Exploration of apathy in people living with neurocognitive disorder

By Clare Burgon, BSc

Thesis submitted to the University of Nottingham

For the degree of Doctor of Philosophy, 23rd January 2023

Acknowledgements

This study would not have been possible without the contribution of the participants, people living with neurocognitive disorders and their caregivers, who kindly gave up their time to take part in the research studies. It has been a true privilege to hear their experiences.

I would like to thank my supervisors, Prof Rowan Harwood, Prof Sarah Goldberg and Dr Veronika van der Wardt for their academic guidance and support but also supporting me through the entirety of this PhD journey. Allowing me to conduct this PhD research whilst continuing to work has enabled me to continue to developing both professionally and personally, with all the changes this has entailed, for this I am truly grateful.

I would like to acknowledge and thank my colleagues across the different sites who were involved in screening, recruiting, and assessing participants for the PrAISED2 study. Thank-you also to my colleagues at the university, in particular Juliette and Rupinder, for their support and encouragement, especially the frequent supportive lunch chats, and when they could no longer happen during lockdown, virtual chats. An extra thank-you is owed to Juliette for proof-reading this work.

I would like to acknowledge and thank Catherine Brewin, Ester Bellavia, Marta Castro Rodriguez, and Kenichi Sakuda and the senior research librarian Jane Grogan for their contributions to the systematic review, Glenys Caswell and Claudio Di Lorito for their feedback on the qualitative work, and Prof Lourens Waldorp for his support and guidance with the network analysis.

Thank-you to the National Institute for Health and Care Research for funding the PrAISED2 study and to the School of Health Sciences and Centre for Rehabilitation & Ageing Research, Unit 3 - Injury, Inflammation and Recovery Sciences, School of Medicine at the University of Nottingham for their tuition fees funding.

Thank-you to the Faculty of Medicine and Health Sciences' Dementia Frail Older People and Palliative Care Patient and Public Involvement group, who provided feedback and guidance throughout the study. In particular, I would like to thank Maureen Godfrey, Marianne Dunlop and Morag Whitworth for giving up their time and offering their keen insights during the systematic review and qualitative study.

To my family, I would like to express never ending gratitude for your support, love and patience. Dad, I know you are no longer able to read this yourself, but I treasure every day we have where you are still your usual self, with your great sense of humour which I don't think will ever leave you. Mum, you always made sure I knew you were proud of me no matter what I did, and I am so grateful for your constant support. I realize the irony that a thesis on dementia has meant that I have been less able to help you with dad, and I look forward to being able to be there for you both more now. George, I can never thank you enough for supporting me through this journey, and all the changes it has seen. You have kept me going through everything. I especially owe you for the extra hours of childcare you have had to put in in the last few months of write up! Eloise, you have taken me on the scenic route on this journey, and what a journey it has

been! Mummy is looking forward to doing less 'tap tap tap' and more 'play', 'cuddles' and 'read it'!

This thesis presents research funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (Reference Number RP-PG-0614- 20007). The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

Statement of Contribution

Topic

I began intending to investigate the relationship between the initiation, planning and organisation and performance components of activities of daily living with executive function and affect in people with dementia. It became clear during my initial literature review that the definition and measure of apathy was not clear, with different criteria for, domains and types of apathy having been recently proposed. This led me to undertake an exploration of apathy in dementia and mild cognitive impairment. I used the opportunity of a large sample of people with dementia and mild cognitive impairment and their carers taking part in the Promoting Activity Independence and Stability in Early Dementia (PrAISED)2 RCT, to collect data for this PhD study.

Systematic review

I identified the lack of systematic reviews of apathy measures for people with dementia and MCI, and the relatively novel COSMIN criteria for systematic reviews of measures which guided this systematic review. With guidance from my supervisors, and a subject librarian, I set the systematic review criteria and developed the search strategy. I conducted the search and screened the titles and abstracts of the texts for eligibility independently. I screened all full-text articles and extracted data independently, and 10% were double-screened whilst 20% of data extraction was checked by collaborators. I rated the studies using COSMIN criteria independently, seeking advice from the COSMIN developers, supervisors, and Patient and Public Involvement members when required. 21% of studies were also double-rated by collaborators (see Chapter 4 for further details). I also wrote this up as a paper which has been published (see publications list).

Network analysis study

As a research assistant within the core research team, I was involved in many aspects of the PrAISED2 RCT, and have co-authored the protocol paper [1]. I was involved in the study set up, and led on the ethics application for the PrAISED2 trial, which I worked on between December 2017 and March 2018. I trained the majority of researchers delivering the research assessments to people with dementia and their carers, and conducted 18 of these screening and baseline assessment visits with participants in the community. This involved screening for eligibility, assessing capacity, taking informed consent, and conducting the questionnaire-based and assessment-based measures. The data was entered by myself and other researchers into a MACRO online database, and once entry was complete, the accuracy of this data entry was checked by myself and other researchers in line with the clinical trials unit's data cleaning plan.

After identifying the method I wished to use for the quantitative study, in January 2019 I attended a five-day course in network analysis for psychology, by the PsychoSystems Project, University of Amsterdam. Once data collection was

complete, I requested the raw data from the clinical trials unit, except for some sum-score measures which were scored by the clinical trials unit statistician. I checked the data for outliers and conducted preliminary analyses independently. I wrote the code and conducted the network analysis, and this was later amended based on meetings and guidance from my collaborator (Prof. Lourens Waldorp). I interpreted the data with guidance from my collaborator.

Qualitative study

With guidance from my supervisors, I devised the qualitative study topic, design, and interview schedule, with feedback from collaborators. I recruited participants to this study, (obtaining details from the PrAISED2 participant log), and conducted the qualitative interviews independently. I transcribed one interview and the remainder were transcribed by a transcription service. I analysed the transcripts independently, taking the codebook, tentative themes, any queries, and points of interest to regular supervision meetings. Collaborators also separately coded some transcripts. See section 6.4 for further details. I also wrote this up as a paper which has been published (see publications list).

Glossary

Abbreviations used throughout the thesis

Abbreviation	Description
AD	Alzheimer's Disease
AES	Apathy Evaluation Scale
ADL	Activities of Daily Living
AMI	Apathy Motivation Index
bADL	basic Activities of Daily Living
BBS	Berg Balance Scale
CANTAB	Cambridge Neuropsychological Test Automated Battery
CFA	Confirmatory Factor Analysis
COSMIN	The Consensus-based Standards for the selection of health Measurement Instruments
COVID-19	Coronavirus 2019
CS	Correlation Stability
CV	Cross Validation
DAD	Disability Assessment for Dementia
DAS	Dimensional Apathy Framework
DSM	Diagnostic and Statistical Manual of Diseases
EBIC	Extended Bayesian Information Criterion
GGM	Gaussian Graphical Model
GP	General Practitioner
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
HADS	Hospital Anxiety and Depression Scale
iADL	Instrumental Activities of Daily Living
ICD	International Classification of Diseases
ICF	International Classification of Functioning, Disability and Health
MCI	Mild Cognitive Impairment
MGM	Mixed Graphical Model
MMSE	Mini Mental State Examination
MoCA	Montreal Cognitive Assessment
NCD	Neurocognitive Disorder
NIHR	National Institute for Health and Care Research
PPI	Patient and Public Involvement
PrAISED2	Promoting Activity Independence and Stability in Early Dementia 2
RCT	Randomised Controlled Trial
SHARE-FI	SHARE Frailty Instrument
TUG	Timed Up and Go
UK	United Kingdom
UVA	Unique Variable Analysis

Abbreviations used in the systematic review

Abbreviation	Description
AD-RD	Alzheimer's Disease and Related Dementias Mood Scale
AES	Apathy Evaluation Scale
AES-12PD	Apathy Evaluation Scale for Parkinson Disease
AES-C	Apathy Evaluation Scale – Clinician version
AES-I	Apathy Evaluation Scale – Informant version
AES-I-16	Apathy Evaluation Scale Informant 16 item version
AES-S	Apathy Evaluation Scale – Self version
AI	Apathy Inventory
AI-C	Apathy Inventory – Clinician version
AI-I	Apathy Inventory Informant version
AI-S	Apathy Inventory Self version
AMI	Apathy Motivation Index
AS	Apathy Scale
AS-HC	Apathy Scale - Home Care version
AS-I	Apathy Scale - Informant version
AS-S	Apathy Scale - Self version
b-DAS	brief-Dimensional Apathy Scale
BMDS	Behavioural and Mood Disturbance Scale
BSSD	Behavioral Syndromes Scale for Dementia
CVS	Content Validity Study
DAIR	Dementia Apathy Interview Rating
DAS	Dimensional Apathy Scale
DAS-I	Dimensional Apathy Scale - Informant version
DAS-S	Dimensional Apathy Scale – Self version
DEX	Dysexecutive Questionnaire
DS	Development Study
FrSBe	Frontal Systems Behavior Scale
FrSBe-11a	Frontal Systems Behavior Scale 11-item apathy subscale
FrSBe-14a	Frontal Systems Behavior Scale 14-item apathy subscale
FrSBe-6a	Frontal Systems Behavior Scale 6-item apathy 101 subscale
GDS-3a	Geriatric Depression Scale 3 item 102 apathy subscale
GDS-3a	Geriatric Depression Scale 3 item apathy subscale
GDS-6a	Geriatric Depression Scale 6 item apathy subscale GIP
GDS-6a	Geriatric Depression Scale 6 item apathy subscale
GIP	Behavioral Rating Scale for Psychogeriatric Inpatients
IMD	Index of Mental Decline
KBCI	Key Behaviors Change Inventory
LARS	Lille Apathy Rating Scale
LARS-C	Lille Apathy Rating Scale - Clinician version
LARS-I	Lille Apathy Rating Scale - Informant version

MDS-UPDRS	Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale
NPI	Neuropsychiatric Inventory
NPI-A	Neuropsychiatric Inventory – Alternative version
NPI-C	Neuropsychiatric Inventory – Clinician version
nr	not reported
OA	Older Adults
PLwNCD	People Living with Neurocognitive Decline
RR	Reviewer Rating
UPDRS	Unified Parkinson's Disease Rating Scale

Abstract

Background

Apathy is highly prevalent in people living with neurocognitive disorder (NCD). What constitutes apathy and how it is experienced by people living with NCD is poorly understood and under-researched, and there are no gold-standard measures of apathy. Apathy is associated with impaired functional ability, worse quality of life, conversion from MCI to dementia, increased care costs and carer burden, and may prevent effective participation in therapy programmes. It is important to understand what constitutes apathy, how to assess it, and identify possible mechanisms for apathy and potential treatment targets, in people living with NCD. However, the typical domain-based approach may be limiting our understanding of apathy. It is important to explore how individual indicators of apathy relate to one another. A broader exploratory approach is needed to understand how to conceptualise apathy, identify possible mechanisms of apathy, including an examination of how apathy is understood and experienced in people living with NCD. This thesis aimed to answer the following research questions: What measures of apathy are available and what is the quality of the evidence for their use with people living with NCD?; How should apathy be characterised and what are its conceptual boundaries?; What are the mechanisms and impact of apathy?

Methods

A comprehensive systematic review of the measurement properties of apathy measures was conducted. The ‘Consensus-based Standards for the Selection of Health Measurement Instruments’ approach was used to identify the available apathy measures and assess their quality of evidence and quality for use people living with NCD, to determine the most appropriate measure to use in a subsequent network analysis. Mixed methods enable the identification of possible mechanisms through quantitative methods, and corroboration and elaboration of these through qualitative methods, and so can be particularly useful for understanding mechanisms and processes. A convergent parallel-databases mixed methods study was conducted. A quantitative network analysis method was used to assess how indicators of apathy relate to one another, to examine the network structure of apathy and depression and to assess the impact of overlapping concepts including activities of daily living, executive functioning, physical impairment, frailty and age on the network. A qualitative interview study was conducted to explore how people living with NCD and their carers experience and understand apathy, how apathy develops, impacts their lives, and what influences this. The qualitative study was analysed using reflexive thematic analysis, whilst the network analysis study was analysed using a (cross-sectional) Gaussian Graphical Model network analysis approach. Participants were people living with NCD and their carers, sampled from a larger randomised controlled trial. 365 participants

were included in the network analysis study, and 16 people living with NCD and 14 carers took part in the qualitative study. Findings were integrated in a narrative discussion.

Results

In the systematic review, fifty-seven publications regarding 18 measures and 39 variations met the eligibility criteria. Few studies involved people living with NCD or carers in the development of the measure. The Apathy Evaluation Scale and Lille Apathy Rating Scale had sufficient content validity, reliability, construct validity, and where applicable, structural validity and internal consistency in people living with NCD.

The network analysis study found that apathy indicators did not cluster into distinct domains, though apathy and depression indicators did cluster separately. Motivation and initiative formed the most important indicators of apathy, whilst energy and lack of cheerfulness were key to linking indicators of depression and apathy. Activities of daily living were independently associated with a number of apathy indicators, in addition to executive function and frailty, linking impairments and apathy.

The qualitative study found that apathy is experienced as an understandable response to the everyday difficulties people with NCD face. It can be understood to be a coping mechanism to preserve identity in the face of declining physical and cognitive abilities and associated threats to competence and autonomy, which is exacerbated by lack of opportunities and social support.

Conclusions

Numerous scales are available to assess apathy, with varying psychometric properties. New measures assessing apathy should be developed with the contribution of people living with NCD and carers, and could forego typical refinement approaches based on factor loadings, to develop a measure more suitable for future network analysis studies. In contrast to existing criteria for apathy, apathy may not be composed of distinct domains, and is proposed to be unavoidably linked to the wider social and environmental context in addition to cognitive and physical impairment. The exclusion of cognitive or physical impairment or lack of opportunity from the diagnostic criteria for apathy is problematic, given the importance of cognitive and physical impairment, social support and opportunity in apathy. Apathy indicators, regardless of the cause, may be improved by reducing task difficulty and effort, and by providing supportive and inclusive environments. Therefore, though apathy may occur as a result of neurodegenerative changes, it should not be considered in isolation from cognitive and physical impairment and the social and environmental context. A portion of this work took place during the

COVID-19 pandemic which restricted the number of participants recruited and measures used.

Publications and Dissemination

Publications

Burgon, C., Goldberg, S. E., Van Der Wardt, V., & Harwood, R. H The Experiences and Understanding of Apathy in People with Neurocognitive Disorders and their Carers: A Qualitative Interview Study. *Age and Ageing* (forthcoming) Statement of contribution: I conceived of the study with supervision from RH, SG, and VvdW. I recruited to and conducted the qualitative interviews. I analysed the data, with support and supervision from RH, SG and VvdW. I drafted the manuscript and all authors edited the text and approved the final manuscript.

Burgon, C., Goldberg, S. E., Van Der Wardt, V., Brewin, C., & Harwood, R. H. (2021). Apathy Measures in Older Adults and People with Dementia: A Systematic Review of Measurement Properties Using the COSMIN Methodology. *Dementia and geriatric cognitive disorders*, *50*(2), 111-123. Statement of contribution: I conceived of the study with supervision from RH, SG, and VvdW. I developed and conducted the search and pre-screening. VvdW, CBr and I screened the studies for eligibility. SG, VvdW and I reviewed the included studies for their risk of bias and quality. I drafted the manuscript and all authors edited the text and approved the final manuscript.

Bajwa RK, Goldberg SE, Van der Wardt V, Burgon C, Di Lorito C, Godfrey M, Dunlop M, Logan P, Masud T, Gladman J, Smith H, Hood-Moore V, Booth V, Das Nair R, Pollock K, Vedhara K, Edwards RT, Jones C, Hoare Z, Brand A, Harwood RH (2019). A randomised controlled trial of an exercise intervention promoting activity, independence and stability in older adults with mild cognitive impairment and early dementia (PrAISED2) - A Protocol. *Trials*. *20*(1):815. doi: 10.1186/s13063-019-3871-9. Statement of contribution: I commented on and provided support with the development of the protocol. I led the preparation of the ethical review application, with RH, SG, VVdW and RB. All authors including myself contributed to management of the programme and approved the final version of this paper.

Presentations

Burgon, C., Network analysis of apathy, presented at the Centre for Rehabilitation & Ageing Research Post Graduate Researchers and Centre for Doctoral Training Student Rehabilitation Research Seminar Series on 3rd November 2022.

Burgon, C., The experiences of apathy in people with dementia and mild cognitive impairment and their carers: A qualitative interview study, presented at the 'Sustaining Wellbeing after a Dementia Diagnosis' event on 13th January 2023.

Poster Presentations

Burgon, C., Goldberg, S., van der Wardt, V., & Harwood, R. H. (2021). The experiences of apathy in people with dementia and mild cognitive impairment and their carers: A qualitative interview study. *Alzheimer's & Dementia*, *17*, e053464. (AAIC21 poster presentation).

Burgon, C., Goldberg, S., van der Wardt, V., & Harwood, R. H. The experiences of apathy in people with dementia and mild cognitive impairment and their carers: A qualitative interview study. Poster and flash presentation presented at School of Health Sciences Postgraduate Allied Health Research Conference 2022.

Burgon, C., Whitworth, M., Bajwa, R. K., van der Wardt, V., Goldberg, S., Harwood, R. (2018). Patient and Public Involvement (PPI) in Reviewing the Validity of Apathy Measures. Poster presented at International Perspectives on Evaluation of PPI in Research on 15th – 16th November 2018 (DOI: 10.13140/RG.2.2.34106.24006).

Patient and Public Involvement

- Presentation and discussion of initial PhD idea at PPI meeting (on 22nd September 2017).
- Meeting with Morag to develop PPI conference abstract (on 24th July 2018).
- Email exchanges regarding PPI in systematic review (November 2018 to January 2019).
- Meeting with Morag, Maureen and Marianne regarding reviewing apathy measures (on 29th August 2019).
- Presentation of qualitative design at PPI meeting (15th February 2019).
- Review of interview schedule by Maureen and Marianne (September 2019)
- Email exchanges and analysis of two transcripts by Marianne and Morag (March 2020).

Additional Funding Awarded

- ARC East Midlands PhD Travel / Research prize 2022, to hold a PPI dissemination event for the work presented in this thesis.
- Bursary from the NIHR School for Primary Care Research (SPCR) and the Institute of Health & Society for travel to and attendance at the International Perspectives on Evaluation of PPI in Research Conference 2018.

Table of Contents

Table of Figures

Table of Tables

Chapter 1. Introduction

1.1. Introduction

The term ‘apathy’ is derived from the Greek ‘apatheia’, meaning “without feeling” but is now understood to refer to a variety of thoughts, feelings and behaviours, including a lack of interest, concern or enthusiasm [2], or indifference [3], lack of energy or unwillingness to act [4]. Apathy may be considered part of the normal human condition and is a non-technical term (i.e. exists outside of scientific and medical terminology). However, it has also be characterised as a symptom, feature, impairment or behavioural disturbance observed in various disorders such as Parkinson’s disease, schizophrenia, depression, and neurocognitive disorders [5–7].

Neurocognitive disorders (NCDs) are primary deficits in cognitive functioning, reflecting a decline from previous abilities, and encompasses various types of dementia and mild cognitive impairment (MCI) [7]. Dementia is an umbrella term for a group of neurological conditions which result in an impairment in cognition and everyday functioning. Mild Cognitive Impairment (MCI) is a term used to describe similar symptoms, but in people whose daily functioning is not currently affected [8]. One in fourteen people aged over 65 are living with dementia, and it affects over 900,000 people in the United Kingdom (UK), which is estimated to rise to 1.6 million by 2040. The total cost of care for dementia in the UK was £34.7 billion in 2019, and is expected to rise [9]. In addition to cognitive symptoms, people with NCD also experience various ‘neuropsychiatric symptoms’ which are disruptions in thoughts, mood and behaviour, for example depression, agitation and delusions [10].

Apathy is often considered the most prevalent [11,12] and persistent [13] neuropsychiatric symptom of NCDs. The exact prevalence of apathy varies depending on the severity and type of NCD, the definition of apathy, and the study method. For example, a systematic review of neuropsychiatric symptoms in dementia revealed that population based studies provided prevalence estimates of between 27% and 57% whilst studies recruiting people attending a memory clinic found the prevalence of apathy was between 42% and 74% [11]. Another review revealed that prevalence estimates for apathy in Alzheimer’s Disease (AD) were between 19% to 92%, 15% to 93% in people with vascular dementia, and much higher in behavioural variant frontotemporal dementia at 62% to 89% [12], with the wide range of figures likely reflective of the criteria used. Apathy is prevalent across all stages of dementia [14], and has also been found across different studies to be present in 3 to 51% of people with MCI [15]. Regardless of the exact prevalence, it is clear that apathy affects many people living with NCD.

In addition to being highly prevalent, apathy is also associated with important outcomes for people living with NCD. It is associated with impaired functional ability in people with dementia, worse quality of life and increased informal care costs in the earlier stages of dementia, weight loss in nursing home residents, increased carer burden, and increased risk of dementia [16–21]. Furthermore, carers may lack understanding of the occurrence of apathy in dementia, despite its negative impacts on their wellbeing and their caregiving role [22,23]. Apathy may prevent effective participation in therapy programmes aimed at improving other symptoms such as functioning [24]. Despite this, treatments for apathy are lacking [25]. It has also been proposed that apathy could be a useful prognostic marker, to detect individuals at risk of dementia at an early stage, potentially enabling better treatment and research [20]. It is therefore not surprising that apathy has been identified as a priority area for dementia research [26].

Despite the importance of apathy in NCDs, disagreements and uncertainty remain about what constitutes apathy (i.e. what should and what should not be considered part of apathy, in other words, what its conceptual, or 'nosological' boundaries are), how it develops and what its effects are [5,27,28]. The relationship between apathy and similar constructs such as depression, executive function and activity require further consideration and elaboration in the context of NCD [22]. More research is needed to better characterise apathy, and produce a framework or model of apathy that could enable a greater understanding of how apathy develops in NCD and what its effects are [22,28,29].

1.2. Aims

1) To determine what measures of apathy are available and their quality for use with people living with NCD
2) To determine how apathy should be characterised and what are its conceptual, or 'nosological', boundaries in people living with NCD
3) To understand the possible mechanisms and impact of apathy in people living with NCD

1.3. Thesis overview

This thesis presents an exploration of the concept of apathy in NCD, featuring a systematic review of apathy measures, and a mixed methods study, with a qualitative and network analysis component.

The following two chapters present a narrative literature review of current understandings of apathy and depression (Chapter 2) and theories of mechanisms of apathy (Chapter 3).

Chapter 4 presents a systematic review of apathy measures. This chapter addresses the first research aim, to determine what apathy measures are available and what is the quality of the evidence for their use with people living with NCD. This was also undertaken to select the best quality apathy measure to be used in the network analysis sub-study.

Chapter 5 outlines the methodology, rationale and aims for the mixed methods study that addresses the final two research aims. This chapter includes a separate rationale for the qualitative and network analysis sub-studies.

Chapter 6 situates the study within the context of the randomised controlled trial alongside which it was conducted, and describes the methods for this. The methods for the network analysis and qualitative sub-studies are described separately. The overall ethics are considered together, and a description of the methods for integration is provided.

Chapter 7 presents the results of the quantitative network analysis sub-study, which addresses the research question regarding how apathy should be characterised, and what are its boundaries and mechanisms. This chapter includes a description of the preliminary analyses, the main network analyses, and stability and sensitivity analyses. The limitations of these analyses are also discussed.

Chapter 8 presents the findings of this qualitative work which addresses the research questions regarding how apathy should be characterised, and what its mechanisms and impacts. This includes a description of the sample characteristics and themes that were constructed. The limitations of the qualitative sub-study are then discussed.

The results of the sub-studies are then integrated and discussed in the final chapter (Chapter 9). An adapted model of apathy is presented to address the final research aim, to determine the possible mechanisms of apathy.

Chapter 2. Apathy: A review of the literature over the last three decades

2.1. Disagreements in the Definition of Apathy

Characterisations of apathy as a neuropsychiatric symptom of dementia frequently failed to define what 'apathy' was [30]. Research into apathy as an independent concept first became popular in the 1990s, when apathy was first defined systematically in a clinical context, and characterised as a lack of motivation [30]. Marin's understanding of apathy as a loss of motivation was highly influential and remains pervasive in current literature, as the most common and agreed upon conceptualisation of apathy to date [29,31]. However, it has been argued that 'lack of motivation' fails to encompass the entirety of the behaviours and symptoms that are described in apathy [24]. Furthermore it has been criticised for being too subjective, as the concept of motivation is a psychological interpretation of behaviour that must be inferred, and thus difficult to assess [24,32].

Apathy has been alternatively defined as a lack of self-initiation [24]. This definition contrasts with others' conceptualisations of apathy that have included lack of response to external prompts [e.g. 33]. However, Mulin and colleagues [34] found symptoms related to responsiveness to prompting to be less frequent than symptoms related to self-initiation. However, it is not clear from this study whether disrupted responsiveness only occurred alongside disturbance in initiation, or whether responsiveness itself is an important indicator of apathy. Stuss and colleagues [24] argue that there is no one definition of apathy, as it is a multifaceted concept. However, it is arguable that such flexibility in a definition of apathy is of limited usefulness to researchers and clinicians alike, who require clear operationalisations of these otherwise ambiguous concepts.

In an attempt to create a clear and objective definition, Levy and Dubois redefined apathy as a "quantitative reduction of self-generated voluntary and purposeful behaviors" [32, p.916]. This has been criticised for reducing motivation to observable activity, when in fact the two can be found separately, with some people experiencing a loss of motivation but preserved performance of behaviours, and vice versa [35]. This definition has been refined to highlight that reduced goal-directed activity can refer to emotional and social activities, not just behavioural [36]. However, it is not clear how some commonly referenced symptoms of apathy, such as emotional blunting, are embodied in this definition.

A review of the definitions of apathy found that various authors have defined apathy as a reduction in or loss of: motivation; interest or concern; goal-directed self-initiated activity; interest and motivation in everyday activities; or

emotion [31]. Additionally, the recent revision of the International Classification of Diseases (ICD) [37] included a specific reference to apathy in dementia, but defined this as indifference or lack of interest [7], which contrasts with previous definitions' emphasis on motivation and activity. Disagreements remain about how apathy should be defined and operationalised, and what distinct dimensions and types of apathy exist, if any. The transdiagnostic attempts to characterise the indicators of apathy into clearly defined domains are summarised in Table 1 and will now be discussed.

Table 1. *Proposed Criteria for apathy*.

Authors	**Definition and Criteria**	**Domains of apathy**
Marin [30,38]	Definition: Lack of motivation. Inclusion: Symptoms from all domains are present. Exclusion: Not caused by intellectual impairment, emotional distress or lack of awareness.	Domains of Reduced: <u>Goal-directed Behaviour</u> Lack of effort, productivity, activity (including social), initiative or persistence; dependency on others to prompt and structure activity. <u>Goal-directed Cognition</u> Lack of interest in learning new things and experiences, lack of self-concern; reduced value attributed to goal directed activity including socialising. <u>Emotional concomitants of goal-directed behaviour</u> Unchanging, euphoric or flat affect; lack of emotional responsiveness to positive or negative events; lack of excitement.
Starkstein and Leentjens [29]	Definition: Lack of motivation compared to previous. Inclusion: Symptoms from all domains are present for four weeks, and for most of the day. Exclusion: Caused by lack of awareness or substances.	Domains of Reduced: <u>Goal-directed behaviour</u> Lack of effort in activity; dependency on others to prompt activity <u>Goal-directed cognition</u> Lack of interest in new things and experiences; lack of self-concern <u>Emotion</u> Flat affect; lack of emotional responsiveness
Robert, Onyike, et al. [33]	Definition: Reduced or loss of motivation, compared to previous. Inclusion: Symptoms from two of the three domains are present for four weeks, most of the time. Exclusion: Caused by: physical disability including sensory and motor impairment; lack of awareness; or substances.	Domains of Reduced spontaneously and/or externally initiated: <u>Goal-directed behaviour</u> Lack of spontaneous initiation of behaviour including communication; Lack of externally initiated behaviour, including responsiveness in conversation and social participation <u>Goal-directed cognition</u> Lack of spontaneous initiation of cognition, including curiosity for news and others, and ideas for activities; Lack of externally initiated cognition, including curiosity for events <u>Emotion</u> Lack of spontaneous initiation of emotion including blunt affect and lack of feeling; Lack

Authors	Definition and Criteria	Domains of apathy
		of externally initiated emotion, including reduced emotional responsiveness to positive or negative events
Robert, Lanctôt et al. [36]	Definition: Quantitative reduction of goal-directed activity in behavioural, cognitive emotional or social dimensions. Inclusion: Symptoms from two of the three domains are present for four weeks, most of the time. Exclusion: Caused by physical disability including sensory and motor impairment, lack of awareness or substances.	Domains of Loss of or reduced: <u>Behaviour or Cognition</u> Reduced amount of activity; reduced persistence in activity; reduced interest or longer time making decisions; reduced interest in events, news, and self <u>Emotion</u> Reduced spontaneous expression of emotion, interest in events or others, expression of emotion in reaction to the environment including positive and negative events, expression of empathy <u>Engagement in social interaction</u> Reduced initiative in suggesting social activities, participation in social activities, or interest in others; Reduced initiation of and increased withdrawal from conversations; Prefers to stay at home
Miller et al. [39]	Definition: None provided within the criteria. Inclusion: At least one symptom from two of three dimensions of apathy are present frequently for at least four weeks. Exclusion: Exclusively caused by other psychiatric disorders, disabilities, lack of awareness, or substances.	Domains of reduced: <u>Initiative</u> Reduced spontaneity, activity or initiation of usual activity, including conversation and social activities <u>Interest</u> Reduced enthusiasm about usual activities, less interested in events, activities or others, reduced participation in activities even when prompted, less persistence in activities <u>Emotional Expression or Responsiveness</u> Reduced expression of spontaneous emotions or affection, reduced expression of emotion in responsive to positive or negative events, reduced concern about the impact of their actions or empathy

All criteria except for Marin's additionally specify that apathy must cause clinically significant impairment and reflect a change in usual behaviour

2.2. Marin's apathy domains

Marin argued that apathy should be understood to be a primary deficit in motivation, characterised by three distinct domains: behavioural, cognitive, and emotional [38]. The behavioural domain of apathy was proposed to be characterised by reduced: activity (including social and leisure activities); initiative; maintenance of behaviour; productivity and effort, and reliance on external prompting to perform tasks [38]. Reduced interest in or concern about self, others or previously relevant activities constituted the cognitive domain. Whereas the emotional domain comprised reduced emotional responsiveness to events, flat or unaltered affect and lack of excitement [38]. The basis on

which Marin proposed these domains has been questioned [40], which will be further discussed below.

