galasko, D. R., Salmon, D. P., & Bondi, M. W. (2014). Subjective Cognitive Complaints Contribute to Misdiagnosis of Mild Cognitive Impairment. *Journal of the International Neuropsychological Society*, 20(8), 836–847. <https://doi.org/10.1017/S135561771400068X>
- Eggink, E., Hafdi, M., Hoevenaar-Blom, M. P., Song, M., Andrieu, S., Barnes, L. E., Birck, C., Brooks, R. L., Coley, N., Ford, E., Georges, J., Groep, A. van der, Gool, W. A. van, Handels, R., Hou, H., Li, D., Liu, H., Lyu, J., Marwijk, H. van, ... Richard, E. (2021). Prevention of dementia using mobile phone applications (PRODEMOS): Protocol for an international randomised

controlled trial. *BMJ Open*, 11(6), e049762.
<https://doi.org/10.1136/bmjopen-2021-049762>

- Eisenberg, D. T. A., Kuzawa, C. W., & Hayes, M. G. (2010). Worldwide allele frequencies of the human apolipoprotein E gene: Climate, local adaptations, and evolutionary history. *American Journal of Physical Anthropology*, 143(1), 100–111. <https://doi.org/10.1002/ajpa.21298>
- Eppinger, B., Walter, M., Heekeren, H., & Li, S.-C. (2013). Of goals and habits: Age-related and individual differences in goal-directed decision-making. *Frontiers in Neuroscience*, 7, 253. <https://doi.org/10.3389/fnins.2013.00253>
- Exalto, L. G., Quesenberry, C. P., Barnes, D., Kivipelto, M., Biessels, G. J., & Whitmer, R. A. (2014). Midlife risk score for the prediction of dementia four decades later. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 10(5), 562–570. <https://doi.org/10.1016/j.jalz.2013.05.1772>
- Fallah, N., Mitnitski, A., Middleton, L., & Rockwood, K. (2009). Modeling the impact of sex on how exercise is associated with cognitive changes and death in older Canadians. *Neuroepidemiology*, 33(1), 47–54. <https://doi.org/10.1159/000211953>
- Farias, S. T., Mungas, D., Reed, B. R., Harvey, D., Cahn-Weiner, D., & DeCarli, C. (2006). MCI is Associated With Deficits in Everyday Functioning. *Alzheimer Disease and Associated Disorders*, 20(4), 217–223. <https://doi.org/10.1097/01.wad.0000213849.51495.d9>
- Faverio, M. (2022, January 13). *Share of those 65 and older who are tech users has grown in the past decade*. Pew Research Center. <https://www.pewresearch.org/fact-tank/2022/01/13/share-of-those-65-and-older-who-are-tech-users-has-grown-in-the-past-decade/>
- Feenstra, H. E. M., Vermeulen, I. E., Murre, J. M. J., & Schagen, S. B. (2017). Online cognition: Factors facilitating reliable online neuropsychological test results. *The Clinical Neuropsychologist*, 31(1), 59–84. <https://doi.org/10.1080/13854046.2016.1190405>
- Ferencz, B., & Gerritsen, L. (2015). Genetics and Underlying Pathology of Dementia. *Neuropsychology Review*, 25(1), 113–124. <https://doi.org/10.1007/s11065-014-9276-3>
- Fernández-Calle, R., Konings, S. C., Frontiñán-Rubio, J., García-Revilla, J., Camprubí-Ferrer, L., Svensson, M., Martinson, I., Boza-Serrano, A., Venero, J. L., Nielsen, H. M., Gouras, G. K., & Deierborg, T. (2022). APOE in the bullseye of neurodegenerative diseases: Impact of the APOE genotype in Alzheimer's disease pathology and brain diseases. *Molecular Neurodegeneration*, 17(1), 62. <https://doi.org/10.1186/s13024-022-00566-4>
- Ferreira, L., Ferreira Santos-Galduróz, R., Ferri, C. P., & Fernandes Galduróz, J. C. (2014). Rate of cognitive decline in relation to sex after 60 years-of-age:

A systematic review. *Geriatrics & Gerontology International*, 14(1), 23–31.
<https://doi.org/10.1111/ggi.12093>

- Fittipaldi, S., Legaz, A., Maito, M., Hernandez, H., Altschuler, F., Canziani, V., Moguilner, S., Gillan, C. M., Castillo, J., Lillo, P., Custodio, N., Avila-Funes, J. A., Cardona, J. F., Slachevsky, A., Henriquez, F., Fraile-Vazquez, M., Cruz de Souza, L., Borroni, B., Hornberger, M., ... Ibanez, A. (2024). Heterogeneous factors influence social cognition across diverse settings in brain health and age-related diseases. *Nature Mental Health*, 2(1), Article 1. <https://doi.org/10.1038/s44220-023-00164-3>
- Flory, J. D., Manuck, S. B., Ferrell, R. E., Ryan, C. M., & Muldoon, M. F. (2000). Memory performance and the apolipoprotein E polymorphism in a community sample of middle-aged adults. *American Journal of Medical Genetics*, 96(6), 707–711. [https://doi.org/10.1002/1096-8628\(20001204\)96:6<707::AID-AJMG1>3.0.CO;2-V](https://doi.org/10.1002/1096-8628(20001204)96:6<707::AID-AJMG1>3.0.CO;2-V)
- Fockler, J., Kwang, W., Ashford, M. T., Flenniken, D., Hwang, J., Truran, D., Mackin, R. S., Jin, C., O'Hara, R., Hallmayer, J. F., Yesavage, J. A., Weiner, M. W., & Nosheny, R. L. (2021). Brain health registry GenePool study: A novel approach to online genetics research. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 7(1), e12118. <https://doi.org/10.1002/trc2.12118>
- Forette, F., Seux, M.-L., Staessen, J. A., Thijs, L., Birkenhäger, W. H., Babarskiene, M.-R., Babeanu, S., Bossini, A., Gil-Extremiera, B., Girerd, X., Laks, T., Lilov, E., Moisseiev, V., Tuomilehto, J., Vanhanen, H., Webster, J., Yodfat, Y., & Fagard, R. (1998). Prevention of dementia in randomised double-blind placebo-controlled Systolic Hypertension in Europe (Syst-Eur) trial. *The Lancet*, 352(9137), 1347–1351. [https://doi.org/10.1016/S0140-6736\(98\)03086-4](https://doi.org/10.1016/S0140-6736(98)03086-4)
- Fornazzari, L., & Fischer, C. E. (2022). Biomarkers and Alzheimer's Disease: What Will the Future Bring for 'The Worried Well'? *Canadian Journal of Neurological Sciences*, 49(1), 5–6. <https://doi.org/10.1017/cjn.2021.126>
- Forsberg, A., Johnson, W., & Logie, R. H. (2019). Aging and feature-binding in visual working memory: The role of verbal rehearsal. *Psychology and Aging*, 34(7), 933–953. <https://doi.org/10.1037/pag0000391>
- Fox, C. A., McDonogh, A., Donegan, K. R., Teckentrup, V., Crossen, R. J., Hanlon, A. K., Gallagher, E., Rouault, M., & Gillan, C. M. (2024). Reliable, rapid, and remote measurement of metacognitive bias. *Scientific Reports*, 14(1), 14941. <https://doi.org/10.1038/s41598-024-64900-0>
- Franco-O'Byrne, D., Gonzalez-Gomez, R., Morales Sepúlveda, J. P., Vergara, M., Ibañez, A., & Huepe, D. (2023). The impact of loneliness and social adaptation on depressive symptoms: Behavioral and brain measures evidence from a brain health perspective. *Frontiers in Psychology*, 14(1096178). <https://doi.org/10.3389/fpsyg.2023.1096178>

- Friedman, J. H., Hastie, T., & Tibshirani, R. (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. *Journal of Statistical Software*, 33, 1–22. <https://doi.org/10.18637/jss.v033.i01>
- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, 137(2), 201–225. <https://doi.org/10.1037/0096-3445.137.2.201>
- Frisoni, G. B., Altomare, D., Ribaldi, F., Villain, N., Brayne, C., Mukadam, N., Abramowicz, M., Barkhof, F., Berthier, M., Bieler-Aeschlimann, M., Blennow, K., Guevara, A. B., Carrera, E., Chételat, G., Csajka, C., Demonet, J.-F., Dodich, A., Garibotto, V., Georges, J., ... Dubois, B. (2023). Dementia prevention in memory clinics: Recommendations from the European task force for brain health services. *The Lancet Regional Health – Europe*, 26. <https://doi.org/10.1016/j.lanepe.2022.100576>
- Fu, C., Hao, W., Shrestha, N., Virani, S. S., Mishra, S. R., & Zhu, D. (2022). Association of reproductive factors with dementia: A systematic review and dose-response meta-analyses of observational studies. *eClinicalMedicine*, 43. <https://doi.org/10.1016/j.eclinm.2021.101236>
- Gagnon, M., Dartigues, J. F., Mazaux, J. M., Dequae, L., Letenneur, L., Giroire, J. M., & Barberger-Gateau, P. (1994). Self-Reported Memory Complaints and Memory Performance in Elderly French Community Residents: Results of the PAQUID Research Program. *Neuroepidemiology*, 13(4), 145–154. <https://doi.org/10.1159/000110373>
- Gallagher, D., Kiss, A., Lanctot, K., & Herrmann, N. (2016). Depressive symptoms and cognitive decline: A longitudinal analysis of potentially modifiable risk factors in community dwelling older adults. *Journal of Affective Disorders*, 190, 235–240. <https://doi.org/10.1016/j.jad.2015.09.046>
- Gallagher, M., & Koh, M. T. (2011). Episodic memory on the path to Alzheimer's disease. *Current Opinion in Neurobiology*, 21(6), 929–934. <https://doi.org/10.1016/j.conb.2011.10.021>
- Galluzzi, S., Pievani, M., Zanetti, O., Benussi, L., The Italian-DIAfN Working Group, Frisoni, G. B., & Di Maria, E. (2022). Disclosure of Genetic Risk Factors for Alzheimer's Disease to Cognitively Healthy Individuals—From Current Practice towards a Personalised Medicine Scenario. *Biomedicines*, 10(12), Article 12. <https://doi.org/10.3390/biomedicines10123177>
- Gatz, M., Reynolds, C. A., Fratiglioni, L., Johansson, B., Mortimer, J. A., Berg, S., Fiske, A., & Pedersen, N. L. (2006). Role of Genes and Environments for Explaining Alzheimer Disease. *Archives of General Psychiatry*, 63(2), 168–174. <https://doi.org/10.1001/archpsyc.63.2.168>

- Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27(15), 2865–2873.
<https://doi.org/10.1002/sim.3107>
- Genin, E., Hannequin, D., Wallon, D., Slegers, K., Hiltunen, M., Combarros, O., Bullido, M. J., Engelborghs, S., De Deyn, P., Berr, C., Pasquier, F., Dubois, B., Tognoni, G., Fiévet, N., Brouwers, N., Bettens, K., Arosio, B., Coto, E., Del Zompo, M., ... Campion, D. (2011). APOE and Alzheimer disease: A major gene with semi-dominant inheritance. *Molecular Psychiatry*, 16(9), 903–907. <https://doi.org/10.1038/mp.2011.52>
- Geraets, A. F. J., & Leist, A. K. (2023). Sex/gender and socioeconomic differences in modifiable risk factors for dementia. *Scientific Reports*, 13(1), 80.
<https://doi.org/10.1038/s41598-022-27368-4>
- Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, 33(2), 271–286.
<https://doi.org/10.1080/13854046.2018.1535662>
- Germine, L., Strong, R. W., Singh, S., & Sliwinski, M. J. (2021). Toward dynamic phenotypes and the scalable measurement of human behavior. *Neuropsychopharmacology*, 46(1), Article 1.
<https://doi.org/10.1038/s41386-020-0757-1>
- Gillan, C. M., Kosinski, M., Whelan, R., Phelps, E. A., & Daw, N. D. (2016). Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife*, 5, e11305. <https://doi.org/10.7554/eLife.11305>
- Gillan, C. M., & Rutledge, R. B. (2021). Smartphones and the Neuroscience of Mental Health. *Annual Review of Neuroscience*, 44(1), 129–151.
<https://doi.org/10.1146/annurev-neuro-101220-014053>
- Gilsanz, P., Mayeda, E. R., Glymour, M. M., Quesenberry, C. P., Mungas, D. M., DeCarli, C., Dean, A., & Whitmer, R. A. (2017). Female sex, early-onset hypertension, and risk of dementia. *Neurology*, 89(18), 1886–1893.
<https://doi.org/10.1212/WNL.0000000000004602>
- Glymour, M. M., Brickman, A. M., Kivimaki, M., Mayeda, E. R., Chêne, G., Dufouil, C., & Manly, J. J. (2018). Will biomarker-based diagnosis of Alzheimer's disease maximize scientific progress? Evaluating proposed diagnostic criteria. *European Journal of Epidemiology*, 33(7), 607–612.
<https://doi.org/10.1007/s10654-018-0418-4>
- Godin, G. (2011). The Godin-Shephard Leisure-Time Physical Activity Questionnaire. *The Health & Fitness Journal of Canada*, 4(1), Article 1.
<https://doi.org/10.14288/hfjc.v4i1.82>
- Gong, J., Harris, K., Peters, S. A. E., & Woodward, M. (2021). Sex differences in the association between major cardiovascular risk factors in midlife and

- dementia: A cohort study using data from the UK Biobank. *BMC Medicine*, 19(1), 110. <https://doi.org/10.1186/s12916-021-01980-z>
- Gorelick, P. B., & Sorond, F. A. (2023). What is brain health? *Cerebral Circulation - Cognition and Behavior*, 6, 100190. <https://doi.org/10.1016/j.cccb.2023.100190>
- Gray, M., Madero, E. N., Gills, J. L., Paulson, S., Jones, M. D., Campitelli, A., Myers, J., Bott, N. T., & Glenn, J. M. (2022). Intervention for a Digital, Cognitive, Multi-Domain Alzheimer Risk Velocity Study: Protocol for a Randomized Controlled Trial. *JMIR Research Protocols*, 11(2), e31841. <https://doi.org/10.2196/31841>
- Greenwood, P. M., Espeseth, T., Lin, M.-K., Reinvang, I., & Parasuraman, R. (2014). Longitudinal change in working memory as a function of APOE genotype in midlife and old age. *Scandinavian Journal of Psychology*, 55(3), 268–277. <https://doi.org/10.1111/sjop.12123>
- Greenwood, P. M., Lambert, C., Sunderland, T., & Parasuraman, R. (2005). Effects of Apolipoprotein E Genotype on Spatial Attention, Working Memory, and Their Interaction in Healthy, Middle-Aged Adults: Results From the National Institute of Mental Health's BIOCARD Study. *Neuropsychology*, 19(2), 199–211. <https://doi.org/10.1037/0894-4105.19.2.199>
- Greenwood, P. M., & Parasuraman, R. (2003). Normal Genetic Variation, Cognition, and Aging. *Behavioral and Cognitive Neuroscience Reviews*, 2(4), 278–306. <https://doi.org/10.1177/1534582303260641>
- Grydeland, H., Vértes, P. E., Váša, F., Romero-Garcia, R., Whitaker, K., Alexander-Bloch, A. F., Bjørnerud, A., Patel, A. X., Sederevičius, D., Tamnes, C. K., Westlye, L. T., White, S. R., Walhovd, K. B., Fjell, A. M., & Bullmore, E. T. (2019). Waves of Maturation and Senescence in Microstructural MRI Markers of Human Cortical Myelination over the Lifespan. *Cerebral Cortex*, 29(3), 1369–1381. <https://doi.org/10.1093/cercor/bhy330>
- Gurvich, C., Hoy, K., Thomas, N., & Kulkarni, J. (2018). Sex Differences and the Influence of Sex Hormones on Cognition through Adulthood and the Aging Process. *Brain Sciences*, 8(9), Article 9. <https://doi.org/10.3390/brainsci8090163>
- Gustavsson, A., Norton, N., Fast, T., Frölich, L., Georges, J., Holzapfel, D., Kirabali, T., Krolak-Salmon, P., Rossini, P. M., Ferretti, M. T., Lanman, L., Chadha, A. S., & van der Flier, W. M. (2023). Global estimates on the number of persons across the Alzheimer's disease continuum. *Alzheimer's & Dementia*, 19(2), 658–670. <https://doi.org/10.1002/alz.12694>
- Hachinski, V., & Avan, A. (2022). From Dementia to Eumentia: A New Approach to Dementia Prevention. *Neuroepidemiology*, 56(3), 151–156. <https://doi.org/10.1159/000525219>

- Hafdi, M., Hoevenaar-Blom, M. P., & Richard, E. (2021). Multi-domain interventions for the prevention of dementia and cognitive decline. *Cochrane Database of Systematic Reviews*, 11. <https://doi.org/10.1002/14651858.CD013572.pub2>
- Han, S. D., & Bondi, M. W. (2008). Revision of the apolipoprotein E compensatory mechanism recruitment hypothesis. *Alzheimer's & Dementia*, 4(4), 251–254. <https://doi.org/10.1016/j.jalz.2008.02.006>
- Harkins, K., Sankar, P., Sperling, R., Grill, J. D., Green, R. C., Johnson, K. A., Healy, M., & Karlawish, J. (2015). Development of a process to disclose amyloid imaging results to cognitively normal older adult research participants. *Alzheimer's Research & Therapy*, 7(1), 26. <https://doi.org/10.1186/s13195-015-0112-7>
- Hastie, T., Friedman, J., & Tibshirani, R. (2001). *The Elements of Statistical Learning*. Springer. <https://doi.org/10.1007/978-0-387-21606-5>
- Hébert, R., Lindsay, J., Verreault, R., Rockwood, K., Hill, G., & Dubois, M.-F. (2000). Vascular Dementia. *Stroke*, 31(7), 1487–1493. <https://doi.org/10.1161/01.STR.31.7.1487>
- Heffernan, M., Andrews, G., Fiatarone Singh, M. A., Valenzuela, M., Anstey, K. J., Maeder, A. J., McNeil, J., Jorm, L., Lautenschlager, N. T., Sachdev, P. S., Ginige, J. A., Hobbs, M. J., Boulamatsis, C., Chau, T., Cobiac, L., Cox, K. L., Daniel, K., Flood, V. M., Guerrero, Y., ... Brodaty, H. (2019). Maintain Your Brain: Protocol of a 3-Year Randomized Controlled Trial of a Personalized Multi-Modal Digital Health Intervention to Prevent Cognitive Decline Among Community Dwelling 55 to 77 Year Olds. *Journal of Alzheimer's Disease*, 70(s1), S221–S237. <https://doi.org/10.3233/JAD-180572>
- Henson, R. N., Suri, S., Knights, E., Rowe, J. B., Kievit, R. A., Lyall, D. M., Chan, D., Eising, E., & Fisher, S. E. (2020). Effect of apolipoprotein E polymorphism on cognition and brain in the Cambridge Centre for Ageing and Neuroscience cohort. *Brain and Neuroscience Advances*, 4, 2398212820961704. <https://doi.org/10.1177/2398212820961704>
- Herlitz, A., Airaksinen, E., & Nordström, E. (1999). Sex differences in episodic memory: The impact of verbal and visuospatial ability. *Neuropsychology*, 13(4), 590–597. <https://doi.org/10.1037/0894-4105.13.4.590>
- Hill, N. L., Bhargava, S., Brown, M. J., Kim, H., Bhang, I., Mullin, K., Phillips, K., & Mogle, J. (2021). Cognitive complaints in age-related chronic conditions: A systematic review. *PLOS ONE*, 16(7), e0253795. <https://doi.org/10.1371/journal.pone.0253795>
- Hill, N. L., Mogle, J., Wion, R., Munoz, E., DePasquale, N., Yevchak, A. M., & Parisi, J. M. (2016). Subjective Cognitive Impairment and Affective

Symptoms: A Systematic Review. *The Gerontologist*, 56(6), e109–e127.
<https://doi.org/10.1093/geront/gnw091>

- Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric Preprocessing for Parametric Causal Inference. *Journal of Statistical Software*, 42(8), 1–28. <https://doi.org/10.18637/jss.v042.i08>
- Ho, J. K., Nation, D. A., & for the Alzheimer's Disease Neuroimaging Initiative. (2018). Neuropsychological Profiles and Trajectories in Preclinical Alzheimer's Disease. *Journal of the International Neuropsychological Society*, 24(7), 693–702. <https://doi.org/10.1017/S135561771800022X>
- Honig, L. S., Barakos, J., Dhadda, S., Kanekiyo, M., Reyderman, L., Irizarry, M., Kramer, L. D., Swanson, C. J., & Sabbagh, M. (2023). ARIA in patients treated with lecanemab (BAN2401) in a phase 2 study in early Alzheimer's disease. *Alzheimer's & Dementia : Translational Research & Clinical Interventions*, 9(1), e12377. <https://doi.org/10.1002/trc2.12377>
- Hsu, S., Gordon, B. A., Hornbeck, R., Norton, J. B., Levitch, D., Loudon, A., Ziegemeier, E., Laforce, R., Chhatwal, J., Day, G. S., McDade, E., Morris, J. C., Fagan, A. M., Benzinger, T. L. S., Goate, A. M., Cruchaga, C., Bateman, R. J., Karch, C. M., & Dominantly Inherited Alzheimer Network (DIAN). (2018). Discovery and validation of autosomal dominant Alzheimer's disease mutations. *Alzheimer's Research & Therapy*, 10(1), 67. <https://doi.org/10.1186/s13195-018-0392-9>
- Hsu, W.-Y., Rowles, W., Anguera, J., Zhao, C., Anderson, A., Alexander, A., Sacco, S., Henry, R., Gazzaley, A., & Bove, R. (2021). Application of an Adaptive, Digital, Game-Based Approach for Cognitive Assessment in Multiple Sclerosis: Observational Study. *Journal of Medical Internet Research*, 23(1), e24356. <https://doi.org/10.2196/24356>
- Huang, Y., Xue, J., Li, C., Chen, X., Fu, H., Fei, T., & Bi, P. (2020). Depression and APOE ϵ 4 Status in Individuals with Subjective Cognitive Decline: A Meta-Analysis. *Psychiatry Investigation*, 17(9), 858–864. <https://doi.org/10.30773/pi.2019.0324>
- Huitfeldt, A. (2016). Is caviar a risk factor for being a millionaire? *BMJ*, 355, i6536. <https://doi.org/10.1136/bmj.i6536>
- Huque, H., Eramudugolla, R., Chidiac, B., Ee, N., Ehrenfeld, L., Matthews, F. E., Peters, R., & Anstey, K. J. (2023). Could Country-Level Factors Explain Sex Differences in Dementia Incidence and Prevalence? A Systematic Review and Meta-Analysis. *Journal of Alzheimer's Disease*, 91(4), 1231–1241. <https://doi.org/10.3233/JAD-220724>
- Hyde, J. S. (2016). Sex and cognition: Gender and cognitive functions. *Current Opinion in Neurobiology*, 38, 53–56. <https://doi.org/10.1016/j.conb.2016.02.007>

- Ihle, A., Bunce, D., & Kliegel, M. (2012). APOE ϵ 4 and cognitive function in early life: A meta-analysis. *Neuropsychology*, 26(3), 267–277. <https://doi.org/10.1037/a0026769>
- IJsselsteijn, W., Tummers-Heemels, A., & Brankaert, R. (2020). Warm Technology: A Novel Perspective on Design for and with People Living with Dementia. In R. Brankaert & G. Kenning (Eds.), *HCI and Design in the Context of Dementia* (pp. 33–47). Springer International Publishing. https://doi.org/10.1007/978-3-030-32835-1_3
- Ilinca, S., & Suzuki, E. (2021). Chapter 12—Gender and socioeconomic differences in modifiable risk factors for Alzheimer’s disease and other types of dementia throughout the life course. In M. T. Ferretti, A. S. Dimech, & A. S. Chadha (Eds.), *Sex and Gender Differences in Alzheimer’s Disease* (pp. 333–360). Academic Press. <https://doi.org/10.1016/B978-0-12-819344-0.00004-1>
- Isaacson, R. S., Ganzer, C. A., Hristov, H., Hackett, K., Caesar, E., Cohen, R., Kachko, R., Meléndez-Cabrero, J., Rahman, A., Scheyer, O., Hwang, M. J., Berkowitz, C., Hendrix, S., Mureb, M., Schelke, M. W., Mosconi, L., Seifan, A., & Krikorian, R. (2018). The clinical practice of risk reduction for Alzheimer’s disease: A precision medicine approach. *Alzheimer’s & Dementia*, 14(12), 1663–1673. <https://doi.org/10.1016/j.jalz.2018.08.004>
- Isernia, S., Cabinio, M., Di Tella, S., Pazzi, S., Vannetti, F., Gerli, F., Mosca, I. E., Lombardi, G., Macchi, C., Sorbi, S., & Baglio, F. (2021). Diagnostic Validity of the Smart Aging Serious Game: An Innovative Tool for Digital Phenotyping of Mild Neurocognitive Disorder. *Journal of Alzheimer’s Disease*, 83(4), 1789–1801. <https://doi.org/10.3233/JAD-210347>
- Iso-Markku, P., Kujala, U. M., Knittle, K., Polet, J., Vuoksima, E., & Waller, K. (2022). Physical activity as a protective factor for dementia and Alzheimer’s disease: Systematic review, meta-analysis and quality assessment of cohort and case–control studies. *British Journal of Sports Medicine*, 56(12), 701–709. <https://doi.org/10.1136/bjsports-2021-104981>
- Jack Jr., C. R., Andrews, J. S., Beach, T. G., Buracchio, T., Dunn, B., Graf, A., Hansson, O., Ho, C., Jagust, W., McDade, E., Molinuevo, J. L., Okonkwo, O. C., Pani, L., Rafii, M. S., Scheltens, P., Siemers, E., Snyder, H. M., Sperling, R., Teunissen, C. E., & Carrillo, M. C. (2024). Revised criteria for diagnosis and staging of Alzheimer’s disease: Alzheimer’s Association Workgroup. *Alzheimer’s & Dementia*, n/a(n/a). <https://doi.org/10.1002/alz.13859>
- Jafari, Z., Kolb, B. E., & Mohajerani, M. H. (2019). Age-related hearing loss and tinnitus, dementia risk, and auditory amplification outcomes. *Ageing Research Reviews*, 56, 100963. <https://doi.org/10.1016/j.arr.2019.100963>

- Jessen, F. (2014). Subjective and objective cognitive decline at the pre-dementia stage of Alzheimer's disease. *European Archives of Psychiatry and Clinical Neuroscience*, 264(1), 3–7. <https://doi.org/10.1007/s00406-014-0539-z>
- Jessen, F., Amariglio, R. E., Buckley, R. F., van der Flier, W. M., Han, Y., Molinuevo, J. L., Rabin, L., Rentz, D. M., Rodriguez-Gomez, O., Saykin, A. J., Sikkes, S. A. M., Smart, C. M., Wolfsgruber, S., & Wagner, M. (2020). The characterisation of subjective cognitive decline. *The Lancet Neurology*, 19(3), 271–278. [https://doi.org/10.1016/S1474-4422\(19\)30368-0](https://doi.org/10.1016/S1474-4422(19)30368-0)
- Jessen, F., Wiese, B., Bickel, H., Eiffländer-Gorfer, S., Fuchs, A., Kaduszkiewicz, H., Köhler, M., Luck, T., Mösch, E., Pentzek, M., Riedel-Heller, S. G., Wagner, M., Weyerer, S., Maier, W., & van den Bussche, H. (2011). Prediction of Dementia in Primary Care Patients. *PLoS ONE*, 6(2), e16852. <https://doi.org/10.1371/journal.pone.0016852>
- Jia, J., Zhao, T., Liu, Z., Liang, Y., Li, F., Li, Y., Liu, W., Li, F., Shi, S., Zhou, C., Yang, H., Liao, Z., Li, Y., Zhao, H., Zhang, J., Zhang, K., Kan, M., Yang, S., Li, H., ... Cummings, J. (2023). Association between healthy lifestyle and memory decline in older adults: 10 year, population based, prospective cohort study. *BMJ*, 380, e072691. <https://doi.org/10.1136/bmj-2022-072691>
- Johar, H., Schaefer, A., & Su, T. T. (2023). Depressive symptoms mediate the longitudinal association between diabetes and subjective cognitive decline. Findings from a semirural multi-ethnic older population in Malaysia. *Preventive Medicine*, 167, 107390. <https://doi.org/10.1016/j.ypmed.2022.107390>
- Jonker, C., Launer, L. J., Hooijer, C., & Lindeboom, J. (1996). Memory Complaints and Memory Impairment in Older Individuals. *Journal of the American Geriatrics Society*, 44(1), 44–49. <https://doi.org/10.1111/j.1532-5415.1996.tb05636.x>
- Josefsson, M., Sundström, A., Pudas, S., Adolfsson, A. N., Nyberg, L., & Adolfsson, R. (2019). Memory profiles predict dementia over 23–28 years in normal but not successful aging. *International Psychogeriatrics*, 1–9. <https://doi.org/10.1017/S1041610219001844>
- Kang, H., Baek, J., Chu, S. H., & Choi, J. (2023). Digital literacy among Korean older adults: A scoping review of quantitative studies. *DIGITAL HEALTH*, 9, 20552076231197334. <https://doi.org/10.1177/20552076231197334>
- Kennedy, J. (1963). Special message to the Congress on the needs of the nation's senior citizens. *Public Papers of the Presidents of the United States: January 1 to November 22*.
- Kesse-Guyot, E., Andreeva, V. A., Lassale, C., Hercberg, S., & Galan, P. (2014). Clustering of Midlife Lifestyle Behaviors and Subsequent Cognitive Function: A Longitudinal Study. *American Journal of Public Health*, 104(11), e170–e177. <https://doi.org/10.2105/AJPH.2014.302121>