Marin attempted to distinguish between the symptom of apathy i.e. impaired motivation caused directly by a disorder, and the syndrome of apathy, which he argued was also an appropriate interpretation of apathy, as it can be characterised as a collection of symptoms, and is not necessarily directly caused by another syndrome [38]. Marin excluded 'reduced consciousness' and symptoms resulting from NCD in his definition of apathy as a syndrome. Reduction in activity may occur due to reduced functional capabilities, and problems with communication may be mistaken for reduced interest and reduced emotional expression. It is not clear how these are realistically separated in practice in the context of NCD. Furthermore, the need to define apathy as either a distinct syndrome or a consequence of cognitive impairment was later regarded as unnecessary [29]. Nevertheless, this understanding of apathy as a larger construct, composed of different indicators from different domains has been highly influential, despite little critical review of the validity of these domains [40].

2.2.1. Validity of Marin's apathy domains

2.2.1.1. Behavioural domain validity

Behavioural domain symptoms are still commonly understood to be indicators of apathy [31], and are highly prevalent in people with apathy [34]. It has been argued however that the higher prevalence of these symptoms may be due to the ease at which others observe them compared to symptoms in other domains, rather than reflective of their importance [34]. For example, in people with NCDs and apathy, reduced general levels of activity have been found to be the most common symptoms ascertained through clinician interviews, with 100% to 96% exhibiting this compared to between 29% and 39% exhibiting reduced empathy, which is a symptom of social apathy that is less readily observed [41]. Of those who did not show apathy, reduced general levels of activity were still common (29% to 47% reporting this), compared to reduced empathy (in 0% to 2% of people). Whilst this suggests that the most observable symptoms of apathy may be being overestimated, it also highlights that, people with apathy did have more symptoms of reduced activity than people without apathy. In other words, the behavioural domain may be a sensitive but not specific indicator of apathy. However, it is worth noting that 'apathetic' individuals were defined as such by a clinician determining that the person was experiencing at least two of the three of the apathy domains. Therefore, they were necessarily more likely to be experiencing the domain symptoms than the healthy comparison group. This highlights the issue of how to research the validity of the apathy domains when this forms the very criteria for inclusion.

2.2.1.2. Validity of the cognitive domain

As with the behavioural domain, it is generally accepted that the cognitive domain symptoms are indicators of apathy [31], and studies have shown they are amongst the most prevalent indicators of apathy [34,41,42].

In addition to its prevalence, the cognitive domain of apathy may have its own important consequences. For example, lack of interest, a key symptom in the cognitive domain, has been found to be the only symptom to predict conversion from MCI to AD, even after controlling for age, educational level and episodic memory [43]. It was suggested this may be due to lack of interest causing reduced leisure participation, which increases risk of dementia [43]. However, the separation of the cognitive and behavioural domains has been criticised for being impossible to assess in practice, given that interest and activity are inextricably linked [40].

2.2.1.3. Validity of the emotional domain

The validity of the emotional domain of apathy in particular has been questioned, as unlike cognition and activity associated with goal-directed behaviour, emotional expression is not necessarily linked to motivation [29]. In people with mild AD, the emotional domain is the least frequently reported aspect of apathy, though at least one symptom from this domain was still reported in 44% of people [42], indicating its relevance to this population. However, the authors only reported the domain prevalence in the overall sample, rather than of those meeting cut-off criteria for apathy, so it is unclear what this may mean for the contribution of emotional domain to clinically relevant apathy.

In a heterogeneous neurodegenerative disorder sample, emotional initiation and responsiveness were found to be the least common symptoms of those meeting criteria for apathy [34], with a prevalence of 56% and 40% respectively. Of those meeting criteria for apathy in a sample of people with NCDs or affective disorders, 73% showed indicators from the emotional domain of apathy, compared to 100% and 97% from the other domains, and the emotional domain had worse sensitivity to clinical apathy compared to the other dimensions [41]. Moreover, the emotional subscale of the Dimensional Apathy Scale (DAS) has been found to have lower correlation to total apathy measured by the Apathy Evaluation Scale [AES; 44] than the other two domains [45]. These studies indicate that the emotional domain is the least common indicator of apathy, and is least related to overall apathy. The emotional domain is also worse than the other two domains at differentiating people with AD from healthy controls, and has differential relationship with other constructs compared to the other two domains [45], which could suggest it is a separate construct. Despite this, it is worth noting that people with AD still experienced significantly more indicators of the emotional domain than

healthy controls [45]. Moreover, whilst the emotional dimension of apathy is the least prevalent, and though prevalence estimates vary across studies (likely due to differing populations and different criteria applied), the literature indicates that it is, nevertheless, frequently experienced by people living with NCD and apathy. What remains unclear is whether the presence of these emotional symptoms should be seen as a necessary criterion of apathy, or a separate phenomenon associated with apathy and cognitive impairment.

2.2.2. Evidence for Marin's dimensional structure

The validity of the separation of Marin's proposed domains of apathy have largely been investigated through factor analysis studies. Many studies using Marin's own scale, the AES, have found that apathy items largely cluster together to form one general 'apathy' factor [44,46–48], which contrasts with Marin's proposal of three distinct dimensions of apathy. Though not all items clustered into a single apathy factor, the other factors identified have been inconsistent, and have included factors of cognitive-behavioural domain [49], friendship [50], insight [44]; insight and social activities [48]; and interest or novelty-seeking [38,46,48]. Thus, it seems even when the studies indicate that apathy is multidimensional, this does not resemble Marin's three domains, nor do they indicate a consistent alternative dimensional structure.

These findings may be the result of an inappropriate method however, as factor analysis typically uses Pearson correlations, which assume each item is measured on a continuous scale [51]. The four-point Likert scoring system commonly used to measure items in these scales violate this assumption. The inappropriate use of factor analysis in this way results in an increased likelihood of multiple factors or 'over factorisation', at least when items are scored on five or less points. This is because the categorisation that occurs in a Likert scale restricts the correlation that can occur between items, resulting in lower correlation [51]. Therefore, the studies assessing the structure of apathy using scales made up of categorical items analysed using methods based on Pearson's correlations, may be overestimating the number of factors. Factor analyses using polychoric correlations do not assume that item level data is continuous, reducing the risk of this over-factorization [52]. When using this approach, it has been found that the AES is unidimensional, in contrast to the two domain structure that was found with traditional principle component analysis using the same data [47].

These studies cast doubt over Marin's proposed distinct dimensions of apathy. Despite this, they have formed the building blocks for subsequent attempts to develop apathy diagnostic criteria including criteria applicable to people living with NCD [29,33,36].

2.3. Clinical criteria for apathy

2.3.1. The first ‘diagnostic’ criteria for apathy

Marin’s criteria for apathy were adapted and refined to propose the first ‘diagnostic’ criteria for apathy [29,53]. Consistent with Marin’s work, these criteria defined apathy as a lack of motivation, underpinned by three domains of reduced goal-directed behaviour; reduced goal-directed cognition, and reduced ‘emotional concomitants’ of goal-directed behaviour [53]. Symptoms of the behavioural domain are reduced effort and a need for prompting from others to initiate activity, symptoms of the cognitive domain are lack of interest in new knowledge and experience and lack of concern about the self, and symptoms of the emotional domain are unchanging affect and reduced emotional reactivity. At least one symptom from each domain was required to meet the criteria, despite the authors noting that the presence of an emotional domain of apathy was contested and uncertain [29]. For the first time, it was proposed that these symptoms must be enduring (for at least four weeks), frequent (for most of the day), and have a negative effect on the individual, through causing significant distress or impaired social or everyday functioning. As with Marin’s definition, it was also made clear that symptoms should not be explained by lack of awareness or physical substances. However, it was deemed unnecessary to make a distinction between apathy as a ‘syndrome’ or a symptom of a disorder such as dementia, as aetiology is irrelevant to our understanding of what constitutes apathy [29]. Therefore, in contrast to Marin’s definition, symptoms due to cognitive impairment are not excluded from this characterisation of apathy.

It is worth noting that these criteria were developed by just a few clinicians, without a formal development process, based on limited literature regarding what constituted apathy. The criteria were supported however by a study that used this to identify the presence of apathy in 37% of people with AD whilst, as expected, no healthy controls met the apathy criteria [53].

2.3.2. Robert et al 2009 diagnostic criteria

To overcome the limitations of criteria developed by just one research group at one location, an expert working group was formed to create new criteria for apathy [33]. This group was composed of experts largely from the Association Française de Psychiatrie Biologique (The French Association of Biological Psychiatry), the European Psychiatric Association, and the European Alzheimer’s Disease Consortium. Experts were all individually sent various drafts of the criteria, then met to discuss the developed criteria at a consensus meeting. These new criteria also defined apathy as a loss of or reduction in motivation, compared to the individual’s previous motivation. Whilst the three domains were kept the same, the cognitive domain description ‘loss of

interest' was replaced by 'lack of curiosity', to emphasise its cognitive nature, avoiding conflation with emotional associations of the word 'interest' [33]. It was also specified that for each domain, problems could be observable in reduced responsiveness to external prompts, or reduced self-initiation [33]. For example, reduced goal-directed behaviour could be indicated by reduced initiation of tasks and conversation (i.e. internally initiated), or reduced responsiveness in conversation and participation (i.e. externally initiated). This contrasts with Levy and Dubois' [32] definition of apathy as reduced self-initiated activity only.

In these criteria, deficits in goal-directed behaviour encompass both non-routine and everyday activities and behaviours [33]. Exclusions were expanded on, as, in addition to diminished consciousness and substances, it was stipulated that the symptoms should not be entirely attributable to motor or sensory impairments. However, identifying whether symptoms of apathy are solely due to sensory or motor impairment may not be straight forward, as physical impairments are likely to interact with various mechanisms in and pathways to apathy [30,54].

Perhaps the biggest change in the criteria was that symptoms from just two of the three domains were necessary to meet the criteria for apathy. Apathy was originally defined as a simultaneous impairment in all three domains [55], so these criteria reflects a significant departure from apathy's original clinical characterisation. Though this is in keeping with other diagnostic criteria formats, in which a given number of listed symptoms are required for a diagnosis, for example diagnosis of depression in the Diagnostic and Statistical Manual of Mental Disorders (DSM) [6], and ICD, this requirement is arbitrary. Despite this, the change and its clinical relevance and implications were not discussed. A study of 319 people with AD found that an additional 30% of people met criteria for apathy when just two of the three domains were required to be present [53]. Thus, it is clear that this change has important implications for our understanding of apathy and its prevalence, and could have important consequences for people living with NCD, for example by determining their inclusion in research programmes with potentially helpful therapies.

2.3.2.1. An Introspective Dimension?

The inclusion of a new 'introspective' dimension was considered in the 2009 criteria [33]. Introspection refers to insight or self-awareness, and is often impaired in people with NCD. This impaired insight, sometimes termed 'anosognosia', can refer to reduced awareness of one's own abilities and capabilities in everyday functioning and cognition, mood or behaviour [56]. Lack of insight has been shown to be closely related to apathy, leading the authors to consider its inclusion as an aspect of apathy [33].

Lack of insight towards apathy symptoms and cognitive ability has been found to be higher in people with AD and apathy, compared to people with AD without apathy [53]. However, in this study, the no-apathy group had depression, so rather than indicate apathy is associated with lack of insight, this finding could be explained by elevated insight or lack of positive bias typically found in people with depression [57]. Nevertheless, lack of insight for cognition, though not for psychological and behavioural problems, has been found to be associated with increased apathy in AD [58]. However, no relationship between lack of insight and apathy was found in people with MCI, [58], so findings may be limited to people with AD only. The relationship between lack of insight and apathy may suggest that lack of insight, at least towards cognitive difficulties, is a part of apathy in people with AD.

Considerations for and uncertainty around the inclusion of this dimension were noted by Robert and colleagues [33], however, as the individual experts' opinions were not reported, it is unclear whether this domain was specifically discussed or how much support the inclusion of this concept received. Although lack of insight has not been included in the criteria for apathy, it has been incorporated into some apathy measures [e.g. 44,59]. Lack of insight was the only subscale of the Lille Apathy Rating Scale [LARS; 59] found to load onto its own factor. This could indicate it is indeed a distinct dimension of apathy, however it may also suggest that it is not part of the apathy construct, though it is related. This is somewhat echoed by Radakovic and Abrahams [60] who suggest that lack of insight occurs across the domains of apathy. However, the finding that this subscale had a particularly low correlation to the overall apathy score (r = .28), and with the other factors (the highest correlation with another factor was r = .13) suggests lack of insight should not be regarded a component of apathy.

An alternative explanation for the relationship between lack of insight and apathy is that cognitive decline results in both apathy and lack of insight, leading to an apparent overlap. However, a longitudinal study demonstrated that lack of insight in people with dementia predicted an increase in apathy, independent of cognition at baseline [61]. Rather than being a by-product of cognitive decline, lack of insight may result in apathy, as a lack of insight into ability to perform activities may result in less success if performing activities, leading to reduced performance and interest [61]. It seems lack of insight may be best seen as a causally relevant construct in the field of apathy, rather than part of the apathy construct itself, and further research could help further our understanding of how apathy occurs.

2.3.3. Robert et al 2018 diagnostic criteria

Whilst Robert and colleagues' [33] proposed criteria based on the consensus of an expert working group, they did not report the methods in sufficient detail

and the disclosure of the results was limited to the finalised criteria. This meant it was difficult to see how much agreement there was on the criteria. However, since then, the authors have developed new criteria, using thorough methods and detailed reporting of results [36]. A Delphi study, surveys and a consensus meeting were conducted with an international consensus group, largely composed of expert apathy and behavioural symptom researchers and healthcare professionals.

In line with Levy and Dubois' [32] aim to operationalise apathy in an objective observable manner, apathy was re-defined as a "quantitative reduction of goal-directed activity either in behavioural, cognitive, emotional or social dimensions in comparison to the patient's previous level of functioning in these areas" [36, p.73]. In contrast to Levy and Dubois' [32] definition however, it highlights that the quantitative observable reduction in activity is not only applicable to goal-directed behaviour, but to three different domains. In keeping the prior diagnostic criteria, one symptom must be present from at least two of the three domains. The domains themselves however, were substantially revised. A new domain of diminished social engagement was included, whilst previous behavioural and cognitive domains were merged into one domain.

2.3.3.1. Social Domain

Prior to this revision, symptoms of reduced social interaction were interspersed throughout the apathy criteria, as both a consequence of apathy (through causing impairment in social functioning), and as part of the behavioural, emotional, and cognitive domains, though indicators such as reduced participation in social events; reduced emotional reaction to others; or reduced curiosity in family affairs [33]. It was argued that recent evidence suggested that impaired social participation should be regarded a specific domain of apathy [36]. A distinct social domain was included, with indicators similar to those previously interspersed throughout the criteria: reduced initiative, participation and interest in social and leisure activities, reduced initiative or increased withdrawal from conversations and a preference for staying at home.

Reasons for the inclusion of social apathy as its own domain were not thoroughly discussed and the limited evidence cited by Robert and colleagues [36] is not sufficiently convincing. Evidence cited to support a distinct social domain of apathy was based on studies that assessed the proposed underlying neurological separation of social apathy, and the distinction of 'prosocial' apathy in experimental studies. The distinct brain areas that are argued to support a distinct dimension of social apathy were thought to be dedicated to processing the expected rewards and thus motivation of others [62], and processing self and others' rewards and pro-social decision-making in monkeys [63]. Whilst reward processing may indeed be relevant to apathy

(see section 3.3.2 for a discussion), these studies tell us little about the possibility and clinical relevance of a social dimension of apathy. The experimental study in healthy younger adults that was cited found that self-reported behavioural apathy was associated with reduced willingness to exert effort for ones' own rewards, whilst self-reported social apathy was associated with reduced willingness to exert effort to reward others [64]. This was argued to suggest that social apathy and behavioural apathy are distinct [36], and social apathy may be characterised by a lack of willingness to put in effort for others, i.e. impaired pro-social behaviour [64]. This characterisation contrasts with that of the criteria for social apathy that were subsequently created, which refers to a lack of social engagement, interest and interaction with others, rather than reduced willingness to help others or diminished altruism [36]. Lack of empathy, which may be linked to impaired pro-social behaviour, was categorised under the emotional domain in the new criteria [36]. Stuss and colleagues [24] also made reference to a 'social apathy', however they characterised this as impairment in both social and self-awareness, or drive to act for the self. This conflicts with Lockwood and colleague's [64] finding that motivation to act for others is dissociated from motivation to act in one's own self-interest. Radakovic and Abrahams [60] proposed that rather than being its own dimension, social apathy is the result of the accumulation of impairments across different domains of apathy, mainly emotional apathy, and self-awareness. These conflicting findings and varied characterisations highlight that further research is needed to understand how emotional apathy, in particular empathy, and pro-social behaviour, relate to social interaction and engagement.

Others have cited evidence regarding the structure of apathy measures to explore whether there is a distinct domain of social apathy (as in [40]). Exploratory and confirmatory factor analyses of the Apathy Motivation Index (AMI) have indicated that social, behavioural and emotional domains load onto separate factors in healthy populations [65], and a recent Confirmatory Factor Analysis (CFA) study of another apathy scale, has supported the separation of an 'executive-initiation' domain and a social-emotional domain (in contrast to an individual-emotional domain) [66]. These studies appear to support the inclusion of a separate social dimension in the apathy criteria.

However, these studies did not always reach satisfactory fit indices, with a Confirmatory Factor Index below the recommended .95 [65,66], and the CFA model [66] also failed to reach a Root Mean Square Error of Approximation above the recommended cut off of .06 [67], suggesting the models were of poor fit. Further analysis of the data led the authors to conclude that behavioural and social apathy domains formed just one type of apathy, whilst emotional type formed another type [65]. This was supported by their finding that these two types of apathy showed differential associations with other clinical factors [65]. These two studies were conducted in healthy adults, the

majority of whom were students, with a mean age of less than 34 years old in both studies, so its applicability to people living with NCD is unknown.

Other investigations of the structure of apathy measures have indicated no such social dimension. For example, in people with Parkinson's disease, the social subscale of the LARS loaded onto the same factor as lack of interest, reduced novelty seeking, and general motivation, aspects of apathy which were deemed to represent lack of curiosity [59], typically understood to be the cognitive domain [33]. However, a factor analytic study of the AES in the same population (people with Parkinson's disease) found one main apathy construct, and a separate factor relating to social aspects [68]. Thus, evidence for a distinct social dimension appears mixed, with little explanation for the differences in findings. Whilst social engagement and participation is certainly relevant to apathy [28], it is less clear whether it should be regarded a consequence of apathy, an aspect of apathy that exists across the different dimensions, or should be regarded its own distinct domain. Future research should consider the role of social apathy [40], and it is argued here, given that social apathy as a dimension is questionable, establishing relationships between individual indicators (rather than whole domains which may or may not be appropriately clustered) could shed light on the role of social apathy indicators and their relation to other symptoms.

2.3.3.2. Merging cognitive and behaviour dimensions

The cognitive and behavioural domains were merged for practical reasons; to keep a three dimensional aspect for the new apathy criteria [36]. This was also justified through the claim that cognitive and behavioural domains are difficult to distinguish clinically [36]. There is some support for the merging of these domains. It has been argued that initiative and interest symptoms, from behavioural and cognitive domain respectively, perpetuate each other, so should not be separated [69], and are indeed difficult to separate in practice [40]. Moreover, items of an initiative and interest scale were found to best be explained by a one factor model, in both people with and without dementia, indicating apathy symptoms did not cluster into cognitive and behavioural domains [69]. Nevertheless, the diagnostic and clinical implications of merging the behavioural and cognitive domains were not sufficiently discussed in the proposal for new apathy criteria. Lack of interest, attributed to the cognitive domain, may be a particularly important aspect of apathy, as it has been found to be the only domain that predicts conversion from MCI to AD [43]. If the cognitive domain has its own distinct clinical implications, it may be inappropriate to merge this with another domain.

Importantly, in the previously proposed diagnostic criteria, one cognitive and one behavioural symptom would have been sufficient to fulfil the proposed criteria for apathy. However, the merging of the two concepts along with the addition of the social domain meant that either a social or emotional element

of apathy is required to be present to meet the criteria. This change is particularly important as these emotional and social domains are the most uncertain, as described above. Emotional symptoms are the least reported symptoms of apathy [42], even in those who fulfilled the previous criteria [34], suggesting that this revision would exclude many people previously thought to experience clinically significant apathy. Despite this, the new criteria have been tested against the 2009 criteria and over 96% concordance was found, with a small number of participants no longer meeting the new criteria for apathy [41]. However, the agreement may have been inflated, as the same rater assessed the same participants using one set of criteria immediately after the other. Robert and colleagues [36] recognised that further work is needed to explore whether these criteria reflect the true structure of apathy.

2.3.4. Miller 2021 criteria for apathy

The criteria described thus far were all designed with a transdiagnostic view, and so were not specific to people living with NCD. The criteria for apathy were revised again, using a Delphi method, in an attempt to make the criteria more relevant to people with dementia [39]. These most recent criteria for apathy reflected something of a reversal of the 2018 criteria by Robert and colleagues [36]. The social domain was re-integrated into the other domains due to a lack of evidence for its distinctiveness, and the behavioural and cognitive domains were separated again, though justification for this was not provided. Some of the language was updated, in an attempt to be more patient-friendly, and this included referring to the behavioural domain as ‘initiative’ and the cognitive domain as ‘interest’. It was stated that indicators must not be the result of lack of opportunity or ability due to cognitive impairment. The authors recognised that the criteria may be an oversimplification of a heterogeneous construct, but highlighted the importance of identifying apathy to enable treatment and research to move forward [39].

2.3.5. Usefulness of criteria for multiple domains of apathy

Diagnostic criteria for apathy have been developed to improve the clinical identification of apathy, and enable research into treatments that could improve the lives of those experiencing apathy and the people who care for them [29,33]. Current criteria stipulate that symptoms from two of the three different domains must be experienced for apathy to be considered clinically relevant. However, justifications for this have not been provided, and given the disagreement regarding what domains apathy is composed of, it is clear more research is needed regarding the validity of these domains if these criteria continue to be used. Furthermore, the clinical relevance of experiencing symptoms from two domains, rather than one, or all three, requires investigation [36].

It is questionable whether it is necessary that diagnostic criteria should apply domains and stipulate that symptoms must be experienced across these domains, if the criteria's primary goal is to identify clinically important apathy. For example, the debate regarding whether there is a distinct social domain of apathy, or if this should be interspersed throughout the criteria, is irrelevant for understanding and treating the individual. Despite this, a Delphi study indicated that an expert working group viewed apathy diagnostic criteria as useful and important, particularly for research purposes, such as the selection of eligible participants to clinical trials, but also for clinical purposes, including helping carers understand apathy and put in place strategies to manage it [36]. However, this was the same working group that was tasked with developing new apathy criteria so responses are not representative of the views of all clinicians, whom may have declined to participate if they did not believe criteria to be useful. The diagnostic criteria thus far have been developed by experts in the field of apathy from clinical, research and industry backgrounds, and regulatory bodies [29,30,33,36], however no such working groups have had representation from people who experience apathy such as people with NCD, or their carers. Further research into what apathy is, and how clinically relevant apathy should be defined, should include the views of people living with NCD and their carers.

Furthermore, Stuss, Van Reekum and Murphy [24] criticised diagnostic criteria for apathy viewing it as a homogenous syndrome, and instead argued that there are different types of apathy. As such, individuals are not expected to present the same symptoms across multiple domains. This is argued to be because disruption in different brain pathways results in different types of apathy, characterised by different symptoms. These were not specifically described in detail but were categorised as the result of disruption to arousal processes, or to executive dysfunction [24]. These processes and further proposed types of apathy are discussed later in section 3.3.

2.4. Conclusions on apathy criteria

Various attempts to outline diagnostic criteria and frameworks for apathy have been made over the past thirty years [32,33,36,38,53,60]. None have been integrated into psychiatric diagnostic manuals such as the DSM-5 [6] or ICD-11 [7], which currently refer to apathy as a behavioural feature of dementia, indicating it is not a recognised diagnostic entity. In the development of the 2018 diagnostic criteria, almost half of those who responded did not agree that the changes made to the dimensions should occur [36], and when these criteria were revised and the social domain was removed, almost 60% felt it should be included as a distinct domain [39]. Frequent changes to apathy criteria and frameworks, and lack of universal agreement highlights that our understanding of what constitutes apathy is still developing. It has been noted that further research is needed to determine the structure of apathy [36].

Whilst apathy is broadly conceptualised as either a disorder of motivation or of diminished goal-directed behaviour in various domains [70], a clear consistent full characterisation of apathy and its indicators is lacking. The usefulness of categorising heterogeneous indicators of apathy into distinct groups of symptom domains of apathy is brought to question. Rather than focus on categorising symptoms of apathy, it may be more useful to aim to understand how each indicator of apathy relates to one another, and what may cause these.

Nevertheless, criteria and frameworks for apathy can be useful in distinguishing symptoms that occur in other disorders that are thought to overlap but ultimately differ from apathy. For example, negative symptoms of schizophrenia, abulia and akinesia are all similar to apathy, but are argued to differ in some important ways. Negative symptoms of schizophrenia are difficult to differentiate from apathy [33], however it has been argued that these represent a more complex set of problems, of which apathy is one [29]. Akinesia is an impairment in which the primary deficit is in movement, as opposed to motivation which is impaired in apathy [30]. However akinetic mutism refers to an impairment in speech and movement, and has been argued to be a very severe form of apathy [24], whilst psychic akinesia is thought to be akin to the auto-activation type [29]. Abulia is understood to be an impairment in transferring motivation to act into action, or a more severe form of motivational disorder, which differs from apathy only by degree of motivational impairment [29]. It has been argued that apathy exists on the lower end of a spectrum of motivational disorders, with abulia being more severe and akinetic mutism more severe still [55]. However, criteria thus far have not made this distinction and it is unclear whether categorically differentiating between these constructs has any clinical relevance. Perhaps one of the most debated and discussed disorders that shares symptoms and overlaps with or encompasses apathy is depression. The relationship between apathy and depression warrants careful consideration as it has important implications for treatments in people with NCD [27].

2.5. Apathy and Depression

2.5.1. The similarity between apathy and depression

Apathy is generally proposed to be distinct from depression, with a recent Delphi survey showing that 100% of experts felt that apathy and depression were distinct constructs [39]. However, apathy shows considerable conceptual overlap with depression, through shared indicators such as loss of interest, reduced activity and fatigue [22,71,72]. Whilst qualitative studies exploring depression or apathy in older adults or those with NCD are lacking, one semi-structured interview study found that adolescents with depression experience loss of emotion and loss of short term motivation similar to apathy [73]. This

overlap could be problematic for distinguishing apathy from depression, which has implications for diagnosis, and treatment of both apathy and depression [27], in particular for people with dementia, in which both apathy and depression are common [11]. Marin [30] argued that the reduced emotional reactivity and emotional blunting present in apathy is inherently incompatible with the increased emotional distress and dysphoria found in depression. A recent book about apathy similarly argued for the distinction between apathy and depression on the basis of apathy being characterised by emotional blunting, and depression being characterised by negative or fluctuating emotions, as well as depression but not apathy containing symptoms of distress, self-criticism and differential reactions to positive versus negative events [71]. Marin [30] further argued that whilst people with apathy are compliant and passive in their activities, people with depression actively avoid activities. Therefore, it was proposed that the loss of motivation that defines apathy is separable from the loss of motivation observed in depression [30]. This may suggest that apathy and depression should not consistently occur together.

2.5.2. Co-occurrence of Apathy and Depression

Apathy and depression can be observed independently [53,74]. In a study of 131 people with AD, apathy was notably more prevalent than depression, and as Marin predicted, apathy was not associated with dysphoria [75]. Therefore, despite overlap in their definitions and criteria, depression and apathy are not unavoidably linked. However, apathy and depression do frequently co-occur, both in people with and without cognitive impairment. One study found that in cognitively intact older adults, as many as 75% of people with clinical depression, indicated by DSM criteria, also had clinically relevant apathy [74]. Similarly, in people with AD, 68% of people with depression also met the criteria for apathy [42]. Starkstein et al. [53] found lower estimates both in older adults with depression, with whom under one third also met criteria for apathy, and in people with AD and depression, with whom 52% also had apathy. Similarly, they later found that 55% of participants with AD and major depression also had apathy [76]. These differences in prevalence of co-occurrence could be attributed to the diagnostic criteria of depression and apathy that were used. The study of older adults which found the greatest co-occurrence utilised a clinical cut off criteria on an apathy scale, rather than clinicians' judgement. Thus this may be the result of shared items in the scales used, rather than co-occurrence of separate diagnoses. Benoit and colleagues' criteria listed a greater number of indicators for depression whilst requiring fewer symptoms to be present than the criteria used by Starkstein et al [53], which could explain the lower estimates in the latter study. However these differences could also be explained by dementia severity, as Benoit and colleagues [42] only included participants with mild dementia, whilst the

average cognitive scores of the sample reported by Starkstein et al. [53] were lower. There is some support for this explanation, as the relationship between apathy and depression has been found to vary as a function of cognitive ability [21].

2.5.3. Evidence for the association and clustering of apathy & depression in cognitive impairment

Studies have indicated that the relationship between apathy and depression in those with intact cognitive ability are weak [65,74], whereas significant associations between apathy and depression are observed in people with MCI [77], or dementia [78]. For example, one study observed a moderate correlation between apathy and depression in people with Mini-Mental State examination (MMSE) scores of 20 or below (indicative of moderate cognitive impairment), but not in people with higher cognitive ability [21]. A factor analysis revealed that apathy loaded onto the same factor as depression in people with MCI and mild dementia, but loaded onto a different factor in moderate to severe dementia, indicating that the overlap between apathy and depression may lessen in more severe cognitive impairment [79]. This further supports the suggestion (in section 2.5.2) that the greater prevalence of co-occurring depression and apathy in people with mild dementia [42] compared to people with more severe cognitive impairment [53] may be due to the relationship between depression and apathy varying as a function of cognition. In other words, cognition may alter the relationship between apathy and depression, and this may not be linear, with apathy and depression showing a stronger relationship in the presence of mild cognitive impairment, and less in people with intact or severe cognitive impairment.

2.5.4. Differential associations between apathy & depression

The dissociation of apathy and depression has been argued for using their different patterns of prevalence across various stages of cognitive impairment, different neurophysiological associations, and different relationships with other factors, suggesting that they may be involved in different mechanisms and pathways and should be considered distinct [24,72]. Whilst the distinction between apathy and depression on this basis ignores complexities and contradictory evidence, and is therefore insufficient, this argument will be briefly discussed.

2.5.4.1. Differential neurobiology

Reviews of the overlap of apathy and depression have argued that they are distinct constructs as they are associated with overlapping but distinct brain pathways, neurotransmitters, and other biomarkers [72,80]. Whilst apathy is associated with impairments in frontal-subcortical circuits involving the anterior cingulate, depression has been typically associated with dysfunction of frontal-

striatal and subcortical circuits involving the limbic system [72,80]. It has also been reported that apathy and depression have different neurochemistry, with apathy being commonly associated with reduced acetylcholine, a key neurotransmitter that is reduced in AD, and depression associated with lower levels of serotonin, dopamine, and norepinephrine [72,80]. This may be further supported by the finding that depression is associated with increased acetylcholine (the reverse of what is found in apathy, suggesting their separation) [81,82]. However, this would suggest a negative association between apathy and depression, which is not found in practice. Furthermore, neurobiological associations with apathy vary and are proposed to differ depending on the different causes and types of apathy [83]. For example, it has recently been proposed that dopaminergic systems are also impaired in apathy, and this is related to reward mechanisms proposed to underly apathy [84]. The role of reward in apathy is discussed in section 3.3.2, but this serves to highlight that the differences in neurobiology underlying apathy and depression are not as clear as previous reviews have proposed.

2.5.4.2. Differential patterns of prevalence and relationship with dementia severity

It has been argued that the distinction between apathy and depression is supported by differences in prevalence at different stages of dementia, suggestive of different mechanisms [72]. For example, studies have suggested that apathy increases with dementia severity [12,85], whereas depression remains relatively stable or is less prevalent later on in the disease [79,86]. However, this difference in prevalence could be due to difficulties of measurement in later stages of dementia, with depression more likely to be underestimated due to fewer outward symptoms [72]. In the same way, it is also plausible that apathy may be increasingly overestimated as dementia progresses, due to the inability to discriminate between loss of motivation and functional ability, further artificially separating the associations of apathy and depression with dementia severity. In contrast, a recent systematic review and meta-analysis found that there were no differences in prevalence of apathy and depression at different stages of dementia or types of dementia, which the authors note may be due to previous studies assessing dementia severity with the MMSE, rather than a measure designed to assess stage of dementia [14].

Nevertheless, apathy and depression may be associated with different risk factors for dementia, as one study found that apathy, but not depression, is associated with greater risk of conversion from MCI to AD [87], suggesting that depression and apathy are associated with different factors of clinical relevance. A meta-analysis has supported this finding that apathy doubles the risk of conversion to dementia from MCI [20]. Though some variation in findings occurred, most studies indicated the risk was increased, and it was deemed there had been little publication bias. However, a meta-analysis of the

role of depression in the conversion to dementia from MCI revealed that depression also predicts conversion to dementia [88]. This study found a smaller effect, and more mixed results however. A handful of studies found that depression actually decreased the risk of conversion to dementia, and the study [89] that suggested that depression offered the biggest reduction in risk was the only study to use the DSM criteria for depression. It might be argued that the inclusion of apathy type items in the depression measures such as the Geriatric Depression Scale used in the studies that found a positive result affected the results of this meta-analysis. However, it appears that most studies assessed depression using the Neuropsychiatric Inventory, which is designed to assess dysphoria rather than the whole spectrum of symptoms of depression [90], so symptoms overlapping with apathy should not have been included. Nevertheless, the authors also commented on the high likelihood of publication bias of these studies, and so the risk of conversion from MCI to dementia in depression may actually be less than reported. Therefore, it appears that apathy does predict the risk of conversion to dementia more than depression does, perhaps indicating they are separate constructs with different clinical relevance in MCI and dementia.

2.5.4.3. Other differential associations

2.5.4.3.1.Executive functioning

As stated above, the argument that depression and apathy have different prevalence in different stages of dementia may be explained by the use of measures that assess cognitive function rather than dementia severity. This suggests that cognition, including executive function, which refers to higher order cognition [91], is differentially associated with depression and apathy. In particular, it has been argued that apathy but not depression is associated with cognition, in particular executive function [92], which has received some support from studies of people with dementia [93,94]. However, others have found no relationship between apathy and cognition [78]. It has been argued that differences in findings may be due to the lack of sensitivity of measures such as the MMSE that were developed for screening rather than comprehensive cognitive assessment, and using other measures, apathy is associated with cognition, but not depression when apathy is controlled for [75]. Depression and apathy may be associated with different types of executive function. In participants with amnestic MCI, apathy was associated with worse verbal fluency (measuring initiation), independently of depression, whereas depression was associated with worse set-shifting, independently of apathy [77]. However, a study of people with AD and MCI found that both apathy and depression were associated with poorer learning, though apathy but not depression was additionally associated with worse episodic memory [95]. The role of executive function in apathy is discussed further in section 3.3.3.

2.5.4.3.2.Lack of insight

The strong relationship between apathy and lack of insight has led to the proposal that it is part of the apathy construct [60]. In people with AD, lack of insight has been found to be significantly worse in people with apathy than people with depression [53]. In fact, people with AD with depression show preserved insight, termed ‘depressive realism’, as rather than holding a negative bias, they lack the ‘usual’ positive bias that is pervasive in people without depression [57]. This may be one way in which apathy and depression differ that warrants further exploration.

2.5.4.3.3.ADL

Activities of daily living (ADL) may also be differentially associated in apathy and depression. In people with dementia, apathy was found to be associated with worse ADL whilst depression was not [96]. However, this may vary depending on whether basic ADL (bADL, such as washing and dressing) or instrumental ADL (iADL, such as managing medication and looking after finances) is being assessed. For example in people with AD living in the community, those meeting the criteria for apathy had significantly worse bADL and iADL, regardless of presence of depression, compared to people without depression or apathy, whereas those meeting criteria for depression but not apathy had no significant difference in bADL, though did also have significantly worse iADL than people without depression or apathy [16]. This could suggest that apathy is distinct from depression, as it is associated with worse bADL whereas depression is not. However, there were no significant differences between those with depression without apathy and those with apathy but no depression in either type of ADL. Furthermore, Zahodne and Tremont [77] found that apathy was associated with significantly worse iADL even when controlling for depression, whereas depression was not.

The differences in findings could be explained by the differing content of measures of ADL. As well as being made up of basic and instrumental functioning, it has been argued that ADL can be separated into initiation, planning, and performance of ADL [97]. When assessed in this way, it was found that in people with questionable dementia, apathy was associated with all aspects of iADL, but was only associated with the initiation of iADL, in people with AD, whereas depression was not associated with any type of iADL in people with AD, though was associated with initiation of iADL in people with questionable dementia [98]. Therefore, the initiation of ADL in particular may underlie the difference in apathy and depressions’ association with ADL, though the effect of this may depend on cognitive status. Therefore it is important to assess the different aspects of ADL when assessing the relationship of ADL with apathy and depression.

2.5.5. The role of dysphoria

2.5.5.1. Diagnostic criteria for depression

Marin [30] proposed that apathy and depression are distinguishable due to the incompatibility of dysphoria and distress in depression, and lack of emotional reactivity in apathy, a claim which is still used to date [70,72]. Whilst, as Marin predicted, apathy may not be associated with dysphoria in people with AD [75], dysphoria is not the only key symptom of depression.

The associations between apathy and depression may be attributed to overlap in their criteria [29,72]. The DSM-V criteria for major depressive disorder requires at least one of the core symptoms of either dysphoria or emotional blunting, i.e. reduced interest or pleasure, should be present [6]. The presence of one of these two key symptoms is also required to meet criteria for depression in AD [99]. Similarly, the ICD-10 lists dysphoria and loss of interest, in addition to fatigue, as key symptoms, two of which must be present in mild to moderate depression [37].

Therefore, dysphoria is not necessary for depression, which can instead be characterised by symptoms also described as part of apathy: emotional blunting (or ‘anhedonia’) and loss of interest.

2.5.5.2. Depression without sadness

Depression not characterised by dysphoric symptoms has been termed 'depression without sadness’ [100], and may explain the association between apathy and depression [72].

2.5.5.2.1. Prevalence of depression without sadness

The prevalence of depression without sadness has been shown to be twice that of depression with sadness in older adults in non-clinical populations [101]. Furthermore, in a clinical sample of people with major depression generally, older adults reported fewer dysphoric symptoms, such as feelings of worthlessness, sadness, and thoughts about suicide, and more symptoms related to loss of interest, compared to adults under 60 [102]. Depression without sadness is also common in people with dementia. Low mood has been found to be present in just 24% of people with mild AD and clinical depression, whereas diminished positive affect was present in as many as 62.9% to 81.3% of this sample [42].

Forsell and colleagues [103] found that a ‘motivation’ factor of a depression measure almost entirely increased with disease severity, whilst the dysphoric factor showed an inverted-U relationship with disease severity, peaking in mild dementia, and this dysphoric factor was more prevalent than the motivational factor in the early stages, but not in moderate to severe dementia. This was somewhat supported by their later study of the individual symptoms from

DSM-IV criteria for depression, which found that the symptom of dysphoria did not have significantly different prevalence across dementia severity groups, whilst the symptom loss of interest was significantly greater in mild and moderate dementia, and were lowest in people with no or questionable dementia [104]. Though Benoit and colleagues [42] did not assess differences across disease status, they found dysphoria was much less prevalent in people with mild AD than diminished positive affect (one of which are required for a diagnosis using criteria developed for diagnosing depression in AD). This difference in findings could be due to the different symptoms assessed, with reduced positive affect appearing in earlier stages than motivational symptoms of loss of interest. Nevertheless, these studies serve to highlight that, in cognitive decline, depression may become increasingly characterised by symptoms that overlap with apathy, such as loss of interest, and less characterised by symptoms proposed to distinguish between apathy and depression. If people living with NCD experience depression without sadness, increasingly more than depression with dysphoric symptoms, the reliance on the presence of dysphoria to distinguish depression from apathy is insufficient. It has also been argued that the relative decrease in dysphoric symptoms in older adults and people living with NCD may be because they are more likely to 'deny' their dysphoric symptoms, rather than not experience them [72]. This highlights the difficulties of assessing internal symptoms with no outward indicators. Nevertheless, whether depression without sadness can and should be distinguished from apathy is still important to explore, particularly as this could affect potential treatment options [70].

2.5.5.2.2.Are depression without sadness and apathy the same?

In older people and people with dementia, depressive symptoms, assessed using DSM-III criteria, clustered into two factors: one was characterised by negative mood, whilst the other was characterised by impaired motivation [103]. In another study, distinct 'dysphoria' and 'withdrawal, apathy, and lack of vigour' factors were found in a different depression measure designed for older adults [105]. In another study of people with AD, the depression items similarly loaded on to two factors, one reflecting apathy, explaining 33% of the variance, and the other reflecting sadness and anxiety, explaining 9% of the variance [94]. Furthermore, the 'apathy' factor in the depression measure was significantly associated with a separate measure of apathy [94]. These studies may suggest that, for older adults and people living with NCD, depression characterised by dysphoric symptoms is qualitatively different from depressive symptoms related to motivation, and thus may reflect two types or domains of depression in these populations.

In the studies above, the factors that were deemed to reflect apathy were associated with significantly worse cognition, whilst the factors reflecting dysphoric symptoms were not [94,105]. This matches the pattern of supposed

differential associations of apathy and depression with cognition [92], discussed in section 2.5.4.3.1. Similarly, anhedonia is associated with substantial increased risk of dementia, whilst dysphoria is not [106], which mirrors risks associated with apathy and depression itself [87]. As discussed previously, these differential patterns of association have been argued to distinguish apathy and depression [72], however here it is demonstrated that they are also found in the different symptom clusters within depression. Therefore, a distinct cluster of symptoms of depression are difficult to separate from, and perhaps even encompass or equate to apathy. In contrast, a longitudinal study found that depression did not increase apathy overtime, causing the authors to conclude that “depression was neither necessary nor sufficient to produce apathy” [76, p.10], and that apathy must be distinct from depression. However, they did not assess depression with and without dysphoria and apathy did predict increase in depression overtime, which could suggest it is a milder or pre-clinical form of depression.

2.5.6. Conclusion on the relationship between apathy and depression

The nosological boundaries between apathy and depression are blurred, and their differences are much less distinct than originally proposed by Marin. The phenomenology of depression appears altered in people living with NCD, so the concepts of apathy and depression require careful reconsideration in the context of NCD [80]. This complexity is the result of overlapping definitions, criteria and indictors of apathy and depression. Whether depression without sadness can be distinguished from apathy relies on whether reduced interest and emotion are key indicators of apathy, and further research is required to ascertain this [29]. More research is needed to explore how depression and apathy relate to one another, and whether they should be regarded as the same construct [22,29]. A symptom-based analysis of apathy and depression in people living with NCD can help shed light on their overlap. Network analysis is a method via which this is possible, and is outlined in sections 5.3. The studies that have used this method thus far are discussed in 5.4, as the method first requires elaboration. First however, the proposed mechanisms of apathy will be discussed.

Chapter 3. Mechanisms of apathy

3.1. Introduction

"Apathy is at the crossroads of several theoretical frameworks" [70, p.16]

Numerous factors are associated with apathy that may complicate the assessment and our understanding of the nosology of apathy but could help explain how apathy occurs. Various explanations for apathy have been proposed, and whilst neurobiological and neurocognitive models and mechanisms specifically underlying apathy have been outlined in detail, social and environmental mechanisms have largely been ignored. Therefore, this chapter will present a discussion of the proposed neurobiological and neurocognitive mechanisms of apathy as well as social and environmental mechanisms proposed to underlie general neuropsychiatric symptoms, and other factors that may influence apathy.

3.2. Neurobiological Mechanisms

It is generally thought that neurological disorders, such as dementia, cause apathy via the changes in the brain that occur as part of the disease process [5,24,107]. If apathy is the result of dementia pathology, it would be expected that apathy increases as the disease progresses (as brain systems become more disrupted) and is associated with neurophysiological changes.

3.2.1. Disease Progression

It has been argued that people with more severe dementia are more likely to experience apathy [12,85]. Though, as described previously (in section 2.5.4.2), this may be due to the use of measures that assess cognition, rather than dementia severity, as this relationship is not supported by a meta-analysis of studies using a dedicated dementia severity assessment tool [14]. Nevertheless, apathy may be associated with progression to dementia or MCI. Evidence suggests there is an increased odds ratio (of between 4.5 to 10.0) of apathy in people with MCI, compared to healthy older adults with no cognitive complaints [108,109], and apathy is more prevalent in people with MCI compared to people with subjective cognitive decline (i.e. self-reported but not detectable on cognitive assessments) [110]. It has been hypothesised that apathy is an early indicator of a more aggressive form of dementia [76]. This is supported by longitudinal studies that have demonstrated that apathy predicts conversion from MCI to AD, increasing the risk of AD by around seven times [87,89]. Furthermore, a meta-analysis of longitudinal studies found that patients presenting at memory clinics with subjective cognitive decline, or MCI, and apathy have around a two-fold increased risk of dementia at follow-up, compared to those without apathy [20]. Therefore, apathy could be interpreted as an early indicator of NCD and disease progression, and

screening for apathy could help identify people at risk of more severe dementia or those who are more likely to progress to dementia. This could also help identify people who need additional support and could help target interventions earlier, highlighting the importance of understanding and assessing apathy, which is further discussed in Chapter 4.

3.2.2. Neurophysiological associations

Though findings vary across disorders and methods [111], apathy in dementia is associated with structural and functional disruption in frontal regions [112], including the anterior cingulate cortex and the orbitofrontal cortex [111,113], and has also been associated with impaired dopaminergic pathways [32]. Furthermore, though findings of pharmacological treatments have been mixed, a Cochrane systematic review has suggested that methylphenidate, typically used to treat attention deficit hyperactivity disorder and narcolepsy, may reduce apathy in people with AD and clinically significant apathy [25], which supports the suggestion of a neurochemical basis. However, whilst apathy may be associated with particular neural circuits involved in motivation and reward, this differs across disorders, and areas outside these circuits are also involved [112]. Furthermore, a comprehensive review of neuroimaging and biomarker studies concluded there were “controversial results regarding neurobiological bases of apathy in AD and MCI, suggesting these correlates remain yet to be clarified” [114, p.646].

3.2.3. Limitations of neurobiological explanation

It is not the aim of this thesis to explore the evidence for the neurobiological underpinnings of apathy in people living with NCD, as whilst apathy may be associated with some common neurobiological structures and functions, this does not tell us why. However, these studies serve to highlight that neurobiological associations are not causal or simple linear relationships. Similarly, Landes and colleagues [22] argue that the relationship between disease progression and apathy is not straightforward, and the presence of apathy at an early stage in cognitive impairment suggests apathy cannot be fully attributed to neurodegenerative processes. Despite this, neurobiological studies are typically interpreted causally: dementia-related brain changes are proposed to result in apathy, by disruption of neural mechanisms that underlie the cognitive processes associated with apathy (e.g. as in Lanctôt and Aleman [71]). It is important to recognise however, that these correlations do not necessarily indicate a causal mechanism.

An alternative hypothesis is that this relationship reflects ‘motivational reserve’ [115]. This theory posits that through practice and use, we build up ‘reserve’, i.e. the preservation of functional ability in the presence of neurodegeneration [116]. The concept of motivational reserve has received relatively little attention, but has received support from a longitudinal study of people with

MCI and AD, which demonstrated that those with better motivational reserve in mid-life were less likely to have AD, and those with better motivational reserve had better cognition at follow-up [115]. Therefore, rather than apathy being a simple marker for neurophysiological changes that will later manifest in cognitive decline, apathy may instead reflect poor motivational reserve, with causal implications for cognition and disease status. Understanding the role of apathy at the earlier stages of the disease will therefore be of particular importance. Furthermore, at the later stages of cognitive decline, it may be more difficult to distinguish apathy from inability, and so research into apathy in people with early NCD may be more beneficial to our understanding of apathy and its relation to neuropsychological changes in dementia.

3.2.4. Conclusion on neurobiological mechanisms

Relying on a purely neurobiological explanation of apathy is deterministic and problematic for the development of the field. Despite claims that our understanding of apathy can be progressed through further neuroimaging studies [112], it is argued here that the identification of neural correlates only enables superficial understanding of apathy. It does not itself help explain the mechanisms via which apathy occurs in people living with NCD. To understand the causes of apathy in NCD, and in turn theorize treatments and test whether interventions work, we need to understand the mechanisms via which apathy occurs and exerts its effects. Neurobiological explanations have however, in recent years, been paired with cognitive explanations of apathy, to offer greater explanation [107].

3.3. Neurocognitive Mechanisms

Neurocognitive explanations of apathy propose that neurodegeneration results in disruption of numerous cognitive processes that underpin motivation and goal-directed behaviour, from option generation and weighing up the various options, to the steps required to select, initiate and complete the behaviour. Disruption to any of these multiple processes could result in apathy, perhaps of different types, depending on the process disrupted [60,71,107]. Whilst there is much discussion surrounding which particular brain areas, neurotransmitters, systems and pathways are disrupted in each process [71,113], this reductionism is not the focus of this thesis, which will instead focus on the proposed processes involved.

3.3.1. Different types of apathy and their different underlying processes

3.3.1.1. Types of apathy and corresponding mechanisms

Levy and Dubois [32,83] have theorised three different processing pathways involved in goal-directed behaviour that, when disrupted, result in different

corresponding types of apathy: auto-activation, cognitive, and emotional-affective (later called ‘affective-motivational’). These have recently been refined to: auto-activation deficit, which can be split into ‘empty mind’ and ‘invigoration deficit’; ‘cognitive inertia’ and ‘amotivation’ [107]. A ‘dimensional framework of apathy’ [60] has also been proposed, in which types of apathy: initiation, executive, and emotional, are all theorised to be underpinned by specific cognitive processes or executive function. These different proposed types of apathy and their proposed descriptions and aetiology are presented together in Table 2.

Table 2. *Proposed mechanisms underlying different types of apathy*

Apathy type	Description / symptoms	Process via which apathy occurs	Brain areas	Citation
BEHAVIOURAL				
Auto-Activation Deficit	Reduced self-initiation of thoughts, activities, and emotions; Intact behavioural response to external prompts;	Dependent on whether this is the result of 'empty mind' or 'invigoration deficit' (see below)	Medial prefrontal cortex; Cognitive and limbic areas of the basal ganglia	[32,83,107]
Sub-process: Empty Mind	Intact non-goal directed (e.g. repetitive) behaviours; Emotional responsiveness is weak and short lived	Sub-process: Inability for signals to reach threshold for action	Multiple areas involved in the other processes theorized to be impaired in different types of apathy	[107]
Sub-process: Invigoration deficit		Sub-process: Impaired: Emotional and cognitive processing; Or integration of the two	Anterior-cingulate – dorsomedial circuit	[107]
Initiation apathy	Impaired initiation of action without prompting	Deficit in 'energization': Initiation; Maintenance	None specified	[60]
EMOTIONAL				
Amotivation / Affective-Motivational/ Emotional-affective	Reduced interest and concern; Impaired emotional recognition and empathy Assessed through items e.g. "does anything interest you; are you concerned about your condition; are you interested in learning new things?"	Impaired reward processing and decision-making, including: Association of emotions with behaviour; Processing of consequence; Reward consumption or processing	Orbital and medial prefrontal cortex areas; Orbito-ventromedial circuit; Limbic and associative areas of the basal ganglia	[32,83]
Emotional apathy	Affective flattening; Indifference; Emotional neutrality/blunting	Impaired behavioural or emotional self-regulation i.e. deficit in integrating emotional and social aspects of behaviour	Left insular cortex	[60]

Apathy type	Description / symptoms	Process via which apathy occurs	Brain areas	Citation
COGNITIVE				
Executive apathy	Lack of curiosity	Impaired executive functioning, particularly: Planning; Attention; Task selection; Maintenance	None specified	[60]
Cognitive Inertia	Reduced goal-directed behaviour; Problems with planning and organising goals; Slow to respond	Impaired executive functioning, particularly: Planning; Set-shifting; Working memory	Dorsolateral prefrontal cortex/ circuit; Cognitive areas of the basal ganglia	[32,83,107]

Note: The different proposed types of apathy have been presented under Marin's apathy domains to facilitate comparison, but do not necessarily correspond to these domains.

The proposed different types of apathy that are caused by different patterns of disruption to these processes do not necessarily agree and sometimes lack clarity. For example, Levy and Dubois describe a type of 'emotional-affective' apathy, but state that it is characterised by reduced interest and concern (typically thought to belong to the cognitive domain), which they argue is due to impaired reward processing and decision-making process i.e. the motivation underlying action [32]. In contrast, the dimensional apathy framework [60] describes a similar type of apathy, characterised by lack of curiosity, which they theorize is the result of impaired executive functioning such as planning, attention and task selection, that are related to executing action (rather than a decision regarding whether to act). This mechanism is similarly described by Levy and Dubois [32], who instead argue it results in cognitive inertia, a type of apathy that has lacked clarity regarding which symptoms it represents, but is stated to be characterised by reduced behaviour and problems with planning and organizing goals and slowness to respond. Furthermore, whilst Levy has described an auto-activation deficit which appears similar to initiation apathy in the dimensional apathy framework, it is stated that this initiation apathy differs from autoactivation deficit, which instead is argued to be the result of impaired motor initiation rather than underlying impairment in executive function [60].

3.3.1.1.1.Correspondence to apathy domains

Additionally, how these proposed types of apathy and processes should be interpreted alongside domain-based understandings of apathy is uncertain. Whilst on the surface the three types of apathy described above may appear to be similar to the domains of apathy defined by Marin [38], their correspondence is uncertain. The dimensional apathy framework drew attention to the uncertain role of lack of insight, i.e. diminished awareness of changes or impairments in mood, behaviour and everyday functioning [117], and its unclear associations with the three domains of apathy [60]. Furthermore, Marin's cognitive domain of apathy refers to reduced thought regarding the self and others, however, for Levy and Dubois [32] cognitive inertia is indicated by poor planning and organisation. Whilst these are not necessarily contradictory, it is unclear how these symptoms are linked. Additionally, whilst cognitive inertia described by Levy and Dubois, and executive apathy described in the dimensional apathy framework, both appear to be the direct result of executive dysfunction, Marin [38] excluded symptoms of apathy that are directly the result of NCD altogether. In particular, the description of executive apathy appears identical to executive dysfunction seen in NCD, so it is unclear how the two phenomena should be distinguished. Furthermore, emotional-affective apathy is said to be indicated by loss of interest, a symptom which instead featured in Marin's [38] cognitive domain. According to Levy [83], auto-activation apathy can result in reduced

cognition, emotion and behaviour. This type of apathy may therefore be more akin to what Marin characterised as apathy itself.

The dimensional apathy framework [60] and Levy and Dubois' models [32,83] further highlight the lack of certainty, detail and clarity in our understanding of the nosology of apathy.