- Keusch, F., Bähr, S., Haas, G.-C., Kreuter, F., Trappmann, M., & Eckman, S. (2022). Non-Participation in Smartphone Data Collection Using Research Apps. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 185(Supplement_2), S225–S245. <https://doi.org/10.1111/rssa.12827>
- Kia, D. A., Zhang, D., Guelfi, S., Manzoni, C., Hubbard, L., Reynolds, R. H., Botía, J., Ryten, M., Ferrari, R., Lewis, P. A., Williams, N., Trabzuni, D., Hardy, J., Wood, N. W., & United Kingdom Brain Expression Consortium (UKBEC) and the International Parkinson's Disease Genomics Consortium (IPDGC). (2021). Identification of Candidate Parkinson Disease Genes by Integrating Genome-Wide Association Study, Expression, and Epigenetic Data Sets. *JAMA Neurology*, 78(4), 464–472. <https://doi.org/10.1001/jamaneurol.2020.5257>
- Killin, L., Abrahams, S., Parra, M. A., & Della Sala, S. (2018). The effect of age on the FCSRT-IR and temporary visual memory binding. *International Psychogeriatrics*, 30(3), 331–340. <https://doi.org/10.1017/S104161021700165X>
- Kim, S., Kim, M. J., Kim, S., Kang, H. S., Lim, S. W., Myung, W., Lee, Y., Hong, C. H., Choi, S. H., Na, D. L., Seo, S. W., Ku, B. D., Kim, S. Y., Kim, S. Y., Jeong, J. H., Park, S. A., Carroll, B. J., & Kim, D. K. (2015). Gender differences in risk factors for transition from mild cognitive impairment to Alzheimer's disease: A CREDOS study. *Comprehensive Psychiatry*, 62, 114–122. <https://doi.org/10.1016/j.comppsy.2015.07.002>
- Kivipelto, M., Ngandu, T., Laatikainen, T., Winblad, B., Soininen, H., & Tuomilehto, J. (2006). Risk score for the prediction of dementia risk in 20 years among middle aged people: A longitudinal, population-based study. *The Lancet Neurology*, 5(9), 735–741. [https://doi.org/10.1016/S1474-4422\(06\)70537-3](https://doi.org/10.1016/S1474-4422(06)70537-3)
- Klunzel, O. H., Boer, A. de, Paes, A. H. P., Seidell, J. C., & Bakker, A. (1997). Sex differences in the pharmacological treatment of hypertension: A review of population-based studies. *Journal of Hypertension*, 15(6), 591.
- Knight, R. G., Tsui, H. S. L., Abraham, W. C., Skeaff, C. M., McMahon, J. A., & Cutfield, N. J. (2014). Lack of effect of the apolipoprotein E ϵ 4 genotype on cognition during healthy aging. *Journal of Clinical and Experimental Neuropsychology*, 36(7), 742–750. <https://doi.org/10.1080/13803395.2014.935706>
- Koelewijn, L., Lancaster, T. M., Linden, D., Dima, D. C., Routley, B. C., Magazzini, L., Barawi, K., Brindley, L., Adams, R., Tansey, K. E., Bompas, A., Tales, A., Bayer, A., & Singh, K. (2019). Oscillatory hyperactivity and hyperconnectivity in young APOE- ϵ 4 carriers and hypoconnectivity in Alzheimer's disease. *eLife*, 8, e36011. <https://doi.org/10.7554/eLife.36011>
- Kohler, J. N., Turbitt, E., & Biesecker, B. B. (2017). Personal utility in genomic testing: A systematic literature review. *European Journal of Human Genetics*, 25(6), 662–668. <https://doi.org/10.1038/ejhg.2017.10>

- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kopp, B. (2011). Neuropsychologists Must Keep Their Eyes on the Reliability of Difference Measures. *Journal of the International Neuropsychological Society*, 17(3), 562–563. <https://doi.org/10.1017/S1355617711000361>
- Kronk, C. A., Everhart, A. R., Ashley, F., Thompson, H. M., Schall, T. E., Goetz, T. G., Hiatt, L., Derrick, Z., Queen, R., Ram, A., Guthman, E. M., Danforth, O. M., Lett, E., Potter, E., Sun, S. D., Marshall, Z., & Karnoski, R. (2021). Transgender data collection in the electronic health record: Current concepts and issues. *Journal of the American Medical Informatics Association : JAMIA*, 29(2), 271–284. <https://doi.org/10.1093/jamia/ocab136>
- Kuhn, M., & Johnson, K. (2013). *Applied Predictive Modeling*. Springer New York. <https://doi.org/10.1007/978-1-4614-6849-3>
- Kunz, L., Schröder, T. N., Lee, H., Montag, C., Lachmann, B., Sariyska, R., Reuter, M., Stirnberg, R., Stöcker, T., Messing-Floeter, P. C., Fell, J., Doeller, C. F., & Axmacher, N. (2015). Reduced grid-cell-like representations in adults at genetic risk for Alzheimer's disease. *Science (New York, N.Y.)*, 350(6259), 430–433. <https://doi.org/10.1126/science.aac8128>
- Kuźma, E., Hannon, E., Zhou, A., Lourida, I., Bethel, A., Levine, D. A., Lunnon, K., Thompson-Coon, J., Hyppönen, E., & Llewellyn, D. J. (2018). Which Risk Factors Causally Influence Dementia? A Systematic Review of Mendelian Randomization Studies. *Journal of Alzheimer's Disease*, 64(1), 181–193. <https://doi.org/10.3233/JAD-180013>
- Kuźma, E., Lourida, I., Moore, S. F., Levine, D. A., Ukoumunne, O. C., & Llewellyn, D. J. (2018). Stroke and dementia risk: A systematic review and meta-analysis. *Alzheimer's & Dementia : The Journal of the Alzheimer's Association*, 14(11), 1416–1426. <https://doi.org/10.1016/j.jalz.2018.06.3061>
- Laczó, J., Andel, R., Vyhnaček, M., Vlček, K., Nedelska, Z., Matoska, V., Gazova, I., Mokrisova, I., Sheardova, K., & Hort, J. (2014). APOE and spatial navigation in amnesic MCI: Results from a computer-based test. *Neuropsychology*, 28(5), 676–684. <https://doi.org/10.1037/neu0000072>
- Lambert, J.-C., Ibrahim-Verbaas, C. A., Harold, D., Naj, A. C., Sims, R., Bellenguez, C., Jun, G., DeStefano, A. L., Bis, J. C., Beecham, G. W., Grenier-Boley, B., Russo, G., Thornton-Wells, T. A., Jones, N., Smith, A. V., Chouraki, V., Thomas, C., Ikram, M. A., Zelenika, D., ... Amouyel, P. (2013). Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nature Genetics*, 45(12), 1452–1458. <https://doi.org/10.1038/ng.2802>

- Lancaster, C., Forster, S., Tabet, N., & Rusted, J. (2017). Putting attention in the spotlight: The influence of *APOE* genotype on visual search in mid adulthood. *Behavioural Brain Research*, 334, 97–104.
<https://doi.org/10.1016/j.bbr.2017.07.015>
- Lancaster, C., Koychev, I., Blane, J., Chinner, A., Chatham, C., Taylor, K., & Hinds, C. (2020). Gallery Game: Smartphone-based assessment of long-term memory in adults at risk of Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, 42(4), 329–343.
<https://doi.org/10.1080/13803395.2020.1714551>
- Lancaster, C., Tabet, N., & Rusted, J. (2017). The Elusive Nature of *APOE* ε4 in Mid-adulthood: Understanding the Cognitive Profile. *Journal of the International Neuropsychological Society*, 23(3), 239–253.
<https://doi.org/10.1017/S1355617716000990>
- Lane, C. A., Barnes, J., Nicholas, J. M., Sudre, C. H., Cash, D. M., Parker, T. D., Malone, I. B., Lu, K., James, S.-N., Keshavan, A., Murray-Smith, H., Wong, A., Buchanan, S. M., Keuss, S. E., Gordon, E., Coath, W., Barnes, A., Dickson, J., Modat, M., ... Schott, J. M. (2019). Associations between blood pressure across adulthood and late-life brain structure and pathology in the neuroscience substudy of the 1946 British birth cohort (Insight 46): An epidemiological study. *The Lancet Neurology*, 18(10), 942–952.
[https://doi.org/10.1016/S1474-4422\(19\)30228-5](https://doi.org/10.1016/S1474-4422(19)30228-5)
- Langbaum, J. B., Karlawish, J., Roberts, J. S., Wood, E. M., Bradbury, A., High, N., Walsh, T. L., Gordon, D., Aggarwal, R., Davis, P., Stowell, C., Trisko, L., Langlois, C. M., Reiman, E. M., & Tariot, P. N. (2019). GeneMatch: A novel recruitment registry using at-home *APOE* genotyping to enhance referrals to Alzheimer's prevention studies. *Alzheimer's & Dementia*, 15(4), 515–524.
<https://doi.org/10.1016/j.jalz.2018.12.007>
- LaPlume, A. A., Anderson, N. D., McKetton, L., Levine, B., & Troyer, A. K. (2022). When I'm 64: Age-Related Variability in Over 40,000 Online Cognitive Test Takers. *The Journals of Gerontology: Series B*, 77(1), 104–117.
<https://doi.org/10.1093/geronb/gbab143>
- LaPlume, A. A., McKetton, L., Anderson, N. D., & Troyer, A. K. (2022). Sex differences and modifiable dementia risk factors synergistically influence memory over the adult lifespan. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1), e12301.
<https://doi.org/10.1002/dad2.12301>
- LaPlume, A. A., McKetton, L., Levine, B., Troyer, A. K., & Anderson, N. D. (2022). The adverse effect of modifiable dementia risk factors on cognition amplifies across the adult lifespan. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1), e12337.
<https://doi.org/10.1002/dad2.12337>

- Largent, E. A., Bhardwaj, T., Abera, M., Stites, S. D., Harkins, K., Lerner, A. J., Bradbury, A. R., & Karlawish, J. (2021). Disclosing Genetic Risk of Alzheimer's Disease to Cognitively Unimpaired Older Adults: Findings from the Study of Knowledge and Reactions to APOE Testing (SOKRATES II). *Journal of Alzheimer's Disease : JAD*, 84(3), 1015–1028. <https://doi.org/10.3233/JAD-210675>
- Leduc-McNiven, K., White, B., Zheng, H., D McLeod, R., & R Friesen, M. (2018). Serious games to assess mild cognitive impairment: 'The game is the assessment'. *Research and Review Insights*, 2(1). <https://doi.org/10.15761/RR.1000128>
- Lee, B. H., Richard, J. E., de Leon, R. G., Yagi, S., & Galea, L. A. M. (2023). Sex Differences in Cognition Across Aging. In C. Gibson & L. A. M. Galea (Eds.), *Sex Differences in Brain Function and Dysfunction* (pp. 235–284). Springer International Publishing. https://doi.org/10.1007/7854_2022_309
- Lee, P.-L., Hsiao, C.-H., & Wang, C.-L. (2013). Physical Activity and Memory Complaints in Middle-Age Americans: Results From the MIDUS Study. *American Journal of Alzheimer's Disease & Other Dementias®*, 28(6), 600–605. <https://doi.org/10.1177/1533317513494744>
- Levine, D. A., Gross, A. L., Briceño, E. M., Tilton, N., Giordani, B. J., Sussman, J. B., Hayward, R. A., Burke, J. F., Hingtgen, S., Elkind, M. S. V., Manly, J. J., Gottesman, R. F., Gaskin, D. J., Sidney, S., Sacco, R. L., Tom, S. E., Wright, C. B., Yaffe, K., & Galecki, A. T. (2021). Sex Differences in Cognitive Decline Among US Adults. *JAMA Network Open*, 4(2), e210169. <https://doi.org/10.1001/jamanetworkopen.2021.0169>
- Lewis, M. J., Spiliopoulou, A., Goldmann, K., Pitzalis, C., McKeigue, P., & Barnes, M. R. (2023). nestedcv: An R package for fast implementation of nested cross-validation with embedded feature selection designed for transcriptomics and high-dimensional data. *Bioinformatics Advances*, 3(1), vbad048. <https://doi.org/10.1093/bioadv/vbad048>
- Li, X., Hildebrandt, A., Sommer, W., Wilhelm, O., Reuter, M., Montag, C., & Zhou, C. (2019). Cognitive Performance in Young APOE ε4 Carriers: A Latent Variable Approach for Assessing the Genotype–Phenotype Relationship. *Behavior Genetics*, 49(5), 455–468. <https://doi.org/10.1007/s10519-019-09961-y>
- Li, Y., Macyszko, J. R., Liu, C.-C., & Bu, G. (2022). ApoE4 reduction: An emerging and promising therapeutic strategy for Alzheimer's disease. *Neurobiology of Aging*, 115, 20–28. <https://doi.org/10.1016/j.neurobiolaging.2022.03.011>
- Liang, Y., Pertzov, Y., Nicholas, J. M., Henley, S. M. D., Crutch, S., Woodward, F., Leung, K., Fox, N. C., & Husain, M. (2016). Visual short-term memory binding deficit in familial Alzheimer's disease. *Cortex*, 78, 150–164. <https://doi.org/10.1016/j.cortex.2016.01.015>

- Lim, F. S., Wallace, T., Luszcz, M. A., & Reynolds, K. J. (2013). Usability of Tablet Computers by People with Early-Stage Dementia. *Gerontology*, 59(2), 174–182. <https://doi.org/10.1159/000343986>
- Lim, Y. Y., Ayton, D., Perin, S., Lavale, A., Yassi, N., Buckley, R., Barton, C., Bruns Jr, L., Morello, R., Pirotta, S., Rosenich, E., Rajaratnam, S. M. W., Sinnott, R., Brodtmann, A., Bush, A. I., Maruff, P., Churilov, L., Barker, A., Pase, M. P., & Group, on behalf of the B. R. (2021). An Online, Person-Centered, Risk Factor Management Program to Prevent Cognitive Decline: Protocol for A Prospective Behavior-Modification Blinded Endpoint Randomized Controlled Trial. *Journal of Alzheimer's Disease*, 83(4), 1603–1622. <https://doi.org/10.3233/JAD-210589>
- Lin, F. R., Pike, J. R., Albert, M. S., Arnold, M., Burgard, S., Chisolm, T., Couper, D., Deal, J. A., Goman, A. M., Glynn, N. W., Gmelin, T., Gravens-Mueller, L., Hayden, K. M., Huang, A. R., Knopman, D., Mitchell, C. M., Mosley, T., Pankow, J. S., Reed, N. S., ... Coresh, J. (2023). Hearing intervention versus health education control to reduce cognitive decline in older adults with hearing loss in the USA (ACHIEVE): A multicentre, randomised controlled trial. *The Lancet*, 0(0). [https://doi.org/10.1016/S0140-6736\(23\)01406-X](https://doi.org/10.1016/S0140-6736(23)01406-X)
- Linari, I., Juantorena, G. E., Ibáñez, A., Petroni, A., & Kamienkowski, J. E. (2022). Unveiling Trail Making Test: Visual and manual trajectories indexing multiple executive processes. *Scientific Reports*, 12(1), Article 1. <https://doi.org/10.1038/s41598-022-16431-9>
- Lindenberger, U., Nagel, I. E., Chicherio, C., Li, S.-C., Heekeren, H. R., & Bäckman, L. (2008). Age-related decline in brain resources modulates genetic effects on cognitive functioning. *Frontiers in Neuroscience*, 2. <https://doi.org/10.3389/neuro.01.039.2008>
- Lindhout, J. E., van Dalen, J. W., van Gool, W. A., & Richard, E. (2023). The challenge of dementia prevention trials and the role of quasi-experimental studies. *Alzheimer's & Dementia*, 19(8), 3722–3730. <https://doi.org/10.1002/alz.13029>
- Lisko, I., Kulmala, J., Annetorp, M., Ngandu, T., Mangialasche, F., & Kivipelto, M. (2021). How can dementia and disability be prevented in older adults: Where are we today and where are we going? *Journal of Internal Medicine*, 289(6), 807–830. <https://doi.org/10.1111/joim.13227>
- Liu, T., Luo, H., Tang, J. Y., & Wong, G. H. (2020). Does lifestyle matter? Individual lifestyle factors and their additive effects associated with cognitive function in older men and women. *Aging & Mental Health*, 24(3), 405–412. <https://doi.org/10.1080/13607863.2018.1539833>
- Livingston, G., Huntley, J., Liu, K. Y., Costafreda, S. G., Selbæk, G., Alladi, S., Ames, D., Banerjee, S., Burns, A., Brayne, C., Fox, N. C., Ferri, C. P., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B.,

- Nakasujja, N., Rockwood, K., ... Mukadam, N. (2024). Dementia prevention, intervention, and care: 2024 report of the *Lancet* standing Commission. *The Lancet*, 404(10452), 572–628. [https://doi.org/10.1016/S0140-6736\(24\)01296-0](https://doi.org/10.1016/S0140-6736(24)01296-0)
- Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., ... Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet*, 396(10248), 413–446. [https://doi.org/10.1016/S0140-6736\(20\)30367-6](https://doi.org/10.1016/S0140-6736(20)30367-6)
- Livingston, G., Sommerlad, A., Orgeta, V., Costafreda, S. G., Huntley, J., Ames, D., Ballard, C., Banerjee, S., Burns, A., Cohen-Mansfield, J., Cooper, C., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Larson, E. B., Ritchie, K., Rockwood, K., Sampson, E. L., ... Mukadam, N. (2017). Dementia prevention, intervention, and care. *The Lancet*, 390(10113), 2673–2734. [https://doi.org/10.1016/S0140-6736\(17\)31363-6](https://doi.org/10.1016/S0140-6736(17)31363-6)
- Logie, R. H., Parra, M. A., & Della Sala, S. (2015). From Cognitive Science to Dementia Assessment. *Policy Insights from the Behavioral and Brain Sciences*, 2(1), 81–91. <https://doi.org/10.1177/2372732215601370>
- Loughrey, D. G., Kelly, M. E., Kelley, G. A., Brennan, S., & Lawlor, B. A. (2018). Association of Age-Related Hearing Loss With Cognitive Function, Cognitive Impairment, and Dementia: A Systematic Review and Meta-analysis. *JAMA Otolaryngology–Head & Neck Surgery*, 144(2), 115. <https://doi.org/10.1001/jamaoto.2017.2513>
- Lövdén, M., Fratiglioni, L., Glymour, M. M., Lindenberger, U., & Tucker-Drob, E. M. (2020). Education and Cognitive Functioning Across the Life Span. *Psychological Science in the Public Interest*, 21(1), 6–41. <https://doi.org/10.1177/1529100620920576>
- Lowndes, G., & Savage, G. (2007). Early Detection of Memory Impairment in Alzheimer's Disease: A Neurocognitive Perspective on Assessment. *Neuropsychology Review*, 17(3), 193–202. <https://doi.org/10.1007/s11065-007-9032-z>
- Lu, K., Nicholas, J. M., Pertzov, Y., Grogan, J., Husain, M., Pavisic, I. M., James, S.-N., Parker, T. D., Lane, C. A., Keshavan, A., Keuss, S. E., Buchanan, S. M., Murray-Smith, H., Cash, D. M., Malone, I. B., Sudre, C. H., Coath, W., Wong, A., Henley, S. M. D., ... Crutch, S. J. (2021). Dissociable effects of APOE ϵ 4 and β -amyloid pathology on visual working memory. *Nature Aging*, 1(11), Article 11. <https://doi.org/10.1038/s43587-021-00117-4>
- Lubben, J. E. (1988). Assessing social networks among elderly populations. *Family & Community Health*, 11(3), 42. <https://doi.org/10.1097/00003727-198811000-00008>

- Lugtmeijer, S., Schneegans, S., Lammers, N. A., Geerligs, L., de Leeuw, F. E., de Haan, E. H. F., Bays, P. M., & Kessels, R. P. C. (2021). Consequence of stroke for feature recall and binding in visual working memory. *Neurobiology of Learning and Memory*, 179, 107387. <https://doi.org/10.1016/j.nlm.2021.107387>
- Lumsden, J., Skinner, A., Woods, A. T., Lawrence, N. S., & Munafò, M. (2016). The effects of gamelike features and test location on cognitive test performance and participant enjoyment. *PeerJ*, 4, e2184. <https://doi.org/10.7717/peerj.2184>
- Lundberg, S., & Lee, S.-I. (2017). *A Unified Approach to Interpreting Model Predictions* (arXiv:1705.07874). arXiv. <https://doi.org/10.48550/arXiv.1705.07874>
- Lyu, J., & Kim, H.-Y. (2016). Gender-Specific Incidence and Predictors of Cognitive Impairment among Older Koreans: Findings from a 6-Year Prospective Cohort Study. *Psychiatry Investigation*, 13(5), 473–479. <https://doi.org/10.4306/pi.2016.13.5.473>
- Maccora, J., Peters, R., & Anstey, K. J. (2021). Gender Differences in Superior-memory SuperAgers and Associated Factors in an Australian Cohort. *Journal of Applied Gerontology*, 40(4), 433–442. <https://doi.org/10.1177/0733464820902943>
- Majoka, M. A., & Schimming, C. (2021). Effect of Social Determinants of Health on Cognition and Risk of Alzheimer Disease and Related Dementias. *Clinical Therapeutics*, 43(6), 922–929. <https://doi.org/10.1016/j.clinthera.2021.05.005>
- Marian, A. J. (2012). Molecular genetic studies of complex phenotypes. *Translational Research*, 159(2), 64–79. <https://doi.org/10.1016/j.trsl.2011.08.001>
- Marinelli, J. P., Lohse, C. M., Fussell, W. L., Petersen, R. C., Reed, N. S., Machulda, M. M., Vassilaki, M., & Carlson, M. L. (2022). Association between hearing loss and development of dementia using formal behavioural audiometric testing within the Mayo Clinic Study of Aging (MCSA): A prospective population-based study. *The Lancet. Healthy Longevity*, 3(12), e817–e824. [https://doi.org/10.1016/S2666-7568\(22\)00241-0](https://doi.org/10.1016/S2666-7568(22)00241-0)
- Markopoulou, K. (2010). Patterns of Cognitive Impairment in Neurological Disease. *Current Psychiatry Reviews*, 6(4), 244–251. <https://doi.org/10.2174/157340010793499404>
- Marmot, M., & Allen, J. J. (2014). Social Determinants of Health Equity. *American Journal of Public Health*, 104(S4), S517–S519. <https://doi.org/10.2105/AJPH.2014.302200>

- Marquard, I. A. (2023). A digital intervention to improve or maintain cognitive function in people at risk for cognitive decline. *Alzheimer's & Dementia*, 19(S19), e071810. <https://doi.org/10.1002/alz.071810>
- Martin, M., & Zimprich, D. (2003). Are Changes in Cognitive Functioning in Older Adults Related to Changes in Subjective Complaints? *Experimental Aging Research*, 29(3), 335–352. <https://doi.org/10.1080/0361073030303722>
- Martin, S., Fleming, J., Cullum, S., Denning, T., Rait, G., Fox, C., Katona, C., Brayne, C., & Lafortune, L. (2015). Exploring attitudes and preferences for dementia screening in Britain: Contributions from carers and the general public. *BMC Geriatrics*, 15(1), 110. <https://doi.org/10.1186/s12877-015-0100-6>
- Mascherek, A., Werkle, N., Göritz, A. S., Kühn, S., & Moritz, S. (2020). Lifestyle Variables Do Not Predict Subjective Memory Performance Over and Above Depression and Anxiety. *Frontiers in Psychology*, 11, 484. <https://doi.org/10.3389/fpsyg.2020.00484>
- McCarrey, A. C., An, Y., Kitner-Triolo, M. H., Ferrucci, L., & Resnick, S. M. (2016). Sex differences in cognitive trajectories in clinically normal older adults. *Psychology and Aging*, 31(2), 166–175. <https://doi.org/10.1037/pag0000070>
- McCool, J., Dobson, R., Whittaker, R., & Paton, C. (2022). Mobile Health (mHealth) in Low- and Middle-Income Countries. *Annual Review of Public Health*, 43(Volume 43, 2022), 525–539. <https://doi.org/10.1146/annurev-publhealth-052620-093850>
- McDermott, K. L., McFall, G. P., Andrews, S. J., Anstey, K. J., & Dixon, R. A. (2017). Memory Resilience to Alzheimer's Genetic Risk: Sex Effects in Predictor Profiles. *The Journals of Gerontology: Series B*, 72(6), 937–946. <https://doi.org/10.1093/geronb/gbw161>
- McGue, M., & Christensen, K. (2002). The Heritability of Level and Rate-of-Change in Cognitive Functioning in Danish Twins Aged 70 Years and Older. *Experimental Aging Research*, 28(4), 435–451. <https://doi.org/10.1080/03610730290080416>
- McKiernan, E. F., Mak, E., Dounavi, M.-E., Wells, K., Ritchie, C., Williams, G., Su, L., & O'Brien, J. (2020). Regional hyperperfusion in cognitively normal APOE ε4 allele carriers in mid-life: Analysis of ASL pilot data from the PREVENT-Dementia cohort. *Journal of Neurology, Neurosurgery & Psychiatry*, 91(8), 861–866. <https://doi.org/10.1136/jnnp-2020-322924>
- McNab, F., Zeidman, P., Rutledge, R. B., Smittenaar, P., Brown, H. R., Adams, R. A., & Dolan, R. J. (2015). Age-related changes in working memory and the ability to ignore distraction. *Proceedings of the National Academy of Sciences*. <https://doi.org/10.1073/pnas.1504162112>