3.3.1.1.2.Evidence for types of apathy

Levy and Dubois' apathy sub-types have been supported by the finding that distinct brain areas are, as predicted, differentially related to these apathy types [35]. Furthermore, in a novel experimental study, in which carers recorded participant's apathy over the course of one week, 'cognitive inertia' was related to executive dysfunction, whilst emotional-affective apathy was related to theory of mind, and auto-activation apathy were related to both dysfunctions [118]. This was argued to support the proposal that auto-activation apathy reflects a summation of emotional and executive dysfunction.

Factor analyses have indicated some support for the dimensional apathy framework, with apathy items loading onto four factors that were condensable into three factors, to reflect the three dimensions described [119]. However, each factor only accounted for a small amount of variance in apathy. Furthermore, a latent class analysis of a measure based on these proposed types of apathy found three distinct types of apathy in people with dementia [45], but these did not match the expected domains. The most common apathy type was labelled executive-initiative apathy despite 37% participants in this group experiencing symptoms from all domains. The remaining two clusters of apathy types were characterised not by domains but by severity: Global apathy, characterised by more symptoms, with symptoms almost exclusively from all three domains (executive, initiation, and emotional) present, and minimal apathy, characterised by fewer symptoms, from fewer domains [45]. The above studies indicate that the types of apathy described by the dimensional apathy framework may not be the most appropriate way to separate and define apathy.

3.3.1.1.3.Summary of mechanisms based on apathy types

The processes outlined above can be summarized by two main mechanisms: the process of creating motivation to act (reward and effort based decision-making), and the processes involved in planning and taking action once motivation is established (executive functioning) [71]. Husain and colleagues have similarly developed neurocognitive explanations of apathy, though not of specific types, outlining the various cognitive processes that underly goal-directed behaviour, proposed to result in general apathy [28,113,120,121]. Due to the lack of clarity surrounding the proposed apathy types, this next

section will focus on these two types of processes, but not necessarily whether they are specific to the proposed different types of apathy.

3.3.2. Decision-making

To produce goal-directed behaviour, a decision of action must be taken. This requires options to be generated and selected (based on their relative value, including consideration of reward, effort, and risk), and learnt from, so that the individual is motivated to act in the future [121].

3.3.2.1. Option generation and attention

Decision-making requires possible options to first be generated, which requires attention and may therefore itself be inseparable from executive functioning [121]. It has been proposed that impairments in option generation could result in apathy. Too many options may be generated, leading to difficulties selecting a single choice or action, resulting in the initiation of none [121]. Alternatively, apathy could be the result of not generating sufficient options: a seeming lack of options to choose from results in the inability to initiate action [120,121].

If apathy was caused by a deficit in option generation alone, then it would be possible for the effects of apathy to be reversed when a task is prompted or options provided, removing the need for option generation. This reversal effect has been reported by some, for example, in Radakovic and Abraham's [60] characterisation of 'initiation-apathy', in which there is a primary deficit in initiation, which is not present when the individual is prompted. However, this does not appear to explain emotional, social or cognitive aspects of apathy. Furthermore, an experimental study concluded that apathy was not the result of an attention deficit impairing decision-making, as decision times do not differ between people with apathy and people without [122].

3.3.2.2. Option Valuation and Selection.

It is proposed that during the decision-making process, subjective values are attributed to the various possible options, based on the weighing up of different estimations of: subjective value of reward if obtained (expected reward); probability or risk; time until outcome (delay), and cost or effort to obtain it [84]. Individuals select options with the highest overall valuation, i.e. reward value relative to delay and cost or effort [84]. It has been proposed that apathy may be a disruption to any of these processes, causing people to choose the less active or effortful options. Indeed, apathy has been found to be associated with brain areas that are theorised to underlie option valuation, cost estimation and the integration of the two to inform decision-making [123].

The option valuation process of decision-making has been investigated in experimental designs in which participants are asked to produce an effortful

response, for example via a hand dynamometer, or repeated pressing of a button, to win a reward. In the former task, the effort required for each trial is typically shown on a screen as a percentage of the participant's maximum grip strength, which is established earlier in the task. Feedback may be displayed, that enables participants to see their performance (effort produced) in relation to a target (effort required). In these tasks, the reward, or maximum possible reward, available for each trial is typically shown on a screen, and often represents real financial reward. Participants are expected to display 'the discount factor', which refers to the reduced likelihood of exerting effort for increasingly higher effortful tasks [84]. It has been proposed that apathy may be the result of oversensitivity to the effort required to do a task, or an underestimation of or insensitivity to anticipated reward [113]. In this experimental model, participants with apathy are indeed less likely to produce effort for reward than those without apathy [122]. However, this does not indicate whether participants show effort-discounting in this task due to reduced sensitivity to expected reward, or over-sensitivity to the expected effort. The evidence that explores this will now be discussed.

3.3.2.2.1.Oversensitivity to effort or insensitivity to anticipated reward?

Anticipatory reward sensitivity, measured using self-report questionnaires, has been shown to be associated with both apathy and effort discounting [124,125]. Effort estimation and perceived task difficulty have also been shown to be associated with apathy [126]. Bonnelle and colleagues [127] attempted to distinguish between these two processes by developing a novel experimental task, in which the reward available and the effort required to obtain the reward varies across trials. Participants can be shown the reward available and effort required, then given a choice of whether to do the trial, or to skip it, and thus have to produce no effort, to obtain a lesser or no reward. The authors found that apathy was not associated with differences in reward sensitivity or effort discounting alone, but was associated with greater effort discounting in the presence of smaller rewards [127]. In other words, apathy was associated with less willingness to exert effort in more effortful tasks when the reward was minimal, but similar willingness to exert effort in more effortful tasks when the reward was higher. However, this study was based on a small number of healthy participants. Nevertheless, a more recent study found that people with apathy were less likely to produce effort for reward, regardless of the effort level required, suggesting that it is not simply an oversensitivity to effort [128]. This effect was not associated with dysphoria, and could not be explained by motor initiation. Moreover, people with Parkinson's disease who performed this task have a reduced likelihood of producing effort for reward, particularly if this reward was low, suggesting a role of reward sensitivity [122]. However, this effect was also itself associated with effort: people with apathy showed the greatest difference from those without apathy in terms of task acceptance when the reward was low, but also

particularly when the effort required was also low. This led the authors to conclude that low effort, everyday tasks which people with apathy are characterised by not doing, do not offer sufficient reward to perform [122].

3.3.2.2.2.*Oversensitivity to negative outcome?*

The above experimental methods assessed individuals' response to reward and effort, with no risk of failure. However, risk of negative outcome also is important in valuation of options. It has been hypothesised that probability estimation is affected in apathy, resulting in risk aversion, and thus lack of activity [121]. However, people with AD and MCI have been found to make worse, more risky, decisions in the Iowa gambling task [129]. Poor performance involving increased risk-taking on this task was associated with the action-initiation domain of apathy [129]. Therefore, the behavioural domain of apathy in particular may be explained by impaired ability to learn from consequences of previous actions, which will now be discussed.

3.3.2.3. Learning: sensitivity to experienced consequences and effort

Interacting with the consequences of actions and learning from these outcomes is essential to this learning process, so apathy may be the result of deficit in interacting with outcomes and integrating this with what was expected (prediction error) [84]. People with Parkinson's disease and apathy have been found to have significantly lower difference in brain activity to wins and losses in a monetary incentive task, compared with healthy controls and people with Parkinson's disease without apathy [130]. This diluted response was associated with subsequent reduced risk taking following an unexpected high win, compared to healthy controls and Parkinson's disease without apathy, who (as expected) took more risks following an unexpected high win. This suggests that apathy is associated with reduced sensitivity to positive and/or negative consequences, and this may impair the learning process of decision-making, which will now be discussed.

3.3.2.3.1.Subjective experience of rewards may be impaired in apathy

In addition to anticipated reward and estimation, the actual subjective experience of reward ('liking') may be impaired in apathy [120]. A study in which participants had to saccade to a target for a stated maximum reward found that people with Parkinson's disease and apathy showed no significant pupil dilation in response to reward, in contrast to healthy controls and people with Parkinson's disease without apathy [131]. Furthermore, though apathy was not associated with speed of saccade once initiated, apathy was associated with a reduced pupil dilation in response to reward. This reduced pupil dilatory response in apathy has been replicated elsewhere [128], suggesting that subjective experience (or 'liking') of reward is impaired in apathy. In contrast, another study found no association between apathy and reward responsiveness [132]. This difference could be explained by the latter

study using a self-report questionnaire to assess participants' theoretical level of responsiveness to reward e.g. level of excitement felt when winning something. These types of subjective generic questionnaires may be inappropriate methods to assess reward sensitivity however, as they may be conflating reward sensitivity with other constructs such as depression. Nevertheless, it has been argued elsewhere that apathy is not associated with the actual subjective experienced outcome [84].

3.3.2.3.2.Learning from positive and negative outcomes

Instead, it was argued that evidence from people with schizophrenia suggests that apathy is associated with inability to learn from positive, but not negative outcome [84]. In contrast, a study in which participants with Huntington's Disease were required to continue pressing a button to win a race which they were covertly designed to fail, found that people with apathy continued for longer than those without apathy, suggesting a lack of sensitivity to or failure to learn from negative outcomes in apathy [132]. These participants also completed a different monetary incentive task and performed worse after losses, but not wins, leading the authors to conclude that apathy is associated with insensitivity to negative but not positive consequences [132]. However, it is not clear whether this reflects impaired learning from or subjective experience of negative outcomes. This does however suggest that there may be disease-dependent differences in mechanisms underlying apathy.

It is of note that the impaired process of insensitivity to or failure to learn from negative outcomes that is described here could present a potential mechanism of impaired insight. In other words, it seems plausible that an impaired ability to learn from or recognise when things go wrong would result in lack of awareness of impaired functioning. Impaired insight is common in people living with NCD, and has also been proposed to result in apathy (as described in section 2.3.2.1), so future research would benefit from understanding how these concepts relate to one another specifically in people living with NCD.

3.3.2.3.3.Subjective experience of effort is not impaired in apathy

Using a similar task to Bonnelle and colleagues, Hartmann and colleagues [124] found that people with apathy did show greater effort discounting, which contrasts with the aforementioned findings [127]. Furthermore, they found this was not associated with participants' subjective experience of effort, i.e. their perceived effort, measured by rating of the effort levels after the effort task. This sensitivity to effort was also not associated with apathy, suggesting effort sensitivity cannot explain greater effort discounting in apathy. This was supported by findings of Le Heron, Plant and colleagues [122] who similarly found that there was no difference in the post-task effort ratings of those with

and without apathy, despite differences in effort discounting. These tasks suggest that subjective experience of effort is not affected in apathy.

3.3.3. Executive function

Once an option is selected, a plan for action needs to be made, then implemented, which requires various executive functions, that could be disrupted in apathy [111]. Executive function can be defined as an umbrella term for higher order cognitive processes involved in decision-making and goal-directed behaviour [91]. Therefore, it is unsurprising that executive dysfunctions, such as impairments in initiating, maintenance, planning, set-shifting, and memory are proposed to result in apathy in people living with NCD, for whom these processes are disrupted as part of their condition [32,119].

Nevertheless, though the dimensional apathy framework [60] and Levy and Dubois [32] cite evidence from neuroimaging and experimental studies, findings for the relationship between executive function and apathy have been mixed in people living with NCD. For example, whilst one study found that attention, inhibition, and set-shifting predicted apathy in people with dementia [133] another found only initiation, assessed by verbal fluency, consistently predicted apathy in people with AD and MCI [134]. This may be explained by the different definitions of apathy used, with only the latter study using clinical cut off criteria, perhaps suggesting that executive function can predict mild, but not more severe apathy. These mixed findings may also be explained by the inappropriate investigation of apathy as a homogenous construct.

In a study that assessed the different domains of apathy separately, people with AD and apathy were found to have less attentional bias for social images compared to people with AD without apathy, and this was associated with the emotional domain of apathy [135]. In another study, performance in an experimental multitasking test (which required various executive functions such as planning, organisation, memory and re-evaluation of performance) was found to be strongly associated with the behavioural domain of apathy, but not emotional or cognitive domains [136]. Interestingly, this finding contradicts the dimensional apathy framework's expected involvement of planning and organisation in cognitive apathy. This may be because the interpretation of the cognitive domain has changed overtime [40], so the type of cognitive apathy proposed by the dimensional apathy framework may not be the same as that measured in this study. An alternative explanation was that this association was driven by prospective memory impairment [136], indicating that people may fail to initiate activities because they forget to, not because they have impaired initiation. The role of prospective memory in initiation of activity, and its potential role in apathy has been acknowledged elsewhere [107,137], though has received little attention in the field of apathy.

Its involvement has been further supported by a study that found a large and medium correlation between prospective memory and lack of initiative and lack of interest respectively, even when controlling for global cognition, working memory, processing speed and mood [138]. This has important treatment implications, as it suggests that prompting may be a sufficient intervention for initiation-apathy. Further research should assess the mechanisms behind initiation of activity in relation to prospective memory and apathy, as well as exploring the different relationships of different executive functions with different domains or indicators of apathy.

3.3.3.1. Initiation

One problem with assessing the relationship with executive function and apathy is that cognitive tests necessarily use various executive functions, making it difficult to determine exactly what it is that is being measured. For example, the relationship between initiation, measured by verbal fluency, and apathy in people with MCI was attenuated when controlling for psychomotor speed, which was thought to be indicative of motor initiation specifically [134]. Therefore, it is possible that the overlap between motor initiation and the initiation component of apathy may drive the overall association between apathy and executive function.

Once an action is planned and decided, it must be physically initiated, so this finding is consistent with proposed mechanism of apathy: disrupted initiation of thought (initiation related to executive function) and motor initiation of action [28,120,121]. Motor initiation and the executive function of initiation may be difficult to separate, however, other studies (described in section 3.3.2.3.1) have not supported the involvement of impaired motor initiation in apathy, as these showed that once a decision to complete the task had been made, there were no differences between people with and without apathy in motor responsiveness (assessed by speed of eye movement to a target) [128,131] and apathy was not associated with amount of effort produced [122].

The issue of conflation is additionally complicated by the overlap of some of the processes proposed to underly apathy with the description of apathy itself. Initiation is proposed to be an executive function that is impaired in people with apathy [32,60,121], whilst at the same time, reduced initiation is interpreted as a symptom or key indicator of apathy [33,36,39], so evidence linking apathy with initiation is perhaps not surprising. Future research should strive to investigate the relationships between particular components of apathy and different executive functions [22].

3.3.3.2. Maintenance

It is also proposed that apathy may result from disrupted task maintenance [113], however one study demonstrated that people with worse apathy are actually more persistent in a task, despite cues indicating task failure [132].

This could suggest that apathy is underpinned by a problem in stopping a task, or 'set-shifting', which also forms an important part of task maintenance [121]. However, this could also be explained by lack of processing or sensitivity to consequence, as discussed above in section 3.3.2.3.

3.3.4. Limitations of neurocognitive models

These models of apathy propose a disruption to internal neurocognitive processes, caused by neuropathology. In particular, apathy is proposed to be the result of a disruption in at least one of the cognitive processes of attention, initiation, reward processing, decision-making, and task maintenance, underpinned by different specific neurobiological changes. Levy even claims that "apathy is only an output syndrome of several different underlying cognitive and neural mechanisms" [107, p. 14].

The International Classification of Functioning, Disability and Health (ICF) is a model of disability and functioning in which impairments, activity, participation, and environmental and individual contextual factors interact [139]. In this model, impairments refer to a problem with physiological or psychological function, or structure of the body, whilst disability refers to impaired functioning, which includes restrictions on or difficulties in participation and activity, including environmental factors, in addition to impairments [139]. It is argued here that within this model, apathy can be understood to be a disability, as it operates at the level of functioning, and when mapped onto this model, it is clear the aforementioned models of apathy only focus on the level of impairment, i.e. neuropathology and executive dysfunction [32,83,119], and offer little insight into levels of activity, participation and context that are important in understanding disability level phenomena [139]. It is important to look beyond neuropathology and executive dysfunction to gain a better understanding of apathy.

Furthermore, much of the evidence for neurocognitive models discussed above were conducted in healthy populations, and those in clinical populations have largely investigated people with schizophrenia or Parkinson's disease. These processes are likely to be additionally disrupted in people living with NCD. For example, learning will be affected by poor memory; option generation affected by impaired executive function; and action and effort will be disrupted by decline in functional ability. Therefore, these processes in the context of NCD deserve special consideration.

3.3.5. Conclusion on neurocognitive models

Apathy may be the result of disruption to various processes in decision-making and other executive functions. In particular, one mechanism that has received support is that apathy may be the result of a reduced willingness to exert effort for lower-level rewards, indicating reduced anticipatory reward

sensitivity. The evidence discussed also suggests that despite being a plausible mechanism for apathy, it appears the sensitivity to expected effort and experience of effort remains intact. Other proposed mechanisms have mixed evidence, typically from different populations.

The aforementioned neurocognitive models of apathy are plausible and useful to our understanding of apathy. However, they were not developed specifically for people with dementia or MCI, so it is unclear how these dimensions and mechanisms apply to people living with NCD. Furthermore, these models consider apathy to be an internal process that happens in isolation, at the level of impairment. Neurocognitive models fail to integrate or thoroughly consider processes and explanations at the level of activity, participation and context that may be important to our understanding of apathy. Existing understandings of the roles of activity, participation and context in apathy will now be discussed.

3.4. Activity and Participation: Activities of Daily Living

Activities of Daily Living (ADLs) are a key aspect of everyday functioning, and encompass basic ADLs (bADLs) such as washing and dressing, or instrumental ADLs (iADLs) such as managing medication and looking after finances [140]. Within the ICF, ADLs may be understood as operating at the level of ‘activity’ and ‘participation’, as this refers to individuals’ difficulties executing a task and participation in everyday life, rather than simply impaired processes underlying this [139].

As noted in section 2.5.4.3.3, both bADL and iADL have been found to be associated with apathy in people living with NCD [16,78,141–144]. Successful independent ADL performance requires initiation of behaviours without prompting [140]. Reduced motivation and impaired self-initiation may result in less frequent ADL performance, subsequently leading to worse performance, suggesting that apathy causes impairment in ADL [77]. This is consistent with criteria for apathy which posit that ‘clinical’ apathy should cause impairment in functioning [33,36,39]. However, impairments in ADL may conversely lead to apathy. As everyday activities become less achievable for people with dementia, they may become frustrated and less likely to start these activities independently [22], resulting in apathy. This has been supported by a study in healthy older adults, which found that self-efficacy was associated with apathy and this was partially mediated by subjective task demand [126]. In contrast, a five year longitudinal study found that ADL did not prospectively predict progression of apathy [145]. However, this study of healthy older adults may not apply to people living with NCD, whose impairments in ADL will be pronounced, and the study had a relatively small sample size, so may have had insufficient power to detect a relationship. In contrast, qualitative interviews with people with Parkinson’s disease indicated that difficulty with

performing previously enjoyable activities resulted in loss of interest and motivation [146]. Loss of interest was seen as a coping mechanism to avoid lamenting the activities that could no longer be performed successfully [146]. Furthermore, in a secondary analysis of a qualitative study with 130 people with mild to moderate dementia, it was found though most people described difficulty initiating ADL, this was often not attributed to difficulties in the performance of ADL [147], suggesting that initiation deficits in apathy may not lead to difficulties in ADL. This emphasises the value of qualitative methods in offering new insights and understandings into the mechanisms behind apathy.

However, the relationship between ADL and apathy is unlikely to be a simple unidirectional effect, and the relationship may be dependent on other factors. For example, one study found that in people with dementia, the relationship between apathy and bADL only remained in people with worse cognitive impairment, whilst the relationship between iADL and apathy was stronger in those with higher cognitive ability [21]. Moreover, in women with AD, it was found that only apathy which was both frequent and enduring was sufficient to predict more rapid decline in iADL [148].

Taken together, these findings indicate an important link between apathy and ADL, however the relationship is complicated by the type and level of both apathy and ADLs, and cognitive ability, so further research is needed to explore these interactions [22].

3.4.1. Age, Physical Impairment and Frailty

ADL is also closely linked to physical impairment. Unlike other neuropsychiatric symptoms in dementia, apathy is associated with older age [90]. This relationship may occur due to the inappropriate conflation of apathy with age-related reduced activity due to loss of mobility and function, caused by frailty and physical or sensory impairments, resulting in the appearance of increased apathy in older age, rather than the actual experience of it [30,54]. If the relationship between age and apathy was entirely attributed to reduced activity associated with age-related impairments, then we would expect to see a relationship between these age-related physical and cognitive impairments and the behavioural domain of apathy (reflecting activity). A longitudinal study of healthy older adults found an increase in overall apathy was driven by changes in the cognitive and emotional domains, whilst behavioural indicators of apathy remained relatively stable [145]. This is contrary to the hypothesis that age-related changes cause apathy via reduced goal-directed activity. Furthermore, whilst everyday functioning was associated with apathy at baseline, physical health was not associated with apathy, and both everyday functioning and executive function were not significantly associated with change in apathy overtime. These results suggest that the relationship between apathy and age is not due to impaired physical health or functioning.

However, the measure of physical health used in this study was crude, and the follow-up time of five years may have been insufficient to see the effects of any age-related changes occuring in a sample in which age-related changes may have already taken place. Nevertheless, the increase in apathy domains of cognition and emotion overtime and their relationship to age and age-related changes of health, everyday functioning, and executive function warrant further investigation. Sensory impairment was not measured in this study, and may be an important factor connecting age and apathy [30]. A large cohort study of people with dementia has found that neuropsychiatric symptoms were more likely to be present in people with hearing or visual impairment [149]. However, this study did not assess the relationship specifically with apathy. This possible mechanism also warrants further exploration.

3.4.1.1. Frailty

Frailty can be considered a syndrome of age-related physiological impairment, causing vulnerability to comorbidities and poor health outcomes [150]. Frailty measured by walking speed has been found to be associated with apathy in older adults [74]. However, walk speed tests involve initiation of movement, so the measure may be necessarily conflated with apathy. Alternative measures of frailty should be used to explore this relationship further. Few studies have assessed the relationship between apathy and frailty using a comprehensive frailty assessment [151], although those that have found that apathy is associated with increased frailty and concluded that apathy may result in frailty through inactivity [151,152]. However, these studies used a repurposed measure of depression to assess apathy, so more studies are needed to assess this relationship using a comprehensive measure of apathy and in particular, looking at the differential associations with the different specific indicators of apathy [151,152].

3.5. Context: Personal and Environmental Factors

The role of personal and environmental factors, is highly significant in any health condition [139], however they have received relatively little attention in the study of apathy thus far [153]. There are two main theoretical frameworks that describe the process via which general neuropsychiatric symptoms occur beyond explanations at the level of impairment: the unmet needs model, and the environmental vulnerability model [154], however these require consideration and elaboration in the context of apathy.

3.5.1. Unmet needs model

The unmet needs models posit that dementia creates a greater gap between an individuals' needs and the fulfilment of these, as it leads to a decreased ability to communicate needs, and greater needs due to a decreased

functional ability [154]. In this model, neuropsychiatric symptoms are theorised to be an attempt of the person living with NCD to meet or communicate the need. For example, if the need is for stimulation, then an individual may show repetitive behaviours in an attempt at achieving greater environmental stimulation, or vocalisations to in an attempt communicate this [154]. It is not immediately clear how relevant this explanation is to apathy, as unlike neuropsychiatric symptoms in which the individual increases their activity to communicate or meet a need, apathy is characterised by decrease in activity, interest and emotion. Furthermore, this model has largely focused on the experience of people with severe dementia, often within a care home setting. It is plausible however that apathy may occur in response to overstimulation, and therefore reflects the response to the unmet need for calmness and slower pace. This is supported by the finding from a qualitative study of people with schizophrenia that negative symptoms, similar to apathy, were understood to be a coping mechanism in response to positive symptoms that caused overstimulation [155].

In contrast, it has also been hypothesised that apathy may result from under-stimulation. It is thought that changes in circumstances such as retirement and care home admission cause a reduction in environmental stimuli that would otherwise encourage and reward behaviour, resulting in apathy [24,54]. Though evidence for this relationship is sparse, interventions involving activities that are tailored to care home residents' interests and abilities improve engagement and affect [156], which may indicate that decreased opportunities for activity result in apathy.

3.5.2. Environmental vulnerability models

Environmental vulnerability models propose that dementia results in greater vulnerability to negative stimuli and situations, so a response, characterised as a neuropsychiatric symptom, is triggered more easily in people with dementia that would not otherwise occur in cognitively able individuals with greater resilience [154]. One type of environmental vulnerability model is the progressively lowered stress threshold model, in which internal (such as fatigue or illness) and external stressors (such as a change in routine, or environment) result in 'anxious' behaviours as people near their stress thresholds, which becomes 'dysfunctional' as the stress threshold is surpassed, both of which can be characterised as neuropsychiatric symptoms [154,157,158]. As dementia progresses, individuals' thresholds at which they can cope with these stressors is increasingly lowered [158]. Support for the progressively lowered stress threshold model has been found through carer education programmes which have shown that applying this model as an intervention can reduce depression and unpredictable behaviours, and reduce carer burden, however less is known about its applicability to other neuropsychiatric symptoms [158]

3.6. A biopsychosocial model

Kales and colleagues [157] proposed a model of neuropsychiatric symptoms that attempted to integrate biological, psychological and social explanations, and this was expanded on recently in the context of apathy in dementia by Massimo and colleagues [153], who created a 'conceptual framework for apathy'. This model posits that there are direct neuropathological effects, as well as indirect stressors that may 'trigger' apathy [153,157]. As proposed by the neurobiological explanations of apathy discussed in section 3.2, impairment in the neurobiological systems that govern behaviours and emotions occurs, which is proposed to directly cause neuropsychiatric symptoms such as apathy. In addition, these neurobiological changes create a vulnerability to stressors, leading to apathy, consistent with the progressively lowered stress threshold model discussed in section 3.5.2. These stressors can be from individual, carer and environmental factors [153].

In the original model for general neuropsychiatric symptoms, individual factors refer to comorbidities or other conditions such as pain, medication side-effects, and other psychological illnesses, as well as personality and unmet needs [157], and predisposition. Carer factors include carer burden and emotional state, lack of education regarding dementia, mismatch of expectations and ability of the person with dementia and problems with communication [157]. Environmental factors include over or under stimulation, lack of activity, structure and routine, and safety issues [157]. These were expanded on specifically for apathy, with risk factors such as type and severity of dementia, genes and occurrence of other neuropsychiatric symptoms being added to the individual factors, and lack of the presence of a reward being added as an environmental factor [153]. Many of the proposed influencing factors have been discussed previously, including: presence of other neuropsychiatric symptoms (i.e. depression, in section 2.5), other comorbidities and reduced ability to interact with the environment (section 3.4.1), unmet needs and lack of stimulation (section 3.5.1), and availability of reward (section 3.3.2).

3.6.1. Limitations of the conceptual framework for apathy

Whilst the biopsychosocial 'conceptual framework for apathy' proposed by Massimo and colleagues [153] attempts a more comprehensive model of apathy, it has received relatively little attention, and fails to integrate the increasingly popular neurocognitive models of decision-making and reward valuation. Many of the factors described in the model are noted as theoretical and have limited evidence, due to a lack of studies particularly regarding carer and environmental factors, with some evidence relying on animal models [153]. This may be due to the difficulties of assessing contextual factors using quantitative methods. Qualitative methods could help explore possible

contextual mechanisms of apathy in people living with NCD. Further investigation and elaboration of this framework is needed, and the neurocognitive models discussed in section 3.3 should be integrated with this framework.

3.7. Conclusion on mechanisms of apathy

When conceptualised as a disorder, made up of various indicators such as reduced emotion, activity, interest, motivation, initiative and social interaction, rather than exclusively a direct symptom of NCD, apathy can be understood at the level of disability, and may be the result of several mechanisms. Most models of apathy thus far have focused on the level of impairment: the mechanisms of impaired executive function, including decision-making processes, caused by neuropathology. However, apathy may also be understood from the perspective of the ICF, to be the result of other mechanisms operating at the level of activity, participation and context [139]. Carers and the environment may fail to meet the needs of people living with NCD, and personal, carer and environmental factors may act as stressors. These stressors are numerous and have little evidence thus far, but might include activity limitations (impaired ADL) and age related physical impairment. These explanations have largely existed in parallel, with neurocognitive models at the level of impairment and social and environmental models focusing on participation and context. Many have recently argued for a transdiagnostic approach to apathy [71,120], however, it is clear that these processes will be uniquely disrupted in NCD. Whilst we may learn from similar phenomena in other disorders, and draw on transdiagnostic research especially when research specific to NCD is lacking, a model for apathy specifically for people living with NCD is needed. Though Massimo and colleagues attempted to develop an integrated conceptual framework for apathy in people with NCD, this requires further elaboration, investigation and integration.

Chapter 4. Systematic review

4.1. Introduction

The above chapters have highlighted what constitutes apathy and how it is experienced by people living with NCD is poorly understood and under-researched [28,153]. High quality measures of apathy suitable for this population are required to further our understanding of apathy, however, no gold-standard apathy measures have been established [159]. This chapter outlines the background, methods, results and discussion of a systematic review of apathy measures. This systematic review was conducted to address the first research aim of this thesis, which was to determine what measures of apathy are available and their quality for use with people living with NCD, as well as to inform the choice of apathy measure that was used later in this thesis. As this work reflects the available evidence at the time, the systematic review has not been updated. The work presented in this chapter has been published [160].

4.2. Background

4.2.1. Rationale

Measurement of apathy is complicated by a lack of enduring consensus regarding an apathy definition or criteria as outlined in Chapter 2, and the overlap of indicators of apathy with that of other disorders [12]. Nevertheless, many measures of apathy exist. Reviews of apathy measures have identified various self-reported, proxy-rated, and clinician interview based measurement tools, that may be generic or disease specific, and may form part of a larger scale, such as one assessing global dysfunction in dementia, or be a dedicated apathy measure [159,161]. Despite this, there is no gold-standard measure of apathy [153,159]. Whilst clinician interview may be the preferred method of identifying apathy, there is no formal diagnostic category for apathy, and quantifiable measurement scales are recommended for research into clinical interventions [27].

Clarke and colleagues [159] and Weiser and Garibaldi [162] have previously reviewed measures of apathy, but these were not systematic reviews, so important studies or measures may have been missed. Two systematic reviews of apathy measures have been published, which examined measures developed for people with neurodegenerative conditions [161] and people with dementia [163]. The first systematic review only included studies that assessed both validity and reliability of a measure. Whilst a scale should indeed be both valid and reliable, evidence for this can come from different studies. The latter review used very limited search criteria and failed to report when the search was conducted. Therefore, important studies regarding the

quality of apathy measures may have been missed in previous reviews. To assess the quality of the eligible studies, both aforementioned systematic reviews used the Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool [164]. QUADAS is designed specifically for studies that assess diagnostic accuracy, and contains reference to physical symptoms. Though both systematic reviews removed QUADAS criteria that was deemed irrelevant for assessing psychiatric symptoms, five of the remaining QUADAS criteria still referred to a 'reference standard'. A reference standard is used to test criterion validity, i.e. the ability of a measure to detect a construct compared to a gold standard [165]. QUADAS criteria that refers to a 'reference standard' should not be used to assess studies of other important measurement properties that do not require use of a comparator measure, such as reliability, structural validity and internal consistency. Furthermore, given that there is no agreed gold-standard criteria for apathy, demonstrated by the drastic change in proposed diagnostic criteria since the systematic reviews were published [33,36,39], it is not currently appropriate to assess criterion validity of apathy measures, as there is no single agreed 'reference standard'.

The Consensus-based Standards for the selection of health Measurement Instruments (COSMIN) programme of work has since published guidance for conducting and reporting systematic reviews of health measures, with methodological quality standards and measurement property quality criteria, that enables a systematic and standardized critical examination of all key measurement properties of measures [166], including content validity [167]. Content validity refers to whether the measure adequately reflects the intended construct, which includes whether its contents is: relevant to the construct; comprehensive (i.e. whether it includes all aspects of the measure); and comprehensible (i.e. whether it is understood as intended by the recipient) [167]. This will be dependent on the population it is being used with. Features of apathy such as diminished goal directed behaviour, may be shared with impairments related to NCD and ageing, such as poor planning and organization, poor mobility, pain, or sensory loss [30]. Apathy measures should avoid including items that could conflate apathy with changes related to NCD and ageing. This can be incorporated into the assessment of a measures' content validity. Content validity of measurement tools is argued to be the most important measurement property, as this affects the other properties, and ultimately, whether a measure is appropriate [167,168]. Given the lack of consensus regarding what constitutes apathy, it is of particular importance to assess content validity of the available measures of apathy.

COSMIN guidelines are compatible with PRISMA guidelines, and provide a modified Grading of Recommendations, Assessment, Development and Evaluation (GRADE) approach, specifically for use in the assessment of measures [166]. A COSMIN systematic review of the availability, evidence for,

and quality of apathy measures for people with NCD is needed to make evidence-based recommendations for measuring apathy in this population. Furthermore, a detailed analysis of how different measures conceptualise apathy will help map the different definitions and understandings of apathy.

4.2.2. Research Objectives

To identify measures of apathy available.

To assess and compare the quality of a variety of measurement properties and characteristics of the identified apathy measures and to analyse the quality of the evidence in people with NCD and older adults.

To assess and compare how the identified measures conceptualise apathy.

4.3. Methods

4.3.1. Protocol and registration

In accordance with the PRISMA guidelines, this systematic review protocol was registered with PROSPERO (ID: CRD42018094390 on 4th June 2018) and published prior to analysis [169].

4.3.2. Searching

4.3.2.1. Inclusion and Exclusion criteria

Inclusion criteria were:

1. Studies that aimed to develop, or to assess the measurement properties of an apathy measure. Studies reporting multi-domain measures or inventories that assess other concepts were only included if they reported the development of or assess the measurement properties of an apathy subscale that was not merged with other concepts.
2. Primary research full-text publications.
3. Participants living in the community. At least the majority (over 50%) of participants were required to live in the community (i.e. not residing in hospitals, nursing homes or care homes).
4. Participants aged 65 or above. Studies that included participants under the age of 65 were included if they had a median or mean age of 65 or older, or if results relevant to this review were reported separately for this age group. Where proxy-reports were used, the age of the participants (i.e. the person who is the focus of the measure) was the age that was required to meet the criteria. Age, rather than cognitive

status, of participants was used as an inclusion criteria, to identify a wider range of studies whose findings are relevant to people with NCD.

Exclusion criteria were:

1. Publications that were not full-text or did not report primary research, such as poster abstracts, review articles, commentaries, letters and editorials. Where these publication types reported the development or validation of an apathy measure, the corresponding author was contacted to determine if a full-text publication was available.
2. Studies of measures designed to assess apathy only towards specific topics or contexts, such as health behaviours (e.g. exercise, healthy eating), academic achievement, and job performance. These were deemed distinct from generalised apathy, such as that seen in cognitive impairment, which was the focus of this review.
3. Studies of apathy measures for disorders that are dependent on external circumstances, such as post-traumatic stress disorder, substance use disorders, and post-natal depression. Measures of apathy developed for these populations are likely to have been developed with these circumstances in mind (e.g. trauma, substance use, pregnancy), and may not be generalisable to people outside of the circumstance for which it was designed.
4. Studies of measures that intend to assess only a component of apathy, such as initiative or emotional responsiveness. It was decided that these measures, whilst possibly useful in the assessment of apathy, did not intend to measure the multifaceted concept of apathy.

No eligibility restrictions were placed on:

1. Year of publication.
2. Language of the publication. Where possible, articles published in a language other than English were screened for eligibility criteria by native speakers. However, due to resource limitations, it was not possible to obtain translations of all articles.
3. Development articles or administration guides for which another study of the measure met the criteria. Eligibility criteria were not applied to these texts, as they provided information on measurement characteristics, independent of population or context (e.g. content of items, response options, scoring etc.), that are required for assessment of content validity.

Where age or residential status of sub-samples (e.g. participants with MCI and participants with dementia) were described separately, yet results were reported together, these characteristics were calculated for the overall sample,

and checked against eligibility criteria. Where characteristics were reported separately, participants described as ‘controls’ were not included in this calculation, because the reference to ‘controls’ indicates that the study does not aim to assess validity or reliability of the measure for this group.

Where information was not sufficient to determine the study’s eligibility, the corresponding author was contacted for further details. Where the author could not provide further information, these studies were included, as the evidence rating criteria incorporates whether the study was conducted in the population of interest within the ‘indirectness’ scoring.

4.3.2.2. Information Sources

MEDLINE (In-Process, Other Non-Indexed Citations and 1946 onwards, via Ovid) EMBASE (1980 onwards, via Ovid), PsycINFO (1806 onwards, via Ovid) and CINAHL (1937 onwards, CINAHL Plus with Full Text) were searched using the specified search strategy on 18th April 2018. The search was re-run in the same databases on 6th May 2020. The reference lists of the included studies, and of any relevant review articles, were examined for relevant publications.

Once the list of included measures was finalised, publications describing the development and administration procedure of each measure were sought. These included full-text research articles, administration guides, and the measures themselves.

4.3.2.3. Search Strategy

The search strategy was developed (by myself, CB) and checked by an expert in the field (RH) and a subject librarian. The COSMIN search strategy for identifying research on the development, validity or reliability of health related outcome measures [170], was incorporated into the search strategy. Keyword searches were applied to titles and abstracts, and medical subject headings were applied where possible. The search strategy was first created for MEDLINE (via Ovid) (see Appendix 1), then the subject headings and syntax were adapted to suit the other databases. For the measurement section of the search strategy, the MEDLINE, EMBASE and CINAHL versions were used for the corresponding databases, and the search strategy for PsycINFO was adapted from the MEDLINE version. The search was limited to adult human participants (aged 18 or 19 and over, dependent on the database). No search limits were applied to study design, date, or language.

4.3.3. Screening and Selection

4.3.3.1. Data management

Data was retrieved, stored, de-duplicated, and sorted using Endnote X8. Data screening and data extraction tables were created and completed in Microsoft

Excel 2016. Duplicate studies were further identified by comparing studies of the same measures by their author and sample size where necessary, to prevent double counting and thus introduction of bias. The search was re-run using the process outlined by Bramer and Bain [171].

4.3.3.2. Selection process

The titles and abstracts of articles were screened to assess whether they met the eligibility criteria, and were included for further review if deemed eligible or if the eligibility was uncertain. All included full text articles were assessed against eligibility criteria (by CB), and a randomly selected 10% of articles was independently assessed by a second reviewer (Cathy Brewin [CBr]), as recommended in circumstances in which resources are limited [172]. Articles for which there was disagreement between reviewers were discussed.

The number of excluded articles were recorded at each stage, and reasons for exclusion of full text articles were recorded. Reviewers were not blinded to journals or author information due to resource limitations.

4.3.4. Data extraction and quality assessment

4.3.4.1. Data collection process

Data extraction was conducted (by CB) into a data extraction table (provided in Appendix 2). Where important information was missing, the corresponding authors were contacted, however it was unfeasible to contact all authors for all missing information given the time elapsed since many of the studies were conducted, and given the amount of missing information.

Ten English language publications were randomly chosen for data extraction checking, by assigning the eligible publications a number and using Excel random number generator. In addition, one non-English language publication was selected for data extraction checking based on availability of a translator. Data extraction of 19% of publications was checked by second reviewers (CBr; VvdW) and no mistakes were found.

4.3.4.2. Data items

For each study included in the review, data relating to study characteristics and methods, participant characteristics, and measurement characteristics and properties were extracted. Measurement properties included that of reliability (including internal consistency, measurement error and test-retest and inter-rater reliability), validity (including structural validity, hypothesis testing for construct validity, and content validity). ‘Hypothesis testing for construct validity’ refers to assessing construct validity through testing for convergent, divergent and known-groups validity, and from here will simply be referred to as ‘construct validity’. Measurement characteristics included stable aspects of the measure such as its content including items, recall period,

response options and scoring system, and other characteristics such as the conceptual model and definition of apathy applied, administration time, licensing and cost information, and mode of administration. This information was ascertained from the included studies, but also from additional sources that did not otherwise meet the inclusion criteria such as original development studies, administration manuals, copies of the measures, and correspondence with authors. Due to resource limitations, information that was only accessible through the purchasing of a licence not already held by the study team was not obtained.

4.3.5. Methods of synthesis and analysis

For each record matching the eligibility criteria, the quality of the study was assessed using the COSMIN risk of bias checklist [173] (see Appendix 3), and then the quality of the result was rated against COSMIN criteria for good measurement properties [166] (see Appendix 4). Publications that reported multiple variants of measures (for example, a study reporting the internal consistency of both a patient-reported and proxy-reported version of the measure) were considered distinct studies for the purposes of this review. Sub-group analyses were not rated separately to avoid double-counting results, however, if the result differed from the total analyses, then the opposing results were rated and reported separately. Risk of bias standards and quality criteria for good measurement properties are specific to each measurement property. Therefore, one publication about one measure can have numerous risk of bias and quality ratings for the different measurement properties it assessed.

Although the COSMIN risk of bias checklist and quality criteria were used to guide assessments, some additional decisions and criteria were required for this review. These are detailed in Appendix 5, and were based on literature, discussions with the review team, and Patient and Public Involvement (PPI) where appropriate. It is important to note that the quality of the methods and results were assessed in relation to the target population of this review (i.e. people with dementia and older adults), rather than the population of the published article. This is because measurement properties are dependent on the population in which they were tested, and may not be transferable to other populations [174].

In line with COSMIN guidelines, the original development study of each measure was included to ascertain the quality of measure development. Where the development study also examined other measurement properties, the methods and results pertaining to these measurement properties were not included in the review unless the study otherwise met the eligibility criteria.

4.3.5.1. Risk of bias in individual studies

Risk of bias in individual studies was examined using the COSMIN risk of bias checklist [173]. Over twenty percent (*N*=12) of risk of bias ratings were independently rated by second reviewers (SG and VvdW). Where there was disagreement, this was discussed between the two raters (SG and I; VvdW and I) and any remaining disagreements were discussed with the third rater for final say.

4.3.5.2. Analysis of results

The content validity of each measure was assessed using the COSMIN methodology [167]. This included evaluating the results of development and eligible content validity studies, and completing an independent reviewer opinion-based rating, which were informed by the content of the measures and administration manuals. The reviewer-opinion ratings of content validity of each measure was performed by myself, and informed by some general principles set by a PPI panel (Appendix 5). Where sufficient information about the measure could not be obtained, content validity rating by the reviewer was not possible.

Four measures deemed to have the best properties after the initial review of all measurement properties were rated again for aspects of content validity relevant to patients (comprehensibility and relevance) by three PPI members. The COSMIN criteria for good content validity that was relevant to patients was sent to the group along with the four measures. A meeting was then held during which the PPI members discussed the measures, going through item by item, considering whether they were appropriately worded and relevant to older adults and people with dementia, whilst keeping in mind the appropriateness and matching of response options. I facilitated this meeting whilst being careful not to provide their own ratings or views on the measures. PPI ratings of the measures were considered in the overall reviewer rating of content validity for these four measures.

Following the initial review of content validity, the results of each remaining measurement property were assessed against the corresponding COSMIN criteria for good measurement properties [166].

In line with COSMIN guidelines, measures with a formative model (i.e. a measure which is made up of items that are not interchangeable as they make up qualitatively different aspects of the overall construct) were not assessed for quality of evidence for structural validity or internal consistency, as measurement theories on which these investigations are based on are only relevant for reflective models (i.e. items are theoretically interchangeable and measure a single construct) [165], though results were still reported. Where it was uncertain what model the measure was based on, but studies of internal

consistency and structural validity had been conducted, a reflective model was assumed and the studies were assessed.

Over twenty percent of studies' results were rated against COSMIN criteria for good measurement properties independently by second raters (SG and VvdW). Using the same procedure as with the evaluation of risk of bias, disagreements were discussed between the two raters (CB and SG; CB and VvdW) and remaining disagreements were discussed with the third rater, to make the final decision.

4.3.5.3. Analysis of conceptualisation of apathy

One aim of this review was to identify how different measures conceptualise apathy. A content comparison table was produced [165], with headers based on the conceptual aspects of apathy from literature, and an 'other' column to allow description of remaining items that did not otherwise fit into the table headers.

4.3.5.4. Synthesis of results

All studies meeting the eligibility criteria were summarised using a narrative synthesis. For each measure, the measurement properties reported in the corresponding studies were summarised, and the quality of these synthesised results was assessed using the criteria for good measurement properties [166]. Results of the different variations of the apathy measures were pooled providing they did not contradict the overall result for the measure. The COSMIN modified GRADE approach was used to assess the quality of the cumulative evidence for content validity [167] and the remaining measurement properties [166] for each measure. COSMIN procedure for the recommendations of measures in systematic reviews (Prinsen et al., 2018) was used to guide the recommendations made.

4.3.5.5. Risk of bias across studies

Publication bias was not assessed as it is not recommended in systematic reviews regarding measurement tools and their properties, as there is no common system or database in place for registering these studies against which to check [166].

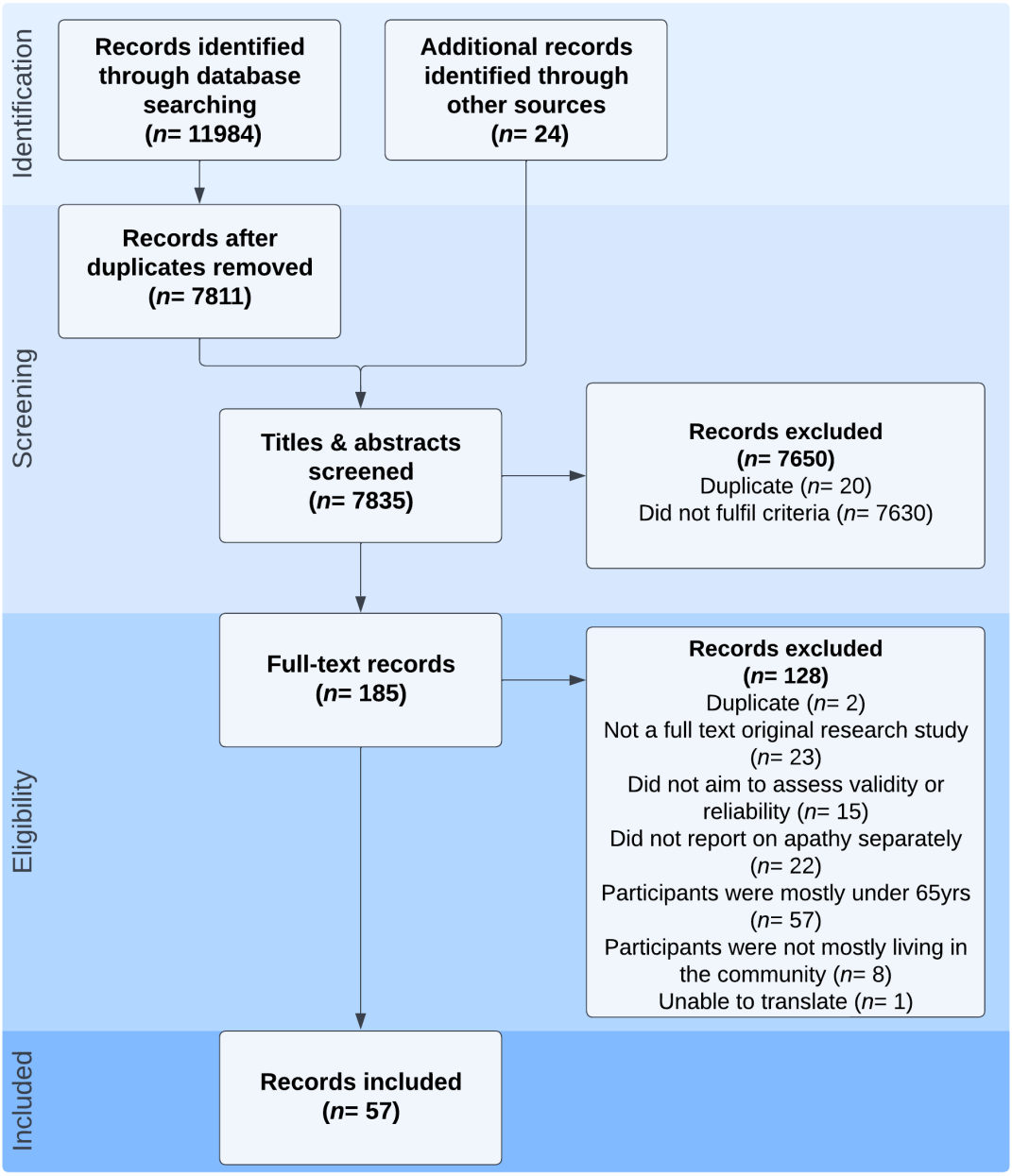
4.4. Results

4.4.1. Study selection

The initial search resulted in 9645 records from MEDLINE (*n*=2872); EMBASE (*n*=3530); CINAHL (*n*=784) and PsycInfo (*n*=2459). The re-executed search resulted in the identification of an additional 2339 records. Following removal of duplicates by searching in Endnote, there were 7811 records to screen, and a further 24 were identified through screening reference lists and visual

searching (Figure 1). Records removed in the initial screening phase (*n*=7650) were largely excluded due to not being a primary research study which aimed to develop or assess the validity or reliability of a measure. Many that did aim to develop or validate a measure were excluded as the measure was not of apathy. Some records were excluded at this phase due to the sample characteristics (age and residential status), and 20 additional records were identified as duplicates. 185 full-text records were assessed against the eligibility criteria.

Figure 1. PRISMA flow diagram



Fourteen texts that required translation were identified. Italian (*n*=2), German (*n*=3), 2 Japanese (*n*=2), Spanish (*n*=3), and Dutch (*n*=2) full-text records were screened by a reviewer for whom that language was their first to third language. French texts (*n*=2) were screened using a translation by Google Translate, as there were no French-speaking reviewers available to perform screening. Of these non-English-language publications, 4 records were initially included and 10 were excluded. Due to resource limitations, all non-English-language texts that met eligibility criteria were then fully translated using Google Translate, and the relevant extracted data was verified with a translator where possible (*n*=1). Following full-text screening, an additional 128 records were excluded. Reasons for exclusion were: the record was not a

full-text or original research study (*n*=23); the aim was not to assess validity or reliability (*n*=15); results were not reported regarding apathy as an independent construct (*n*=22); participants were mostly under 65 years old (*n*=57); participants were not mostly living in in the community (*n*=8); the record was a duplicate not previously identified (*n*=2). A full-text translation of a Japanese publication [79] that met eligibility criteria was not possible, so this was later excluded.

57 publications of 18 distinct measures (and 39 variations) were identified as meeting eligibility criteria (detailed in Appendix 6). The 18 measures identified were eleven subscales taken from global scales designed to assess other or less specific constructs (such as dementia severity, and neuropsychiatric symptoms), and seven dedicated measures of apathy. The global scales with an apathy subscale were: Alzheimer's Disease and Related Dementias mood scale (ADRD) [175,176]; Behavioural and Mood Disturbance Scale (BMDS) [177]; Behavioral Syndromes Scale for Dementia (BSSD) [178]; Dysexecutive Questionnaire (DEX) [179]; Frontal Systems Behavior Scale (FrSBe) [180]; Geriatric Depression Scale (GDS) [181,182]; Behavioural Rating Scale for Geriatric Inpatients (GIP) [183]; Index of Mental Decline (IMD) [184]; Key Behaviours Change Inventory (KBCI) [185]; Neuropsychiatric Inventory (NPI) [90]; and the Unified Parkinson's Disease Rating Scale (UPDRS) [186]. The apathy specific measures were: The Apathy Evaluation Scale (AES) [44]; Apathy Inventory (AI) [187]; Apathy Motivation Index (AMI) [65]; Starkstein Apathy Scale (AS) [188]; Dementia Apathy Interview and Rating (DAIR) [144]; Dimensional Apathy Scale (DAS) [119]; and the Lille Apathy Rating Scale (LARS) [59]. An overview of these measures and their characteristics is provided in Table 3.

Table 3. *Overview of measures and measure characteristics*

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
AD-RD [175,176]	1 [175]	a) Mood b) Moderate to severe AD c) Research or clinical	n/a	Interviewer-judgement, informed by observation and patient and carer interview	Unable to obtain for this review, suggesting not publicly available.	7 days	5	Items rated for frequency on Likert scale (1 to 5, all options described)
AES [44]	9 [44,46–48,50,189–192]	a) Apathy b) People with various clinical disorders or apathy, (with MMSE over 10 for patient reported version) c) Clinical	AES-C	Clinician-rated based on semi-structured interview with patient and observations. Bachelor level raters can conduct with 4-6 hours experience.	10 – 20 minutes; Measure and administration guide publicly available	4 weeks	18	Items rated on Likert scale (1 to 4; all options described), and quantifiable items rated 1 to 4 based 0, 1-2, 2-3, 3 or more quantifiable instances. Requires verbal or nonverbal evidence of intensity. Total score is sum of item scores. Range 18 to 72.
			AES-I	Informant-report via paper and pencil	10 – 20 minutes; Measure and administration guide publicly available	4 weeks	18	Likert scale (1 to 4; all options described). Total score is sum of item scores. Range 18 to 72.
			AES-I (16 item versions)	Informant-report via paper and pencil		4 weeks	16	Likert scale (1 to 4; all options described). Total score is sum of item scores. Range 18 to 64
			AES-S	Self-report via interview (recommended) or paper and pencil	10 – 20 minutes; Measure and administration guide publicly available	4 weeks	18	Likert scale (1 to 4; all options described). Total score is sum of item scores. Range 18 to 72.

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
			AES-12PD	Self-report		4 weeks	12	Likert scale (1 to 4; all options described). Total score is sum of item scores. Range 18 to 48.
AI [187]	3 [187,193,194]	a) Apathy b) Older adults with brain disorders c) Clinical	Clinician	Clinician opinion based on observations, and participant and informant answers to the AI when available.	At least 20 minutes; Publicly available.	Since beginning of the disease, last clinical assessment, or other defined time period e.g. last four weeks.	3	Likert scale (0 to 4; 3 options described) Total score is the sum of item scores. Range 0 to 12
			Informant	Informant-report via interview	Publicly available	Since beginning of the disease or an otherwise specified time point	3	Screening questions: (Yes=0 or No) with follow-up questions rated on Likert scale (Frequency: 1 to 4; Severity: 1 to 3; all options described) Item score is Frequency x Severity. Range 0 to 12. Total score is the sum of items scores. Range 0 to 36.
			Self	Self-report via interview	Publicly available	Since beginning of the disease or an otherwise specified time point	3	Screening questions: 0="Yes"; "No" with follow up question rated on a visual scale (1 to 12; end-points described). Total score is the sum of item scores. Range 0 to 36.
AMI [65]	1 [195]	a) Apathy b) Healthy Adults	Self	Self-report via paper & pencil		2 weeks	18	Likert Scale (0 to 4; all options described).

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
		c) Research						Total score is sum of item scores. Range 0 to 72.
AS [188]	8 [49,188,196–201]	a) Apathy b) Parkinson's Disease c) Clinical	Self	Self-report via interview	Publicly available	4 weeks	14	Likert Scale (0 to 3; all options described). Total score is sum of item scores. Range 0 to 42.
			Self (13 item version)	Self-report via interview	Publicly available	4 weeks	13	Likert scale: (0 to 3; all options described). Total score is sum of item scores. Range 0 to 39
			Informant	Informant report via interview	10 minutes; Publicly available	4 weeks	14	Likert Scale (0 to 3; all options described). Total score is sum of item scores. Range 0 to 42.
			'Home Care'	Self-report via paper and pencil.	-	4 weeks	11	Likert scale: (0 to 3; all options described). Total score is sum of item scores. Range 0 to 33
BMDS [177]	1 [177]	a) Neuropsychiatric symptoms (behaviour & mood disturbances) b) Dementia c) Research	n/a	Informant report via interview		-	11	Likert scale (0 to 4; all options described) Total score is sum of item scores. Range 0 to 44.
BSSD [178]	1 [178]	a) Neuropsychiatric symptoms (behavioural syndromes in AD)	n/a	Clinician-judgement based on information from interview with	20-30 minutes for the whole scale (not just apathy subscale)	1 week	7	Likert scale (0 to 6; all options described).

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
		b) AD c) Clinical		informant and informed by clinician observations				Total score is not specified but presumable sum of item scores.
DAIR [144]	1 [144]	a) Apathy b) Dementia (mild-moderate) c) Research and clinical	n/a	Interviewer-judgement based on informant reports. In person or over the phone.	30 minutes; Publicly available	1 month	16	Main items rated on Likert scale by informant: (0 to 3; all options described) with follow-up questions to determine if this was a change in apathy rated by the interviewer (no change; increase; decrease) Total score is sum of all items reflecting a change (more apathetic), divided by the number of items completed.
DAS [119]	5 [45,202–205]	a) Apathy b) Neurodegenerative diseases specifically with motor disability c) Research and clinical	DAS Informant	Informant reported via online or paper and pencil	5 minutes; Publicly available	1 month	24 (8 per subscale)	Likert scale (0 to 3; all options described). 'Executive', 'Initiation' and 'Emotional' subscales are scored by summing all items in subscale. Range 0 to 24. Total score is the sum of the subscale scores. Range 0 to 72.
			DAS Self	Self-reported via online or paper and pencil	Publicly available	1 month	24 (8 per subscale)	Likert scale (0 to 3; all options described). 'Executive', 'Initiation' and 'Emotional' subscales are scored by summing all items in subscale. Range 0 to 24.