- McWhirter, L., Ritchie, C., Stone, J., & Carson, A. (2020). Functional cognitive disorders: A systematic review. *The Lancet. Psychiatry*, 7(2), 191–207. [https://doi.org/10.1016/S2215-0366\(19\)30405-5](https://doi.org/10.1016/S2215-0366(19)30405-5)
- Meador, N., King, K., Moe-Byrne, T., Wright, K., Graham, H., Petticrew, M., Power, C., White, M., & Sowden, A. J. (2016). A systematic review on the clustering and co-occurrence of multiple risk behaviours. *BMC Public Health*, 16(1), 657. <https://doi.org/10.1186/s12889-016-3373-6>
- Messner, D. A. (2011). Informed Choice in Direct-to-Consumer Genetic Testing for Alzheimer and Other Diseases: Lessons from Two Cases. *New Genetics and Society*, 30(1), 59–72. <https://doi.org/10.1080/14636778.2011.552300>
- Mielke, M. M., Aggarwal, N. T., Vila-Castelar, C., Agarwal, P., Arenaza-Urquijo, E. M., Brett, B., Brugulat-Serrat, A., DuBose, L. E., Eikelboom, W. S., Flatt, J., Foldi, N. S., Franzen, S., Gilsanz, P., Li, W., McManus, A. J., van Lent, D. M., Milani, S. A., Shaaban, C. E., Stites, S. D., ... for the Diversity and Disparity Professional Interest Area Sex and Gender Special Interest Group. (2022). Consideration of sex and gender in Alzheimer's disease and related disorders from a global perspective. *Alzheimer's & Dementia*, 18(12), 2707–2724. <https://doi.org/10.1002/alz.12662>
- Minen, M. T., Gopal, A., Sahyoun, G., Stieglitz, E., & Torous, J. (2021). The Functionality, Evidence, and Privacy Issues Around Smartphone Apps for the Top Neuropsychiatric Conditions. *The Journal of Neuropsychiatry and Clinical Neurosciences*, 33(1), 72–79. <https://doi.org/10.1176/appi.neuropsych.19120353>
- Miniawi, S. E., Orgeta, V., & Stafford, J. (2022). Non-affective psychotic disorders and risk of dementia: A systematic review and meta-analysis. *Psychological Medicine*, 52(15), 3323–3335. <https://doi.org/10.1017/S0033291722002781>
- Mishra, J., Anguera, J. A., & Gazzaley, A. (2016). Video Games for Neuro-Cognitive Optimization. *Neuron*, 90(2), 214–218. <https://doi.org/10.1016/j.neuron.2016.04.010>
- Mitchell, A. J., Beaumont, H., Ferguson, D., Yadegarfar, M., & Stubbs, B. (2014). Risk of dementia and mild cognitive impairment in older people with subjective memory complaints: Meta-analysis. *Acta Psychiatrica Scandinavica*, 130(6), 439–451. <https://doi.org/10.1111/acps.12336>
- Mittermaier, M., Venkatesh, K. P., & Kvedar, J. C. (2023). Digital health technology in clinical trials. *Npj Digital Medicine*, 6(1), 1–2. <https://doi.org/10.1038/s41746-023-00841-8>
- Mogle, J. A., Hill, N., & McDermott, C. (2017). Subjective Memory in a National Sample: Predicting Psychological Well-Being. *Gerontology*, 63(5), 460–468. <https://doi.org/10.1159/000466691>

- Mondadori, C. R. A., de Quervain, D. J.-F., Buchmann, A., Mustovic, H., Wollmer, M. A., Schmidt, C. F., Boesiger, P., Hock, C., Nitsch, R. M., Papassotiropoulos, A., & Henke, K. (2007). Better Memory and Neural Efficiency in Young Apolipoprotein E ϵ 4 Carriers. *Cerebral Cortex*, 17(8), 1934–1947. <https://doi.org/10.1093/cercor/bhl103>
- Mortamais, M., Ash, J. A., Harrison, J., Kaye, J., Kramer, J., Randolph, C., Pose, C., Albala, B., Ropacki, M., Ritchie, C. W., & Ritchie, K. (2017). Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility. *Alzheimer's & Dementia*, 13(4), 468–492. <https://doi.org/10.1016/j.jalz.2016.06.2365>
- Muhammad, T., Sekher, T. V., & Srivastava, S. (2022). Association of objective and subjective socioeconomic markers with cognitive impairment among older adults: Cross-sectional evidence from a developing country. *BMJ Open*, 12(8), e052501. <https://doi.org/10.1136/bmjopen-2021-052501>
- Mukadam, N., Sommerlad, A., Huntley, J., & Livingston, G. (2019). Population attributable fractions for risk factors for dementia in low-income and middle-income countries: An analysis using cross-sectional survey data. *The Lancet Global Health*, 7(5), e596–e603. [https://doi.org/10.1016/S2214-109X\(19\)30074-9](https://doi.org/10.1016/S2214-109X(19)30074-9)
- Mukadam, N., Wolters, F. J., Walsh, S., Wallace, L., Brayne, C., Matthews, F. E., Sacuiu, S., Skoog, I., Seshadri, S., Beiser, A., Ghosh, S., & Livingston, G. (2024). Changes in prevalence and incidence of dementia and risk factors for dementia: An analysis from cohort studies. *The Lancet Public Health*, 9(7), e443–e460. [https://doi.org/10.1016/S2468-2667\(24\)00120-8](https://doi.org/10.1016/S2468-2667(24)00120-8)
- Muniz-Terrera, G., Watermeyer, T., Danso, S., & Ritchie, C. (2019). Mobile cognitive testing: Opportunities for aging and neurodegeneration research in low- and middle-income countries. *Journal of Global Health*, 9(2), 020313. <https://doi.org/10.7189/jogh.09.020313>
- Ngandu, T., Lehtisalo, J., Solomon, A., Levälahti, E., Ahtiluoto, S., Antikainen, R., Bäckman, L., Hänninen, T., Jula, A., Laatikainen, T., Lindström, J., Mangialasche, F., Paajanen, T., Pajala, S., Peltonen, M., Rauramaa, R., Stigsdotter-Neely, A., Strandberg, T., Tuomilehto, J., ... Kivipelto, M. (2015). A 2 year multidomain intervention of diet, exercise, cognitive training, and vascular risk monitoring versus control to prevent cognitive decline in at-risk elderly people (FINGER): A randomised controlled trial. *The Lancet*, 385(9984), 2255–2263. [https://doi.org/10.1016/S0140-6736\(15\)60461-5](https://doi.org/10.1016/S0140-6736(15)60461-5)
- Nianogo, R. A., Rosenwohl-Mack, A., Yaffe, K., Carrasco, A., Hoffmann, C. M., & Barnes, D. E. (2022). Risk Factors Associated With Alzheimer Disease and Related Dementias by Sex and Race and Ethnicity in the US. *JAMA Neurology*, 79(6), 584–591. <https://doi.org/10.1001/jamaneurol.2022.0976>

- Nichols, E., Merrick, R., Hay, S. I., Himali, D., Himali, J. J., Hunter, S., Keage, H. A. D., Latimer, C. S., Scott, M. R., Steinmetz, J. D., Walker, J. M., Wharton, S. B., Wiedner, C. D., Crane, P. K., Keene, C. D., Launer, L. J., Matthews, F. E., Schneider, J., Seshadri, S., ... Vos, T. (2023). The prevalence, correlation, and co-occurrence of neuropathology in old age: Harmonisation of 12 measures across six community-based autopsy studies of dementia. *The Lancet Healthy Longevity*, 4(3), e115–e125. [https://doi.org/10.1016/S2666-7568\(23\)00019-3](https://doi.org/10.1016/S2666-7568(23)00019-3)
- Nichols, E. S., Wild, C. J., Owen, A. M., & Soddu, A. (2021). Cognition across the Lifespan: Investigating Age, Sex, and Other Sociodemographic Influences. *Behavioral Sciences*, 11(4), Article 4. <https://doi.org/10.3390/bs11040051>
- Nichols, E., Steinmetz, J. D., Vollset, S. E., Fukutaki, K., Chalek, J., Abd-Allah, F., Abdoli, A., Abualhasan, A., Abu-Gharbieh, E., Akram, T. T., Hamad, H. A., Alahdab, F., Alanezi, F. M., Alipour, V., Almustanyir, S., Amu, H., Ansari, I., Arabloo, J., Ashraf, T., ... Vos, T. (2022). Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: An analysis for the Global Burden of Disease Study 2019. *The Lancet Public Health*, 7(2), e105–e125. [https://doi.org/10.1016/S2468-2667\(21\)00249-8](https://doi.org/10.1016/S2468-2667(21)00249-8)
- Nicosia, J., Aschenbrenner, A. J., Balota, D. A., Sliwinski, M. J., Tahan, M., Adams, S., Stout, S. S., Wilks, H., Gordon, B. A., Benzinger, T. L. S., Fagan, A. M., Xiong, C., Bateman, R. J., Morris, J. C., & Hassenstab, J. (2023). Unsupervised high-frequency smartphone-based cognitive assessments are reliable, valid, and feasible in older adults at risk for Alzheimer's disease. *Journal of the International Neuropsychological Society*, 29(5), 459–471. <https://doi.org/10.1017/S135561772200042X>
- Norman, D. (2019, May 8). *I wrote the book on user-friendly design. What I see today horrifies me*. Fast Company. <https://www.fastcompany.com/90338379/i-wrote-the-book-on-user-friendly-design-what-i-see-today-horrifies-me>
- Norton, D. J., Parra, M. A., Sperling, R. A., Baena, A., Guzman-Velez, E., Jin, D. S., Andrea, N., Khang, J., Schultz, A., Rentz, D. M., Pardiella-Delgado, E., Fuller, J., Johnson, K., Reiman, E. M., Lopera, F., & Quiroz, Y. T. (2020). Visual short-term memory relates to tau and amyloid burdens in preclinical autosomal dominant Alzheimer's disease. *Alzheimer's Research & Therapy*, 12(1), 99. <https://doi.org/10.1186/s13195-020-00660-z>
- Norton, M. C., Dew, J., Smith, H., Fauth, E., Piercy, K. W., Breitner, J. C. S., Tschanz, J., Wengreen, H., Welsh-Bohmer, K., & Investigators, for the C. C. (2012). Lifestyle Behavior Pattern Is Associated with Different levels of Risk for Incident Dementia and Alzheimer's Disease: The Cache County Study. *Journal of the American Geriatrics Society*, 60(3), 405–412. <https://doi.org/10.1111/j.1532-5415.2011.03860.x>
- Nucci, D., Sommariva, A., Degoni, L. M., Gallo, G., Mancarella, M., Ntarelli, F., Savoia, A., Catalini, A., Ferranti, R., Pregliasco, F. E., Castaldi, S., &

- Gianfredi, V. (2024). Association between Mediterranean diet and dementia and Alzheimer disease: A systematic review with meta-analysis. *Aging Clinical and Experimental Research*, 36(1), 77. <https://doi.org/10.1007/s40520-024-02718-6>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.
- Obuchowski, N. A., Lieber, M. L., & Wians, F. H., Jr. (2004). ROC Curves in Clinical Chemistry: Uses, Misuses, and Possible Solutions. *Clinical Chemistry*, 50(7), 1118–1125. <https://doi.org/10.1373/clinchem.2004.031823>
- O'Donoghue, M. C., Murphy, S. E., Zamboni, G., Nobre, A. C., & Mackay, C. E. (2018). APOE genotype and cognition in healthy individuals at risk of Alzheimer's disease: A review. *Cortex*, 104, 103–123. <https://doi.org/10.1016/j.cortex.2018.03.025>
- OECD Data. (2022). *Adult education level*. OECD. <https://www.oecd.org/en/data/indicators/adult-education-level.html>
- Öhman, F., Hassenstab, J., Berron, D., Schöll, M., & Papp, K. V. (2021). Current advances in digital cognitive assessment for preclinical Alzheimer's disease. *Alzheimer's & Dementia : Diagnosis, Assessment & Disease Monitoring*, 13(1), e12217. <https://doi.org/10.1002/dad2.12217>
- Ojakäär, T., & Koychev, I. (2021). Secondary Prevention of Dementia: Combining Risk Factors and Scalable Screening Technology. *Frontiers in Neurology*, 12, 2004. <https://doi.org/10.3389/fneur.2021.772836>
- Pandya, S. Y., Clem, M. A., Silva, L. M., & Woon, F. L. (2016). Does mild cognitive impairment always lead to dementia? A review. *Journal of the Neurological Sciences*, 369, 57–62. <https://doi.org/10.1016/j.jns.2016.07.055>
- Pankratz, V. S., Roberts, R. O., Mielke, M. M., Knopman, D. S., Jack, C. R., Geda, Y. E., Rocca, W. A., & Petersen, R. C. (2015). Predicting the risk of mild cognitive impairment in the Mayo Clinic Study of Aging. *Neurology*, 84(14), 1433–1442. <https://doi.org/10.1212/WNL.0000000000001437>
- Paradise, M. B., Glozier, N. S., Naismith, S. L., Davenport, T. A., & Hickie, I. B. (2011). Subjective memory complaints, vascular risk factors and psychological distress in the middle-aged: A cross-sectional study. *BMC Psychiatry*, 11(1), 108. <https://doi.org/10.1186/1471-244X-11-108>
- Parial, L. L., Lam, S. C., Ho, J. Y. S., Suen, L. K. P., & Leung, A. Y. M. (2021). Public knowledge of the influence of modifiable cardiovascular risk factors on dementia: A systematic literature review and meta-analysis. *Aging & Mental Health*, 25(8), 1395–1409. <https://doi.org/10.1080/13607863.2020.1786801>

- Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17*(2), 299–320.
- Parra, M. A., Abrahams, S., Fabi, K., Logie, R., Luzzi, S., & Sala, S. D. (2009). Short-term memory binding deficits in Alzheimer's disease. *Brain, 132*(4), 1057–1066. <https://doi.org/10.1093/brain/awp036>
- Parra, M. A., Abrahams, S., Logie, R. H., Méndez, L. G., Lopera, F., & Della Sala, S. (2010). Visual short-term memory binding deficits in familial Alzheimer's disease. *Brain, 133*(9), 2702–2713. <https://doi.org/10.1093/brain/awq148>
- Parra, M. A., Gazes, Y., Habeck, C., & Stern, Y. (2024). Exploring the Association between Amyloid- β and Memory Markers for Alzheimer's Disease in Cognitively Unimpaired Older Adults. *The Journal of Prevention of Alzheimer's Disease, 11*(2), 339–347. <https://doi.org/10.14283/jpad.2024.11>
- Passell, E., Strong, R. W., Rutter, L. A., Kim, H., Scheuer, L., Martini, P., Grinspoon, L., & Germine, L. (2021). Cognitive test scores vary with choice of personal digital device. *Behavior Research Methods, 53*(6), 2544–2557. <https://doi.org/10.3758/s13428-021-01597-3>
- Pedersen, M. K., Díaz, C. M. C., Wang, Q. J., Alba-Marrugo, M. A., Amidi, A., Basaiaawmoit, R. V., Bergenholtz, C., Christiansen, M. H., Gajdacz, M., Hertwig, R., Ishkhanyan, B., Klyver, K., Ladegaard, N., Mathiasen, K., Parsons, C., Rafner, J., Villadsen, A. R., Wallentin, M., Zana, B., & Sherson, J. F. (2023). Measuring Cognitive Abilities in the Wild: Validating a Population-Scale Game-Based Cognitive Assessment. *Cognitive Science, 47*(6), e13308. <https://doi.org/10.1111/cogs.13308>
- Pennington, C., Duncan, G., & Ritchie, C. (2021). Altered awareness of cognitive and neuropsychiatric symptoms in Parkinson's disease and Dementia with Lewy Bodies: A systematic review. *International Journal of Geriatric Psychiatry, 36*(1), 15–30. <https://doi.org/10.1002/gps.5415>
- Pennington, C., Hodges, J. R., & Hornberger, M. (2011). Neural Correlates of Episodic Memory in Behavioral Variant Frontotemporal Dementia. *Journal of Alzheimer's Disease, 24*(2), 261–268. <https://doi.org/10.3233/JAD-2011-101668>
- Pérès, K., Helmer, C., Amieva, H., Matharan, F., Carcaillon, L., Jacqmin-Gadda, H., Auriacombe, S., Orgogozo, J.-M., Barberger-Gateau, P., & Dartigues, J.-F. (2011). Gender Differences in the Prodromal Signs of Dementia: Memory Complaint and IADL-Restriction. A Prospective Population-Based Cohort. *Journal of Alzheimer's Disease, 27*(1), 39–47. <https://doi.org/10.3233/JAD-2011-110428>
- Peters, R., Booth, A., Rockwood, K., Peters, J., D'Este, C., & Anstey, K. J. (2019). Combining modifiable risk factors and risk of dementia: A systematic review

and meta-analysis. *BMJ Open*, 9(1), e022846.
<https://doi.org/10.1136/bmjopen-2018-022846>

Peters, R., Xu, Y., Fitzgerald, O., Aung, H. L., Beckett, N., Bulpitt, C., Chalmers, J., Forette, F., Gong, J., Harris, K., Humburg, P., Matthews, F. E., Staessen, J. A., Thijs, L., Tzourio, C., Warwick, J., Woodward, M., Anderson, C. S., & Dementia rlsk REduCTion (DIRECT) collaboration. (2022). Blood pressure lowering and prevention of dementia: An individual patient data meta-analysis. *European Heart Journal*, 43(48), 4980–4990.
<https://doi.org/10.1093/eurheartj/ehac584>

Pew Research Center. (2024). *Mobile Fact Sheet*. Pew Research Center.
<https://www.pewresearch.org/internet/fact-sheet/mobile/>

Polsinelli, A. J., Logan, P. E., Lane, K. A., Manchella, M. K., Nemes, S., Sanjay, A. B., Gao, S., & Apostolova, L. G. (2023). APOE ϵ 4 carrier status and sex differentiate rates of cognitive decline in early- and late-onset Alzheimer's disease. *Alzheimer's & Dementia*, 19(5), 1983–1993.
<https://doi.org/10.1002/alz.12831>

Ponds, R. W., Van Boxtel, M., & Jolles, J. (2010). Age-Related Changes in Subjective Cognitive Functioning. *Educational Gerontology*, 26(1), 67–81.
<https://doi.org/10.1080/036012700267402>

Post, M. W. (2016). What to Do With “Moderate” Reliability and Validity Coefficients? *Archives of Physical Medicine and Rehabilitation*, 97(7), 1051–1052. <https://doi.org/10.1016/j.apmr.2016.04.001>

Proust-Lima, C., Amieva, H., Letenneur, L., Orgogozo, J.-M., Jacqmin-Gadda, H., & Dartigues, J.-F. (2008). Gender and education impact on brain aging: A general cognitive factor approach. *Psychology and Aging*, 23(3), 608–620.
<https://doi.org/10.1037/a0012838>

Psychology Software Tools Inc. (2012). *E-Prime* (Version 2.0) [Computer software].

R Core Team. (2023a). *R: A Language and Environment for Statistical Computing* (Version 4.3.1.) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>

R Core Team. (2023b). *R: A Language and Environment for Statistical Computing* (Version 4.3.2) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>

R Core Team. (2024). *R: A Language and Environment for Statistical Computing* (Version 4.4.1) [Computer software]. R Foundation for Statistical Computing. <https://www.R-project.org/>

- Radloff, L. S. (1977). The CES-D Scale: A Self-Report Depression Scale for Research in the General Population. *Applied Psychological Measurement*, 1(3), 385–401. <https://doi.org/10.1177/014662167700100306>
- Ranson, J. M., Rittman, T., Hayat, S., Brayne, C., Jessen, F., Blennow, K., van Duijn, C., Barkhof, F., Tang, E., Mummery, C. J., Stephan, B. C. M., Altomare, D., Frisoni, G. B., Ribaldi, F., Molinuevo, J. L., Scheltens, P., Llewellyn, D. J., Abramowicz, M., Altomare, D., ... on behalf of the European Task Force for Brain Health Services. (2021). Modifiable risk factors for dementia and dementia risk profiling. A user manual for Brain Health Services—Part 2 of 6. *Alzheimer's Research & Therapy*, 13(1), 169. <https://doi.org/10.1186/s13195-021-00895-4>
- Reid, G., Vassilev, P., Irving, J., Ojakäär, T., Jacobson, L., Lawrence, E. G., Barnett, J., Tapparel, M., & Koychev, I. (2024). The usability and reliability of a smartphone application for monitoring future dementia risk in ageing UK adults. *The British Journal of Psychiatry*, 224(6), 245–251. <https://doi.org/10.1192/bjp.2024.18>
- Reitz, C., Tang, M.-X., Schupf, N., Manly, J. J., Mayeux, R., & Luchsinger, J. A. (2010). A Summary Risk Score for the Prediction of Alzheimer Disease in Elderly Persons. *Archives of Neurology*, 67(7), 835–841. <https://doi.org/10.1001/archneurol.2010.136>
- Reuben, A., Moffitt, T. E., Abraham, W. C., Ambler, A., Elliott, M. L., Hariri, A. R., Harrington, H., Hogan, S., Houts, R. M., Ireland, D., Knodt, A. R., Leung, J., Pearson, A., Poulton, R., Purdy, S. C., Ramrakha, S., Rasmussen, L. J. H., Sugden, K., Thorne, P. R., ... Caspi, A. (2022). Improving risk indexes for Alzheimer's disease and related dementias for use in midlife. *Brain Communications*, 4(5), fcac223. <https://doi.org/10.1093/braincomms/fcac223>
- Riaz, M., Huq, A., Ryan, J., Orchard, S. G., Tiller, J., Lockery, J., Woods, R. L., Wolfe, R., Renton, A. E., Goate, A. M., Sebra, R., Schadt, E., Brodtmann, A., Shah, R. C., Storey, E., Murray, A. M., McNeil, J. J., & Lacaze, P. (2021). Effect of APOE and a polygenic risk score on incident dementia and cognitive decline in a healthy older population. *Aging Cell*, 20(6), e13384. <https://doi.org/10.1111/accel.13384>
- Richiardi, L., Pizzi, C., & Pearce, N. (2013). Commentary: Representativeness is usually not necessary and often should be avoided. *International Journal of Epidemiology*, 42(4), 1018–1022. <https://doi.org/10.1093/ije/dyt103>
- Riedel, B. C., Thompson, P. M., & Brinton, R. D. (2016). Age, APOE and sex: Triad of risk of Alzheimer's disease. *The Journal of Steroid Biochemistry and Molecular Biology*, 160, 134–147. <https://doi.org/10.1016/j.jsbmb.2016.03.012>
- Ritchie, C. W., Bridgeman, K., Gregory, S., O'Brien, J. T., Danso, S. O., Dounavi, M.-E., Carriere, I., Driscoll, D., Hillary, R., Koychev, I., Lawlor, B., Naci, L.,

- Su, L., Low, A., Mak, E., Malhotra, P., Manson, J., Marioni, R., Murphy, L., ... on behalf of the PREVENT dementia programme group. (2024). The PREVENT dementia programme: Baseline demographic, lifestyle, imaging and cognitive data from a midlife cohort study investigating risk factors for dementia. *Brain Communications*, 6(3), fcae189. <https://doi.org/10.1093/braincomms/fcae189>
- Ritchie, C. W., Molinuevo, J. L., Truyen, L., Satlin, A., Van der Geyten, S., & Lovestone, S. (2016). Development of interventions for the secondary prevention of Alzheimer's dementia: The European Prevention of Alzheimer's Dementia (EPAD) project. *The Lancet Psychiatry*, 3(2), 179–186. [https://doi.org/10.1016/S2215-0366\(15\)00454-X](https://doi.org/10.1016/S2215-0366(15)00454-X)
- Ritchie, C. W., & Ritchie, K. (2012). The PREVENT study: A prospective cohort study to identify mid-life biomarkers of late-onset Alzheimer's disease. *BMJ Open*, 2(6), e001893. <https://doi.org/10.1136/bmjopen-2012-001893>
- Ritchie, C. W., Wells, K., & Ritchie, K. (2013). The PREVENT research programme – A novel research programme to identify and manage midlife risk for dementia: The conceptual framework. *International Review of Psychiatry*, 25(6), 748–754. <https://doi.org/10.3109/09540261.2013.869195>
- Ritchie, K., Carrière, I., Howett, D., Su, L., Hornberger, M., O'Brien, J. T., Ritchie, C. W., & Chan, D. (2018). Allocentric and Egocentric Spatial Processing in Middle-Aged Adults at High Risk of Late-Onset Alzheimer's Disease: The PREVENT Dementia Study. *Journal of Alzheimer's Disease*, 65(3), 885–896. <https://doi.org/10.3233/JAD-180432>
- Ritchie, K., Carrière, I., Su, L., O'Brien, J. T., Lovestone, S., Wells, K., & Ritchie, C. W. (2017). The midlife cognitive profiles of adults at high risk of late-onset Alzheimer's disease: The PREVENT study. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 13(10), 1089–1097. <https://doi.org/10.1016/j.jalz.2017.02.008>
- Ritchie, K., Ritchie, C. W., Yaffe, K., Skoog, I., & Scarmeas, N. (2015). Is late-onset Alzheimer's disease really a disease of midlife? *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 1(2), 122–130. <https://doi.org/10.1016/j.trci.2015.06.004>
- Roberts, J. S. (2000). Anticipating response to predictive genetic testing for Alzheimer's disease: A survey of first-degree relatives. *The Gerontologist*, 40(1), 43–52. <https://doi.org/10.1093/geront/40.1.43>
- Rocca, W. A., Mielke, M. M., Vemuri, P., & Miller, V. M. (2014). Sex and gender differences in the causes of dementia: A narrative review. *Maturitas*, 79(2), 196–201. <https://doi.org/10.1016/j.maturitas.2014.05.008>
- Rock, P. L., Roiser, J. P., Riedel, W. J., & Blackwell, A. D. (2014). Cognitive impairment in depression: A systematic review and meta-analysis.

Psychological Medicine, 44(10), 2029–2040.
<https://doi.org/10.1017/S0033291713002535>

- Rodriguez, F. S., Schroeter, M. L., Arélin, K., Witte, A. V., Baber, R., Burkhardt, R., Engel, C., Löffler, M., Thiery, J., Villringer, A., Luck, T., & Riedel-Heller, S. G. (2017). APOE e4-genotype and lifestyle interaction on cognitive performance: Results of the LIFE-Adult-study. *Health Psychology*, 37(2), 194. <https://doi.org/10.1037/hea0000515>
- Roh, H. W., Ryu, H., Jeong, S., Han, J., Park, B., Moon, S. Y., Choi, S. H., Son, S. J., & Hong, C. H. (2022). The effectiveness of a motivational enhancement smartphone application promoting lifestyle improvement for brain health: A randomized controlled trial. *PLOS ONE*, 17(6), e0267806. <https://doi.org/10.1371/journal.pone.0267806>
- Röhr, S., Pabst, A., Baber, R., Engel, C., Glaesmer, H., Hinz, A., Schroeter, M. L., Witte, A. V., Zeynalova, S., Villringer, A., Löffler, M., & Riedel-Heller, S. G. (2022). Social determinants and lifestyle factors for brain health: Implications for risk reduction of cognitive decline and dementia. *Scientific Reports*, 12(1), 12965. <https://doi.org/10.1038/s41598-022-16771-6>
- Rohrer, J. D., Guerreiro, R., Vandrovicova, J., Uphill, J., Reiman, D., Beck, J., Isaacs, A. M., Authier, A., Ferrari, R., Fox, N. C., Mackenzie, I. R. A., Warren, J. D., de Silva, R., Holton, J., Revesz, T., Hardy, J., Mead, S., & Rossor, M. N. (2009). The heritability and genetics of frontotemporal lobar degeneration. *Neurology*, 73(18), 1451–1456. <https://doi.org/10.1212/WNL.0b013e3181bf997a>
- Rouault, M., Seow, T., Gillan, C. M., & Fleming, S. M. (2018). Psychiatric Symptom Dimensions Are Associated With Dissociable Shifts in Metacognition but Not Task Performance. *Biological Psychiatry*, 84(6), 443–451. <https://doi.org/10.1016/j.biopsych.2017.12.017>
- Rudolph, J. E., Zhong, Y., Duggal, P., Mehta, S. H., & Lau, B. (2023). Defining representativeness of study samples in medical and population health research. *BMJ Medicine*, 2(1), e000399. <https://doi.org/10.1136/bmjmed-2022-000399>
- Russell, D., Peplau, L. A., & Ferguson, M. L. (1978). Developing a measure of loneliness. *Journal of Personality Assessment*, 42(3), 290–294. https://doi.org/10.1207/s15327752jpa4203_11
- Rusted, J. M., Evans, S. L., King, S. L., Dowell, N., Tabet, N., & Tofts, P. S. (2013). APOE e4 polymorphism in young adults is associated with improved attention and indexed by distinct neural signatures. *NeuroImage*, 65, 364–373. <https://doi.org/10.1016/j.neuroimage.2012.10.010>
- Rutledge, R. B., Smittenaar, P., Zeidman, P., Brown, H. R., Adams, R. A., Lindenberger, U., Dayan, P., & Dolan, R. J. (2016). Risk Taking for

- Potential Reward Decreases across the Lifespan. *Current Biology*, 26(12), 1634–1639. <https://doi.org/10.1016/j.cub.2016.05.017>
- Sabayan, B., Doyle, S., Rost, N. S., Sorond, F. A., Lakshminarayan, K., & Launer, L. J. (2023). The role of population-level preventive care for brain health in ageing. *The Lancet Healthy Longevity*, 4(6), e274–e283. [https://doi.org/10.1016/S2666-7568\(23\)00051-X](https://doi.org/10.1016/S2666-7568(23)00051-X)
- Sachdev, P. S., Lipnicki, D. M., Crawford, J., Reppermund, S., Kochan, N. A., Trollor, J. N., Draper, B., Slavin, M. J., Kang, K., Lux, O., Mather, K. A., & Brodaty, H. (2012). Risk Profiles for Mild Cognitive Impairment Vary by Age and Sex: The Sydney Memory and Ageing Study. *The American Journal of Geriatric Psychiatry*, 20(10), 854–865. <https://doi.org/10.1097/JGP.0b013e31825461b0>
- Salk, R. H., Hyde, J. S., & Abramson, L. Y. (2017). Gender differences in depression in representative national samples: Meta-analyses of diagnoses and symptoms. *Psychological Bulletin*, 143(8), 783–822. <https://doi.org/10.1037/bul0000102>
- Sallis, H. M., Palmer, T., Tilling, K., Davey Smith, G., & Munafò, M. R. (2023). Using allele scores to identify confounding by reverse causation: Studies of alcohol consumption as an exemplar. *International Journal of Epidemiology*, 52(2), 536–544. <https://doi.org/10.1093/ije/dyac165>
- Salzman, T., Sarquis-Adamson, Y., Son, S., Montero-Odasso, M., & Fraser, S. (2022). Associations of Multidomain Interventions With Improvements in Cognition in Mild Cognitive Impairment: A Systematic Review and Meta-analysis. *JAMA Network Open*, 5(5), e226744. <https://doi.org/10.1001/jamanetworkopen.2022.6744>
- Santamaria-Garcia, H., Sainz-Ballesteros, A., Hernandez, H., Moguilner, S., Maito, M., Ochoa-Rosales, C., Corley, M., Valcour, V., Miranda, J. J., Lawlor, B., & Ibanez, A. (2023). Factors associated with healthy aging in Latin American populations. *Nature Medicine*, 29(9), Article 9. <https://doi.org/10.1038/s41591-023-02495-1>
- Sapkota, S., Bäckman, L., & Dixon, R. A. (2017). Executive function performance and change in aging is predicted by *apolipoprotein E*, intensified by *catechol-O-methyltransferase* and *brain-derived neurotrophic factor*, and moderated by age and lifestyle. *Neurobiology of Aging*, 52, 81–89. <https://doi.org/10.1016/j.neurobiolaging.2016.12.022>
- Saykin, A. J., Wishart, H. A., Rabin, L. A., Santulli, R. B., Flashman, L. A., West, J. D., McHugh, T. L., & Mamourian, A. C. (2006). Older adults with cognitive complaints show brain atrophy similar to that of amnesic MCI. *Neurology*, 67(5), 834–842. <https://doi.org/10.1212/01.wnl.0000234032.77541.a2>
- Scarabino, D., Gambina, G., Broggio, E., Pelliccia, F., & Corbo, R. M. (2016). Influence of family history of dementia in the development and progression

- of late-onset Alzheimer's disease. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics: The Official Publication of the International Society of Psychiatric Genetics*, 171B(2), 250–256.
<https://doi.org/10.1002/ajmg.b.32399>
- Scharfen, J., Peters, J. M., & Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence*, 67, 44–66.
<https://doi.org/10.1016/j.intell.2018.01.003>
- Schiepers, O. J. G., Köhler, S., Deckers, K., Irving, K., O'Donnell, C. A., van den Akker, M., Verhey, F. R. J., Vos, S. J. B., de Vugt, M. E., & van Boxtel, M. P. J. (2018). Lifestyle for Brain Health (LIBRA): A new model for dementia prevention. *International Journal of Geriatric Psychiatry*, 33(1), 167–175.
<https://doi.org/10.1002/gps.4700>
- Schliep, K. C., Barbeau, W. A., Lynch, K. E., Sorweid, M. K., Varner, M. W., Foster, N. L., & Qeadan, F. (2022). Overall and sex-specific risk factors for subjective cognitive decline: Findings from the 2015–2018 Behavioral Risk Factor Surveillance System Survey. *Biology of Sex Differences*, 13, 16.
<https://doi.org/10.1186/s13293-022-00425-3>
- Schneegans, S., & Bays, P. M. (2019). New perspectives on binding in visual working memory. *British Journal of Psychology*, 110(2), 207–244.
<https://doi.org/10.1111/bjop.12345>
- Schott, J. M., Bartlett, J. W., Fox, N. C., Barnes, J., & Investigators, for the A. D. N. I. (2010). Increased brain atrophy rates in cognitively normal older adults with low cerebrospinal fluid A β 1-42. *Annals of Neurology*, 68(6), 825–834.
<https://doi.org/10.1002/ana.22315>
- Seblova, D., Fischer, M., Fors, S., Johnell, K., Karlsson, M., Nilsson, T., Svensson, A. C., Lövdén, M., & Lager, A. (2021). Does Prolonged Education Causally Affect Dementia Risk When Adult Socioeconomic Status Is Not Altered? A Swedish Natural Experiment in 1.3 Million Individuals. *American Journal of Epidemiology*, 190(5), 817–826. <https://doi.org/10.1093/aje/kwaa255>
- Self, W. K., & Holtzman, D. M. (2023). Emerging diagnostics and therapeutics for Alzheimer disease. *Nature Medicine*, 29(9), 2187–2199.
<https://doi.org/10.1038/s41591-023-02505-2>
- Shteyler, V. M., Clarke, J. A., & Adashi, E. Y. (2020). Failed Assignments—Rethinking Sex Designations on Birth Certificates. *The New England Journal of Medicine*, 383(25), 2399–2401.
<https://doi.org/10.1056/NEJMp2025974>
- Shute, V. J., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior*, 63, 106–117.
<https://doi.org/10.1016/j.chb.2016.05.047>

- Siedlecki, K. L., Falzarano, F., & Salthouse, T. A. (2019). Examining Gender Differences in Neurocognitive Functioning Across Adulthood. *Journal of the International Neuropsychological Society*, 25(10), 1051–1060. <https://doi.org/10.1017/S1355617719000821>
- Simons, R. L., Ong, M. L., Beach, S. R. H., Lei, M.-K., Philibert, R., & Mielke, M. (2023). Direct and Indirect Effects of Socioeconomic Status and Discrimination on Subjective Cognitive Decline: A Longitudinal Study of African American Women. *The Journals of Gerontology: Series B*, 78(5), 799–808. <https://doi.org/10.1093/geronb/gbad029>
- Sindi, S., Kåreholt, I., Ngandu, T., Rosenberg, A., Kulmala, J., Johansson, L., Wetterberg, H., Skoog, J., Sjöberg, L., Wang, H.-X., Fratiglioni, L., Skoog, I., & Kivipelto, M. (2021). Sex differences in dementia and response to a lifestyle intervention: Evidence from Nordic population-based studies and a prevention trial. *Alzheimer's & Dementia*, 17(7), 1166–1178. <https://doi.org/10.1002/alz.12279>
- Sindi, S., Toopchiani, S., Barbera, M., Håkansson, K., Lehtisalo, J., Rosenberg, A., Stephen, R., Udeh-Momoh, C., & Kivipelto, M. (2021). Chapter 10—Sex and gender differences in genetic and lifestyle risk and protective factors for dementia. In M. T. Ferretti, A. S. Dimech, & A. S. Chadha (Eds.), *Sex and Gender Differences in Alzheimer's Disease* (pp. 269–308). Academic Press. <https://doi.org/10.1016/B978-0-12-819344-0.00013-2>
- Singh, S., Strong, R. W., Jung, L., Li, F. H., Grinspoon, L., Scheuer, L. S., Passell, E. J., Martini, P., Chaytor, N., Soble, J. R., & Germine, L. (2021). The TestMyBrain Digital Neuropsychology Toolkit: Development and Psychometric Characteristics. *Journal of Clinical and Experimental Neuropsychology*, 0(0), 1–10. <https://doi.org/10.1080/13803395.2021.2002269>
- Singh-Manoux, A., Dugravot, A., Shipley, M., Brunner, E. J., Elbaz, A., Sabia, S., & Kivimäki, M. (2018). Obesity trajectories and risk of dementia: 28 years of follow-up in the Whitehall II Study. *Alzheimer's & Dementia*, 14(2), 178–186. <https://doi.org/10.1016/j.jalz.2017.06.2637>
- Singh-Manoux, A., & Marmot, M. (2005). High blood pressure was associated with cognitive function in middle-age in the Whitehall II study. *Journal of Clinical Epidemiology*, 58(12), 1308–1315. <https://doi.org/10.1016/j.jclinepi.2005.03.016>
- Sipilä, P. N., Heikkilä, N., Lindbohm, J. V., Hakulinen, C., Vahtera, J., Elovainio, M., Suominen, S., Väänänen, A., Koskinen, A., Nyberg, S. T., Pentti, J., Strandberg, T. E., & Kivimäki, M. (2021). Hospital-treated infectious diseases and the risk of dementia: A large, multicohort, observational study with a replication cohort. *The Lancet Infectious Diseases*, 21(11), 1557–1567. [https://doi.org/10.1016/S1473-3099\(21\)00144-4](https://doi.org/10.1016/S1473-3099(21)00144-4)

- Sliwinski, M. J., Mogle, J. A., Hyun, J., Munoz, E., Smyth, J. M., & Lipton, R. B. (2018). Reliability and Validity of Ambulatory Cognitive Assessments. *Assessment*, 25(1), 14–30. <https://doi.org/10.1177/1073191116643164>
- Small, B. J., Rosnick, C. B., Fratiglioni, L., & Bäckman, L. (2004). Apolipoprotein E and Cognitive Performance: A Meta-Analysis. *Psychology and Aging*, 19(4), 592–600. <https://doi.org/10.1037/0882-7974.19.4.592>
- Smittenaar, P., Rutledge, R. B., Zeidman, P., Adams, R. A., Brown, H., Lewis, G., & Dolan, R. J. (2015). Proactive and Reactive Response Inhibition across the Lifespan. *PLOS ONE*, 10(10), e0140383. <https://doi.org/10.1371/journal.pone.0140383>
- Solomon, A., Turunen, H., Ngandu, T., Peltonen, M., Levälahti, E., Helisalmi, S., Antikainen, R., Bäckman, L., Hänninen, T., Jula, A., Laatikainen, T., Lehtisalo, J., Lindström, J., Pajala, S., Stigsdotter-Neely, A., Strandberg, T., Tuomilehto, J., Soininen, H., & Kivipelto, M. (2018). Effect of the Apolipoprotein E Genotype on Cognitive Change During a Multidomain Lifestyle Intervention: A Subgroup Analysis of a Randomized Clinical Trial. *JAMA Neurology*, 75(4), 462–470. <https://doi.org/10.1001/jamaneurol.2017.4365>
- Steele, O. G., Stuart, A. C., Minkley, L., Shaw, K., Bonnar, O., Anderle, S., Penn, A. C., Rusted, J., Serpell, L., Hall, C., & King, S. (2022). A multi-hit hypothesis for an APOE4-dependent pathophysiological state. *European Journal of Neuroscience*, 56(9), 5476–5515. <https://doi.org/10.1111/ejn.15685>
- Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P., Vuksimaa, E., & and the Reserve, Resilience and Protective Factors PIA Empirical Definitions and Conceptual Frameworks Workgroup. (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*, 16(9), 1305–1311. <https://doi.org/10.1016/j.jalz.2018.07.219>
- Stiles, J., & Jernigan, T. L. (2010). The Basics of Brain Development. *Neuropsychology Review*, 20(4), 327–348. <https://doi.org/10.1007/s11065-010-9148-4>
- Stone, A. A., Schneider, S., Smyth, J. M., Junghaenel, D. U., Couper, M. P., Wen, C., Mendez, M., Velasco, S., & Goldstein, S. (2024). A population-based investigation of participation rate and self-selection bias in momentary data capture and survey studies. *Current Psychology*, 43(3), 2074–2090. <https://doi.org/10.1007/s12144-023-04426-2>
- Stone, A. A., Schneider, S., Smyth, J. M., Junghaenel, D. U., Wen, C., Couper, M. P., & Goldstein, S. (2023). Shedding light on participant selection bias in Ecological Momentary Assessment (EMA) studies: Findings from an

internet panel study. *PLOS ONE*, 18(3), e0282591.
<https://doi.org/10.1371/journal.pone.0282591>

- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*, 3rd ed. Oxford University Press.
- Sundermann, E. E., Barnes, L. L., Bondi, M. W., Bennett, D. A., Salmon, D. P., & Maki, P. M. (2021). Improving Detection of Amnesic Mild Cognitive Impairment with Sex-Specific Cognitive Norms. *Journal of Alzheimer's Disease: JAD*, 84(4), 1763–1770. <https://doi.org/10.3233/JAD-215260>
- Sundermann, E. E., Biegon, A., Rubin, L. H., Lipton, R. B., Landau, S., Maki, P. M., & Initiative, for the A. D. N. (2017). Does the Female Advantage in Verbal Memory Contribute to Underestimating Alzheimer's Disease Pathology in Women versus Men? *Journal of Alzheimer's Disease*, 56(3), 947–957. <https://doi.org/10.3233/JAD-160716>
- Sundermann, E. E., Edmonds, E. C., Delano-Wood, L., Galasko, D. R., Salmon, D. P., Rubin, L. H., Bondi, M. W., & Initiative, for the A. D. N. (2018). Sex Influences the Accuracy of Subjective Memory Complaint Reporting in Older Adults. *Journal of Alzheimer's Disease*, 61(3), 1163–1178. <https://doi.org/10.3233/JAD-170425>
- Sundermann, E. E., Katz, M. J., & Lipton, R. B. (2017). Sex Differences in the Relationship between Depressive Symptoms and Risk of Amnesic Mild Cognitive Impairment. *The American Journal of Geriatric Psychiatry*, 25(1), 13–22. <https://doi.org/10.1016/j.jagp.2016.08.022>
- Suthaharan, S. (2016). Support Vector Machine. In S. Suthaharan (Ed.), *Machine Learning Models and Algorithms for Big Data Classification: Thinking with Examples for Effective Learning* (pp. 207–235). Springer US. https://doi.org/10.1007/978-1-4899-7641-3_9
- Teki, S., Kumar, S., & Griffiths, T. D. (2016). Large-Scale Analysis of Auditory Segregation Behavior Crowdsourced via a Smartphone App. *PLOS ONE*, 11(4), e0153916. <https://doi.org/10.1371/journal.pone.0153916>
- Thompson, L. I., Harrington, K. D., Roque, N., Strenger, J., Correia, S., Jones, R. N., Salloway, S., & Sliwinski, M. J. (2022). A highly feasible, reliable, and fully remote protocol for mobile app-based cognitive assessment in cognitively healthy older adults. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 14(1), e12283. <https://doi.org/10.1002/dad2.12283>
- Tian, Y. E., Cropley, V., Maier, A. B., Lautenschlager, N. T., Breakspear, M., & Zalesky, A. (2023). Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. *Nature Medicine*, 1–11. <https://doi.org/10.1038/s41591-023-02296-6>

- Tolppanen, A.-M., Solomon, A., Kulmala, J., Kåreholt, I., Ngandu, T., Rusanen, M., Laatikainen, T., Soininen, H., & Kivipelto, M. (2015). Leisure-time physical activity from mid- to late life, body mass index, and risk of dementia. *Alzheimer's & Dementia: The Journal of the Alzheimer's Association*, 11(4), 434-443.e6. <https://doi.org/10.1016/j.jalz.2014.01.008>
- Tombaugh, T. N. (2004). Trail Making Test A and B: Normative data stratified by age and education. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, 19(2), 203–214. [https://doi.org/10.1016/S0887-6177\(03\)00039-8](https://doi.org/10.1016/S0887-6177(03)00039-8)
- Tong, H. L., Quiroz, J. C., Kocaballi, A. B., Fat, S. C. M., Dao, K. P., Gehringer, H., Chow, C. K., & Laranjo, L. (2021). Personalized mobile technologies for lifestyle behavior change: A systematic review, meta-analysis, and meta-regression. *Preventive Medicine*, 148, 106532. <https://doi.org/10.1016/j.ypmed.2021.106532>
- Tong, T., Chignell, M., Tierney, M. C., & Lee, J. (2016). A Serious Game for Clinical Assessment of Cognitive Status: Validation Study. *JMIR Serious Games*, 4(1), e7. <https://doi.org/10.2196/games.5006>
- Trachtenberg, A. J., Filippini, N., & Mackay, C. E. (2012). The effects of APOE-ε4 on the BOLD response. *Neurobiology of Aging*, 33(2), 323–334. <https://doi.org/10.1016/j.neurobiolaging.2010.03.009>
- Trevis, K. J., McLachlan, N. M., & Wilson, S. J. (2016). Cognitive Mechanisms in Chronic Tinnitus: Psychological Markers of a Failure to Switch Attention. *Frontiers in Psychology*, 7, 1262. <https://doi.org/10.3389/fpsyg.2016.01262>
- Tromp, D., Dufour, A., Lithfous, S., Pebayle, T., & Després, O. (2015). Episodic memory in normal aging and Alzheimer disease: Insights from imaging and behavioral studies. *Ageing Research Reviews*, 24, 232–262. <https://doi.org/10.1016/j.arr.2015.08.006>
- Tu, S., Wong, S., Hodges, J. R., Irish, M., Piguet, O., & Hornberger, M. (2015). Lost in spatial translation – A novel tool to objectively assess spatial disorientation in Alzheimer's disease and frontotemporal dementia. *Cortex*, 67, 83–94. <https://doi.org/10.1016/j.cortex.2015.03.016>
- Tucker-Drob, E. M. (2019). Cognitive Aging and Dementia: A Life Span Perspective. *Annual Review of Developmental Psychology*, 1, 177–196. <https://doi.org/10.1146/annurev-devpsych-121318-085204>
- Turrini, S., Wong, B., Eldaief, M., Press, D. Z., Sinclair, D. A., Koch, G., Avenanti, A., & Santarnecchi, E. (2023). The multifactorial nature of healthy brain ageing: Brain changes, functional decline and protective factors. *Ageing Research Reviews*, 88, 101939. <https://doi.org/10.1016/j.arr.2023.101939>
- Uchida, Y., Sugiura, S., Nishita, Y., Saji, N., Sone, M., & Ueda, H. (2019). Age-related hearing loss and cognitive decline—The potential mechanisms

- linking the two. *Auris Nasus Larynx*, 46(1), 1–9.
<https://doi.org/10.1016/j.anl.2018.08.010>
- Valladares-Rodriguez, S., Fernández-Iglesias, M. J., Anido-Rifón, L., Facal, D., & Pérez-Rodríguez, R. (2018). Episodix: A serious game to detect cognitive impairment in senior adults. A psychometric study. *PeerJ*, 6, e5478.
<https://doi.org/10.7717/peerj.5478>
- Valladares-Rodríguez, S., Pérez-Rodríguez, R., Anido-Rifón, L., & Fernández-Iglesias, M. (2016). Trends on the application of serious games to neuropsychological evaluation: A scoping review. *Journal of Biomedical Informatics*, 64, 296–319. <https://doi.org/10.1016/j.jbi.2016.10.019>
- Valladares-Rodriguez, S., Perez-Rodriguez, R., Facal, D., Fernandez-Iglesias, M. J., Anido-Rifon, L., & Mouriño-Garcia, M. (2017). Design process and preliminary psychometric study of a video game to detect cognitive impairment in senior adults. *PeerJ*, 5, e3508.
<https://doi.org/10.7717/peerj.3508>
- van Dalen, J. W., Brayne, C., Crane, P. K., Fratiglioni, L., Larson, E. B., Lobo, A., Lobo, E., Marcum, Z. A., Moll van Charante, E. P., Qiu, C., Riedel-Heller, S. G., Röhr, S., Rydén, L., Skoog, I., van Gool, W. A., & Richard, E. (2022). Association of Systolic Blood Pressure With Dementia Risk and the Role of Age, U-Shaped Associations, and Mortality. *JAMA Internal Medicine*, 182(2), 142–152. <https://doi.org/10.1001/jamainternmed.2021.7009>
- van Duinkerken, E., & Ryan, C. M. (2020). Diabetes mellitus in the young and the old: Effects on cognitive functioning across the life span. *Neurobiology of Disease*, 134, 104608. <https://doi.org/10.1016/j.nbd.2019.104608>
- Van Patten, R. (2021). Introduction to the Special Issue - Neuropsychology from a distance: Psychometric properties and clinical utility of remote neurocognitive tests. *Journal of Clinical and Experimental Neuropsychology*, 43(8), 767–773.
<https://doi.org/10.1080/13803395.2021.2021645>
- van Rijsbergen, M. W. A., Mark, R. E., de Kort, P. L. M., & Sitskoorn, M. M. (2014). Subjective Cognitive Complaints after Stroke: A Systematic Review. *Journal of Stroke and Cerebrovascular Diseases*, 23(3), 408–420.
<https://doi.org/10.1016/j.jstrokecerebrovasdis.2013.05.003>
- Vanderschaeghe, G., Dierickx, K., & Vandenberghe, R. (2018). Review of the Ethical Issues of a Biomarker-Based Diagnoses in the Early Stage of Alzheimer's Disease. *Journal of Bioethical Inquiry*, 15(2), 219–230.
<https://doi.org/10.1007/s11673-018-9844-y>
- VandeVrede, L., Gibbs, D. M., Koestler, M., La Joie, R., Ljubenkov, P. A., Provost, K., Soleimani-Meigooni, D., Strom, A., Tsoy, E., Rabinovici, G. D., & Boxer, A. L. (2020). Symptomatic amyloid-related imaging abnormalities in an APOE ε4/ε4 patient treated with aducanumab. *Alzheimer's & Dementia*:

Diagnosis, Assessment & Disease Monitoring, 12(1), e12101.
<https://doi.org/10.1002/dad2.12101>

- Varjadic, A., Mantini, D., Demeyere, N., & Gillebert, C. R. (2018). Neural signatures of Trail Making Test performance: Evidence from lesion-mapping and neuroimaging studies. *Neuropsychologia*, 115, 78–87.
<https://doi.org/10.1016/j.neuropsychologia.2018.03.031>
- Veldsman, M., Tai, X.-Y., Nichols, T., Smith, S., Peixoto, J., Manohar, S., & Husain, M. (2020). Cerebrovascular risk factors impact frontoparietal network integrity and executive function in healthy ageing. *Nature Communications*, 11(1), Article 1. <https://doi.org/10.1038/s41467-020-18201-5>
- Ventry, I. M., & Weinstein, B. E. (1983). Identification of elderly people with hearing problems. *ASHA*, 25(7), 37–42.
- Verhagen, S. J. W., Daniëls, N. E. M., Bartels, S. L., Tans, S., Borkelmans, K. W. H., Vugt, M. E. de, & Delespaul, P. A. E. G. (2019). Measuring within-day cognitive performance using the experience sampling method: A pilot study in a healthy population. *PLOS ONE*, 14(12), e0226409.
<https://doi.org/10.1371/journal.pone.0226409>
- Vikbladh, O. M., Meager, M. R., King, J., Blackmon, K., Devinsky, O., Shohamy, D., Burgess, N., & Daw, N. D. (2019). Hippocampal Contributions to Model-Based Planning and Spatial Memory. *Neuron*, 102(3), 683–693.e4.
<https://doi.org/10.1016/j.neuron.2019.02.014>
- Villemagne, V. L., Burnham, S., Bourgeat, P., Brown, B., Ellis, K. A., Salvado, O., Szoek, C., Macaulay, S. L., Martins, R., Maruff, P., Ames, D., Rowe, C. C., & Masters, C. L. (2013). Amyloid β deposition, neurodegeneration, and cognitive decline in sporadic Alzheimer's disease: A prospective cohort study. *The Lancet Neurology*, 12(4), 357–367.
[https://doi.org/10.1016/S1474-4422\(13\)70044-9](https://doi.org/10.1016/S1474-4422(13)70044-9)
- Vinke, E. J., de Groot, M., Venkatraghavan, V., Klein, S., Niessen, W. J., Ikram, M. A., & Vernooij, M. W. (2018). Trajectories of imaging markers in brain aging: The Rotterdam Study. *Neurobiology of Aging*, 71, 32–40.
<https://doi.org/10.1016/j.neurobiolaging.2018.07.001>
- Vos, S. J. B., van Boxtel, M. P. J., Schiepers, O. J. G., Deckers, K., de Vugt, M., Carrière, I., Dartigues, J.-F., Peres, K., Artero, S., Ritchie, K., Galluzzo, L., Scafato, E., Frisoni, G. B., Huisman, M., Comijs, H. C., Sacuiu, S. F., Skoog, I., Irving, K., O'Donnell, C. A., ... Köhler, S. (2017). Modifiable Risk Factors for Prevention of Dementia in Midlife, Late Life and the Oldest-Old: Validation of the LIBRA Index. *Journal of Alzheimer's Disease*, 58(2), 537–547. <https://doi.org/10.3233/JAD-161208>
- Walkowiak, S., Coutrot, A., Hegarty, M., Velasco, P. F., Wiener, J. M., Dalton, R. C., Hölscher, C., Hornberger, M., Spiers, H. J., & Manley, E. (2023).

- Cultural determinants of the gap between self-estimated navigation ability and wayfinding performance: Evidence from 46 countries. *Scientific Reports*, 13(1), Article 1. <https://doi.org/10.1038/s41598-023-30937-w>
- Wang, A.-Y., Hu, H.-Y., Ou, Y.-N., Wang, Z.-T., Ma, Y.-H., Tan, L., & Yu, J.-T. (2023). Socioeconomic Status and Risks of Cognitive Impairment and Dementia: A Systematic Review and Meta-Analysis of 39 Prospective Studies. *The Journal of Prevention of Alzheimer's Disease*, 10(1), 83–94. <https://doi.org/10.14283/jpad.2022.81>
- Wang, S., Zheng, X., Huang, J., Liu, J., Li, C., & Shang, H. (2024). Sleep characteristics and risk of Alzheimer's disease: A systematic review and meta-analysis of longitudinal studies. *Journal of Neurology*, 271(7), 3782–3793. <https://doi.org/10.1007/s00415-024-12380-7>
- Wang, X., Zhou, W., Ye, T., Lin, X., Zhang, J., & Initiative, for A. D. N. (2019). Sex Difference in the Association of APOE4 with Memory Decline in Mild Cognitive Impairment. *Journal of Alzheimer's Disease*, 69(4), 1161–1169. <https://doi.org/10.3233/JAD-181234>
- Weichbroth, P. (2020). Usability of Mobile Applications: A Systematic Literature Study. *IEEE Access*, 8, 55563–55577. IEEE Access. <https://doi.org/10.1109/ACCESS.2020.2981892>
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D., Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., ... Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, 80(11_supplement_3), S54–S64. <https://doi.org/10.1212/WNL.0b013e3182872ded>
- Weissberger, G. H., Nation, D. A., Nguyen, C. P., Bondi, M. W., & Han, S. D. (2018). Meta-analysis of cognitive ability differences by apolipoprotein e genotype in young humans. *Neuroscience & Biobehavioral Reviews*, 94, 49–58. <https://doi.org/10.1016/j.neubiorev.2018.08.009>
- Wen, C., Hu, H., Ou, Y.-N., Bi, Y.-L., Ma, Y.-H., Tan, L., & Yu, J.-T. (2021). Risk factors for subjective cognitive decline: The CABLE study. *Translational Psychiatry*, 11(1), Article 1. <https://doi.org/10.1038/s41398-021-01711-1>
- Whelan, R., Barbey, F. M., Cominetti, M. R., Gillan, C. M., & Rosická, A. M. (2022). Developments in scalable strategies for detecting early markers of cognitive decline. *Translational Psychiatry*, 12(1), Article 1. <https://doi.org/10.1038/s41398-022-02237-w>
- Wightman, D. P., Jansen, I. E., Savage, J. E., Shadrin, A. A., Bahrami, S., Holland, D., Rongve, A., Børte, S., Winsvold, B. S., Drange, O. K., Martinsen, A. E., Skogholt, A. H., Willer, C., Bråthen, G., Bosnes, I., Nielsen, J. B., Fritsche, L. G., Thomas, L. F., Pedersen, L. M., ...