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
								Total score is the sum of the subscale scores. Range 0 to 72.
			b-DAS	Informant reported via online or paper and pencil	< 5 minutes; Publicly available	1 month	9 (3 per subscale)	Likert scale (0 to 3; all options described). 'Executive', 'Initiation' and 'Emotional' subscales are scored by summing all items in subscale. Range 0 to 9. Total score is the sum of the subscale scores. Range 0 to 27. (an awareness deficit rating is also present but not included in the total score)
DEX [179]^	1 [206]	-	-	-	Unable to obtain for this review, suggesting not publicly available.	-	-	-
FrSBe [180]^	2 [207,208]	-	FrsBe (14 item version)	-	Unable to obtain for this review, suggesting not publicly available.	-	14	-
			FrsBe (11 item version)	-	Unable to obtain for this review, suggesting not publicly available.	-	11	-
			FrsBe (6 item version)	-	Unable to obtain for this review,	-	6	-

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
					suggesting not publicly available.			
GDS [181,182]	2 [209,210]	a) Depression b) Older adults c) Clinical screening	GDS-3a	Self-reported via paper and pencil (interviewer administered if required)	-	1 week	3	Responses (Yes/No) that indicate depression are scored 1. Total score is sum of items. Range 0 to 3
			GDS-6a	Self-reported via paper and pencil (interviewer administered if required)	-	1 week	6	Responses (Yes/No) that indicate depression are scored 1. Total score is sum of items. Range 0 to 3
GIP [183]^	1 [211]	-	GIP-subscale	-	Unable to obtain for this review, suggesting not publicly available.	-	-	-
			GIP-domain	-	Unable to obtain for this review, suggesting not publicly available.	-	-	-
			GIP-9a (subscale of the GIP-28)	-	Requires translating	-	9	-
IMD [184]	1 [184]	a) 'Mental decline' or 'impairment' b) Older adults, particularly with dementia	n/a	Informant reported	-	-	3	Items are rated using categories that are associated with weighted scores depending on the item.

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
		c) Research. (Possibly also for clinical evaluation of progression but should not be used for diagnosis)						0="Absent"; 2/3="Mild-moderate / discontinuous symptoms"; 4/5/6="Severe / continuous symptoms" Total score is sum of item scores. Range 0 to 15
KBCI [185,212]	1 [213]	a) Behaviour change b) Traumatic Brain Injury c) Clinical and research	KBCI-8a	Informant reported via paper and pencil	Unable to obtain for this review, suggesting not publicly available.	Not reported	8	Likert scale (all options described) Total score is the sum item scores but the scores attributed to the Likert scale and therefore also the range is unspecified.
			KBCI-10a	Informant reported via paper and pencil	-	Not reported	10	Likert scale (all options described) Total score is the sum item scores but the scores attributed to the Likert scale and therefore also the range is unspecified.
LARS [59]	3 [214–216]	a) Apathy b) Parkinson's Disease c) Clinical and research	Clinician	Interviewer-judgement informed by patient self-report and interviewer observations during the interview with the patient	10 minutes; Instructions and measure available as supplementary material with the original article	4 weeks	33	Four items are based on 3 or 5 point Likert scales (all options described) For the remaining items, patient responses are categorised by the interviewer as 1 or -1 (all options described). Items are scored 0 if they are rated 'N/A' or the interviewer was not able to categorise the reply.

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
								Total score is the sum item scores. Range -36 to 36.
			Informant	Interviewer-judgement informed by informant-responses during the interview with the informant	-	4 weeks	33	Five items are based on 3 or 5 point Likert scales (all options described) For the remaining items, informant responses are categorised by the interviewer as 1 or -1 (all options described). Items are scored 0 if they are rated 'N/A' or the interviewer was not able to categorise the reply. Total score is the sum item scores. Range -36 to 36.
NPI [90]	12 [90,217–227]	a) Neuropsychiatric symptoms b) Dementia c) Research and clinical	NPI (total)	Informant rated via interview	10 minutes for whole measure (not just apathy subscale); Full measure and manual available at no cost for all noncommercial research and clinical purposes	1 month (and represents a change from behaviour before the illness)	1 (but rated for frequency and severity)	Screening question (Yes=0; No), with follow-up questions using Likert scales, regarding severity (1 to 3; all options described) and frequency (1 to 4; all options described). Total score is Frequency x Severity (a distress rating is also present but not included in total score)
			NPI-A	Informant rated via interview	-	1 month (and represents a change from	-	Each item is rated for frequency on the same Likert scale as the original NPI.

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
						behaviour before the illness)		Total score is the sum of frequency scores. (Severity is also rated for the overall domain as per the original NPI procedure, but not included in the total score)
			NPI-C	Clinician-judgement, informed by information from the NPI with an informant and patient as well as other relevant information about the patient. Clinicians must have a minimum of two years experience of NPSs in people with dementia	No cost for non-comercial research. Permission is required prior to use.	4 weeks	11	Each item is scored individually by informants, employing the Likert method as the original NPI, regarding frequency, severity and distress. Total score is the summation of frequency and severity item scores. A clinical rating method is also required: Each item is also rated by a clinician based on their clinical impressions, informed by the interview with the patient and informant, clinical notes and other carers, rated on Likert scale (0 to 3). Total score is the sum of these clinician rated item scores. Two separate total scores are obtained: one from the informant, one from the clinician.

Measure	N of studies meeting criteria†	Original intended… a) construct b) target population c) context	Version	Mode of administration	Feasibility (administration time; availability)	Recall Period	N of items	Scoring and Response options*
UPDRS [186]^	4 [228–231]	-	UPRDS	-	10-20 minutes for whole scale (not apathy-subscale); Difficult to obtain	-	1	Likert scale (0 to 4; all options described). No total score calculation required as only 1 item present.
			MDR-UPDRS	Rater-judgement informed by interview with patient and / or informant	Certification program for the whole measure (not apathy-subscale)	1 week	1	Likert scale (0 to 4; all options described). No total score calculation required as only 1 item present.

AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; AES, Apathy Evaluation Scale; AI, Apathy Inventory; AS, Apathy Scale; BMDS, Behaviour and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; DAIR, Dementia Apathy Interview and Rating; DAS, Dimensional Apathy Scale; DEX, Dysexecutive Questionnaire; FrSBe, The Frontal Systems Behavior Scale; GDS, Geriatric Depression Scale; GIP, The Behavioural Rating Scale for Geriatric Inpatients; IMD, Index of Mental Decline; KBCI, Key Behaviors Change Inventory; LARS, Lille Apathy Rating Scale; NPI, The Neuropsychiatric Inventory (NPI-A, NPI-Alternative; NPI-C, NPI-Clinician Rating Scale; UPDRS, Unified Parkinson's Disease Rating Scale (MDS-UPDRS, Movement Disorder Society Revision of the Unified Parkinson's Disease Rating Scale).

† Number does not include development article where development article did not meet the inclusion criteria, even if it was later assessed for purposes of content validity

* Reverse coding is not included here

^ Unable to obtain development article for rating

- Unable to obtain information

Whilst nine of the included publications also detailed the development of the measure, nine additional studies reporting on the remaining measures' development were then included to aid review of content validity and measurement characteristics such as mode of administration, number of items and subscales, scoring guides and target population [181,212,185,59,176,119,182,232,233].

The content of the included apathy measures will be discussed in the next section. Results will then be discussed in relation to each measurement property (section 4.4.3), followed by presentation of the synthesised results per measure (section 4.4.4), as advised by the COSMIN standards for reporting systematic reviews [166].

The measurement properties assessed and study characteristics are reported for each publication in Appendix 6. The results and risk of bias ratings for development and content validity studies, reviewer-rated content validity, and results and risk of bias of studies of the remaining measurement properties are reported in Appendix 6. No studies assessed cross-cultural validity.

4.4.2. Content of apathy measures

Table 4 details which symptoms or 'indicators' of apathy were included in the eligible measures. Measures included a variety of aspects of apathy, namely domains of behaviour (activity, initiation and persistence), cognition (interest and novelty seeking), emotion (internal feeling and outward expression), social (activity, interest in, or initiation of socialising), and general motivation or drive. Other concepts that were included that did not seem to fall into these expected categories were cognitive ability, having friends, disinhibition, insight, personal care, energy, compliance, and creativity.

Table 4. *Content of apathy measures*

	Cognitive			**Behavioural**					**Emotional**		**Motivation**	**Other**
Domain:	Interest (general)	Interest (social)	Novelty seeking	Activity (general)	Activity (social)	Initiation (general)	Initiation (social)	Persistence	Emotional expression (external)	Emotional (internal)	Motivation, or drive	Other concepts included
AES	✓	✓	✓	✓	-	✓	-	✓	-	✓	✓	Insight; Friendship.
AI	✓	✓	-	-	-	✓	✓	-	✓	-	-	
AMI	✓	✓	-	✓	✓	✓	✓	✓	-	✓	-	Disinhibition; Cognition; Decision-making.
AS*	✓	-	✓	✓	-	✓	-	-	-	✓	✓	Insight; Energy.
BMDS-a	-	✓	-	✓	-	-	✓	✓	-	-	-	Cognition.
BSSD-a	✓	✓	-	-	-	✓	-	-	✓	-	-	Personal care.
DAIR	✓	✓	-	✓	-	✓	✓	✓	✓	✓	-	
DAS	✓	✓	✓	✓	✓	✓	-	✓	✓	✓	✓	Cognition; Empathy.
bDAS	✓	✓	✓	-	-	✓	-	✓	-	✓	-	Cognition.
DEX-a	-	✓	-	-	-	-	-	✓	-	✓	✓	Cognition; Disinhibition.
FrSBE-a	✓	-	-	✓	-	✓	✓	✓	✓	-	✓	Personal care.
FrsBE-6a	✓	-	-	✓	-	✓	-	-	✓	-	✓	Personal care.
FrsBE-11a	✓	-	-	✓	-	✓	✓	-	✓	-	✓	
GDS-3a	✓	-	✓	✓	-	-	-	-	-	-	-	Energy.
GDS-6a	-	-	✓	✓	✓	✓	-	-	-	✓	-	Energy.
GIP-9a	-	✓	-	✓	-	-	-	-	-	-	-	Compliance; Involvement in personal care;

	Cognitive			Behavioural					Emotional		Motivation	Other
												Energy; Cognition; Disinhibition.
IMD-a	✓	-	-	✓	-	-	-	-	-	✓	✓	
KBCI-10a	✓	-	-	✓	-	✓	-	-	-	✓	✓	Creativity; Energy.
LARS	✓	✓	✓	✓	-	✓	✓	✓	✓	✓	✓	Insight; Friendship.
MDS-UPDRS-a	✓	✓	-	-	-	-	-	-	-	-	-	
NPI-a**	✓	✓	-	-	-	✓	-	-	-	-	✓	
NPI-C-a***	✓	✓	-	✓	-	-	✓	-	✓	✓	-	
UPDRS-a	-	-	-	-	-	✓	-	-	-	-	✓	

*The main item questions are the only content that is considered here, as it is not clear how the questions in parentheses after each item are scored.
**The screening questions are the only content that is considered here, as if the respondent does not answer yes to theses, then they score 0 (i.e. no apathy).
***The screening questions are not considered here as they do not contribute to the score and the NPI-C advice is to administer all the questions on the construct of interest, regardless of screening responses.
Abbreviations: AES Apathy Evaluation Scale; AI, Apathy Inventory; AMI, Apathy Motivation Index; AS, Apathy Scale; b-DAS, basic Dimensional Apathy Scale; BMDS-a, Behavioural and Mood Disturbance Scale apathy subscale; BSSD-a, Behavioral Syndromes Scale for Dementia apathy subscale; DAIR, Dementia Apathy Interview Rating; DAS, Dimensional Apathy Scale; DEX-a, Dysexecutive Questionnaire apathy subscale; FrSBe-a, Frontal Systems Behavior Scale- apathy subscale; FrsBE-6a, Frontal Systems Behavior Scale 6-item apathy subscale; FrsBE-11a, Frontal Systems Behavior Scale 11-item apathy subscale; GDS-3a, Geriatric Depression Scale 3-item apathy subscale; GDS-6a, Geriatric Depression Scale 6-item apathy subscale; GIP-9a, Behavioral Rating Scale for Psychogeriatric Inpatients 9-item apathy subscale; IMD-a, Index of Mental Decline apathy subscale; KBCI-10a, Key Behaviors Change Inventory 10-item apathy subscale; LARS, Lille Apathy Rating Scale; MDS-UPDRS-a, Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale apathy subscale; NPI, Neuropsychiatric Inventory; NPI-a, Neuropsychiatric Inventory apathy subscale; NPI-C-a, Neuropsychiatric Inventory Clinician apathy subscale; UPDRS, Unified Parkinson's Disease Rating Scale apathy subscale

4.4.3. Results and Risk of Bias of individual studies

4.4.3.1. Measurement and study characteristics

Very few studies reported measurement characteristics of floor and ceiling effects, minimal important change or difference, response shift or missing items. As a result, these are not reported here, due to inability to compare these properties across measures. Sample sizes for individual studies ranged from 6 [219] to 665 [50]. Many studies did not report on the age range of participants, but the mean age of participants from eligible studies ranged from 65 [196] to 83 [190]. Of the studies that met the eligibility criteria, participants were formed of healthy older adults, and people with various types of dementia, MCI, Parkinson's disease, depression, stroke, and other neurocognitive disorders. Of the additional studies included only for the purposes of assessing the development of the measure, participants were patients, carers and professionals.

4.4.3.2. Content validity

Literature pertaining to development was obtained for all measures except the DEX, GIP, and FrSBe. Just four additional content validity studies were available that met the eligibility criteria. Both development and content validity studies often failed to provide sufficient detail, did not involve patients, carers, or members of the public, or did not report a systematic process through which items were produced or refined. As such, the included publications offered little evidence for content validity, with no study exceeding doubtful methodological quality. Most studies had indeterminate results regarding relevance, due to inadequate or poorly reported method. The remaining studies [144,185,212] suggested inconsistent evidence for relevance, largely due to the lack of justification for the response options and recall period. Studies addressing comprehensiveness and comprehensibility were all indeterminate as even when it was clear that participants were involved in assessing these properties, it was not clear what aspects of the measure participants were consulted about (e.g. items, response options, recall period, instructions).

Reviewer ratings of the content of measures were possible for most measures, however the ADRD and GIP and some aspects of the DEX, FrSBe, IMD and KBCI were indeterminate as the full list of items, or key aspects of the administration instruction were missing. Seven measures were rated sufficient, and eight were rated as having inconsistent content validity in both older adult and dementia populations, whilst the DAIR was rated as having sufficient content validity in a dementia population, but inconsistent for older adults. No measures were found to have insufficient content validity.

Content validity of the AES, AS, DAS and LARS was assessed by three PPI members as these measures had the best evidence for good validity and reliability. All four measures were judged by PPI members as having less than 85% of items appropriately worded, despite positive feedback for some measures (in particular the DAS). Inappropriate wording of items was largely due to questions being too broad, vague, or ambiguous, and using technical language or complex sentences. The AS was seen as potentially insensitive due to its direct questioning style (such as “Do you put much effort into things?”). Statements to which people could agree or disagree were preferred as this was viewed as less interrogational. Most measures had at least 85% of items that were relevant, and most had response options that were deemed appropriate and to match the items.

4.4.3.3. Structural validity and Internal consistency

Twenty seven studies of structural validity across 16 publications met the inclusion criteria [44,46–50,144,189,196,198,200,203,204,208,216,221]. For many apathy measures (AD-RD, FrSBe, Geriatric Depression Scale [GDS; 181,182], KCBI, AES, AS, Apathy Inventory [AI; 187], DAIR) it was uncertain whether they were based on a formative or reflective model, so a reflective model was assumed. The AMI, DAS and LARS are based on a formative model, as indicated by the presence of distinct dimensional subscales, and therefore the results of structural validity (*N*=3) or internal consistency (*N*=10) of these measures are reported but not rated. Just three publications assessed structural validity using the preferred method of confirmatory factor analysis or item response theory analysis [198,200,204], enabling three studies to have very good methodological quality, whilst seven had adequate, nine had doubtful quality, and five were inadequate.

Internal consistency was assessed by 31 publications [44,45,47–50,144,178,187–191,194–196,198–200,202–205,207–209,214,216,221,222,225] and was considered a valid assessment (i.e. the measure was based on a reflective model) for 38 studies[1]. Nineteen studies were rated very good, 17 doubtful, and 2 inadequate. As there was no determinate evidence for sufficient structural validity of the AI, AS, FrSBe, GDS, or Neuropsychiatric Inventory [NPI; 90], conclusions could not be drawn about internal consistency of these measures.

4.4.3.4. Reliability and measurement error

Methodological quality of reliability studies was mostly limited as a result of not using the optimal statistical method, such as the use of Kappa rather than weighted Kappa, or Pearson or Spearman correlation instead of ICC. Where

[1] Note that some publications presented more than one study

the most appropriate method was used, the model or formula of ICC or weighted Kappa was often not reported.

It was not possible to draw conclusions from all but one study of measurement error, due to lack of appropriate anchor-based estimates of Minimal Important Change (MIC) for any of the measures. No studies reporting measurement error referred to MIC and despite its inclusion in the systematic review search strategy, no studies assessing MIC were found.

4.4.3.5. Construct validity

The most commonly assessed measurement property was construct validity, with one hundred and eighty studies of construct validity meeting eligibility criteria from 45 publications [44–50,90,144,178,184,187–193,195,197–199,201–203,206,207,209,210,213–216,218–220,222–226,228–231]. Sixty-four studies were of very good quality, 47 adequate, 36 doubtful, and 33 inadequate. Twenty-one of these had indeterminate results due to not reporting sufficient information. The vast majority of studies failed to state a-priori hypotheses, and the majority of studies reported *p* values but not effect sizes.

4.4.4. Synthesis of results

A synthesis of the results and GRADE for content validity is provided in Table 5, and for the remaining measurement properties in Table 6. These results are further discussed below.

Table 5. *Content validity summary of findings, quality rating and GRADE for each measure*

Measure	**Summary of findings**			**Quality rating & GRADE**
	Development study	Content validity study	Researcher rating	
AD-RD	Indeterminate			?
AES	Indeterminate		Sufficient	Sufficient (1; 1)
AI	Indeterminate		Inconsistent	Inconsistent (1; 1)
AMI	Indeterminate		Sufficient	Sufficient (1; 1)
AS	Indeterminate	Indeterminate	Sufficient	Sufficient (1; 1)
BMDS	Indeterminate		Inconsistent	Inconsistent (1; 1)
BSSD	Indeterminate		Sufficient	Sufficient (1; 1)
DAIR	Inconsistent		Inconsistent (OA); Sufficient (NCD)	Inconsistent (1; 2)
DAS	Indeterminate		Sufficient	Sufficient (1; 1)
DEX			Inconsistent	Inconsistent (1; 1)
FrSBE		Indeterminate	Inconsistent	Inconsistent (1; 1)
GDS	GDS-3a: Indeterminate GDS-6a: Indeterminate		Inconsistent Sufficient	Inconsistent (1; 1) Sufficient (1; 1)
GIP	Inconsistent			Inconsistent (1; 1)

IMD	Indeterminate			?
KBCI	Inconsistent		Inconsistent	Inconsistent (1; 1)
LARS	Indeterminate		Sufficient	Sufficient (1; 1)
NPI	Indeterminate		Inconsistent	Inconsistent (1; 1)
NPI-A				
NPI-C		Indeterminate	Sufficient	Sufficient (1; 1)
UPDRS	Indeterminate		Inconsistent	Inconsistent (1; 1)

GRADE Key: 1, Very low; 2, Low; 3, Moderate; 4, High quality;
Abbreviations: AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; AES Apathy Evaluation Scale; AI, Apathy Inventory; AMI, Apathy Motivation Index; AS, Apathy Scale; BMDS, Behavioural and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; DAIR, Dementia Apathy Interview Rating; DAS, Dimensional Apathy Scale; DEX, Dysexecutive Questionnaire; FrSBe, Frontal Systems Behavior Scale; GDS Geriatric Depression Scale; GIP, Behavioral Rating Scale for Psychogeriatric Inpatients; IMD, Index of Mental Decline; KBCI, Key Behaviors Change Inventory; LARS, Lille Apathy Rating Scale; NPI, Neuropsychiatric Inventory; NPI-A, Neuropsychiatric Inventory Alternative; NPI-C, Neuropsychiatric Inventory Clinician; UPDRS, Unified Parkinson's Disease Rating Scale; OA, Older Adults; NCD, People Living with Neurocognitive Decline

Table 6. *Overall findings and GRADE per measurement property (excluding content validity)*

Measure	Structural Validity		Internal Consistency		Reliability		Measurement error		Construct validity	
	Summary of findings	Quality & GRADE	Summary of findings Cronbach's α	Quality & GRADE	Summary of findings	Quality & GRADE	Summary of findings	Quality & GRADE	Summary of findings N hypotheses confirmed / tested (%)	Quality & GRADE
AD-RD					r=.72	+ (1; 1)				
AES	1 main apathy factor with smaller factors of various description	+ (3; 3)	.86 to .95	+ (3; 3)	r/ICC= .72 to .94	+ (3; 2)	SEM= 2.7 to 2.9	?	48/69 (70%)	+† (3^; 3^)
AI			.83 to .96	?	Kappa/ICC = .96 to .99	+ (2; 2)			5/8 (67%)	+/-† (3; 4)
AMI			.86	*					0/2 (0%)	- (2; 1)
AS	1 to 3 factors	+/- (4; 4)	.69 to .94	?	r/ICC=.78 to .90	+ (1; 1)	SEM= 2.34	?	8/12 (67%)	+†† (2^; 3^)
BMDS					r=.90	+ (1; 1)				
BSSD					ICC= .65 to .85	+/- (1; 1)			½ (50%)	-† (1^; 3^)
DAIR	1 factor	+ (1; 3)	.89	+ (2; 3)	r=.85	+ (1; 1)	100% agreement	+ (1; 1)	¾ (75%)	+ (2; 4)
DAS	3 factors: cognitive; behavioural; emotional	*	.81 to .93	*	ICC=.84	+ (1; 1)			10/13 (77%)	+ (3; 4)
DEX					ICC=.93	+ (1; 1)			2/4 (50%)	+/- (3; 4)
FrSBE	1 factor	?	.80 to .88	?					4/5 (80%)	+ (1; 2)
GDS			.51	?					GDS-3a: 0/2 (0%)	GDS-3a: - (3; 2)

Measure	Structural Validity		Internal Consistency		Reliability		Measurement error		Construct validity	
	Summary of findings	Quality & GRADE	Summary of findings Cronbach's α	Quality & GRADE	Summary of findings	Quality & GRADE	Summary of findings	Quality & GRADE	Summary of findings N hypotheses confirmed / tested (%)	Quality & GRADE
									GDS-6a: 3/3 (100%)	GDS-6a: + (3; 2)
GIP					ICC= .72 to .83	+ (1; 1)	MDD= 2.8 to 3.8	?		
IMD									3/3 (100%)	+ (1; 2)
KBCI									6/7 (86%)	+ (2; 1)
LARS	4 factors	*	.81 to .87	*	r/ Kappa / ICC = .93 to 1.00	+ (2; 3)			11/13 (85%)	+ (4; 4)
NPI			.82 to .83	?	r/ rs / ICC= .53 to .99	+ (1; 2)			1/5 (20%)	- (4; 4)
NPI-A	1 factor	+ (1; 3)	.91	OA: ? PLwNCD: + (3)						
NPI-C					ICC= .87	+ (2; 3)			1/2 (50%)	+/- (3; 4)
UPDRS										

Blank cells indicate no eligible studies or results.
Quality of measurement property: +, Sufficient; +/-, Inconsistent; -, Insufficient, ? Indeterminate.
GRADE rating : 1, Very low; 2, Low; 3, Moderate; 4, High quality. Quality of evidence rating in parentheses first indicates quality of evidence for older adults, then people living with NCD.
* not applicable due to formative model.
†Greater emphasis placed on results of better quality (sub)studies
†† Greater emphasis placed on studies of convergent validity
^ Marked down for inconsistency
Abbreviations: AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; AES Apathy Evaluation Scale; AI, Apathy Inventory; AMI, Apathy Motivation Index; AS, Apathy Scale; BMDS, Behavioural and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; DAIR, Dementia Apathy Interview Rating; DAS, Dimensional Apathy Scale; DEX, Dysexecutive Questionnaire; FrSBe, Frontal Systems Behavior Scale; GDS-3a, Geriatric Depression Scale 3 item apathy subscale; GDS-6a, Geriatric Depression Scale 6 item apathy subscale; GIP, Behavioral Rating Scale for Psychogeriatric Inpatients; ICC, Intraclass Correlation Coefficient ;IMD, Index of Mental Decline; KBCI, Key Behaviors Change Inventory; LARS, Lille Apathy Rating Scale; MDS-UPDRS, Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale; MDD, Minimal

Detectable Difference;NPI, Neuropsychiatric Inventory; NPI-A, Neuropsychiatric Inventory Alternative; NPI-C, Neuropsychiatric Inventory Clinician; OA, Older Adults; PLwNCD, People Living With NCD; SEM, Standard Error of Measurement; UPDRS, Unified Parkinson's Disease Rating Scale

4.4.4.1. Apathy specific measures

4.4.4.1.1.AES

The development and piloting of the AES was described by Marin and colleagues [44], however there was insufficient detail reported to make conclusions regarding content validity from this. Reviewer ratings of the relevance, comprehensiveness, and comprehensibility of the AES found there was sufficient content validity.

There was moderate evidence for sufficient structural validity. This limited the quality of evidence for sufficient internal consistency to be moderate also (as in COSMIN standards, the internal consistency evidence is limited by whether there is good evidence for uni-dimensionality). There was moderate to low evidence for sufficient reliability, except where the AES-S was conducted with people with AD, where test-retest reliability (r=.44) was insufficient (Marin et al., 1991). One study of measurement error was conducted, however there was insufficient information available to draw conclusions. Across a range of convergent, divergent and known-group validity studies, hypotheses were met in only 70% of cases, however, the overall construct validity was rated sufficient, as this was consistent with the results of the higher quality studies. Evidence was marked down for inconsistency.

4.4.4.1.2.DAS

The DAS was developed deductively from other apathy and related scales, and the item pool was refined by the authors, though the exact process by which this occurred was not clear enough for COSMIN standards, and therefore the content validity is indeterminate from the development study [119]. Reviewer ratings, determined that the DAS, including b-DAS, had sufficient content validity. This was the preferred scale (of AES, AS, DAS and LARS) amongst PPI members, who judged that the items were generally well worded and relevant, and that the format of statements rather than direct questions felt less intrusive than other measures.

The DAS subscales' internal consistency (Cronbach's α) ranged from .81 to .90 for the executive domain, .76 to .90 for the initiation domain, and .54 to .71 for the emotional domain. However, there is no evidence for the unidimensionality of the individual subscales, so conclusions could not be made. There was very low evidence for sufficient test-retest reliability (ICC=.93), however this evidence came from a single study of the b-DAS so conclusions may not be generalisable to the longer version. There was moderate to high-quality evidence of sufficient construct validity.

4.4.4.1.3.LARS

There was no apparent systematic process via which items of the LARS were elicited with professionals or participants. A later pilot study asked about the relevance and comprehensibility of a Spanish version, however there was insufficient detail reported to determine validity [216]. Reviewer ratings of content validity however found that there was sufficient content validity.

One study found that internal consistency of the LARS domains ranged from .93 to .94 [215]. However, there was no evidence for the unidimensionality of each subscale, so no conclusions could be made. There was low to moderate evidence of sufficient reliability, and high-quality evidence for sufficient construct validity of the LARS.

4.4.4.1.4.DAIR

The development of the DAIR involved participation from people with dementia and their carers, and the measure was further piloted, though it was unclear what and whether the participants were asked about comprehensibility and comprehensiveness, so these aspects of content validity were indeterminate. The development study was of doubtful methodological quality, and indicated inconsistent evidence for relevance. The reviewer rating concluded that there was sufficient content validity for people with dementia, but inconsistent validity for older adults due to references to an 'illness' i.e. dementia. It is worth noting that this may be somewhat problematic in assessing people with a range of severity of NCD, as people with MCI in particular, who by definition are not significantly impacted in daily function, may not consider their condition an 'illness', though it is likely that this is easily altered with a small wording change. Some items, when combined with the response options, became confusing or created a double negative (e.g. 'Does s/he no longer seem to react to things…' 'No, never…').

There was very low to moderate evidence for sufficient structural validity, which also limited the evidence for sufficient internal consistency. There was very low evidence for sufficient test-retest reliability, very low evidence for sufficient measurement error, and low to high quality evidence for sufficient construct validity.