- Posthuma, D. (2021). A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nature Genetics*, 53(9), Article 9. <https://doi.org/10.1038/s41588-021-00921-z>
- Wilks, H., Aschenbrenner, A. J., Gordon, B. A., Balota, D. A., Fagan, A. M., Musiek, E., Balls-Berry, J., Benzinger, T. L. S., Cruchaga, C., Morris, J. C., & Hassenstab, J. (2022). Sharper in the morning: Cognitive time of day effects revealed with high-frequency smartphone testing. *Journal of Clinical and Experimental Neuropsychology*, 43(8), 1–13. <https://doi.org/10.1080/13803395.2021.2009447>
- Wisdom, N. M., Callahan, J. L., & Hawkins, K. A. (2011). The effects of apolipoprotein E on non-impaired cognitive functioning: A meta-analysis. *Neurobiology of Aging*, 32(1), 63–74. <https://doi.org/10.1016/j.neurobiolaging.2009.02.003>
- Wolters, F. J., Chibnik, L. B., Waziry, R., Anderson, R., Berr, C., Beiser, A., Bis, J. C., Blacker, D., Bos, D., Brayne, C., Dartigues, J.-F., Darweesh, S. K. L., Davis-Plourde, K. L., de Wolf, F., Debette, S., Dufouil, C., Fornage, M., Goudsmit, J., Grasset, L., ... Hofman, A. (2020). Twenty-seven-year time trends in dementia incidence in Europe and the United States. *Neurology*, 95(5), e519–e531. <https://doi.org/10.1212/WNL.00000000000010022>
- World Health Organization. (2015). *World Report on Ageing and Health*. World Health Organization. <https://www.who.int/publications/i/item/9789241565042>
- World Health Organization. (2019). *Risk reduction of cognitive decline and dementia: WHO guidelines*. World Health Organization. <https://iris.who.int/handle/10665/312180>
- Wu, Q. (2016). Subjective cognitive impairment of older adults: A comparison between the US and China. *International Journal of Methods in Psychiatric Research*, 25(1), 68–75. <https://doi.org/10.1002/mpr.1499>
- Xiong, L. Y., Wood Alexander, M., Wong, Y. Y., Wu, C.-Y., Ruthirakuhan, M., Edwards, J. D., Lanctôt, K. L., Black, S. E., Rabin, J. S., Cogo-Moreira, H., & Swardfager, W. (2024). Latent profiles of modifiable dementia risk factors in later midlife: Relationships with incident dementia, cognition, and neuroimaging outcomes. *Molecular Psychiatry*, 1–11. <https://doi.org/10.1038/s41380-024-02685-4>
- Yarkoni, T., & Westfall, J. (2017). Choosing Prediction Over Explanation in Psychology: Lessons From Machine Learning. *Perspectives on Psychological Science*, 12(6), 1100–1122. <https://doi.org/10.1177/1745691617693393>
- Yasen, A. L., Raber, J., Miller, J. K., & Piper, B. J. (2015). Sex, but not Apolipoprotein E Polymorphism, Differences in Spatial Performance in

- Young Adults. *Archives of Sexual Behavior*, 44(8), 2219–2226.
<https://doi.org/10.1007/s10508-015-0497-1>
- Yassuda, M. S., Carthery-Goulart, M. T., Cecchini, M. A., Cassimiro, L., Fernandes, K. D., Baradel, R. R., Garcia, R. B., Nitrini, R., Della Sala, S., & Parra, M. A. (2020). Free Recall of Bound Information Held in Short-Term Memory is Unimpaired by Age and Education. *Archives of Clinical Neuropsychology*, 35(2), 165–175. <https://doi.org/10.1093/arclin/acz015>
- Yates, J. A., Clare, L., & Woods, R. T. (2017). Subjective memory complaints, mood and MCI: A follow-up study. *Aging & Mental Health*, 21(3), 313–321. <https://doi.org/10.1080/13607863.2015.1081150>
- You, J., Zhang, Y.-R., Wang, H.-F., Yang, M., Feng, J.-F., Yu, J.-T., & Cheng, W. (2022). Development of a novel dementia risk prediction model in the general population: A large, longitudinal, population-based machine-learning study. *eClinicalMedicine*, 53, 101665. <https://doi.org/10.1016/j.eclinm.2022.101665>
- Zhang, M. (2023). Older people's attitudes towards emerging technologies: A systematic literature review. *Public Understanding of Science*, 32(8), 948–968. <https://doi.org/10.1177/09636625231171677>
- Zhang, T., Liu, S., Zhang, Y., Guan, Y., Wang, X., Zhao, L., Shi, Z., Yue, W., Zhang, Y., Liu, S., & Ji, Y. (2017). Apolipoprotein E ε4 Allele Is Associated with Subjective Cognitive Decline: A Meta-Analysis. *Neuroepidemiology*, 49(3–4), 165–173. <https://doi.org/10.1159/000482018>
- Zokaei, N., Giehl, K., Sillence, A., Neville, M. J., Karpe, F., Nobre, A. C., & Husain, M. (2017). Sex and APOE: A memory advantage in male APOE ε4 carriers in midlife. *Cortex*, 88, 98–105. <https://doi.org/10.1016/j.cortex.2016.12.016>
- Zokaei, N., Grogan, J., Fallon, S. J., Slavkova, E., Hadida, J., Manohar, S., Nobre, A. C., & Husain, M. (2020). Short-term memory advantage for brief durations in human APOE ε4 carriers. *Scientific Reports*, 10(1), Article 1. <https://doi.org/10.1038/s41598-020-66114-6>
- Zullo, L., Clark, C., Gholam, M., Castelao, E., von Gunten, A., Preisig, M., & Popp, J. (2021). Factors associated with subjective cognitive decline in dementia-free older adults—A population-based study. *International Journal of Geriatric Psychiatry*, 36(8), 1188–1196. <https://doi.org/10.1002/gps.5509>
- Zygouris, S., Giakoumis, D., Votis, K., Doumpoulakis, S., Ntovas, K., Segkouli, S., Karagiannidis, C., Tzovaras, D., & Tsolaki, M. (2015). Can a Virtual Reality Cognitive Training Application Fulfill a Dual Role? Using the Virtual Supermarket Cognitive Training Application as a Screening Tool for Mild Cognitive Impairment. *Journal of Alzheimer's Disease*, 44(4), 1333–1347. <https://doi.org/10.3233/JAD-141260>

8 Appendices

8.1 Appendix I: Supplemental Information for Chapter 3

8.1.1 Age distribution in the final sample

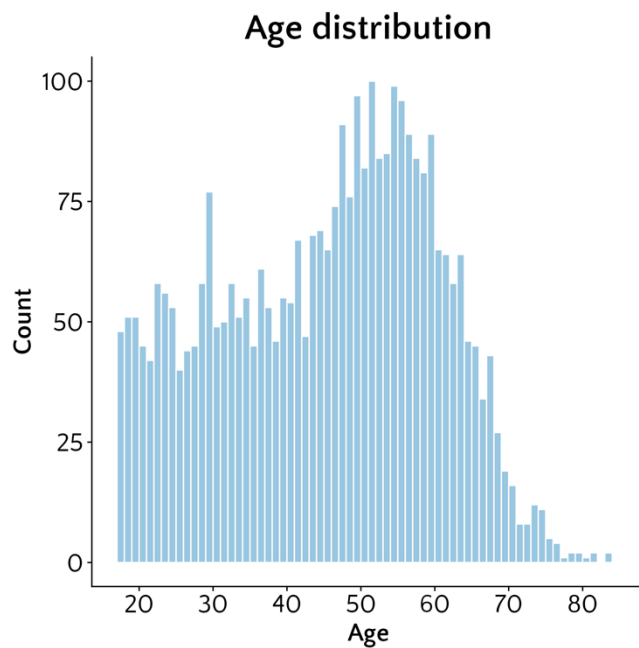


Figure S1. Age histogram showing the number of participants per years of age in the final sample ($N = 3,327$).

8.1.2 Analysis of sample representativeness with regards to SES

Socioeconomic status (SES) can drive the co-occurrence of dementia risk factors (Meader et al., 2016; Röhr et al., 2022), making SES a likely confound in online epidemiological research. To assess potential bias arising from SES, we evaluated how representative¹ the current study sample is with respect to SES within

¹ Note that the term “representativeness” is used here to describe the degree to which the sample reflects expected characteristics of the target population, such as educational stratification. This differs from statistical representativeness, which requires truly random sampling from the target population. Representativeness in the current sense is necessary but not sufficient to ensure that study results will be representative (Richiardi et al., 2013; Rudolph et al., 2023).

participants' respective countries, using educational achievement as a proxy. Participants in the total sample ($N = 3,327$) came from 54 different countries. However, 94.14% of the total sample ($n = 3,132$) came from only five countries: the United Kingdom ($n = 1,650$), the United States ($n = 936$), Ireland ($n = 443$), Canada ($n = 57$), and Germany ($n = 46$). For feasibility, the representativeness analyses were restricted to these five most frequent countries.

Neureka data on educational achievement were recoded into three categories aligned with OECD classifications: 1) below upper secondary, 2) upper secondary or post-secondary non-tertiary, and 3) tertiary (OECD Data, 2022). To ensure comparability with OECD national estimates for 2022 (OECD Data, 2022), we further restricted the Neureka dataset to adults aged 25–64, resulting in $n = 2,550$ (i.e., 76.65% of the total sample). Within each country, the distribution of education levels in the Neureka sample was then nominally compared to the respective 2022 OECD percentages. The results are presented in Table S1.

Compared to OECD data, individuals with tertiary education were overrepresented in the Neureka sample, with the most striking differences observed among German participants (80% in Neureka vs. 33% in OECD data). The only exception was Canada, where a smaller proportion of participants had tertiary education compared to national estimates (54% vs. 63%). Conversely, individuals with less than upper secondary education were underrepresented in all countries, with the exception of the United States (9% in Neureka vs. 8% in OECD data). These findings suggest that the Neureka sample may not be fully representative of the general population in terms of educational attainment, which could contribute to SES-related sampling bias. This overrepresentation of highly educated individuals was not surprising, given that participation in a smartphone-based study may be more accessible or appealing to those with higher levels of education. This is even more likely the case for participants whose first language is not English (e.g., German nationals).

Table S1. Education data for the five most frequent countries in Neureka data, relative to OECD 2022 indicators. Only participants aged 25–64 were included to allow for direct comparison, resulting in $n = 2,550$, i.e., 76.65% of the total sample.

Country (n in Neureka)	Education*	Neureka data (collected in 2020–2023) n (%) per group	OECD indicator data (OECD, 2022) % per group
United Kingdom ($n = 1,348$)	Below upper secondary	109 (8.09%)	18.56%
		340 (25.22%)	30.12%
	Upper secondary	899 (66.69%)	51.32%
	Tertiary		
United States ($n = 800$)	Below upper secondary	79 (9.88%)	8.16%
		184 (23.00%)	41.82%
	Upper secondary	537 (67.13%)	50.02%
	Tertiary		
Ireland ($n = 325$)	Below upper secondary	14 (4.30%)	12.37%
		82 (25.23%)	33.25%
	Upper secondary	229 (70.46%)	54.39%
	Tertiary		
Canada ($n = 41$)	Below upper secondary	5 (12.20%)	6.84%
		14 (34.15%)	30.42%
	Upper secondary	22 (53.66%)	62.74%
	Tertiary		
Germany ($n = 36$)	Below upper secondary	1 (2.78%)	16.46%
		6 (16.67%)	51.02%
	Upper secondary	29 (80.56%)	32.52%
	Tertiary		

NOTE. *Neureka education was mapped onto OECD education categories as follows: 1) Below upper secondary = No formal education or Lower secondary; 2) Upper secondary = Upper secondary; 3) Tertiary = University/College degree or equivalent, Master's degree, or PhD.

8.1.3 Sensitivity analysis excluding non-native speakers

We have replicated the main analyses of the paper in a subset of $n = 2,996$ participants who reported having English as their first language. This analysis did not reveal a substantially different pattern of results, although the association between diabetes and working memory was rendered non-significant at $p < .0038$ (see Figure S2 below), probably due to the reduction in sample size. Additionally, the association between history of stroke and binarised cognitive flexibility (panel C) was now significant, unlike in the full sample analyses. Overall, these results illustrate that the inclusion of non-native speakers did not have a substantial effect on the results.

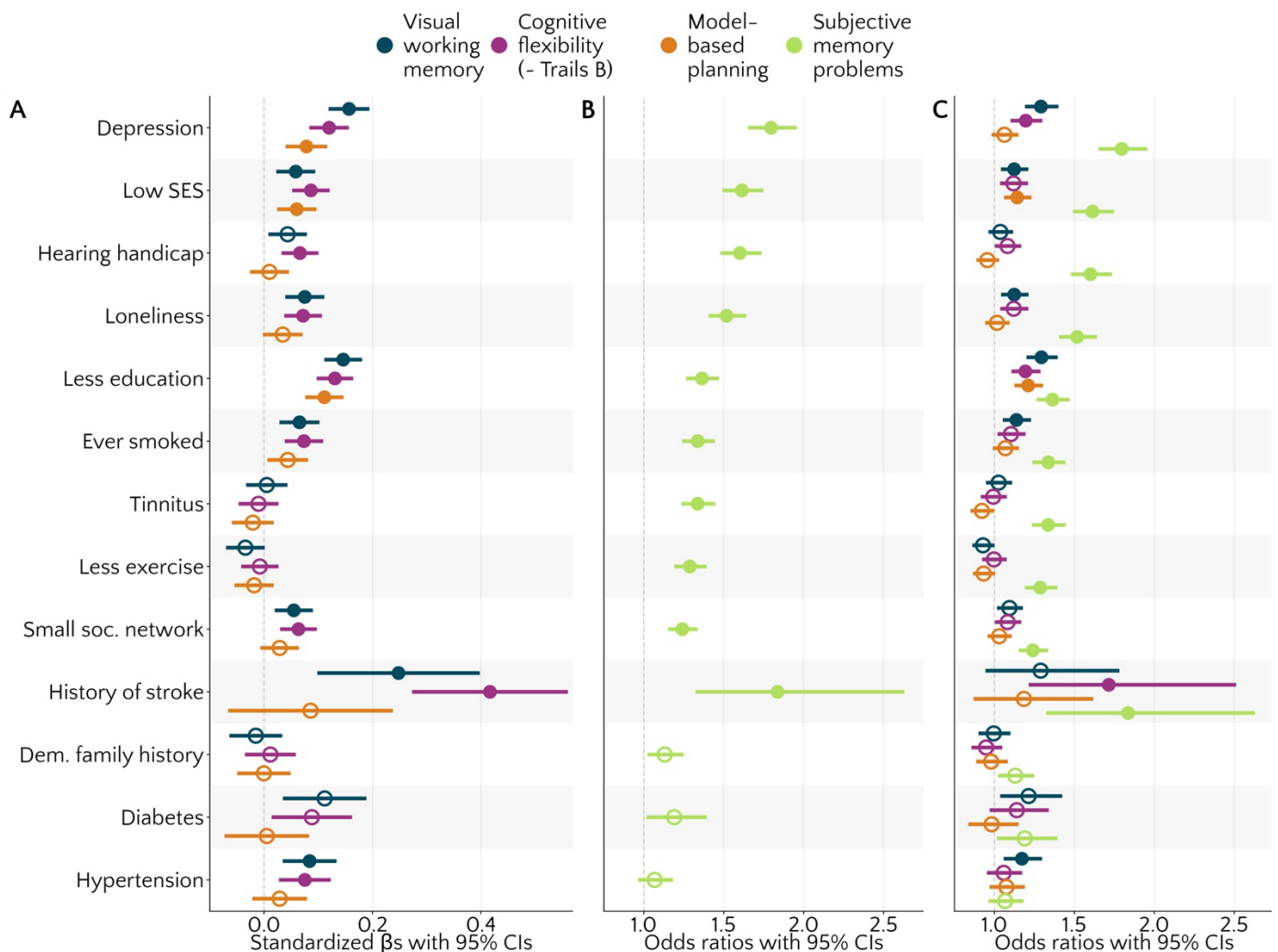


Figure S2. Associations between cognitive measures and risk factors, controlling for age and gender, in the subset of native-speaking participants ($n = 2,996$). Filled circles represent values significant at $p < .0038$, i.e., after applying Bonferroni

correction per dependent variable. Higher estimates indicate worse cognitive outcomes. Circles and lines represent (A) standardised beta estimates with 95% confidence intervals or (B–C) odds ratios with 95% confidence intervals. In (C), objective cognitive scores were binarised to enable direct comparison of odds ratios with subjective memory problems.

8.1.4 Neureka Recruitment Strategies

Participants were recruited for *Neureka* through a variety of methods since the app's release in June 2020 and the recruitment strategies have been changing over time. Broadly speaking, the ads are targeted at people interested in contributing to research of mental health, brain health, or both. Irrespective of the recruitment strategy, potential participants are brought to the *Neureka* page in Google Play Store (Android) or in App Store (iOS), where they can learn more about the app and browse screenshots before deciding whether to download the app. The contents of the store pages are described in more detail below, followed by an overview of the most common recruitment strategies.

8.1.4.1 App Store Description and Screenshots

The following description is shown to users of both Google Play Store (Android) and App Store (iOS). See Figure S3 for screenshots attached to this description.

Description

Neureka is a collection of brain games, quizzes, surveys and challenges that allows you to solve cognitive puzzles and help in cutting edge scientific research. 450 million people suffer from mental health conditions worldwide and a further 50 million are currently living with dementia. Using gamified cognitive tests the Neureka application aims to allow you the user to contribute and help in the global fight against dementia.

The Neureka App has the following fun games and challenges to play

Risk Factors - Play this fun set of games against the clock and complete questionnaires to help us understand how to prevent dementia!

The Science

There is a growing body of knowledge that up to 30% of cases of dementia could be prevented through what are called “modifiable risk factors”. We want to understand why this is the case - to unpack the complex relationship that exists between mind, body and environment so we can figure out new ways to fight dementia.

In this challenge you will play three fun games that require different types of thinking (cognitive) processes. Prior scientific research has suggested changes in these cognitive processes may predict people’s risk of dementia years before diagnosis.

By completing this challenge you will help us understand how potential risk factors might confer risk for developing dementia and in doing so, pave the way towards new and effective interventions that can keep us all healthy for longer!

Multi-Mood - Log your moods over 8 weeks and help detect changes in your mood!

The Science

Major depression affects over 300 million people worldwide at any given moment. Approximately 15% of all people will develop depression at some point in their lives.

In MultiMood, you rate your mood twice a day over 8-weeks and complete a depression questionnaire once a week. Research has shown that distinct changes in mood occur before entering a period of depression. Through MultiMood, you’ll be helping us understand how interactions between different aspects of mood contribute to the risk of experiencing a depressive episode.

Memory Match - Play this memory matching game where you identify shapes and numbers against the clock !

The Science

Research suggests that difficulties linking different forms of memory together (e.g. shapes and colours) is impaired very early on in dementia (potentially 10 years before diagnosis). It is thought that this is because the parts of the brain that link these memories together is one of the earliest to be affected by the disease

Star Racer - Collect points against the clock by selecting the numbers and letters as fast as you can !

The Science

This game assesses aspects of your mental flexibility (your ability to switch between numbers and letters) and your ability to maintain two trains of thought simultaneously (keeping track of where you are in the number & letter sequences).

These abilities are known as 'executive functions' and prior research has repeatedly shown that they decline 2-3 years prior to dementia diagnosis.

Cannon Blast - Shoot the cannon and collect diamonds as you move through levels collecting as many points as you can!

The Science

Research suggests that the ability to anticipate the consequences of future decisions and make optimal choices based on them (known as model-based planning) is impaired in older adults. It is thought that one reason for age-related decline in this type of decision making is atrophy of an area of the brain called the hippocampus. Dementia results in early and severe hippocampal atrophy and therefore the decision-making abilities tracked by Canon Blast may be sensitive to early signs of dementia.

Download Neureka today, play fun games and challenges and help us fight dementia and mental illness.

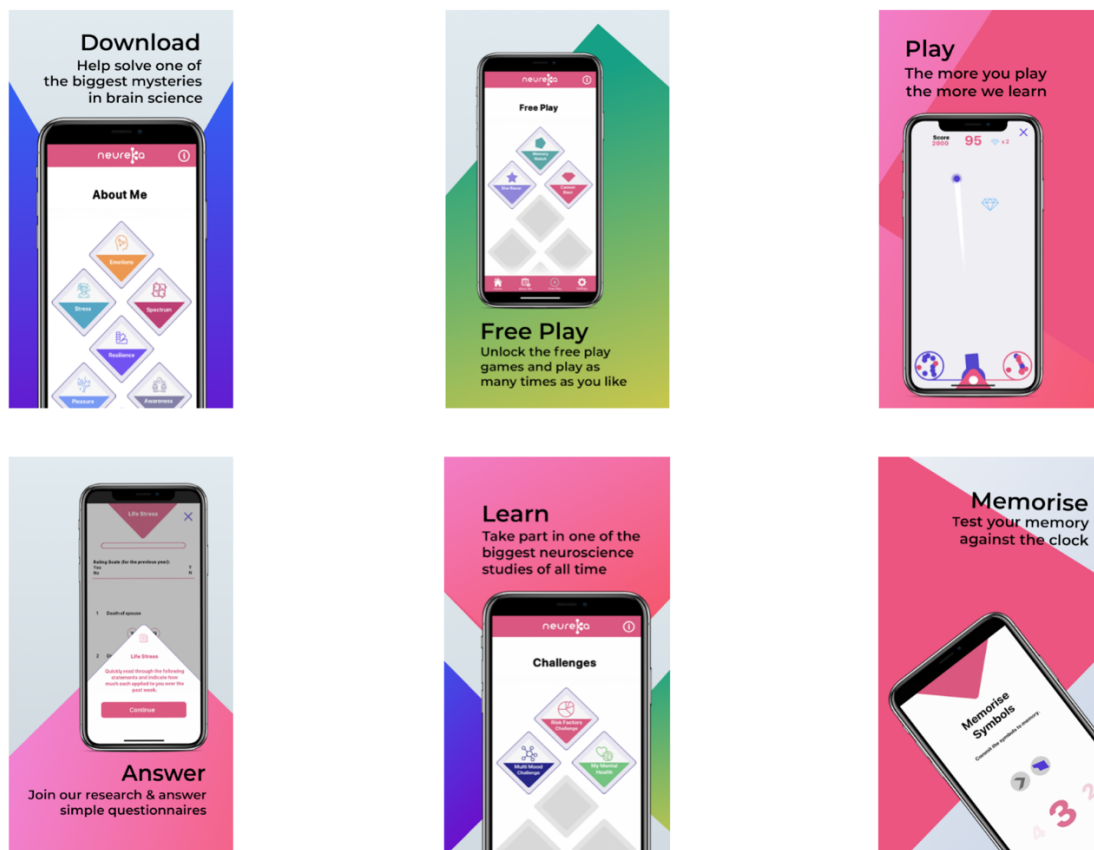


Figure S3. In-app screenshots as presented in app stores.

8.1.4.2 Common Recruitment Strategies

Ads on Google Play Store. Responsive search ads were presented to the users of Google Play Store, targeting mainly users based in Ireland and the UK, irrespective of age and gender. In this type of ads, multiple headlines and descriptions were shown to users in an order optimized based on algorithms to improve ad performance. For example, the most-shown ad in 2024 (Figure S4) was presented over 1.2 million times and had a click-through rate of 1.96%. Clicking on the ad brought users to the *Neureka* page in Google Play Store. However, note that the click-through rate does not indicate how many times the *Neureka* app was downloaded as a result.

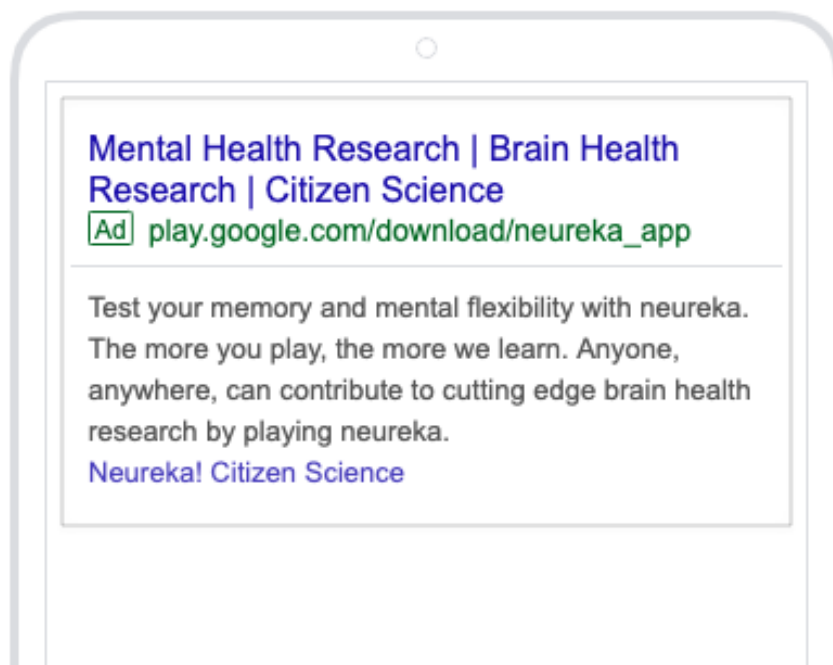


Figure S4. An example of a Google Play Store app used to attract users to download the app.

The headlines and descriptions in these ads included the following terms:

- Mental Health Research
- Brain Health Research
- Citizen Science
- Support Depression Research
- Support Anxiety Research

- Support Neuroscience Research
- Support Dementia Research
- Play Brain Games
- Complete Challenges
- Play Games to Help Research
- Become A Citizen Scientist
- Support Science Research
- Play Games
- Support Psychology Research
- Track Your Mood
- Test your memory and mental flexibility with neureka. The more you play, the more we learn
- Anyone, anywhere, can contribute to cutting edge brain health research by playing neureka.
- By playing games & completing challenges, you help us fight dementia & mental illness.
- Help scientists solve the biggest mysteries in brain health by playing games on neureka!

CenterWatch database of clinical trials. Another important recruitment strategy was a profile on the CenterWatch website, which is a US-based database of clinical trials for study volunteers. See the Study Summary below. The full profile can be found at the following link: <https://www.centerwatch.com/clinical-trials/listings/TX291222/the-neureka-project>

CenterWatch Study Summary

The Neureka Project aims to develop new ways of detecting disorders of the brain early. From mental health conditions to dementia, we want to find out who is at risk and why, so we can develop preventative interventions that can be used the world over.

To do this, we are taking research out of the lab and into your phone, through the Neureka app. On the app, you can complete science challenges, play games and tells us about your thoughts, feeling, experiences and lifestyle. Every minute you spend on the app helps researchers move the needle forward in the global fight against disorders of the mind. Download the neureka app for free today by searching 'neureka' on the App Store or Google Play Store or use the following link <<http://onelink.to/csrw86>>.

Check out www.neureka.ie for more!

CenterWatch Study Description

Using a smartphone/tablet to complete a series of science challenges and self-report questionnaires. Participation may be ongoing on a rolling daily, weekly, or monthly basis depending on which science challenges the participants engage with.

The SciStarter project. *Neureka* was also advertised on the SciStarter website. This is a platform listing and promoting Citizen Science projects from around the world. *Neureka* became a member of an affiliate programme at SciStarter, meaning that active SciStarter participants can receive credit in the programme by participating in *Neureka* and connecting their data, and *Neureka* receives more visibility on the platform. See the *Neureka* profile description below. Note that the most recent version of this description that was available at the time of writing advertised another app challenge, which was focused on mental health rather than dementia. The full SciStarter profile can be found at the following link:
<https://scistarter.org/the-neureka-project>

SciStarter Description:

Project Goal: Improve mental health and dementia research

Description: The Neureka Project aims to develop new ways of detecting disorders of the brain. Mental health conditions affect 1 in 4 people around the world. The personal, societal and economic burden of mental disorders continues to rise globally. Through decades of research, we have identified important risk factors for mental health, like early stress or neglect, and have built an enormous body of knowledge regarding cognitive and brain changes that go hand in hand with mental illness. Unfortunately, this research has typically taken a high level view - studying groups of people and looking for associations with diagnostic categories like depression. This means that we really struggle to help people in the real world. If we take any individual person, we cannot predict if and when they will develop mental illness or what will cause their specific illness. Predicting who will respond to what treatment is also a challenge. Remarkably, about HALF of people who start an antidepressant won't get better.

At Neureka, we aim to change that. We want to conduct the largest and most detailed study of YOU that has ever been attempted. By playing games and telling Neureka how you feel, day by day, we aim to figure out how mental health issues evolve over time. Neureka wants to know why some people are more resilient than others, how our mood and behaviour relate to one another and the environment and how they can kick-start vicious cycles that are difficult to break free from.

WHAT'S NEW ON NEUREKA:

Participate in Neureka by taking part in Pattern Wise, a new feature that allows you to track and monitor your habits over time. Do you find yourself checking social media on your phone over and over and OVER again, even though you know there is nothing new to see? Have you ever found it difficult to stick to the limit of that one pint? Perhaps you buy new things that you don't really need, even though you planned on saving your money? Maybe you diet or exercise or wash your hands in a rigid way that is sometimes harmful to you?

These are all examples of compulsive habits that affect our everyday lives. But what we don't know, is how and why compulsive habits develop and change over time. With Pattern Wise, you can track any behaviour and zoom in on your moment-to-moment experiences. This digital tool can capture and graphically represent how your chosen behaviour, thoughts, mood and experiences interact over time.

Hundreds of people have chosen to track their habits with Pattern Wise. Join them today to get a valuable insight in to your own behaviour!

8.1.5 Validation of gamified cognitive measures

We developed gamified smartphone-based versions of three cognitive tests that have been linked, directly or indirectly, to incipient cognitive decline and Alzheimer's dementia. In brief, compared to the traditional tests, our newly developed games take less time to complete, aim to be more engaging, and can be fully self-administered via smartphone while maintaining a standardised presentation of stimuli. Validation of the visual short-term memory task ('Memory Match') was presented in Chapter 2 of this thesis. A detailed validation of the task that assesses model-based planning ('Cannon Blast') has already been published (Donegan et al., 2023) and the methodology and results are summarised briefly below. For the cognitive flexibility task ('Star Racer'), we describe its validation here. In each case, we compared the new tasks with their traditional counterparts

in an in-lab study and further assessed the validity and reliability in a large sample crowdsourced via the *Neureka* app.

8.1.5.1 Star Racer

Method

Participants. We conducted a separate in-lab validation study, in which paid participants completed Star Racer together with the more traditional version of the task, the Trail-Making Test (TMT; Bowie & Harvey, 2006). We recruited a convenience sample at Trinity College Dublin, Ireland, through posters on campus and online advertisements and from the Greater Dublin area via online advertisements in the Dublin volunteer centre. Of the originally recruited $N = 45$ participants, $n = 41$ had complete data for both Star Racer and TMT. The final $n = 41$ Star Racer/TMT participants were aged 18–68 ($M = 26.2 \pm 9.9$) and consisted of 56% female participants.

Second, we used data crowdsourced from unpaid ‘citizen scientists’ (i.e., volunteer members of the general public) who have completed Star Racer within the ‘Risk Factors Science Challenge’ module in the *Neureka* app since the app’s release in June 2020 until January 2023. After removing incomplete datasets, participants who left and re-entered the app during gameplay, and Star Racer response times that exceeded cut-off times (i.e., over 100s on Star Racer A and over 300s on Star Racer B, see more in the Materials section below), the final sample consisted of $N = 5,986$ participants. Participants were aged 18–85 ($M = 44.5 \pm 14.7$). With regards to gender, 3,841 (64.2%) were cisgender female, 108 (1.8%) were non-cisgender, and 18 (0.3%) preferred not to state their gender. The participants came from countries, with the United Kingdom, United States, Ireland, Canada, and Germany being the most prevalent. Additionally, a subset of $n = 67$ participants completed a part of the ‘Free Play’ module in the *Neureka* app that was comparable to the ‘Risk Factors’ version of each of the games, providing us with data for a test-retest reliability estimate.

Procedure.

Study 1 (in-lab validation study). Participants were administered Star Racer and TMT in random order. Star Racer was administered on a Google Pixel 3a phone. After completing the tasks, participants provided their demographic data, then they were debriefed and reimbursed.

Study 2 (citizen scientists). Star Racer was included in the 'Risk Factors Science Challenge' module in the *Neureka* app and administered remotely to volunteer citizen scientists alongside self-report lifestyle and health questionnaires. The Risk Factors challenge is further described in the main part of this paper. Importantly, the order of presentation of the tasks was randomised and completion of the entire challenge could be distributed across time. In the current analyses, we included the first complete take on Star Racer in the Risk Factors challenge. Additionally, a modified version of Star Racer was available in the 'Free Play' section of the *Neureka* app as a standalone game, where the participants could choose which difficulty level they complete and how many times. We used performance on completed Free Play takes and compared them to Risk Factors takes to estimate test-retest reliability of each task.

Both studies were approved by the Trinity College Dublin School of Psychology Ethics Committee and informed consent was obtained from all participants prior to participation.

Materials.

Traditional trail-making test. The Trail Making Test (TMT) is a pen-and-paper cognitive test, commonly used to detect neuropsychological impairment both in research and clinical settings (Bowie & Harvey, 2006). It consists of two parts that require the participant to connect 25 labelled circles in numerical order (TMT A) or numerical and alphabetical order, alternating between numbers and letters (i.e., 1–A–2–B–etc.; TMT B). The main variable of interest is the total time to complete each part. TMT A taps mainly into processing speed and visual search, whereas TMT B taps mainly into cognitive flexibility and executive function (Bowie & Harvey, 2006). We administered the traditional trail-making test according to

standard instructions (Bowie & Harvey, 2006) — including two practice runs and both A and B versions.

Star Racer. Participants can play Star Racer within the Risk Factors challenge or in a 'Free Play' section of the app. In Risk Factors, where most participants first encounter the game, it begins with a set of self-paced tutorial screens and practice runs (see Figure S5A). In this tutorial, participants receive static screens illustrating the task and then complete a short practice round of trails A and B, with just eight stars. Once the participants tap 'Start Game', this is followed by six runs of the main task (see Figure S5B). Of note, the first two runs of A and B mirror closely the original layout of A and B forms and are as such are referred to as 'hard-coded'. The remaining 4 runs have randomly generated star locations (i.e., 'random-coded'), thus setting up the task for repeated administration with reduced learning effects.

In Free Play, there are three difficulty levels that participants can choose from — two easier levels with fewer stars ('Easy' with 8 stars and 'Medium' with 15 stars) and the third one with the same number of stars as in the Risk Factors version of Star Racer ('Hard', 25 stars). The 'Hard' level consists of one run of each version A or B, with randomly generated star locations. Unlike in Risk Factors, the task instructions and practice runs are not presented automatically at the start of gameplay, but they are available for re-visiting within the 'Free Play' section.

In both the Risk Factors and the Free Play version of the task, we applied exclusion criteria as specified in the main paper to deal with inattentive responders: We excluded all runs of version A that exceeded 100 seconds, and all runs of version B that exceeded 300 seconds. Performance was calculated as the mean time to complete the remaining runs for each participant. These cut-offs were based on approximately double the median completion times for the oldest and least educated group in a normative study of the traditional task (Tombaugh, 2004).

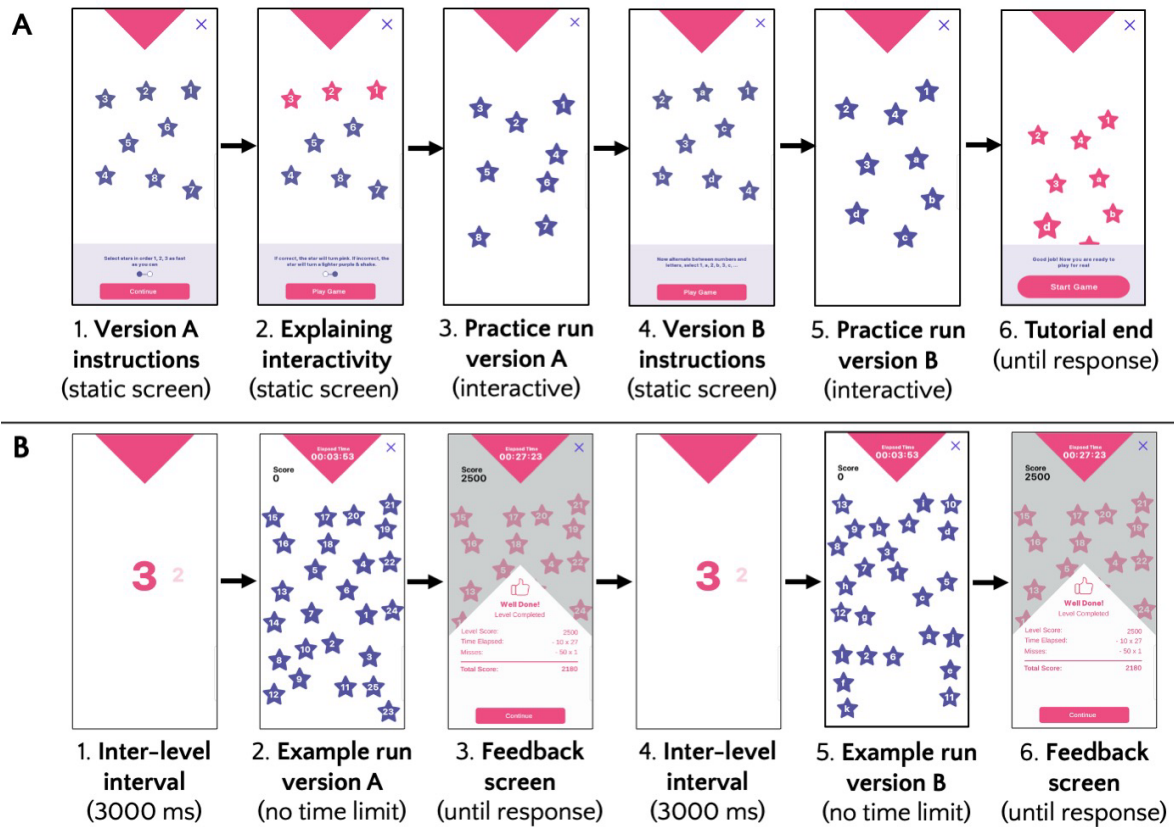


Figure S5. Task structure of Star Racer. **(A)** Self-paced instruction screens and practice runs that appear at the start of Star Racer within the Risk Factors challenge. **(B)** Example of Star Racer runs (first A, then B version) with hard-coded star positions.

Analyses.

Validity. In the in-lab validation study sample, we estimated convergent validity by calculating correlations of the new task and its different conditions with its more traditional version. We also ran a series of two-sample t-tests as applicable across both samples (i.e., in-lab and citizen scientist) to compare each task's conditions among themselves and establish the extent to which each task behaves similarly to its traditional version.

Internal consistency. Internal consistency was calculated for A and B versions separately as Cronbach's α of the three runs in each version, using the data from the citizen scientists. We did not have a comparator for the traditional task, as it is a paper and pen assessment with only one measure of total time to complete. Additionally, we computed bootstrap 95% confidence intervals for Cronbach's α based on 1000 samples.

Test-retest reliability. We calculated test-retest reliability as a Pearson correlation between overall performance on Star Racer (separately for the A and B versions) in the Risk Factors section and the equivalent subset of Free Play data. We only included participants who completed the Risk Factors and the Free Play versions of the games within 30 days from each other. The median distance between assessments was 1 day. As the Free Play version of the game had fewer runs compared to the full Risk Factors version and participants could complete the various difficulty levels in whatever order they preferred, we only used data from participants who completed runs in a number and difficulty that was equivalent to the Risk Factors version of the game, and only if they completed all these equivalent Free Play runs within the same day. Mean completion time of each participant was calculated by averaging completion time on their first 3 attempts at the 'Hard' difficulty level in the Free Play section.

Results

Similarly to the traditional TMT ($t(40) = -12.66, p < .001$), completion times in Star Racer were larger in version B compared to version A consistently across task conditions (Figure S6A) — both in runs with hard-coded ($t(40) = -12.84, p < .001$) and random star positions ($t(40) = -8.86, p < .001$) of the Risk Factors version, as well as the Free Play version of Star Racer (i.e., always random star positions; $t(1539) = -55.935, p < .001$). In the in-lab validation sample, performance on Star Racer A or B was correlated with completion time on the corresponding A or B version of TMT — both when split by condition (Figure S6B), as well as when taking into account mean completion time across all 3 runs of Star Racer A ($r(39) = .47, p = .002$) and Star Racer B ($r(39) = .63, p < .001$; see Figure S6C). Cronbach's alpha was α [95% CI] = .82 [.81, .83] for Star Racer A and .77 [.76, .79] for Star Racer B, suggesting fair to good internal consistency. The test-retest reliability of mean completion time was also fair to good ($r(65) = .87, p < .001$ for Star Racer A, $r(66) = .72, p < .001$ for Star Racer B).

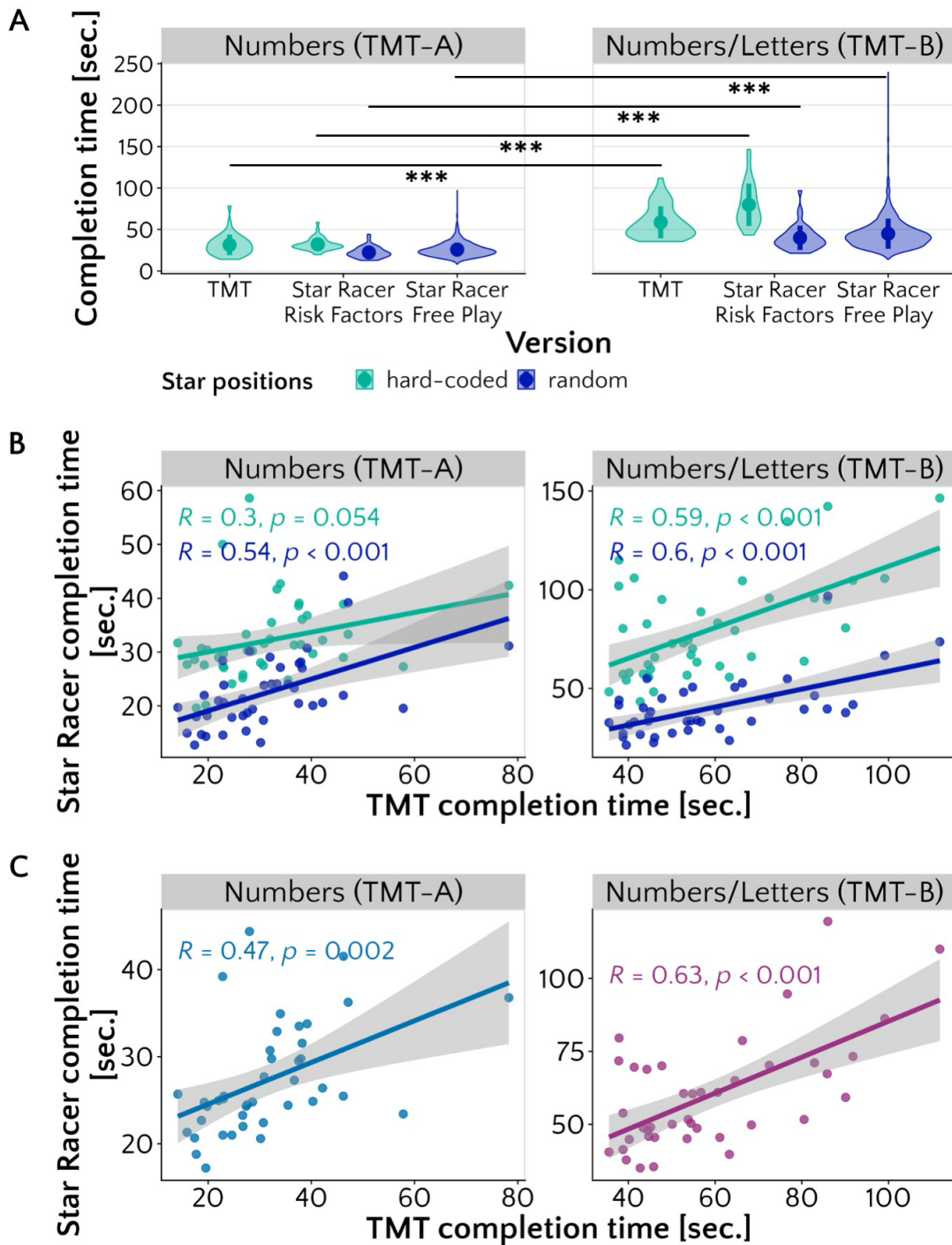


Figure S6. (A) Comparison of completion times on versions A and B in hard difficulty level (25 stars), from left to right: (i) TMT, in-lab sample; (ii) Star Racer in Risk Factors, only runs with hard-coded star positions (i.e., layout directly comparable to TMT), in-lab sample; (iii) Star Racer in Risk Factors, average performance across two runs with random star positions, in-lab sample; (iv) Star Racer in Free Play (where all runs had random star positions), citizen scientist sample. **(B)** Correlation between Star Racer and TMT completion times, split by version (A/B) and trial type (hard-coded/random star positions). **(C)** Correlation between mean completion time in the Risk Factors version of Star Racer and TMT completion time, split by version (A/B).

8.1.5.2 Cannon Blast

Cannon Blast is a gamified version of the ‘Two-Step Reinforcement Learning Task’ (Daw et al., 2011a). It includes key elements of the original task’s structure (i.e., drifting rewards, a probabilistic transition structure), wrapped up in diamond shooting game (see Figure S7). In this game, users attempt to strike diamonds, which are sometimes moving around the screen or partially obstructed, by firing from one of two containers on the screen. Each container has a mix of purple and pink balls; one has 80% pink balls and the other 80% purple, corresponding directly to the probability that a ball of that colour will be released — this is what we refer to as ‘task structure’. This means that someone can intentionally choose to increase their chances of firing a pink or purple ball. Crucially, in this task, some balls explode upon firing and therefore cannot reach their target. This is not at random, but rather is partially predictable from the colour of the ball, whereby the chances that a pink/purple ball will explode drifts slowly and independently over the course of the task. This means that a person can reduce their chances of getting a bad ball by tracking which ball is currently bad and choosing the container least likely to produce it. This is the signature of model-based planning on the task. To operationalise this, data were analysed using a well-established procedure — using hierarchical logistic regression (HLR) models, which are mixed effects models implemented with the lme4 package in R (Bates et al., 2015). The model tests if participants’ choice behaviour in the first stage state (coded as switch: 0 and stay: 1, relative to their previous choice) was influenced by IVs tracking (i) whether that ball was good or bad on the last trial (coded as bad: -1 and good:1) and (ii) whether the last trial was a trial where the ball produced from the container was the one expected by the explicit probability (‘transition’ coded as uncommon: -1 and common: 1), and (iii) their interaction. Within-participants factors (main effect of reward, transition and their interaction) were modelled as random effects. Model-based index (MBI) is quantified as the interaction between Reward (good vs. dud ball) and Transition (common vs uncommon ball colour appearing from the chosen container). Individual estimates of the MBI were extracted and a single value for each participant was brought forward for the main analysis.

A prior paper validated this task in detail (Donegan et al., 2023). In brief, there was a moderate positive association between MBI derived from *Cannon Blast* and the Traditional task ($r = .40$, $p = .002$). Split-half reliability for MBI were high similar for both the traditional ($r = .81$, 95% CI [.70, .88], $p < .001$) and *Cannon Blast* ($r = .78$, 95% CI [.66, .87], $p < .001$). The test-retest reliability was $r(423) = .63$ assessed over a variable interval (median 4 days).

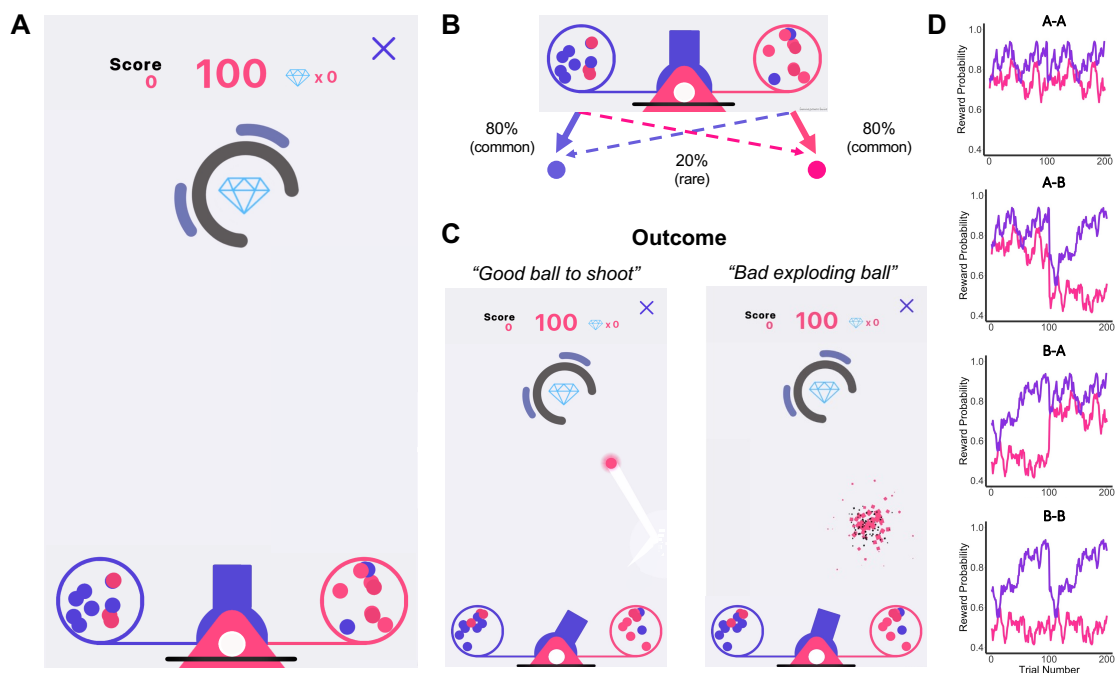


Figure S7. Task structure of Cannon Blast, a smartphone game to assess model-based planning. **(A)** In this game, participants' goal is to shoot as many diamonds as possible before their total number of shots (100 per block) runs out. To do so, they must aim a central cannon and then select which circular container to draw from. **(B)** Purple and pink balls dynamically bounce around each of the flanked containers which depict the probability of a pink or purple ball being released. For example, the left container displays 8 purple balls and releases a purple ball 80% of the time ('common' transition) and displays 2 pink balls, giving a pink ball on 20% of trials ('rare' transition). **(C)** The purple and pink balls have different values that dynamically change throughout the game. The value of the ball is defined as the probability of it being a 'good ball', i.e., one that remains intact after firing (rewarding trial), or a 'dud ball' (non-rewarding trial) that explodes shortly after being fired, and therefore cannot reach the diamond. **(D)** We included 2 drifting reward probabilities (A, B) that quantitatively differed on various metrics. Participants were randomly assigned a reward drift set at each block leading to four distinct drift set combinations (A-A, A-B, B-A, B-B). Figure and legend reproduced with permission from Donegan et al., 2023, published in *Communications Psychology* under the CC BY 4.0 licence (<https://creativecommons.org/licenses/by/4.0/>).

8.1.6 Associations between risk factor measures

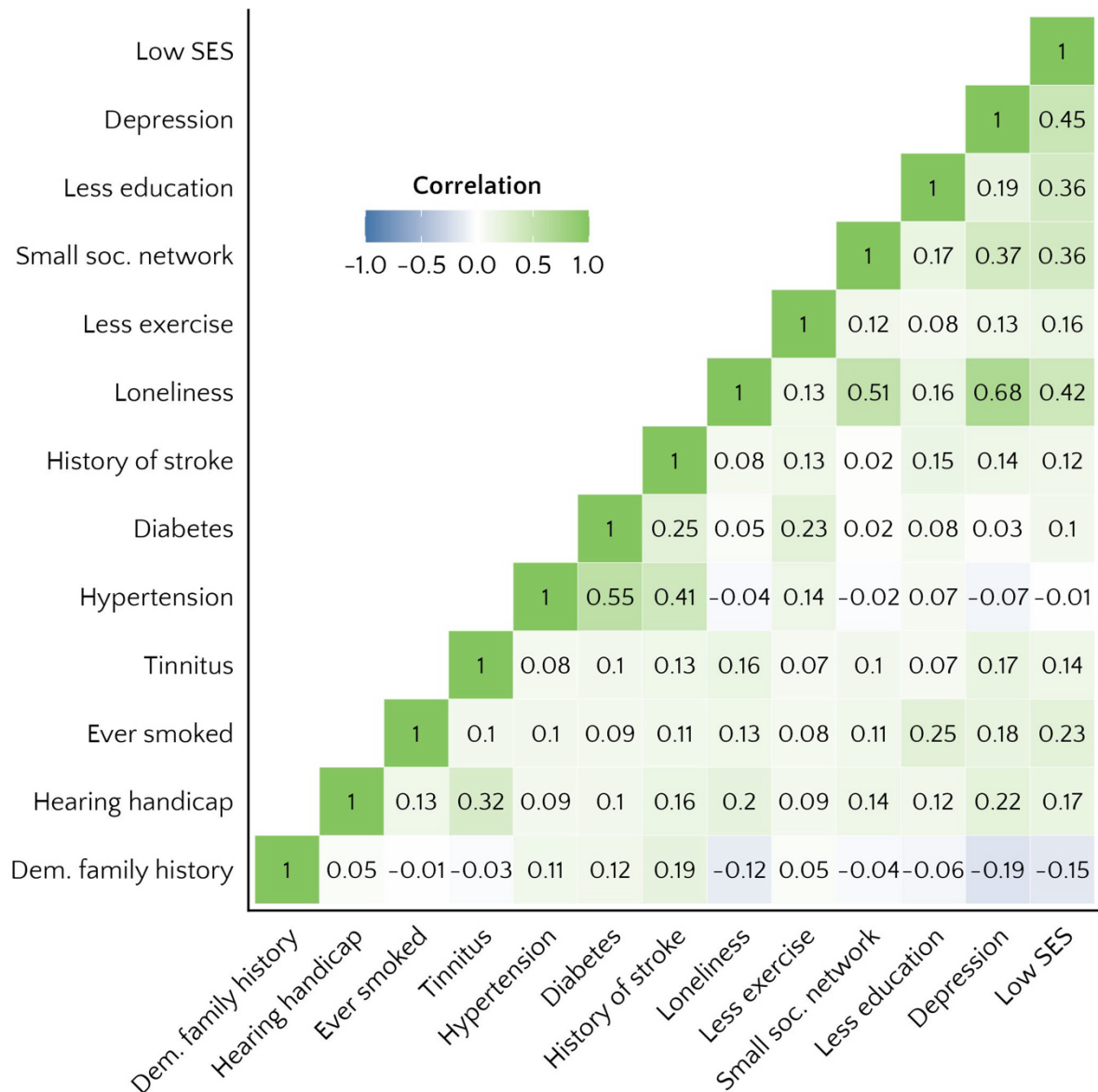


Figure S8. Correlation matrix of all biopsychosocial factors assessed in the current study (calculated in the final sample, $N = 3,342$). Numbers represent Pearson product-moment correlation coefficients for pairs of continuous variables, polyserial correlations for pairs of continuous and categorical variables, and polychoric correlations for pairs of categorical variables.

8.1.7 Results for non-cisgender participants

Compared to cisgender men, non-cisgender individuals had significantly better visual working memory (β [95% CI] = -0.21 [-0.34, -0.09]; $p = .001$; see Figure

S9A), cognitive flexibility (β [95% CI] = -0.16 [-0.28, -0.04]; p = .008; see Figure S9B), and model-based planning (β [95% CI] = -0.14 [-0.26, -0.02]; p = .028; see Figure S9C), but also a significantly higher likelihood to report subjective memory problems (OR [95% CI] = 1.64, [1.28, 2.12], p < .001; see Figure S9D).