4.4.4.1.5.AI

The development of the AI did not report a systematic eliciting of items, and therefore conclusions regarding the content validity of the AI cannot be drawn from this study. Reviewer ratings of content validity found that there was inconsistent evidence of content validity for all versions of the AI, largely due to poor comprehensibility due to double-barrelled questions and lack of clear wording, likely due to translation issues (e.g. "do you make decisions and initiatives?").

There was low evidence for sufficient reliability of the AI and internal consistency was indeterminate. Across convergent, divergent and known group studies there was moderate to high quality evidence for inconsistent construct validity in older adult and dementia populations respectively.

4.4.4.1.6.AS

Though two professionals were involved in eliciting and modifying items to develop the AS (confirmed via email with S. Starkstein, MD PhD [sergio.starkstein@uwa.edu.au] in October 2018), there was no systematic procedure by which professionals or participants elicited items or developed and reviewed the measure components, so the results of the development study are indeterminate. A study of the Portuguese version of the AS-I [197] asked participants about the comprehensibility of the measure, and whilst no modifications were reportedly required, it was not clear which aspects carers were consulted about, and thus the quality is indeterminate. The reviewer rating concluded that there was sufficient content validity, though it is worth acknowledging that this was the least preferred measure that PPI members assessed, as they felt some questions were not appropriately worded or too direct, though sometimes these comments were related to sub-questions (e.g. “do you think anything is wrong with you?” from the item “Are you concerned about your condition?”), and it is not clear how sub-questions should be used when administrating the measure.

Despite the high-quality studies included, the AS had inconsistent results regarding its structural validity, and as such the internal consistency was also indeterminate. The AS-HC however had moderate to low evidence for sufficient structural validity and internal consistency in older adult and dementia populations respectively. There was very low evidence for sufficient reliability of the AS. Measurement error was assessed but there was insufficient information to draw conclusions. Across a range of convergent, divergent and known-group validity studies, hypotheses were met in only 67% of studies, however, the overall construct validity was deemed sufficient, as this was consistent with the results of the higher quality studies, and evidence was marked down for inconsistency.

4.4.4.1.7.AMI

There was no systematic item elicitation or development of the AMI items, so the result of the development study was indeterminate. Though the reviewer rating concluded there was sufficient content validity of the AMI, 22% of items were deemed inappropriate, as they were too conflated with cognition or disinhibition (e.g. “I get things done when they need to be done, without requiring reminders from others”). Nevertheless, 78% were relevant, and all items were otherwise comprehensive and comprehensible.

Internal consistency was calculated separately for the behavioural-activation, social motivation and emotional sensitivity subscales, however conclusions could not be drawn as it is unknown whether the subscales are unidimensional in this population. There was low to very low evidence of insufficient construct validity of the AMI.

4.4.4.2. Global measures with an apathy subscale

Apathy subscales were present in eleven global measures designed to assess a variety of constructs (such as dementia severity, and neuropsychiatric symptoms). Measurement properties were only reviewed if they were regarding the apathy subscale, not the overall global measure.

4.4.4.2.1.NPI

The items of the NPI did not appear to be elicited through a systematic process with professionals or participants. Two Delphi studies with professionals have been conducted to assess content validity of the NPI [90] and NPI-C [220], however the results of both studies were indeterminate, as insufficient detail was reported for assessment via COSMIN standards. Reviewer ratings determined that content validity of the original NPI apathy subscale (NPI-a) was inconsistent, as the emotional domain was missing from the screening questions, which could mean the presence of apathy would go undetected when using the screening questions as described in the NPI instructions. In contrast, the NPI-C advises that all items on the construct of interest should be rated, regardless of answers to screening questions, so the NPI-C apathy subscale's (NPI-C-a) content validity was sufficient. The NPI-A apathy subscale (NPI-A-a) could not be rated as it was not possible to obtain the full measure and the description of included items was not sufficiently clear to determine the actual items included.

No studies aimed to assess the structural validity of the NPI-a, and as such the results of internal consistency studies were indeterminate. However, it is worth nothing that the NPI-a is made up of just two rateable items and so is likely to be unidimensional. The NPI-A-a however was investigated for structural validity and was found to be made up of one factor, indicating sufficient structural validity with moderate to very low evidence. There was moderate evidence for sufficient internal consistency for the NPI-A-a in people with dementia. There was very low to low evidence of reliability of the original NPI-a for older adult and dementia populations respectively. The NPI-C-a had better evidence, with low and moderate evidence for sufficient inter-rater reliability for older adult and dementia populations respectively. There were no studies of reliability for the NPI-A-a. None of the included studies assessed measurement error for the NPI-a or its various versions, though one study found 97.9% agreement in apathy frequency and 89.4% agreement for apathy severity [90]. There was high quality evidence of insufficient construct validity

of the NPI-a. For the NPI-C-a, there was moderate to high quality evidence of inconsistent construct validity.

4.4.4.2.2.BSSD

The BSSD was developed and piloted with healthcare professionals, however the process was not described thoroughly enough to be able to rate the quality of content validity, which is therefore indeterminate. The researcher rating indicated there was sufficient content validity. However, caution is to be taken as only 71% of items were deemed relevant to apathy, as some measured more physical problems caused by disorders such as Parkinson's disease (e.g. "Has his voice been more flat in tone or does it fluctuate?").

There was very low evidence for inconsistent reliability for face-to-face administration, with insufficient reliability when administered via phone-call. Construct validity findings were mixed, but the overall construct validity was rated insufficient, as this was consistent with the result of the better-quality study, and evidence was marked down for the inconsistency. However, results should be interpreted with caution, as no studies of convergent validity were included, which is a superior indicator of construct validity than divergent or known-group validity [173].

4.4.4.2.3.GDS

The development of the GDS involved professionals in eliciting, selecting and refining items, though the method for this did not appear to be systematic. The authors also piloted the GDS with people with and without depression, and found it to be an acceptable measure, though methods by which this was conclusion was reached are unclear, so the content validity is indeterminate from these studies. Reviewer ratings found inconsistent content validity for the GDS-3a, due to inclusion of items too conflated with physical ability, and lack of comprehensiveness. Despite inclusion of items that could be conflated with physical ability and dysphoria, the GDS-6a was deemed to have sufficient validity, as comprehensiveness and comprehensibility were sufficient (as per COSMIN guidelines).

There was moderate to low evidence for insufficient construct validity of the GDS-3a. In contrast, the GDS-6a had moderate to low evidence of sufficient construct validity, however this evidence was from divergent validity studies only, so results should be interpreted with caution.

4.4.4.2.4.DEX

The DEX was developed as part of the behavioural assessment of the dysexecutive syndrome test battery, which could not be obtained for this review, and therefore could not assess the content validity from the development study. An apathy factor (DEX-a) was described [206], however, so reviewer rating of the content validity of DEX-a was possible for relevance

and comprehensiveness. There was inconsistent content validity as only 63% of items were deemed relevant to apathy, with other items being related to cognitive impairment or disinhibition (e.g. “abstract thinking problems”).

There was very low evidence for sufficient test-retest reliability of the DEX-a , and moderate to high quality evidence of inconsistent construct validity.

4.4.4.3. Measures with little evidence

The following measures are presented together as they had either no evidence for or inconsistent content validity and evidence for only one other measurement property, indicating little is known about the quality of these measures.

4.4.4.3.1.KBCI

Elicitation, reduction and rewording of items for the KBCI was performed with input from patients, carers and professionals. However, although patients and carers were asked about comprehensiveness, this was not done for the final revision of the scale. Furthermore, there was insufficient detail reported to determine whether participants were specifically asked about the comprehensibility, of all aspects of the scale, or just the items. Therefore, despite a relatively thorough development process, results of the development study were indeterminate. Reviewer ratings found that there was inconsistent content validity of the KBCI-10a, due to some items possibly not being relevant to older adults and people with dementia (e.g. “has a lot of get-up-and-go”), and others lacking clear comprehensibility (e.g. “is enterprising”).

There was low to very low evidence of construct validity, however no convergent validity studies met the criteria, so results regarding construct validity should be interpreted with caution.

4.4.4.3.2.AD-RD

An apathy subscale of the AD-RD (AD-RD-a) was described by Tappen & Williams [175], but it was not possible to obtain the full list of items and instructions, so reviewer content validity ratings were not possible. The development of the AD-RD was based on interviews with carers to elicit items [176], and professional feedback to modify these items was later described by Tappen and Williams [175]. However, in the former instance, the carers were asked about mood generally, rather than apathy specifically, and in the latter, it was unclear what aspects of the measure professionals were asked about or what alterations were made.

There was very low evidence for sufficient test-retest reliability, and no studies for the remaining measurement properties met the criteria.

4.4.4.3.3.IMD

Items of the IMD were developed from literature and by professionals, though there was no apparent systematic process for this, and participants were not involved, therefore the results of the development study are indeterminate. It was not possible to obtain a copy of the full instructions, including recall period, so reviewer ratings were not possible for relevance and comprehensibility. However, comprehensiveness was deemed sufficient, and all items were relevant to apathy.

There was very low to low evidence of sufficient construct validity of the IMD-a, however there were no convergent validity studies that met the criteria, so results should be interpreted with caution.

4.4.4.3.4.UPDRS

The Unified Parkinson's Disease Rating Scale (UPDRS) was developed by a panel of professionals [186]. The 1-item motivation-initiation assessment was designed to screen for apathy. The scale, including the wording of the apathy item, was later revised by a further panel of experts, who reportedly surveyed participants, to select the items to be included, to produce the Movement Disorder Society-UPDRS [MDS-UPDRS 232,233]. Neither version included sufficient detail on its development or content validity to assess the quality using COSMIN guidelines. Reviewer ratings determined that there was inconsistent content validity for both the UPDRS and MDS-UPDRS apathy item, as the single item measure had insufficient comprehensiveness.

There was moderate quality evidence for inconsistent construct validity.

4.4.4.3.5.FrSBe

A rigorous process of development of the FrSBe has been referenced by others [207] but it was not possible to obtain the original development study, and thus unable to include this as evidence for content validity. Carvalho [207] investigated the content validity of the FrSBe-a and an 11-item version (FrSBe-11a). They judged that 86% and 82% of items respectively had sufficient agreement between participants' interpretations and the intended interpretation of the subscale. However, the comprehensibility of the scale instructions and response options was not investigated, so content validity from these studies was indeterminate. Reviewer ratings were not possible for comprehensibility as the full wording could not be obtained. In all three versions of the FrSBE considered, none had at least 85% of items relevant to apathy, due to items related to personal hygiene that could be conflated with other conditions.

One study investigated the structural validity of the apathy subscale, finding just one factor, however the appropriate statistical results were not reported, leaving the structural validity and subsequent internal consistency

indeterminate. There was very low to low evidence of construct validity, however no convergent validity studies met the criteria for this review, so results regarding construct validity should be interpreted with caution.

4.4.4.4. Measures that are not recommended

As in the above section, the following measures had little evidence for their quality. In addition, they were deemed poor measures of apathy due to insufficient content validity.

4.4.4.4.1.GIP

It was not possible to obtain the development study of the GIP, nor a copy of the original GIP measure. Therefore the only apathy subscale that could be obtained (GIP-9a) was rated, which was taken from the short version of the GIP (GIP-28), as described by de Jonghe and colleagues [234]. Content validity was inconsistent, however, it is worth noting that comprehensiveness was insufficient, as the emotional domain of apathy was missing, and only 44% of items were deemed relevant to apathy, with one item being easily conflated with cognitive problems (“can explain things clearly”). Therefore, the GIP-9a is not recommended to assess apathy in people living with NCD.

4.4.4.4.2.BMDS

The BMDS [BMDS; 177] was developed from the literature and authors opinions, though no systematic development process was described, so content validity is indeterminate from the development study. The reviewer rating concluded that there was inconsistent content validity of the apathy subscale (BMDS-a), though it is important to note that only 55% of items of were actually deemed relevant to apathy. Many of the items directly assessed cognition rather than apathy (such as “gets mixed up about where he/she is”). Items and response options created confusing double negatives, and the emotional domain of apathy was missing from this measure. Therefore, the BMDS-a is not recommended for the assessment of apathy in people living with NCD.

4.5. Discussion

4.5.1. Summary of the content of apathy measures

The LARS and DAS encompassed the most aspects of apathy, and assessed apathy across cognitive, behavioural, emotional and social domains. This is perhaps not surprising, as the LARS and DAS were both dedicated apathy measures, and the longest measures included in this review. The AES, AMI and DAIR also encompassed a broad range of content and measured apathy across all domains. Some measures included items that appeared to measure constructs beyond apathy, such as cognitive ability or energy. Inclusion of

these items may be appropriate for a healthy population, however when applied to older adults and people with dementia, the measures may be assessing their age-related impairments or disease state, rather than apathy. Measures which were subscales of a global measure of another construct were those which covered the fewest aspects of apathy and were more likely to fail to include all domains of apathy.

4.5.2. Summary of evidence and recommendations

According to COSMIN guidelines, measures should be recommended if they have sufficient content validity, at least low-level evidence for sufficient internal consistency, and no high-quality evidence for insufficient properties in the population of interest. The AES, AMI, AS, DAS, GDS-6a, LARS and NPI-C were all judged to have sufficient content validity, but the AS, GDS-6a and NPI-C did not have evidence for sufficient internal consistency. The AES had sufficient internal consistency. The AMI, DAS and LARS were based on a formative model, so internal consistency was not applicable. COSMIN guidelines do not advise how to recommend studies of measures based on a formative model. Therefore, the AES was the only measure that met the COSMIN criteria for a recommended measure. The NPI-a was the only measure to meet COSMIN criteria for a measure that should not be recommended for use. All other measures fell in category B, i.e. a measure that could be recommended, depending on further research. However, it is argued here that the BMDS and GIP are not appropriate measures to assess apathy in people living with NCD due to inclusion of items that are not relevant and would conflate apathy with cognition.

Despite the numerous studies of measurement properties identified by this review, evidence across all measurement properties was often of low or very low quality. Only one measure (LARS) had high quality evidence for at least one sufficient measurement property (construct validity) in both older adults and people with NCD. The DAIR had high quality evidence for sufficient construct validity, in people with dementia only. No single measure had overwhelmingly superior measurement properties. The AES, DAS, and LARS were however all judged to have sufficient content validity, reliability, construct validity, and structural validity and internal consistency where applicable.

The LARS appeared to be the measure with the most consistent best evidence for good measurement properties. However, the LARS may have less desirable measurement characteristics, as both the self and informant versions involve interviewer ratings, as well as respondent reports, and was the largest scale found by the review, with 33 items assessing apathy, so requires more resources and could be burdensome. The AES had the second most consistent quality evidence across measurement properties, and may have preferable measurement characteristics as there are versions that do not

require trained raters, and are based on less items. The DAS is a promising measure, with good evidence for sufficient measurement properties, with the exception of reliability, and desirable measurement characteristics, as the measure does not require interviewers and a short version is available. The DAS was also the preferred measure of PPI members in the PPI review of the four aforementioned apathy measures.

It is worth considering that the AES was based on a reflective model, whilst the LARS and DAS were based on formative models and therefore structural validity and internal consistency were not applicable for the latter two measures. Researchers may have different reasons for preferring to use a measure based on a formative or reflective model. Many researchers require a measure that assesses one construct using equal and interchangeable items, that are independent of each other, and all caused by the construct of interest (i.e. reflective model), as this is the basis for classical test theory [165]. Whilst a number of measures met the criteria for unidimensionality in this review, in reality, various structural validity studies of the same measure reported various different factor loadings, with unexpected item groupings even in the same population. For example, though the AES had the best evidence for unidimensionality, the included studies found the AES separated into two or three factors, with different items loading onto differently characterised factors across studies (apathy, novelty seeking, insight and socialising, interest, task completion, and friendship). This highlights that sum-scores may not currently be appropriate for the study of apathy. This is further suggested by the various ways in which measures defined and assessed apathy, emphasizing how it is unclear how apathy should be characterised and divided. Instead, it may be better to view apathy not as one disorder or category, but as a set of tendencies and behaviours that are causally linked in varying ways. This is termed the network perspective and is discussed in the next chapter.

4.5.3. Strengths & Limitations

This systematic review applied a wider search strategy and eligibility criteria than previous systematic reviews, resulting in the inclusion of a larger number of studies, allowing more evidence to contribute to the results.

This is the first systematic review of apathy measures to use guidelines to assess study methodological quality and results criteria that were specifically designed for studies of patient reported outcome measures. This built on previous systematic reviews that used criteria designed to assess studies of criterion validity only. In lieu of guidelines for clinician rated and informant reported measures, the COSMIN guidelines were applied to these measures also, whilst clinician observation-only measures were excluded. This was applicable to all measurement properties, however, as the guidelines were

developed for patient reported measures specifically, the required evidence for content validity emphasised the involvement of patients. This set a high standard for many studies where sometimes this was arguably not necessary. For example, the guidelines argue that involvement of patients is essential in the development of a patient reported outcome measure. However, if the measure is based on clinician ratings, it may not have been necessary to penalise studies for not including the views of patients, carers or members of the public. Nevertheless, clinician-rated measures typically involve standardised questions with patients or carers, so their understanding of these questions is still key to good content validity.

This study only included measures which were based on patient or informant reports or interviewer or clinician ratings. There has been increasing literature investigating how technology can be used to assess apathy. For example, accelerometers and other experimental methods have recently been used to assess apathy [235–237]. These are however limited to certain aspects of apathy such as motor activity, and be better suited to reviews of diagnostic accuracy when compared against a gold standard measure of apathy once this has been developed.

One of the aims of this review was to explore how apathy was defined, however part of the eligibility criteria was that studies must be of a measure of apathy. Studies were excluded if they only measured part of apathy e.g. emotional blunting, or a similar but different concept e.g. negative symptoms. Therefore the findings of this aspect of the review are limited to the initial definition of apathy, informed by conceptualisations by Marin [38] and Robert and colleagues [36]. Nevertheless, a variety of content was still found amongst the included apathy measures.

To my knowledge, this is the first review of apathy measures to engage PPI members. PPI members helped set criteria for and guide decisions regarding content validity, reliability and GRADE (see Appendix 5), and were involved in reviewing a proportion of measures for content validity. When reviewing measures, PPI members noticed issues with the measures that had not been anticipated by the reviewer, though they sometimes found it difficult to decide whether the issue was inappropriate wording or lack of relevance to the population. Members instead preferred to make suggestions for alternative wording. This indicates that including patient or public representatives in the development of measures is highly feasible, and would offer an invaluable insight that is lacking from many of the development studies seen in this review.

Previous systematic reviews of apathy measures used QUADAS which was designed only to assess studies of diagnostic accuracy, and applied these to studies of a variety of measurement properties. COSMIN on the other hand provides guidance for a variety of measurement properties, allowing a

comprehensive review of apathy measures. No measurement properties were found by this review that were not anticipated by the COSMIN guidelines. This is the first review to systematically assess the structural validity and content validity of apathy measures. Content validity is particularly important as it determines what is actually being measured [168]. Additionally, this review assessed measurement properties of apathy measures for both a dementia and older adult population. A measure that only has good properties in one population is not suited to trials that wish to assess apathy over the course of disease, from a pre-clinical state. This review was also being conducted to select the best measure via which to assess apathy in a study where participants were aged 65 or over, and had a diagnosis of dementia or MCI (further details provided in section 6.2), therefore it was essential that the measure selected was suitable for this population.

Whilst the COSMIN guidelines aim to reduce subjectivity, there was a need for further decisions to be made beyond the guidelines and criteria. As a result, the review process became iterative, as during the process of the review, a scenario would arise for which there was no clear guidance: a decision was made by the reviewer, then everything previously reviewed that could be affected by this decision was re-assessed to maintain consistency. Whilst this need for subjective decision-making would likely be true of any systematic review guidelines for rating studies of measurement properties, this was particularly perpetuated by the presence of “any other methodological flaws” which could see a very good study rated as very poor. Future amendments of the COSMIN guidelines may benefit from anticipating these other methodological flaws and state them specifically as part of risk of bias assessment, rather than leaving this to researchers to decide. The issues around this can be minimised by duplicating all rating activities, however, due to the large number of studies found by this review, this was impractical. The duplication of review for a portion of the included studies did however help discussions around what these flaws may be, limiting subjective decisions.

COSMIN quality criteria are binary and risk of bias guidelines are based on a ‘worse score counts’ method, which risks over simplifying the complexities of the true measurement properties and research evidence. Prinsen and colleagues [166] note that there are no current guidelines for upgrading studies, suggesting that this could be an option in the future, which would somewhat account for this problem. An alternative for assessing quality of measures could have been that used by Radakovic and colleagues [161] which rated each result on a scale of four to six possible scores depending on the measurement property being assessed. However this does not appear to have been developed in a systematic way, unlike COSMIN criteria that were created following a Delphi procedure.

It is worth noting that the majority of development and content validity studies were poorly reported, and the COSMIN guidelines set a high standard for these studies that they failed to live up to. As a result, content validity was largely determined entirely by reviewer ratings of the measure itself, meaning that content validity findings are not evidence-based, but decided by the review team. Bias in reviewer ratings was minimised by following COSMIN guidelines for direct reviewer rating of measures, and creating additional criteria, informed by PPI, that could be followed for all measures. PPI members also reviewed four of the measures directly. Using the reviewer rating technique ensured a validity rating was possible, even in the presence of insufficient information from the development and content validity studies.

There is a clear need for better more detailed reporting of the development of measures. Participants should be involved, and professionals should adopt a clear and transparent method for developing and reviewing items. This would prevent bias in rating of content validity. The high standards set by the COMSIN guidelines and criteria were also perhaps sometimes too high and unattainable. For example in a development study, a lack of justification of recall period and response options can prevent the results of a development study being rated as sufficient, yet these aspects represent a small part of the scale, and are rarely provided by even the best quality studies. A weighted criteria for content validity which places greater emphasis on the items may be preferable.

4.6. Conclusion

A number of measures that assess apathy of varying quality are available. In general, these measures assess indicators of apathy related to behavioural, cognitive, emotional and social domains, in addition to general motivation and drive. The development of measures was generally poor, due to lack of transparency and a systematic approach eliciting and refining items and developing the other measurement aspects such as recall period and response options. Future development of measures should include a clear and systematic approach at all stages, and involve patients or members of the public as well as professionals to ensure good content validity. Despite generally poor development, the COSMIN approach enabled conclusions to be made about content validity even in the absence of development and content validity studies. This review indicates that the LARS has good measurement properties, though could be less feasible to use in studies with limited resources, and may be more burdensome for participants. The DAS, in particular the resource efficient b-DAS, is a promising measure that requires more research into its properties, particularly reliability. The AES has evidence for good measurement properties for people living with NCD and is easy and relatively quick to administer. As a result, the AES was selected as the tool to

measure apathy in the network analysis study, methods for which are described in Chapter 6.

Chapter 5. Methodology, Rationale and Aims

5.1. Introduction

What constitutes apathy and how it is experienced by people living with NCD is poorly understood and under-researched [28,153]. It is important to explore the relationship between specific indicators of apathy, depression, and other relevant factors [69]. It has been argued here that apathy should be understood within the ICF model, in which impairments, activity, participation, and environmental and individual contextual factors interact [139,238]. Neurocognitive models discussed in section 3.3 only focus on the level of impairment, whilst attempts to integrate biological and psychological explanations with a wider social context have been insufficiently developed and largely ignored. A broader exploratory approach to apathy is needed to understand how to conceptualise it, and identify possible mechanisms to produce a conceptual model of apathy that is needed to enable further investigations [107].

I began this PhD with limited scrutiny of the positivist assumptions that underlie much of apathy research. However, throughout my studies, in particular, while conducting the systematic review (described in Chapter 4), and studying for the Philosophy of Research module, I became increasingly aware that this positivist epistemology and traditional quantitative approach was not appropriate. Whilst apathy was proposed to be a unidimensional observable construct, the systematic review and other literature discussed above highlighted that apathy research thus far has been inconsistent and only explains some of the variance in apathy some of the time [69]. The items included in measures are selected based on quantitative consistency, yet results frequently do not replicate. Consistent attempts to study apathy as a unidimensional disorder, distinct from related concepts such as depression, have not been successful. The systematic review of apathy measures additionally revealed that apathy was measured in varied ways, using various items, but all deemed to measure the homogenous construct of apathy. This similarly occurs in measures of depression, which have received criticism for this [239]. It became increasingly clear that apathy was not successfully studied as a unidimensional singular construct. It is likely that attempting to work within definitive boundaries between apathy and the depressive syndrome is an impossible and unhelpful task, and a more flexible approach is required.

Epistemology refers to the theory of what constitutes knowledge, and thus will shape the methodology and methods used in this thesis [240]. This chapter presents the epistemology that underlies this research and the rationale for the methods chosen. A brief explanation of one of the methods, network

analysis, is also provided as this is a novel method within this field and requires elaboration. Methods are reported in full in Chapter 6.

5.2. Epistemology

5.2.1. Positivism and The Disease Model

The positivist epistemology is the implicit, “nocturnal philosophy” [241, p.102], that is pervasive in ‘scientific’ and social research. For positivism, the aim of science is to produce generalisable laws and predict new phenomena [242]. Causality, according to positivism is akin to a ‘black box’: though input and output can be observed, what occurs inside is not observable, and therefore can never be known [243], so causes are instead understood in terms of empirical regularities [244]. Empirical regularities are used to generate universal laws, that are enduring across time and space, providing the required conditions are met [245]. However, positivist explanations of phenomena rely on describing the conditions in which it occurred; thus prediction and explanation are circular as they both rely on empirical regularities [246].

Psychopathology research has traditionally been based on a disease model, which is underpinned by positivism. The disease model posits that symptoms co-occur because they have a shared underlying cause, i.e. a disorder that exists independently of the symptoms it produces [247]. In this model, symptoms are equal indicators of the disease, and as such do not causally interact and are considered interchangeable.

The study of apathy is dominated by this positivist epistemology and disease model. Apathy is consistently understood in relation to its observable ‘symptoms’, such as reduced goal-directed activities, and attempts have also been made to reconceptualization apathy purely as a loss of, or reduction in, goal-directed activity [32] to provide a directly observable and thus measurable concept, consistent with the goal of positivism. Positivism and the disease model predicts that disorders are distinct and observable, and thus would expect their aetiology and treatment to be specific, and diagnoses to be consistent [247].

Findings consistently fail to meet predictions of the disease model, indicating a problem with its underlying epistemology [247]. The psychiatric positivist approach assumes the ontology of psychological disorders, and conflates knowledge of the world (e.g. the diagnostic labels we use) with the reality of the world (e.g. the independent existence of these labels) [247]. Whilst this disease model holds for physical disorders, such as cancer, where a tumour exists independently of a set of symptoms that it produces, it is problematic for psychological disorders, whose evidence of existence relies on the symptoms themselves [247,248]. Diagnostic criteria, such as those proposed for apathy,

require that individuals experience a certain set of symptoms for psychopathology such as apathy to be identified. It cannot be experienced, identified, or diagnosed independently of the symptoms.

5.2.2. Critical Realism and the Network Perspective

In contrast, critical realism distinguishes between our theories of the world, referred to as ‘the transitive dimension’, and how the world really is, ‘the intransitive dimension’ [249]. Critical realism argues that the conflation of these two dimensions is an ‘epistemic fallacy’ [250]. From a critical realist perspective, psychological disorders and categories can be considered a label of the transitive dimension. Although these psychological labels can have real consequences, and are used to refer to real experiences and behaviours or symptoms within the intransitive dimension, they are not necessarily themselves ontologically ‘real’ [247].

Furthermore, viewing symptoms as independent, exchangeable indicators of a disorder is not wholly consistent with commonly accepted clinical and theoretical understandings of how psychological disorders develop and are sustained [251]. In people with psychological disorders, symptoms are thought to interact and reinforce one another. For example, in apathy, difficulty in performing activities may lead to loss of interest, which may in turn lead to reduced activity [146].

The ‘network perspective’ on the other hand does not require that symptoms are independent of each other. Instead, symptoms are proposed to causally influence one another [248,252]. From the network perspective, it has been argued that symptoms constitute rather than reflect a disorder, and as such may be better termed ‘indicators’, and so disorders are considered to be the clustering together of symptoms in a densely connected network [248,253,254]. Boundaries between disorders are blurred, as there are common ‘bridge’ symptoms that are present in different disorders, connecting them to one another, resulting in co-occurrence [248]. This causal interaction of symptoms is consistent with many clinicians’ understanding of diseases, such as depression which is frequently understood and treated by clinicians on the basis of its symptoms having differing importance and interacting with one another [239].