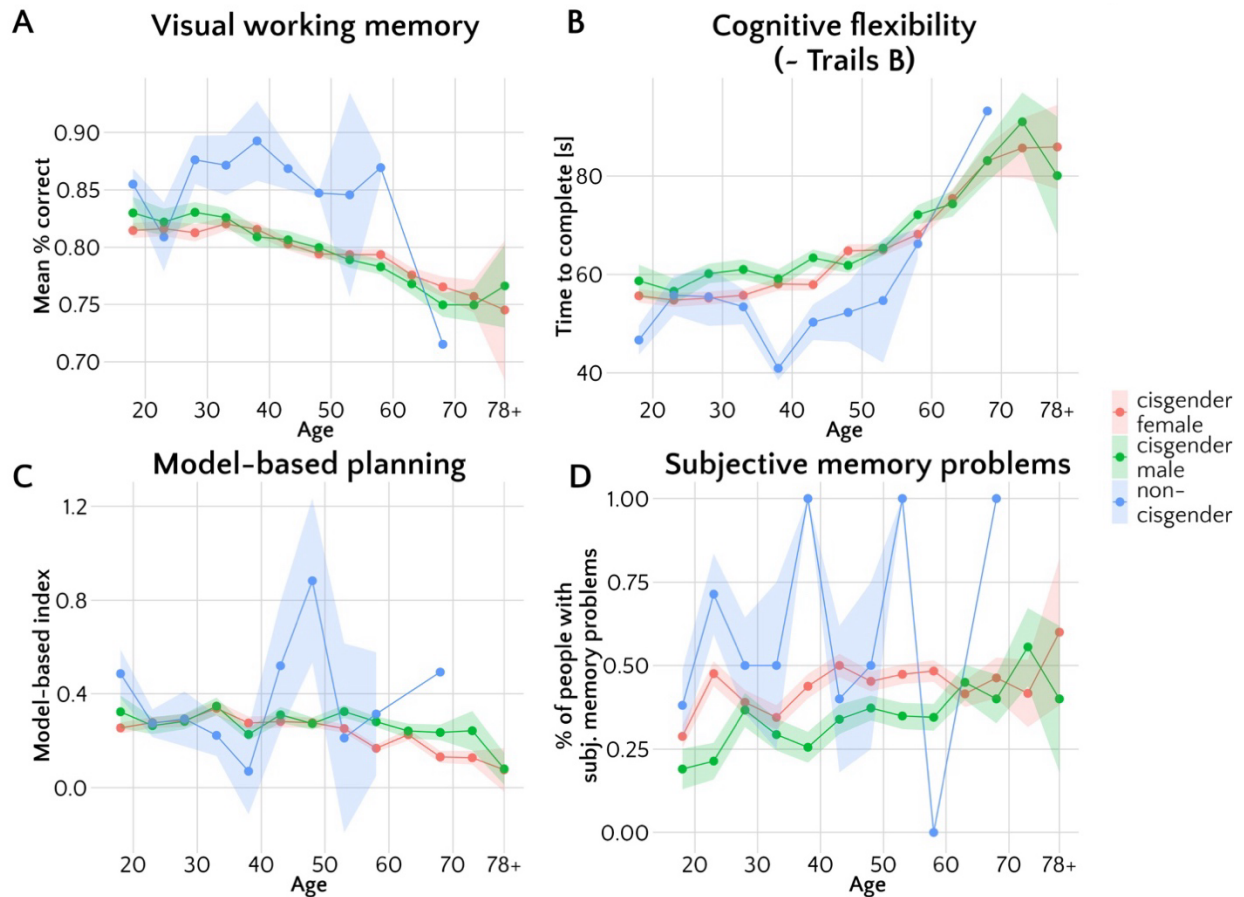


Figure S9. Associations of cognitive measures with age and gender, including cisgender females and males and ‘non-cisgender’ participants (i.e., participants who identified as either non-binary, transgender female, or transgender male). Points correspond to mean raw scores on (A) visual working memory, (B) cognitive flexibility, (C) model-based planning, and (D) mean proportion of participants with memory problems. The means were calculated per 5-year bins, split by gender. Error bars represent standard errors (A–C) or standard errors of proportion (D). Note that due to the very small number of non-cisgender participants ($n = 67$, i.e., 2% of the total sample), not all age groups were represented and calculating means or standard errors was not always possible.

8.1.8 Comparison of effect magnitude for subjective vs. objective cognition

Table S2. Associations of risk factors with binarised cognitive outcomes ($n = 3,327$), controlling for gender and age, expressed as estimates from logistic regressions for all DVs. Highlights indicate results significant after Bonferroni correction ($p < .0038$).

Risk factor	Visual working memory		Cognitive flexibility (~ Trails B)		Model-based planning		Subjective memory problems	
	OR [CI _{95%}]	p	OR [CI _{95%}]	p	OR [CI _{95%}]	p	OR [CI _{95%}]	p
Depression	1.28 [1.18, 1.38]	< .001	1.16 [1.07, 1.25]	< .001	1.05 [0.98, 1.14]	.166	1.82 [1.68, 1.97]	< .001
Low SES	1.14 [1.06, 1.23]	< .001	1.11 [1.03, 1.2]	.005	1.11 [1.03, 1.19]	.005	1.65 [1.53, 1.78]	< .001
Hearing handicap	1.04 [0.97, 1.11]	.327	1.07 [0.99, 1.15]	.068	0.96 [0.89, 1.03]	.236	1.59 [1.47, 1.71]	< .001
Loneliness	1.12 [1.04, 1.2]	.003	1.09 [1.01, 1.17]	.028	1.02 [0.95, 1.1]	.500	1.52 [1.41, 1.64]	< .001
Less education	1.33 [1.24, 1.43]	< .001	1.18 [1.09, 1.27]	< .001	1.2 [1.12, 1.29]	< .001	1.39 [1.3, 1.5]	< .001
Less exercise	1.03 [0.96, 1.11]	.357	1.06 [0.98, 1.14]	.147	1.01 [0.94, 1.08]	.761	1.41 [1.31, 1.52]	< .001
Ever smoked	1.14 [1.05, 1.22]	.001	1.08 [1, 1.17]	.039	1.07 [1, 1.15]	.067	1.33 [1.24, 1.44]	< .001
Tinnitus	1.02 [0.94, 1.1]	.609	1 [0.93, 1.08]	.971	0.92 [0.85, 0.99]	.037	1.33 [1.23, 1.43]	< .001
Small soc. network	1.1 [1.02, 1.18]	.012	1.08 [1.01, 1.16]	.032	1.04 [0.97, 1.12]	.232	1.27 [1.18, 1.36]	< .001

Risk factor	Visual working memory		Cognitive flexibility (~ Trails B)		Model-based planning		Subjective memory problems	
	OR [CI _{95%}]	<i>p</i>	OR [CI _{95%}]	<i>p</i>	OR [CI _{95%}]	<i>p</i>	OR [CI _{95%}]	<i>p</i>
History of stroke	1.32 [0.98, 1.81]	.074	1.63 [1.18, 2.33]	.004	1.08 [0.8, 1.45]	.614	1.84 [1.34, 2.6]	< .001
Dem. family history	0.98 [0.89, 1.08]	.663	0.95 [0.86, 1.05]	.305	0.99 [0.9, 1.09]	.893	1.12 [1.01, 1.23]	.024
Diabetes	1.21 [1.04, 1.42]	.014	1.16 [0.99, 1.35]	.072	0.97 [0.83, 1.13]	.678	1.18 [1.01, 1.38]	.035
Hyper-tension	1.17 [1.06, 1.29]	.002	1.06 [0.96, 1.17]	.244	1.06 [0.96, 1.18]	.217	1.08 [0.97, 1.19]	.146

8.1.9 Smoking history x age interaction

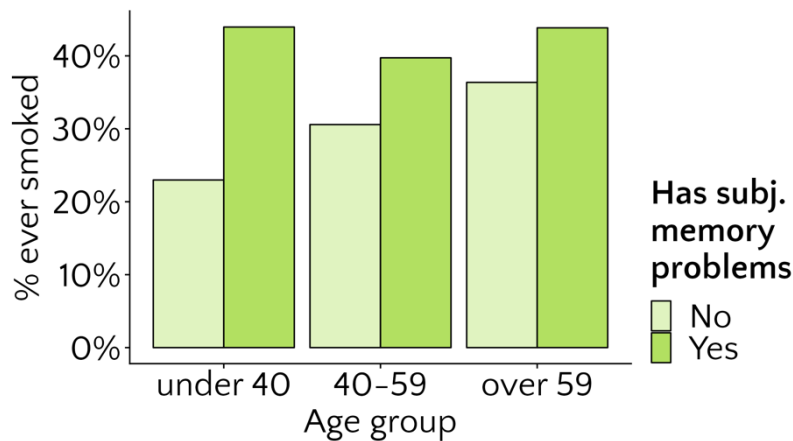


Figure S10. Associations of smoking with subjective memory problems modified by age. For plotting purposes, age was grouped into three categories: young adults (under 40; $n = 1,563$), middle aged adults (40–59; $n = 1,136$), and older adults (over 59; $n = 628$).

8.1.10 Interactions between age and risk factors

Table S3. Interactions of risk factors with age, regressed on cognitive outcomes, controlling for gender and age ($n = 3,327$). Estimates come from logistic regressions (subj. memory problems) or linear regressions (all other DVs). Highlights indicate significance at $p < 0.0015$.

Cognitive outcome	Interaction	β (SE)	t/z	p	OR [95% CI]
Visual working memory	Age * Small soc. network	-0.04 (0.02)	-2.50	.013	N/A
	Age * Low SES	-0.04 (0.02)	-2.27	.023	N/A
	Age * Loneliness	-0.05 (0.02)	-2.63	.009	N/A
	Age * Less education	-0.02 (0.02)	-1.34	.180	N/A
	Age * Hypertension	0 (0.03)	-0.06	.950	N/A
	Age * History of stroke	-0.13 (0.07)	-1.92	.055	N/A
	Age * Ever smoked	0 (0.02)	-0.15	.879	N/A
	Age * Diabetes	-0.09 (0.05)	-1.74	.083	N/A
	Age * Depression	-0.02 (0.02)	-1.17	.241	N/A
	Age * Small soc. network	-0.02 (0.02)	-1.22	.221	N/A

Cognitive outcome	Interaction	β (SE)	t/z	p	OR [95% CI]
Cognitive flexibility (~ Trails B)	Age * Low SES	0 (0.02)	0.20	.839	N/A
	Age * Loneliness	-0.04 (0.02)	-2.40	.016	N/A
	Age * Less education	-0.02 (0.02)	-0.93	.350	N/A
	Age * Hypertension	0.06 (0.03)	2.08	.037	N/A
	Age * History of stroke	-0.08 (0.07)	-1.17	.244	N/A
	Age * Hearing handicap	0.01 (0.02)	0.40	.687	N/A
	Age * Ever smoked	0.02 (0.02)	0.95	.343	N/A
	Age * Depression	-0.04 (0.02)	-2.28	.023	N/A
Model-based planning	Age * Low SES	0 (0.02)	0.08	.935	N/A
	Age * Less education	-0.03 (0.02)	-1.56	.118	N/A
	Age * Depression	-0.02 (0.02)	-1.31	.189	N/A
Subjective memory problems	Age * Tinnitus	-0.06 (0.04)	-1.60	.110	0.9 [0.87, 1.01]
	Age * Small soc. network	-0.07 (0.04)	-1.96	.050	0.93 [0.86, 1]
	Age * Low SES	-0.05 (0.04)	-1.28	.200	0.95 [0.88, 1.03]
	Age * Loneliness	-0.01 (0.04)	-0.37	.712	0.99 [0.92, 1.06]
	Age * Less exercise	-0.03 (0.04)	-0.88	.380	0.97 [0.89, 1.04]
	Age * Less education	-0.07 (0.04)	-1.90	.057	0.93 [0.86, 1]
	Age * History of stroke	-0.01 (0.15)	-0.09	.932	0.99 [0.73, 1.34]
	Age * Hearing handicap	-0.1 (0.04)	-2.65	.008	0.9 [0.84, 0.97]
	Age * Ever smoked	-0.14 (0.04)	-3.65	< .001	0.87 [0.81, 0.94]
	Age * Depression	0.03 (0.04)	0.87	.386	1.03 [0.96, 1.12]

8.2 Appendix II: Supplemental Information for Chapter 4

8.2.1 Recruitment process

Between January and December 2022, all current participant of each PREVENT site received a generic invitation email ('Generic Invitation', see section 8.2.1.1) and a study information sheet ('Study Information Sheet', see section 8.2.1.2) from their respective study site coordinator, to which they were encouraged to reply if interested in study participation. A second round of similar general invitations ('Generic Reminder', see section 8.2.1.3) went out later during the year. These initial invitations and reminders were scattered across the year based on each site's availability and other ongoing add-on studies, so as not to overburden participants with participation requests.

Each participant who replied to the initial invitation received further information about the study via email from their site coordinator ('Personalized Invitation', see section 8.2.1.4). They were also provided with their unique ID and a link to an online consent form administered through the Qualtrics platform. This consent form included general information about the study, as well as instructions on how to connect data between *Neureka* and PREVENT using the unique ID ('Participation Instructions', see section 8.2.1.5). Participants could also opt in for having the instructions sent to their email and complete the study at a time of their own choosing.

8.2.1.1 Generic invitation

SUBJECT LINE: Neureka! A new research project affiliated with PREVENT

Good morning, [replace with greeting of your choice]

The PREVENT Dementia team would like to invite you to participate in **a new study** that can be completed on your smartphone in the comfort of your own home!

Brain scans and genetic screening are expensive and therefore not available to everyone. To try and solve this issue, **PREVENT Dementia has partnered with the Neureka project** to see if we can understand a person's brain health using a free smartphone app.

I have attached the **study information**. If this sounds like something you would be interested in, please **reply to this email** for further details on how to participate.

Best wishes,

[Insert name of PREVENT Coordinator]

8.2.1.2 Study Information Sheet

PREVENT x Neureka Study Information Sheet

Brain scans and genetic screening are expensive and therefore not available to everyone. To try and solve this issue, PREVENT has partnered with the Neureka project to see if we can understand a person's brain health using a free smartphone app.

If you decide to take part, we will ask you to play some games and answer some questionnaires on the Neureka app and we will link the data from this smartphone app with the data from your participation in the PREVENT study.

What will happen if I agree to take part?

This study will entail downloading the smartphone application "Neureka" on your phone. Neureka is a smartphone application for citizen science research in mental health that is developed and maintained by the Gillan Lab at Trinity College Dublin.

We will ask you to complete the "Risk Factors Challenge" in the app which will take **~45 minutes to complete**. This challenge asks you to play games that test your cognitive abilities and answer questions about your health, experiences and lifestyle.

The Neureka app is free to download. To register with the Neureka app you will need to provide an email address and some basic demographic information. Please be advised that Neureka will never share your email with unauthorised third parties, email addresses are required so that you can sign into the app across different devices and so that we can contact you in future about your contribution to science and any future studies you may want to participate in. They are stored in encrypted format.

After registering, you will need to link your study participation to PREVENT via the 'Study ID' feature in the settings section of the Neureka app. Your Study ID is a 5-digit code that will be provided by your PREVENT study coordinator. **Data you submit to the Neureka app will be linked to the other data you have provided as part of PREVENT.**

Once you have successfully linked your study ID, you may proceed to complete the 'Risk Factors' challenge. This challenge contains a series of gamified cognitive tasks and questionnaires that will take ~45 minutes to complete. You are free to participate more broadly with the Neureka project, but that data will not be provided to PREVENT and will be stored and managed in accordance to the procedures and policies of the Neureka team.

How long will it take to participate?

It will take you approximately 1 hour to participate, including time taken to install the app. The download and registration process will take <15 minutes. It will take ~45

minutes to complete the Risk Factors challenge. As mentioned previously, you are free to complete other sections of the app at your own discretion.

What data will I be sharing?

Data you submit to the Neureka app will be linked to the other data you provided as part of the PREVENT study. This is done using a “study ID” that we will ask you to enter into the Neureka app at the time of download. Data acquired in Neureka includes your email address (this will not be shared with anyone outside of the Neureka team), demographic information and self-report questionnaire and cognitive data.

On top of that, your email address will be collected in the consent form if you decide to have participation instructions sent to your email.

What will happen to my data?

We will use the data you provide us within the app for academic research only. As part of the PREVENT study, you will be asked to enter a Study ID in the settings section of the Neureka app. By inputting a valid Study ID you agree for us to share your data from PREVENT with the Neureka research team. Moreover, PREVENT will acquire your data from the ‘Risk Factors’ challenge in the Neureka app.

Data will be shared in pseudoanonymised format. This means that no identifiable information will be shared by either study. Your data will be connected together using a unique ID. We will publish the findings that your contribution has helped generate in scientific journals and make them available for you to read and share online at www.neureka.ie, all publications will be based on anonymised data no identifiable data would be published.

If you share your email address through the consent form to have participation instructions sent to your email, this will be only used for the purpose of sending the instructions and not shared with anyone outside the Neureka team. Your email address will be kept entirely separate to any information submitted via the Neureka app as part of this study.

Can I access my data?

You are also free to access data that you have submitted through Neureka through making a request via the ‘GDPR > Download My Data’ section of the app. A member of the Neureka research team will send your raw data to you via email. Under data protection legislation, you will have access to any data you provide through the app for up to 10 years after you have provided it, up until the point at which it has been fully anonymised and we can no longer recover your unique contribution.

Is it confidential?

Yes. We will not disclose any personally identifiable data (e.g. your email address) to any third party. However, there are legal limits that apply to confidentiality in the Republic of Ireland where the Neureka project is based, specifically if you provide unsolicited information to the research team pertaining to risk of harm to yourself or another person, this additional information is not protected by confidentiality and may be disclosed to the relevant authorities. Secondly, any information that you provide (solicited or unsolicited) may

be disclosed as part of a legal process or police investigation in the Republic of Ireland, without your permission being sought.

Is it secure?

All data you provide through the app is encrypted before it is submitted over the internet. The data you send us through this app is stored and processed in accordance with EU General Data Protection Regulation (GDPR) with our procedures for storing and processing data having undergone and passed review by the data protection officer at Trinity College Dublin.

How do I withdraw (i.e., get out) of taking part in the research?

If you decide that you don't want to take part in the research then you should go to the "settings"->"GDPR" screen (accessed through the settings button in the top right of the app) and deselect "participate in research". This will stop any new data from your device being submitted to our server. Your use of the app won't be affected by this. If you change your mind you can reselect "participate in research" at any time.

How much Mobile Data (i.e. smartphone access to the internet without WiFi) does the app use?

Not very much. The average user will use less than 3 megabytes in total during their time using the app. This doesn't include the data used to download or update the app.

Are there any risks involved in participating?

The risks associated with participating in this research are minimal. However, it is possible that you might find the sensitive nature of some of the questions in questionnaires that you might choose to complete upsetting. If you experience upset or distress we encourage you to contact your General Practitioner/Doctor. We also provide information on support services in a dedicated tab within the app.

Are there any benefits to participating?

By participating in this research, you are helping researchers at PREVENT and the Neureka project move the needle forward in the global fight against disorders of the mind. This collaboration will allow us to gain a comprehensive insight into both biological and cognitive markers of brain health. It is our hope that this collaboration will help us develop novel preventative interventions that can be used the world over – without your participation this would not be possible.

Will the research team take responsibility for diagnosing app users with dementia or mental health disorders?

No. The app is not diagnostic and users will not be provided with any processed scores, or individualised predictions on risk of dementia or mental health disorders based on the data that they provide through the app. While we hope that one day, some combination of these measures might help us to detect illness early – the science is simply not there yet.

Who is running this study?

The study is being conducted by the PREVENT dementia research team in partnership with the Gillan Lab based at the Global Brain Health Institute at Trinity College Dublin in the Republic of Ireland. The Neureka has received ethical approval from the School of Psychology Research Ethics Committee at Trinity College Dublin.

In taking part and providing consent, you are agreeing to participate in a study that handles your data in line with the Republic of Ireland Data Protection Act 2018 which is in line with the EU General Data Protection Regulation (GDPR) 2016/679. For information on Data Protection law in Ireland, see <https://dataprotection.ie/docs/A-guide-to-your-rights-Plain-English-Version/r/858.htm>.

Who can I contact if I have questions, comments or complaints about this study?

[contact details of both Neureka and PREVENT Team]

8.2.1.3 Generic Reminder

SUBJECT LINE: Neureka! Final call for participants in a remote study affiliated with PREVENT

Good morning, [replace with greeting of your choice]

The PREVENT Dementia team would like to invite you to participate in **an add-on study** that can be completed on your smartphone in the comfort of your own home!

Brain scans and genetic screening are expensive and therefore not available to everyone. To try and solve this issue, **PREVENT Dementia has partnered with the Neureka project** [insert hyperlink: <https://www.neureka.ie/post/test-your-brain-health-one-game-at-a-time>] to see if we can understand a person's brain health using a free smartphone app.

If you have already taken part, **thank you so much for your help**. Each and every completion counts.

If you have not participated yet, **there is still time for you to do so!** I have attached the **study information**. If you are interested, please **reply to this email** for further details on how to participate.

Best wishes,

[Insert name of PREVENT Coordinator]

8.2.1.4 Personalized Invitation

SUBJECT LINE: Neureka! A new research project affiliated with PREVENT

Dear Participant,

Thank you for your interest in this study. In this project, we will ask you to play some games and answer some questionnaires on the Neureka app and we will link the data from this smartphone app with the data from your participation in the PREVENT study.

To participate in the study, click on the following link:

[INSERT PERSONAL QUALTRICS LINK]

Your unique Study ID (**you will need this for sign-up**): [INSERT STUDY ID]

Once you have signed up through the link above, the whole study takes about 45 minutes to complete.

Note that you will need a smartphone with access to internet to be able to participate.

Thank you for your time!

Best wishes,

[Insert name of PREVENT Coordinator]

8.2.1.5 Participation Instructions

PREVENT x The Neureka Project: Instructions for participation

Thank you for agreeing to participate in the Neureka project!

To set off, you will need:

- these **instructions**
 - your **smartphone**
 - your **Study ID**, which was provided to you in the invitation email by your PREVENT study coordinator
 - **about 10 minutes** to download the app and get set up
 - **about 45 minutes** to complete the in-app challenge Risk Factors
- Note: You can also complete this challenge later and not necessarily in one sitting, although doing so in one sitting is preferable.*

Important: Please follow these instructions until the very end, otherwise we won't be able to use your data!

Step 1: Download the Neureka app onto your mobile device.

Neureka is free to download. It is currently adapted only for smartphones, not for tablets or computers.

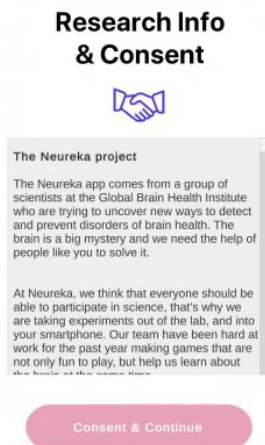
One way to download it is to scan the following QR code with your phone:

[QR code]

If that does not work for you, go to the *App Store* or *Google Play Store* on your phone, search for "neureka" and download the app right away!

Step 2: Join Neureka.

Once you have downloaded the app and read the in-app information, you will once again be asked for your consent.



**Research Info
& Consent**



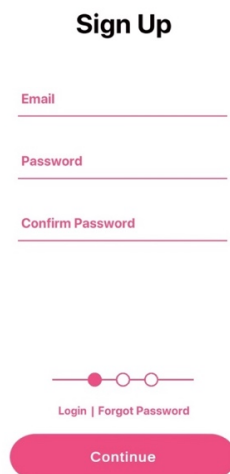
The Neureka project

The Neureka app comes from a group of scientists at the Global Brain Health Institute who are trying to uncover new ways to detect and prevent disorders of brain health. The brain is a big mystery and we need the help of people like you to solve it.

At Neureka, we think that everyone should be able to participate in science, that's why we are taking experiments out of the lab, and into your smartphone. Our team have been hard at work for the past year making games that are not only fun to play, but help us learn about the brain at the same time.

[Consent & Continue](#)

After that, you can sign up to Neureka:



Sign Up

[Login](#) | [Forgot Password](#)

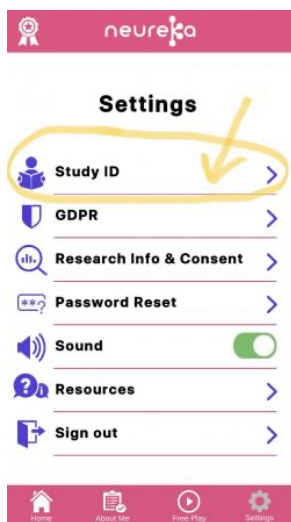
[Continue](#)

Step 3: Link your Neureka account to your PREVENT data.

Next, link your research participation though clicking on the 'Settings' icon on the bottom right of the screen.



Select 'Study ID' and **enter your unique 5 digit study ID code** that was provided by your clinical coordinator in the invitation email.

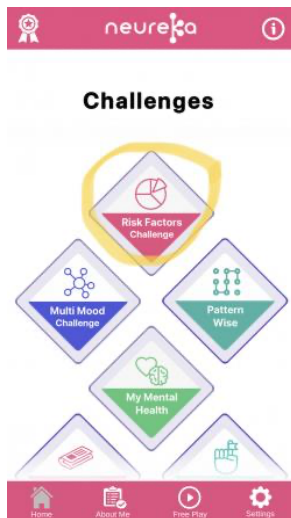


If you have not received your study ID, please contact your PREVENT study coordinator, who sent you the initial invitation to this study.

Step 4: Start playing Neureka!

Once you have successfully entered your study ID please proceed to the 'home' screen.

Please complete the 'Risk Factors Challenge'.



This challenge contains 3 gamified cognitive tasks and a variety of questionnaires. It will take **approximately 45 minutes to complete**.

You are welcome to complete as many additional science challenges and questionnaires at your leisure!

If you have any trouble with downloading Neureka or linking your study ID, please **contact Anna at prevent@neureka.ie** and we can look into this further.

Thank you again and we look forward to your participation!

8.2.2 Associations between variables

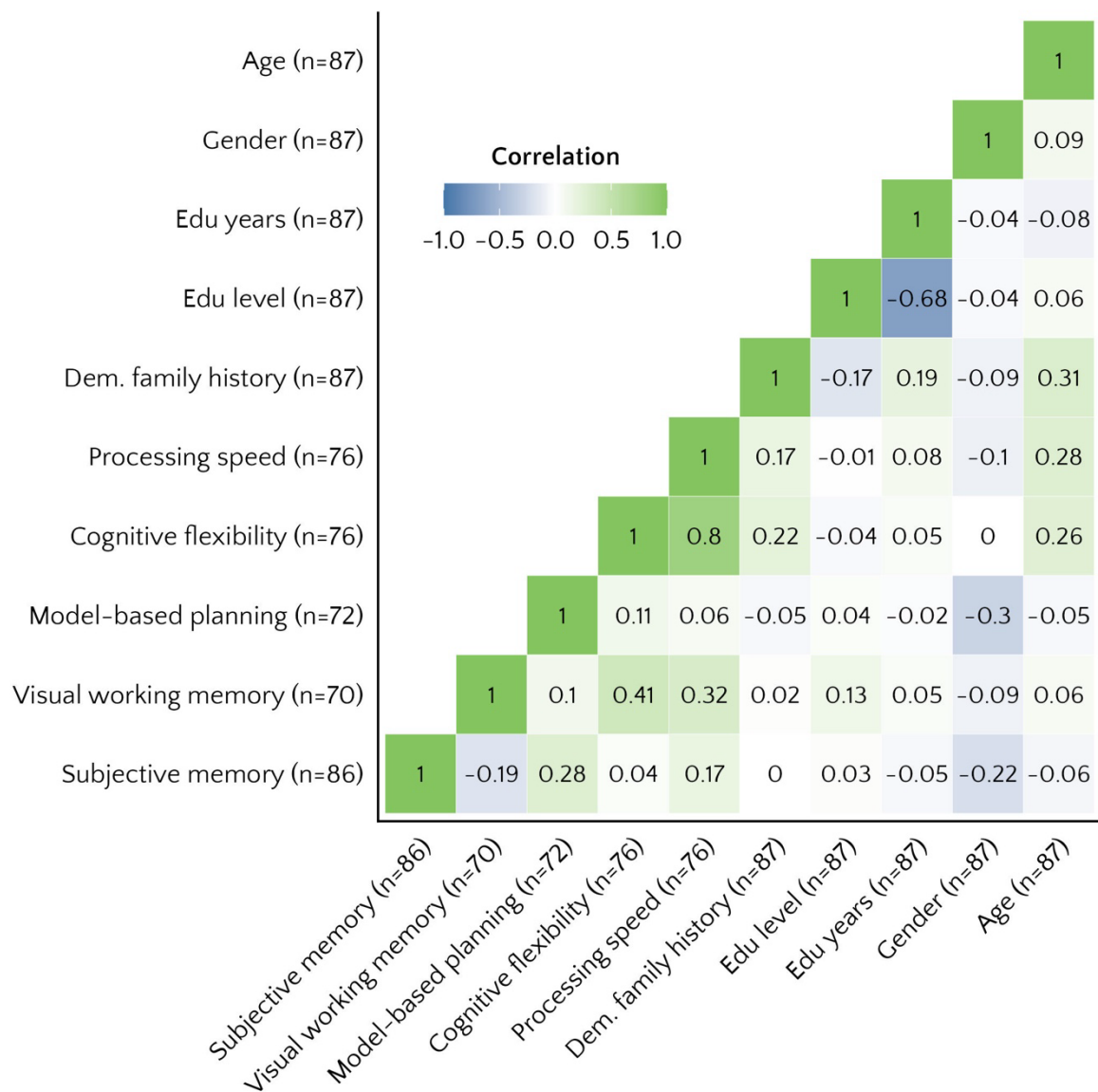


Figure S9: Univariate associations between demographic characteristics and cognitive measures. Numbers represent Pearson product-moment correlation coefficients for pairs of continuous variables, polyserial correlations for pairs of continuous and categorical variables, and polychoric correlations for pairs of categorical variables.

8.2.3 Sensitivity analyses

We ran sensitivity analyses to disentangle which features of the visual working memory task were driving the positive associations with APOE- ϵ 4. Based on previous studies that found a positive effect on a visual working memory task assessing object-location binding (Lu et al., 2021; Zokaei et al., 2017, 2020), we identified two additional measures in our task that tap into different possible types of error participants could make: misbinding error and identification error. In our study, misbinding error was calculated as the proportion of all selections in which participants selected a correct shape of incorrect colour (see Figure S12, panel A). Identification error was calculated as the mean proportion of selections in which participants made any other errors, i.e., cases where participants' erroneous selections were driven by incorrectly identified symbols irrespective of the selected colour (see Figure S12, panel B). We ran the same set of models as in the main analyses, predicting both misbinding and identification errors from APOE- ϵ 4 status and controlling for the same covariates (age, gender, years of education, and family history of dementia). Compared to APOE- ϵ 4 non-carriers, carriers made significantly fewer identification errors (β [bootstrapped 95% CI] = -0.32 [-0.55, -0.09]; $t = -2.60$; $p = .012$), but there were no significant differences in misbinding errors (β [bootstrapped 95% CI] = -0.08 [-0.39, 0.13]; $t = -0.63$; $p = .532$). Although based on a small sample and a different visual working memory paradigm, the present pattern of results corroborates the findings of previous studies (Lu et al., 2021; Zokaei et al., 2020), which reported APOE- ϵ 4 carriers outperformed non-carriers on a visual working memory binding task and found that this was driven by impaired memory for object identity. Similar to what we report here, the previous studies did not find a significant difference in object-location binding errors based on APOE status. This is interesting, given that visual short-term memory binding, but not memory for single features, has been suggested as a sensitive marker for AD (Cecchini et al., 2023; Parra et al., 2009). Given the small sample size here, the present findings warrant replication by larger studies, but they hint at overall better visual working memory performance in APOE- ϵ 4 carriers, but not an effect on binding as such.

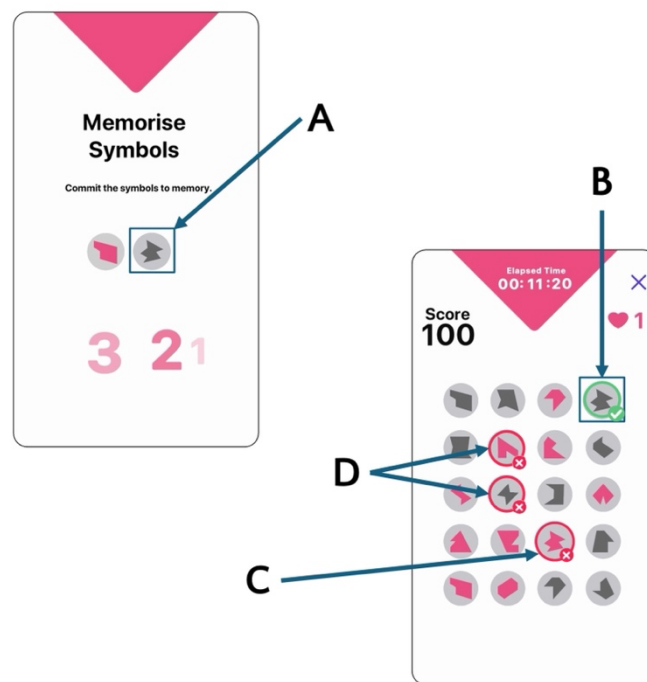


Figure S12: Illustration of different error types possibly driving the differences in visual working memory performance between APOE- ϵ 4 carriers and non-carriers. In this example, the participant is trying to find the memorised shape (**A**) in the search grid. Correct selection requires the participants to tap (**B**). A misbinding error (**C**) occurs when the participant selects the correct shape of an incorrect colour. Any other erroneous selection (**D**) would be classified as an identification error. Note: Arrows and squares added in this example.

8.3 Appendix III: Supplemental Information for Chapter 5

8.3.1 Data distribution by sex

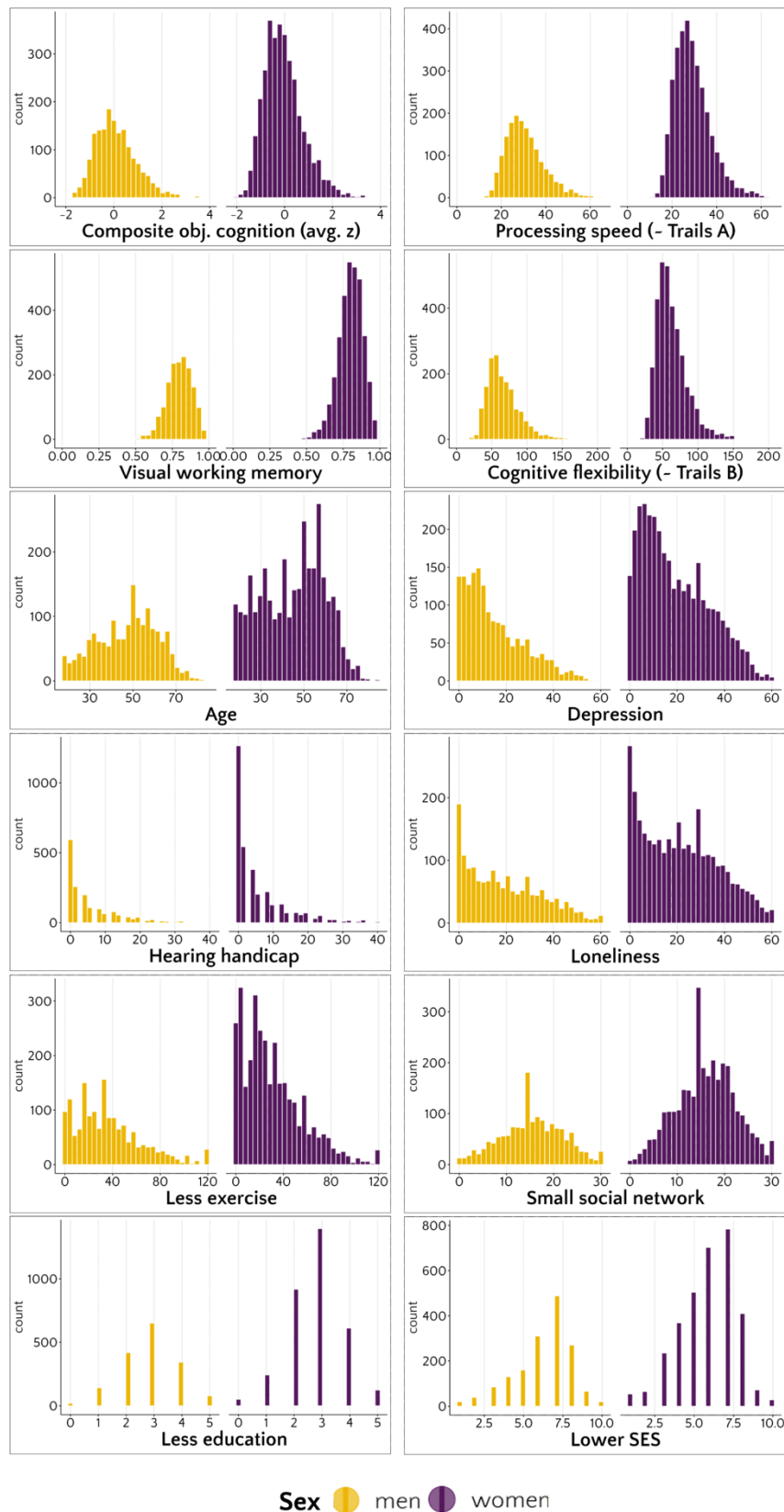


Figure S13. Histograms showing data distribution of each variable split by sex.

8.3.2 Associations between predictors

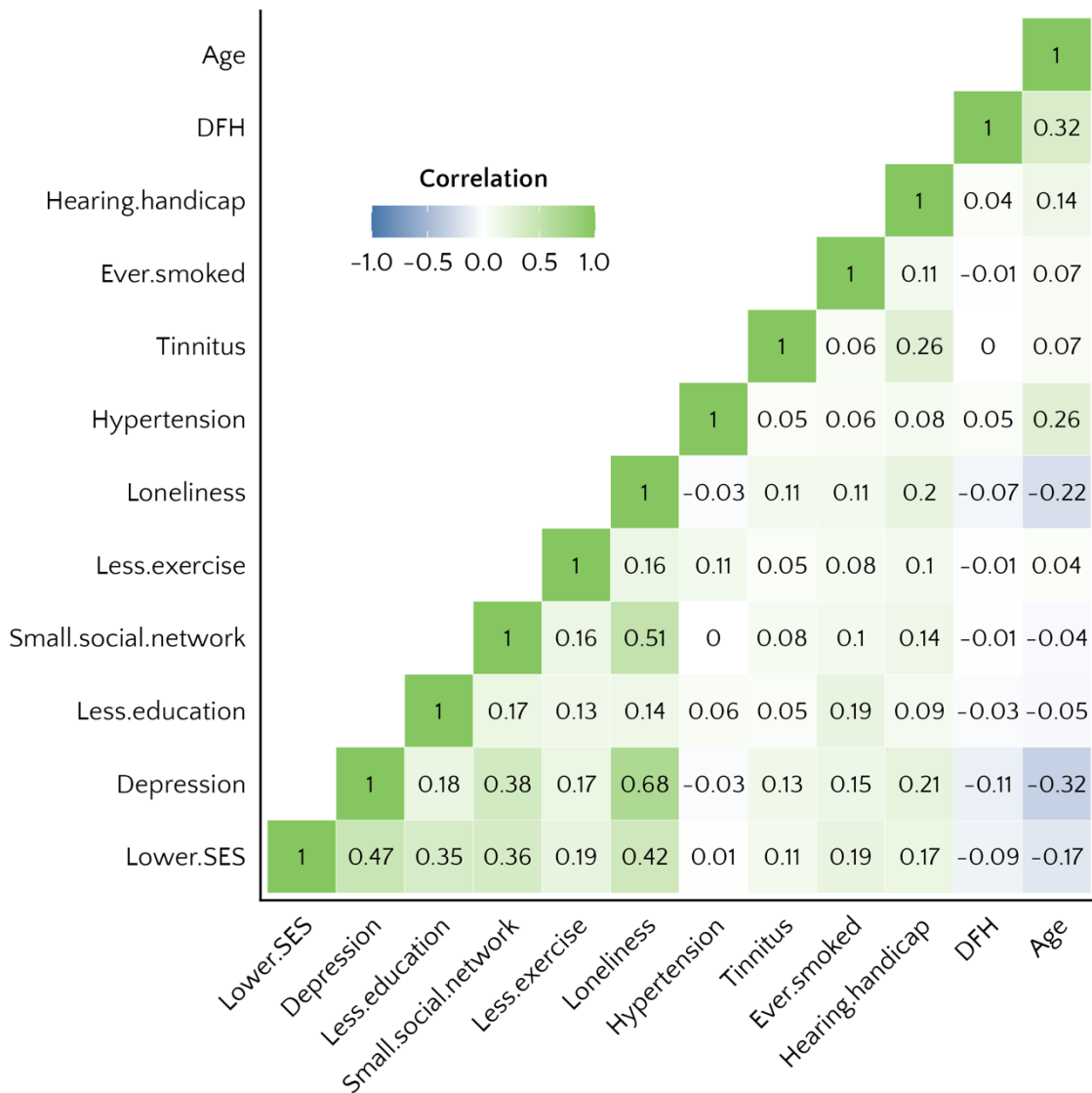


Figure S14: Correlations between predictors in the total sample ($N = 4,995$). Numbers represent Pearson product-moment correlation coefficients for pairs of continuous variables, polyserial correlations for pairs of continuous and categorical variables, and polychoric correlations for pairs of categorical variables.

8.3.3 Sensitivity analyses in a sex-balanced subset of the data

As the ratio of women to men was 66.9% to 33.1% in the total sample, we reran the same series of analyses in a sex-balanced subset of the data. We used the MatchIt package v.4.5.5 (D. E. Ho et al., 2011) to down-sample women so that $n = 1,323$ for both women and men in the training data, matching them on age. The resultant mean age of women ($M (\pm SD) = 47.43 (\pm 14)$) in this subset of the training data was approximately the same as that of men ($M (\pm SD) = 47.45 (\pm 14.04)$). Using the same approach, we also generated a subset of the hold-out data so that $n = 328$ and mean age $M (\pm SD) = 47.25 (\pm 13.80)$ for both women and men.

8.3.3.1 Model training performance

Performance of all models fitted in the sex-balanced training subset is shown in Table S4. Benchmark model with sex and age alone explained 14.3% of variance in objective cognition and sex and age alone also predicted subjective memory problems with an AUC = .588. These numbers were slightly higher compared to the main analyses in the total unbalanced training sample, where $R^2 = 12.7\%$ and AUC = .570. After including risk factors, the models explained 16.6% of variance in objective cognition and predicted subjective memory problems with an AUC = .720. Permutation analyses show that this was a significant increase in model performance in both cases, as the effects did not overlap with the permutation-based null distribution (both $p < .001$; see Figure S15, panels A and B). Adding sex interactions to the model did not significantly increase model performance for objective cognition ($R^2 = 16.7\%$, $p = .288$; see Figure S15, panel C) or for subjective memory problems (AUC = .716; $p = .463$; see Figure S15, panel D).

Table S4: Model performance results from nested repeated cross-validation predicting objective vs. subjective cognition in a sex-balanced subset of the training set ($n = 2,646$).

Model	Objective cognition			Subjective memory problems			
	Regression model performance			Classification model performance			
	MAE	R ²	RMSE	AUC	Accuracy	Balanced accuracy	Log-Loss
Benchmark (sex + age)	0.608	.143	0.769	.588	.609	.522	0.659
Sex + age + risk factors	0.600	.166	0.758	.720	.673	.629	0.596
Sex + age + risk factors + sex interactions	0.600	.167	0.758	.716	.674	.630	0.599

Note: R² = R squared; RMSE = Root mean square error; MAE = Mean absolute error; AUC = Area under the classification curve, LogLoss = Logarithmic loss.

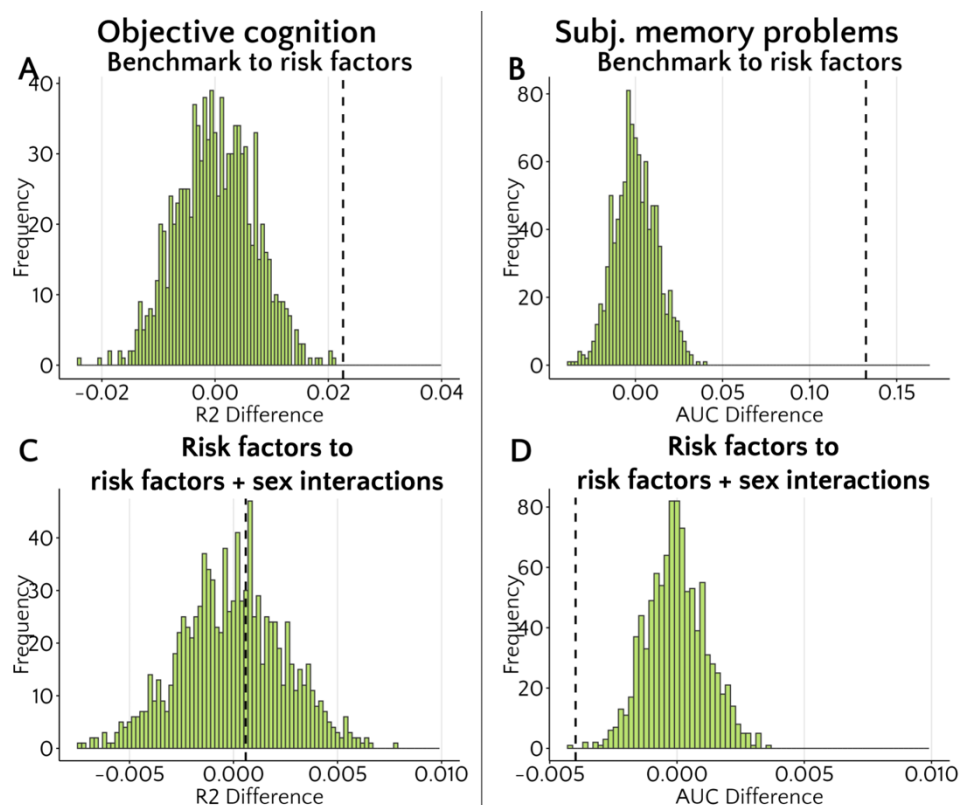


Figure S15: Permutation analysis of between-model differences in performance, based on the sex-balanced subset of the training sample ($n = 2,646$). Bars represent the null distribution of 1000 permuted performance differences between models. Dashed lines represent true performance differences between models, specifically: (A): R² difference between the benchmark model and the risk factors model; (B): AUC difference between the benchmark model and the risk factors model; (C): R² difference between the risk factors model and the

interaction model; (D): AUC difference between the risk factors model and the interaction model.

8.3.3.2 Predictor importance

In the benchmark models, both sex and age were retained as predictors, with final hyperparameters being $\alpha = .0527$ and $\lambda = 0.0001$ for objective cognition and $\alpha = .0001$ and $\lambda = 0.0001$ for subjective memory problems, which means a prevalently ridge, respective fully ridge model with the lowest possible penalty parameter, i.e., a model very close to simple linear/logistic regression without regularisation. The mean SHAP values for the benchmark models (Figure S16, panel A) show that being a woman was linked to better objective cognition, but higher likelihood of subjective memory problems, and higher age was linked to both objectively and subjectively worse cognition. Age was more important for objective cognition than sex. This pattern for age and sex held when brought forward to the other models as covariates.

In the risk factors models, all 13 predictors were retained, with final hyperparameters being $\alpha = .0001$ and $\lambda = 0.0001$ for both objective cognition and subjective memory problems, which means a fully ridge model with the lowest possible penalty parameter. The mean SHAP values of the risk factors models are presented in Figure S16, panel B. The most important predictors of worse objective cognition were age, lower education, and depression. *Better* objective cognition in this model was predicted by being a woman, but also by tinnitus, less exercise, smaller social network, and lower SES. The most important predictors of having subjective memory problems were depression, hearing handicap, closely followed by lower SES, age, being a woman, having less exercise, and then other predictors. The low absolute mean SHAP value for hypertension suggests low importance of this predictor for the model. Notably, smaller social network was the only predictor linked to a *lower* likelihood of subjective memory problems.

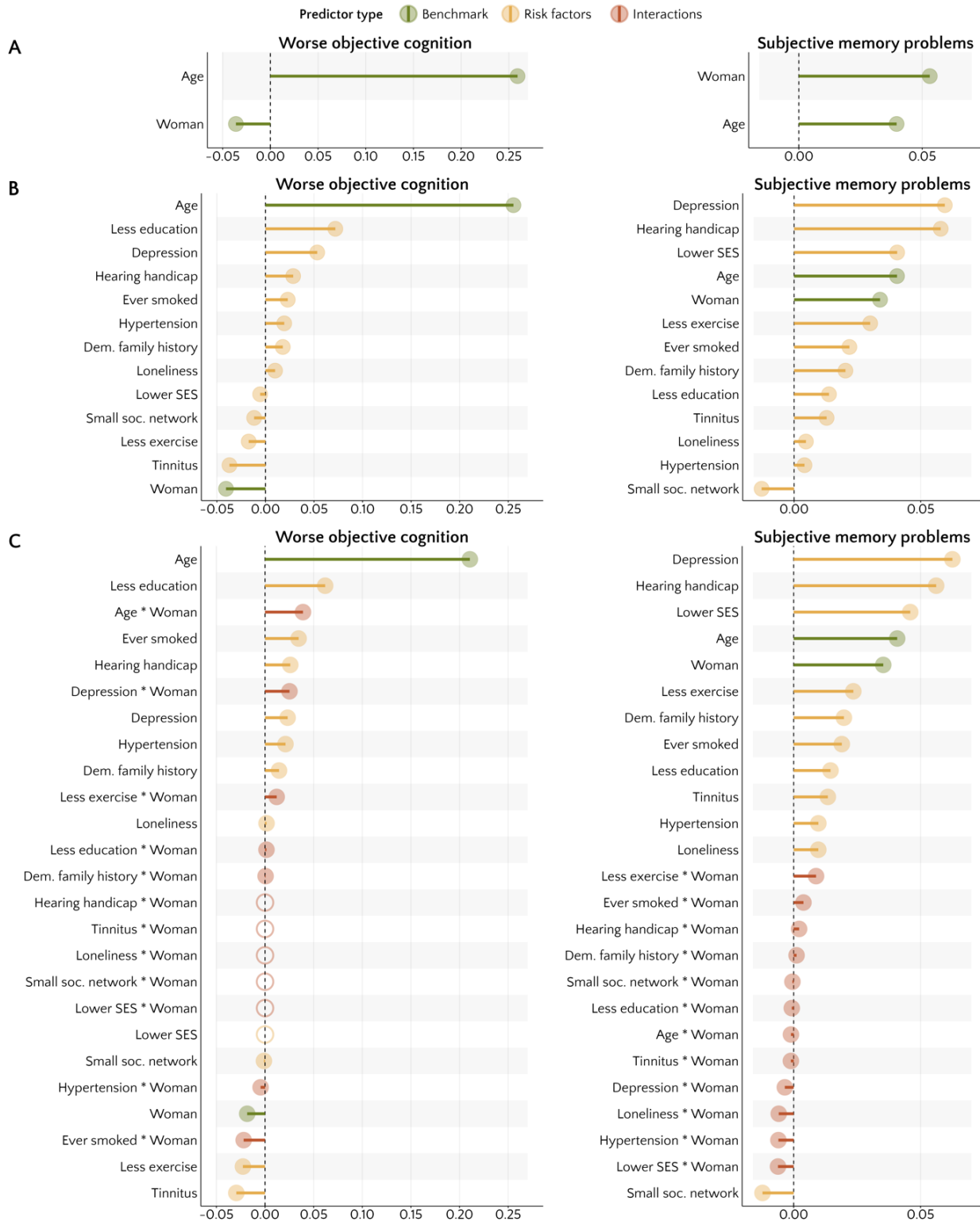


Figure S16: Ranked mean SHAP values for (A) benchmark models; (B) risk factors models without interactions; (C) risk factors models with sex interactions, shown separately for the objective cognition regression models (left) and subjective memory problems classification models (right). Higher absolute values indicate higher predictor importance for the model. Positive SHAP values indicate the predictor is associated with worse cognitive outcomes (e.g., higher depressive symptoms are linked to lower objective cognition and to a higher likelihood of

subjective memory problems. Predictors not retained in the final model are depicted as empty circles.

In the risk factors model with interactions, six predictors were omitted for objective cognition and all predictors were retained for subjective memory problems. Final hyperparameters again suggest a prevalently ridge ($\alpha = .1053$), prevalently low penalty model ($\lambda = 0.0527$) for objective cognition and a fully ridge ($\alpha = .0001$), very low penalty ($\lambda = 0.0001$) model for subjective memory problems. The mean SHAP values of the interaction models are presented in Figure S16, panel C. The pattern of variable importance is similar to that of models without interactions and closely resembles that in the main analyses, with most interaction terms ranking rather low in importance for both objective cognition and subjective memory problems. The most important predictors of objective cognition were age, less education, and the interaction of age and sex, with depression alone notably moving down the ranks.

8.3.3.3 Model testing performance

We evaluated the performance of these models in a sex-balanced hold-out subset (see Table S5 for detailed performance metrics). Benchmark model with sex and age alone explained 11.8% of variance in objective cognition and sex and age alone also predicted subjective memory problems with an AUC = .552. After including risk factors, the model explained 13.5% of variance in objective cognition, but this increase in performance was not significant when comparing to a permutation-based null distribution ($p = .064$; see Figure S17, panel A). The equivalent risk factors model predicted subjective memory problems with an AUC = .716, which was a significant increase in model performance, as the effects did not overlap with the permutation-based null distribution ($p = .064$; see Figure S17, panel B). Adding sex interactions to the model did not significantly increase model performance for either objective cognition ($R^2 = 13.2\%$, $p = .288$; see Figure S17, panel C) or for subjective memory problems (AUC = .719; $p = .463$; see Figure S17, panel D).

Table S5: Model performance results from nested repeated cross-validation predicting objective vs. subjective cognition in a sex-balanced subset of the hold-out set ($n = 656$).

Model	Objective cognition			Subjective memory problems			
	Regression model performance			Classification model performance			
	MAE	R^2	RMSE	AUC	Accuracy	Balanced accuracy	Log-Loss
Benchmark (sex + age)	0.602	.118	0.775	.552	.590	.489	0.664
Sex + age + risk factors	0.594	.135	0.767	.716	.686	.630	0.591
Sex + age + risk factors + sex interactions	0.593	.132	0.768	.719	.683	.628	0.590

Note: R^2 = R squared; RMSE = Root mean square error; MAE = Mean absolute error; AUC = Area under the classification curve, LogLoss = Logarithmic loss.

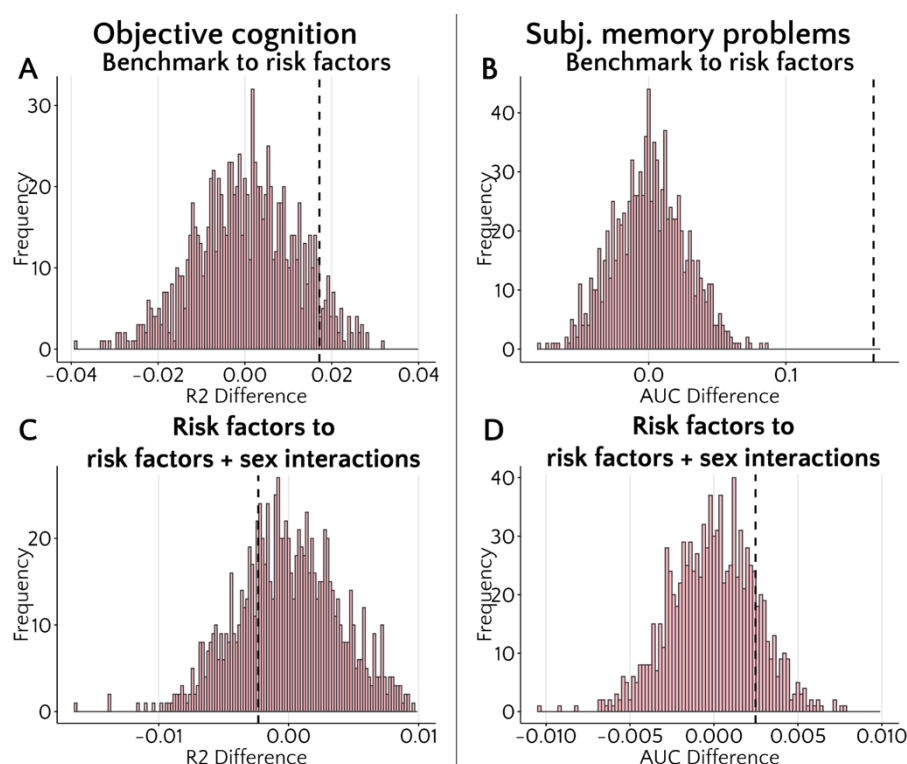


Figure S17: Permutation analysis of between-model differences in performance, based on the sex-balanced subset of the hold-out sample ($n = 656$). Bars represent the null distribution of 1000 permuted performance differences between models. Dashed lines represent true performance differences between models, specifically: **(A)**: R^2 difference between the benchmark model and the risk factors model; **(B)**: AUC difference between the benchmark model and the risk factors model; **(C)**: R^2 difference between the risk factors model and the

interaction model; (**D**): AUC difference between the risk factors model and the interaction model.