Whilst the network perspective has sometimes described disorders as ‘emergent’ of symptoms [254], disorders are not necessarily claimed to be ontologically real. Indeed their lack of ontological reality is suggested by numerous authors, who contrast the description and diagnosis of psychological disorders, that cannot be described independently of their symptoms, with physical disorders, that can be detected separately [254]. Therefore, it is argued that the network perspective avoids the epistemic fallacy, as disorders can be understood to exist within the transitive

dimension, whilst symptoms, such as lack of energy, can be understood to exist within the intransitive dimension.

Methods used to study concepts such as apathy and depression, and whether this is consistent with the network perspective, will now be discussed.

5.2.2.1. Justification of Quantitative methods

Positivism maintains that researchers can achieve 'objective', un-bias, value-free observation [244]. Quantification is often seen as synonymous with positivism, [255] as quantitative methods such as standardized questionnaires are seen to remove possible sources of bias [256], and enable replicability and standardisation, that, for positivists, makes generalisations possible [257]. Statistics enable the reduction of variance to a single measure of central tendency that is used to represent the sample, supporting predictions and generalisations [243].

For critical realists, research should aim to produce detailed understanding and explanations of mechanisms, in contrast to positivism which aims to create predictions and generalisable laws [256]. It has been argued that quantitative methods, in particular inferential statistics, are incompatible with critical realism because: (1) Statistical methods are reductionist and contradict the open system ontology, as a discrete number of variables are used to summarize complex information [258,259]; (2) Correlations are frequently used to make causal conclusions, yet regardless of how frequently an association is observed, this provides no indication of causation itself [260]. However, it is argued that these are misdirected criticisms of positivism, and quantitative methods are consistent with a critical realist epistemology because: (1) whilst statistical analysis may apply a closed system method, this does not necessitate the assumption of a closed system reality [258]. If results are acknowledged to be indicative of, but not equal to the complex reality which they represent, then the reality of the world need not be assumed to be equal to the closed system of a research study; (2) In critical realism, the world contains 'demi-regularities', in which some patterns of associations can be identified. Rather than seek empirical regularities from which generalised predictions are made, demi-regularities can be explored to identify locations at which to theorize possible mechanisms [261]. It is important here that observations are not equated to reality itself, and correlations are not equated to causation. Whilst positivism may have mistakenly equated quantified data with facts, this is not inherent in quantitative methods, in which data can be treated as 'ficts' (i.e. possible relations that are worthy of further explanation) [258]. A critical realist approach emphasises that our knowledge of the world exists in the transitive dimension, and should not be equated with reality. Research necessarily takes place within an open system, and because there is an external reality that cannot be directly accessed, researchers must apply cautious abstraction of constructs and their relations [262].

This supports a careful exploration of the construct of apathy. Quantitative (extensive) approaches can be used to identify demi-regularities, whilst qualitative (intensive) approaches can be used to formulate possible mechanisms that might explain these observed relationships [256]. These two methods can be integrated in a mixed methods design, consistent with critical realism [263], in order to further our understanding of apathy in NCD.

5.2.2.2. Problems with traditional quantitative methods in apathy research

Though I have argued that a network perspective is required for the study of apathy, and have justified the use of quantitative methods in its exploration, I will now outline reasons why traditional applications of statistical methods are inappropriate here, and describe an alternative quantitative method.

The study of psychological constructs has traditionally involved the use of statistical methods such as factor analysis and item response theory in a way that is consistent with the common cause model [239,251]. Traditionally, these methods are used to assess whether items relate to one another sufficiently to be considered indicators of one common cause. This is used to validate the use of a single sum-score, which is used to represent an individual's position on the common cause or 'latent trait', such as a psychological disorder. Shared variance between items is assumed to be the result of the latent trait, and the remaining variance is considered residual error [251]. As such, the relationships between individual symptoms are not explored. This sum-score is then used to assess how the latent trait (an assumed ontologically 'real' disorder) relates to other variables, such as risk factors for the disorder. Using this approach, the shared variance of the predictors with the dependent variable is assessed, but the relationships between the symptoms themselves or their differential relationships with the dependent variables are ignored [239]. Whilst these methods are not problematic in themselves, in psychopathology research they can result in the inappropriate reduction of complex heterogeneous phenomena to a single sum score, resulting in diluted or contradictory findings across studies, and a failure to explore which symptoms are important both within disorders and in connecting different disorders [239,248].

Measurement of apathy using standardized questionnaires that rely on external indicators of apathy is common, and measures are used to calculate an apathy score to which a threshold to indicate 'abnormal' apathy is applied [162]. Apathy is proposed to be composed of domains [30,36,39], and the distinction between apathy and depression has been previously argued based on appealing to the supposedly distinct symptoms they separately produce [30,38,54,71].

As outlined above, the network perspective views symptom interaction as highly relevant and important, as it is thought that symptoms interact with one another, and constitute a disorder which exists in the transitive dimension (i.e. is not ontologically real) [254]. An alternative quantitative method that is consistent with this perspective has recently begun to gain popularity in the field of psychopathology [264]: ‘network analysis’. As this is a novel method in this field, the method will be outlined before the small number of studies using this method to explore apathy are discussed.

5.3. Network analysis, an explanation and rationale

Network analysis is a method that enables the production of network models (that can be visualised as maps) to describe relationships between nodes [248]. Network analysis has been used in various fields, such as sociology, physics, and neuroimaging [265]. Nodes (circles on the map) represent small unitary constructs, such as individuals, as in a social network map. Edges (lines on the map) are used to represent relationships between nodes [248]. Networks can be assessed for their structure, for example, take a social network map of relationships between students, where nodes are students, and edges represent a ‘friendship’ on the social network platform Facebook. The network structure can be assessed to see which particular students are Facebook friends, but also who has the most Facebook friends, which students are linked through a mutual friend, and how far ‘removed’ students are from each other.

Network maps can be directed or undirected, signed or unsigned, and weighted or unweighted [266]. An undirected network is composed of edges with no indication of direction, whilst in a directed network, the edges are arrows that point from one node to another, to indicate the direction of a causal relationship. Edges in a directed network can also display bidirectional relationships through double-headed arrows. For example, a directed social network map could indicate which person ‘follows’ the other on Twitter (a social network platform in which individuals choose to follow each other and do not need to mutually agree to do so). An unsigned network does not indicate the sign of the relationship (typically indicated by a black line), whilst in a signed network, the edges are colour coded, to represent a positive (typically indicated by a green or blue colour) or negative relationship (typically indicated by a red colour). For example, in the aforementioned Facebook example, a green edge could indicate a ‘friendship’, whilst a red edge could indicate where someone has been ‘blocked’. In an unweighted network, edges are all the same thickness, and do not indicate strength of a relationship, whereas in a weighted network, the thickness of the edges are representative and proportionate to the strength of a relationship. For example, a weighted social network map could be produced by enabling the edge weight to represent the length of time that they have been Facebook friends, or how

many ‘likes’ they have given each other, relative to the other relationships in the network.

5.3.1. Psychopathology networks

Network analysis is a novel method for psychopathology research [251], and has gained popularity in the last few years [264]. In social network models such as the social media examples above, edges can be objectively ascertained i.e. the presence or absence of, or length of time of a Facebook friendship is directly observable. However, in psychopathology networks, where the nodes often represent symptoms, edges (e.g. relationships between psychological symptoms) cannot be directly observed. For critical realists, symptom interactions may operate at the level of ‘real’ but are not directly observable. Edges in psychopathology networks are estimated based on observations, such as patient reported symptoms measured by an item or just a few items of a questionnaire [239]. Observations are typically of a group of participants or participants over multiple time points [248].

5.3.1.1. Pairwise Markov Random Field network models

Pairwise Markov Random Field network models are a type of model in which (undirected) edges reflect the strength of the ‘conditionally dependent’ relationships, i.e. the relationship that remains when controlling for all the other variables [266,267]. Estimation of Pairwise Markov Random Field models can be performed using partial correlations with continuous data, or log-linear relationships with ordinal data. As with any statistical method however, these conditional dependencies do not equate to causality: whilst it may occur due to direct (of either direction or bi-directional) causal mechanism, it may also be the result of not accounting for a common effect or common cause [268]. Conditional dependency psychometric networks, such as partial correlation networks therefore offer an exploratory method of identifying *potential* direct relationships between individual symptoms for future research and theorisation of explanatory mechanisms [269,270].

5.3.1.2. How does network analysis differ from traditional methods?

A typical method of assessing the relationship between apathy with other constructs such as depression may be to assess these constructs using a regression analysis, and a typical method of assessing the clustering of symptoms of apathy may be factor analysis, item response theory methods or structural equation modelling. Whilst the statistical methods behind network analysis and these traditional approaches may be similar, their use and assumptions differ.

Factor analysis, item response theory and structural equation modelling assume that the items in a measure relate due to the presence of an underlying latent variable, and they are used to assess the shared variance of

these items [239]. In contrast, the network models discussed here assume that items directly associate with one another, rather than interactions being due to the product of a latent trait, and are used to identify the conditional dependent relationships; i.e. the unique variance between each pair of nodes [239,271,272]. As the unique variance, rather than shared variance, is what is of interest, the relationship between individual symptoms can be assessed.

Similarly, traditional psychopathological research has used regression analyses to assess apathy and depression as latent traits measured by a sum-score, against dependent variables, such as cognition and ADL [e.g. 74,77]. In contrast, network analysis enables the assessment of which symptoms are driving these interactions, which may help develop a better understanding of how these relationships occur, and enable the development of more targeted and thus more efficient treatment methods [239].

5.3.2. Interpreting Psychopathology Networks

In a psychopathology Pairwise Markov Random Field network, the network of symptoms is the disorder itself, and distinct clusters of symptoms are indicative of distinct disorders [248,266]. The centrality of nodes can indicate which symptoms are most, and most strongly, connected to other symptoms. This has practical implications as a symptom that is peripheral to the network may be of little interest to further research, but a symptom that has numerous and strong connections with other symptoms will likely be of relevance to helping reduce the symptoms experienced [248]. Similarly, bridge centrality can be estimated, which assesses centrality of a node only taking into account its connection with nodes from other groups [273]. This is of interest as it could indicate what symptoms might explain comorbidity, and which symptoms may be useful targets for treatment to manage multiple disorders [273].

5.4. Rationale for Network Analysis Methods

Before causal mechanisms can be hypothesised and tested, relationships that occur between symptoms across individuals, indicating candidate locations for explanatory mechanisms, first need to be identified [251]. As discussed, traditionally, quantitative research into apathy has relied on factor analysis or item response theory to examine the dimensionality of apathy, and regression analysis to explore how apathy relates to other concepts. However, network analysis could offer different insights, and has been used in recent years to assess comorbidity or boundaries of disorders such as depression, identify bridge symptoms and assess symptom centrality and network connectivity [274].

When I first decided to use this method for this thesis, network analysis was “just taking its first steps in the field of neurodegenerative disease” [275,

p.1089] and no studies assessing apathy with network analysis had been conducted. To my knowledge, there are now four studies that have included apathy in a network analysis, in which apathy has been investigated in the context of depression or as a single node in a neuropsychiatric symptom network.

Of the studies conducted thus far, two used the GDS-3a (a 3-item measure of apathy originally intended to assess depression) [276,277], one used the NPI-apathy (as a single binary node) [278], and another used both [275]. The study that used two measures of apathy found that whilst the GDS-3a items connected relatively strongly, they did not form a cluster, nor did they connect to the NPI-apathy node, and were connected to various depression nodes [275]. However, they did find distinct clusters of 'lack of positive affect' (from both apathy and depression measures) and 'depressive symptoms and decline', which may represent the distinction between depression without sadness, and depression with sadness that was discussed in section 2.5.5.2. In contrast, another study found that the items of the GDS-3a were not well connected to each other, nor the GDS depression items in the network [276]. Conversely, a similar study using the same measures found that items of the GDS-3a clustered together but were also highly connected to the GDS depression items [277]. This difference was explained as a result of cultural differences regarding older adults perception of activity in South Korea compared to America [276]. Alternatively, this could be due to the differences in nodes, as the former study only assessed the items of the GDS [276], whilst the latter assessed the items of the GDS in addition to nodes that assessed 'dementia at follow-up' and ADL [277]. In other words, when controlling for functional ability and cognitive decline, symptoms of apathy may cluster together more. Regardless, these studies indicated that apathy does not form a clear cluster that is distinct from depression.

These studies highlight the importance of network analysis, as they show findings that would not have been revealed by a sum-score approach, such as the identification of central symptoms [276] and key symptoms for predicting outcomes such as dementia progression [277]. However, they also highlight issues with the measures used. Further research using network analysis is required to enhance our understanding of how symptoms of apathy and depression and ADL interact [278]. This should use a dedicated measure of apathy [277], and a different measure of various symptoms of depression [276]. This is what the present network analysis sub-study aims to achieve.

5.5. Rationale for Qualitative Methods

Whilst quantitative methods can be used to identify relationships or demi-regularities, qualitative methods enable a deeper understanding as to why these might occur, facilitating model development [279]. Critical realism

recognises qualitative methods as a way of identifying possible causal explanations, whilst also acknowledging these are not necessarily complete nor correspond with the 'true' mechanisms at work [256].

There is a need to understand the individual reactions and emotions that people experience in dementia [280]. Little is known about the influence of carers and the environment on apathy [153], and the perspectives of carers and people living with dementia or MCI is under researched [281]. Qualitative methods are needed to further our understanding of apathy [85,145], yet studies exploring these lived experiences have been lacking thus far [282]. At the time of starting this PhD, there were no published studies exploring apathy in people with dementia, MCI or their carers, and just one qualitative study had been published that explored apathy in a neurodegenerative disease, which was Parkinson's disease [146]. This novel study shed new light on apathy in various ways. For example, it revealed participants' awareness and frustration with their own apathy, in contrast to the frequent claim that apathy is not itself distressing due to the inherently impaired emotional processing that accompanies it [146]. However, this study was conducted in males with Parkinson's disease, who are likely to experience apathy differently from people with dementia or MCI, carers, and females.

Three more studies investigating apathy using qualitative methods have recently been published: one with four people with Parkinson's disease, one carer and healthcare professionals [283], one with six people with AD [284], and a sister study with the same sample's six carers [285]. Together, these studies highlighted the complex nature of apathy and the need to understand it within the wider social and environmental context. Though studies of people with Parkinson's disease can provide interesting insights into apathy, the experience of people with Parkinson's disease may differ from that of people living with dementia and MCI. For example, societal perceptions of the two disorders may differ, cognitive symptoms, though present in Parkinson's disease, are not the main feature, and the amount of support or information from services could differ. Furthermore, of the two studies of people with AD and their carers, lack of diversity was noted as a limitation, as all participants were from one local area, were White British, had the same diagnosis, had a co-resident spousal carer, and only included one female with AD and one male carer. It is important to conduct further qualitative studies to understand the lived experience of apathy for a diverse group of people and their carers.

5.6. Critical Realism and Mixed Methods

The aim of research for critical realism is to develop explanations of the mechanisms underlying phenomena [256]. Mixed methods are particularly useful for research questions in which there is a need to corroborate and explain results [279], so mixed methods is particularly suited to critical realist

research. Mixed methods were chosen for this research, as it enables a detailed description of apathy in NCD in addition to the identification of possible mechanisms.

Eastwood et al. [263] outlined a three stage critical realist method of explanatory theory building. This process begins with an emergent phase, in which the phenomenon is thoroughly described and relationships are identified using both quantitative and qualitative data. Next, there is a construction phase, in which the mixed methods findings are integrated and further triangulated with previous literature. Tentative theories are compared and contrasted to identify the best possible explanations. Then a final confirmatory phase is used to develop hypotheses from these explanations, and test these with new studies.

The present study will use a critical realist informed mixed methods approach to answer the following research questions:

1. How should apathy be characterised and what are its boundaries in people living with NCD?
2. What are the possible mechanisms and impact of apathy in people living with NCD?

5.7. Overall aims and objectives

Table 7 presents the overall research questions, aims, and objectives in an implementation matrix, as recommended by Creswell and Plano Clark [279] for mixed methods research. This does not include the first aim of this thesis, which was to determine what measures of apathy are available and their quality for use with people living with NCD, as this was addressed by the systematic review, outlined in the previous chapter, and did not form part of the mixed methods study.

Table 7. *Implementation Matrix for this mixed methods study*

Aim 1: To characterise apathy and its boundaries:		
Research questions	Objectives	Methods
What constitutes apathy?	To determine how apathy indicators might influence each other	Network analysis; Qualitative
	To determine which apathy indicators are most important	Network analysis; Qualitative;
	To assess if there are distinct domains within apathy	Network analysis
How do apathy and depression indicators relate?	To assess if apathy and depression form distinct domains	Network analysis
	To determine how depression and apathy indicators relate to each other	Network analysis
How do people living with NCD and their carers understand apathy?	To explore participants' understandings and beliefs about apathy (including reduction in activity, interests, emotions and social engagement)	Qualitative
Aim 2: To understand the possible mechanisms and impact of apathy:		
Research questions	Objectives	Methods
What are the mechanisms of apathy indicators (including reduced motivation, activity, initiation, interest, concern, emotional reactivity, and social engagement)?	To determine how external factors relate to apathy indicators	Network analysis
	To explore experiences of and explanations for apathy indicators	Qualitative
	To develop a model of apathy	Produced from the integrated mixed methods
How does apathy impact the everyday lives of people with NCD and their carers?	To explore how apathy impacts the everyday lives of carers	Qualitative
	To explore how apathy impacts the everyday lives of people living with NCD	Qualitative

Chapter 6. Methods

6.1. Introduction

This chapter presents the methods for the overall PrAISED2 study, the separate network analysis and qualitative methods used particularly for this PhD thesis, a consideration of the overall ethics of this study, and a comment on how this work will be integrated.

6.2. Design and Procedure

This work was situated within the context of a large, pragmatic, parallel, multi-centre Randomised Controlled Trial (RCT) to test the effectiveness of an intervention which aimed to improve activities of daily living at 12 months (PrAISED2; ISRCTN registration:14015320670), methods for which are outlined in a published protocol article [1]. For my PhD work, I used the opportunity of a large sample of people with NCD and their carers, which would not otherwise have been possible to obtain in a standalone PhD project.

6.2.1. Setting and Participants

Participants were dyads of people with dementia or MCI and a carer taking part in the PrAISED2 study. Patient inclusion criteria for the PrAISED2 study were:

- Aged 65 or over
- Diagnosis of any dementia (except dementia with Lewy Bodies) or MCI
- Montreal Cognitive Assessment (MoCA) score of 13 to 25 (inclusive)
- Has a family member, friend, or other carer who knows them well and is willing and able to participate
- Able to communicate in English
- Has sufficient dexterity, sight, and hearing to take part in cognitive testing
- Able to walk without human help
- Has capacity to consent to the study

Exclusion criteria were:

- Plans to move out of area or having a life expectancy of less than twelve months.

- Comorbid conditions that would prevent participation (for example severe: pain, breathlessness, psychosis, Parkinson's disease or other severe neurological disease)

6.2.2. Sample size and Recruitment

The target sample size was set by the study's clinical trials unit team at 368 participants, which was deemed sufficient to detect changes in ADL, to a medium effect size, for the purposes of the PrAISED2 study aims. Sample size specific to the network analysis sub-study is discussed later (in section 6.3.1.1).

Participants were recruited into the PrAISED2 study from five different localities in England: Nottinghamshire, Derbyshire, Lincolnshire, Bath and North East Somerset and Oxfordshire. Participants were recruited via memory assessment services (including associated support groups), the Join Dementia Research register, and General Practitioner (GP) practices. Join Dementia Research is an online platform for people interested in taking part in dementia research, which matches its members to research projects. Researchers could screen and contact potential participants from the Join Dementia Research network directly. Participants recruited via memory assessment services were referred to a researcher after having expressed interest in the research, and the researcher contacted them to provide further details and discuss the study. Where GP practices sent letters to eligible patients registered at their practice, patients expressed their interest via phone call or postal reply directly to the research team.

6.2.3. Procedure

For all recruitment methods, a member of the research team discussed the study and screened for eligibility where possible prior to arranging a research visit. Interested participants (patients and carers) were provided with an information sheet. A short summary version was also included for ease of understanding. If the participant was still interested in taking part after having read the information, a consent and baseline visit was arranged (at least 24 hours from being given the information sheet). The consent process is outlined in section '6.5. Ethics'

Researchers (including myself) visited patients and carers together in pairs, to explain the study, answer any questions, and if the participant was still interested, to proceed with consenting the patient and carer to the study. The researchers formally assessed patient capacity prior to them consenting to the study. After informed consent was taken, the baseline assessments were performed. Patients were also provided with a wrist-worn accelerometer to wear for a week.

When the researcher returned from the visit, they used an online randomisation system, managed by the clinical trials unit, to randomly allocate patients to either a control or active therapy programme.

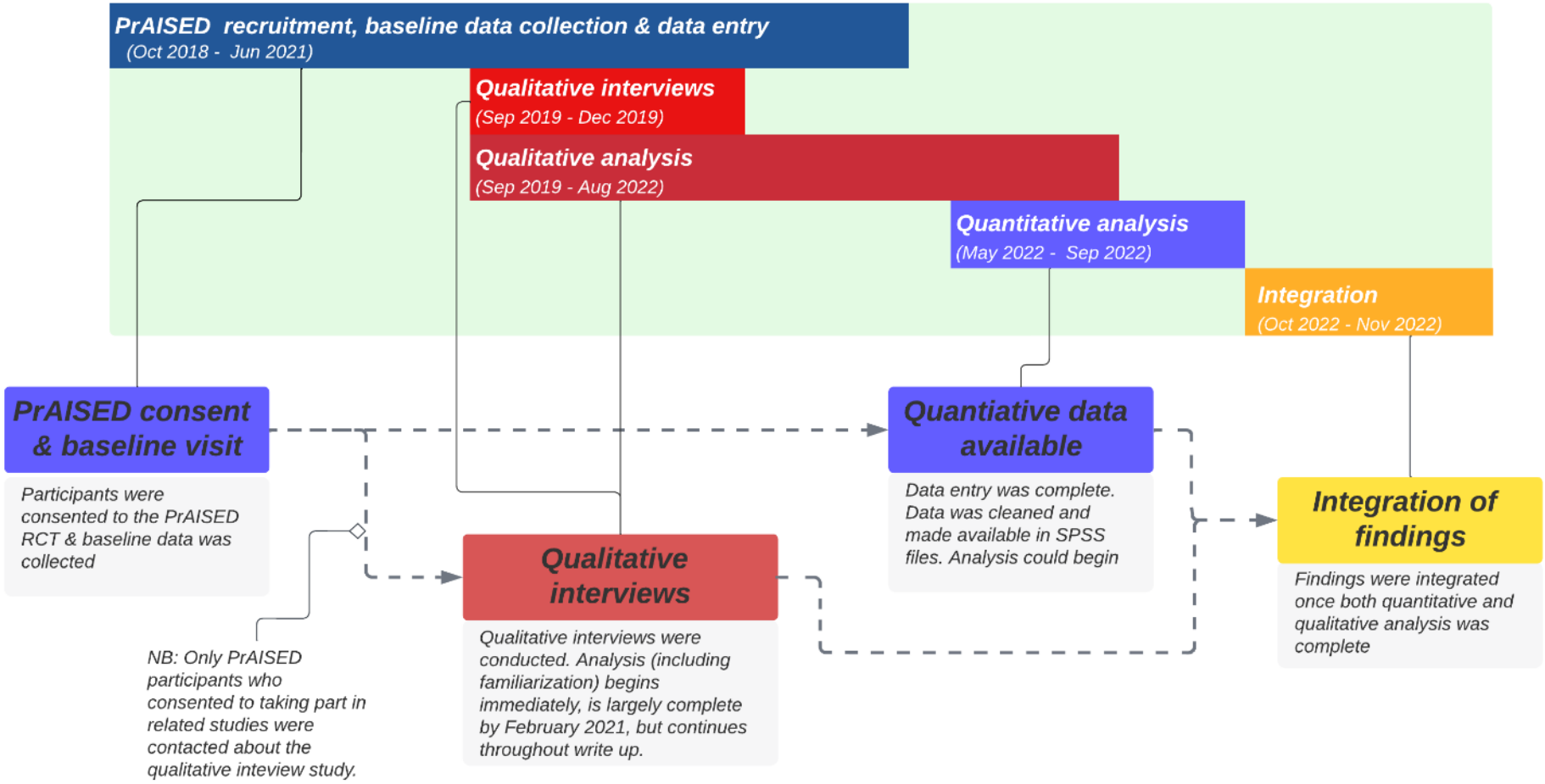
Participants and carers were visited at home by researchers to collect baseline (0 months) and follow-up (12 months) data, though some follow-up data was collected remotely (see section 9.5.1). Assessments were of ADL, memory and executive functioning, balance, mobility, frailty, falls history, physical activity, quality of life (including health-related), mood, personality, fear of falling, service use (including hospital admissions), carer strain and carer health-related quality of life. Participants were visited for a baseline assessment, and the same assessments were repeated twelve months later (at follow-up). Data was entered into an online database, 'MACRO', by researchers across the five sites. This database was held by Bangor University's 'North Wales Organisation for Randomised Trials in Health' clinical trials unit, who were responsible for the management, cleaning, and analysis of data for the PrAISED2 RCT. The clinical trials unit identified outliers in the data to be checked against hard copies, and a random 5% of all data for full checking against hard copies was also identified by the clinical trials unit, as per their analysis plan. This data checking was performed by researchers across the individual sites, including myself. Quantitative data was obtained for the present analysis by requesting the data from the clinical trials unit. Data was provided in its raw form in SPSS files.

6.2.4. Mixed methods design

The above section discussed the methods for the PrAISED2 RCT. This next section will discuss the methods specific to this thesis.

A convergent parallel-databases mixed methods design was chosen for this thesis, as this offered a method via which to collect complementary data that is considered equally important, whilst also maximising time efficiency [279]. The design was not fully parallel, as the qualitative sample is nested within the network analysis sample. Data was collected concurrently overall (as quantitative recruitment and baseline data collection occurred whilst qualitative recruitment and data collection took place), though this was sequential on an individual level (as all participants completed quantitative data collection before being invited to take part in the qualitative study). Data analysis was also performed separately. Figure 2 shows the mixed methods research process for this thesis.

Figure 2. Mixed methods process chart of key research stages.



This mixed methods approach was guided by the critical realist theory construction process described by Eastwood et al. [263], outlined previously. This thesis relates to the first two phases of this process: the emergent phase, and the construction phase; as the confirmatory phase was considered beyond the scope of this PhD. It does however differ from the proposed methods by Eastwood [263] as the emergent phase is not entirely inductive, as previous literature informs this phase, for example, diagnostic criteria literature informed the interview questions in the qualitative study.

Integration is essential to the success of mixed methods research, which enables the findings to go beyond that of separate network analysis and qualitative studies [279]. In order to ensure it is possible to integrate these two studies in this way, guidance by Yin [286] was followed: in both studies the research questions overlap; individual participants form the basis of analysis; the qualitative sample is nested within the network analysis sample and data were collected on specific domains of apathy in both studies; and side-by-side comparison of results in a narrative discussion will facilitate discussion on convergence, divergence, agreement and discrepancies between the two methods [279].

The results were then used to produce a tentative model. This helps analysis move beyond description and instead focus on mechanisms [287].

6.3. Network analysis sub-study Methods

The methods specific to the network analysis sub-study will now be described.

6.3.1. Design, Setting and Participants

This network analysis sub-study used a cross-sectional study design. All participants from the PrAISED2 RCT (methods for which were described in section 6.2) baseline time-point were selected for the analysis. To prepare for this novel analysis method, I attended a five-day course at the University of Amsterdam, and sought advice from an expert in the field (further details outlined in ‘Statement of Contribution’).

6.3.1.1. Sample size calculation

In network analysis, as in similar analyses such as regression, determining whether there is sufficient power to detect relationships is effected by their size, so sample size calculations would require information from a similar previously estimated network [270]. This is the first study to conduct network analysis of the varied apathy indicators in people living with NCD, so a sufficiently similar estimated network is not available from which to draw the information required. Studies with 81, 91 and 306 participants using 20, 8 and 22 nodes (i.e. variables) respectively have previously been published [288–290], tentatively suggesting that 368 participants is a reasonable sample size

for this analysis with 18 nodes in the first phase. Though later phases of analysis involve the addition of nodes, making the detection of relationships by chance increasingly likely, I aimed to simultaneously reduce the number of nodes, by removing redundant items (see section 6.3.4.1.5). Furthermore, power analysis is based on the positivist tradition of hypothesis testing, so the concept of power is less relevant to exploratory studies such as these [291].

6.3.2. Measures

Measures used in this study were selected from the PrAISED2 RCT assessment battery. The inclusion of an apathy measure was made for the purposes of this PhD, however the choice of other measures was limited to what was available in the PrAISED2 RCT.

It is of particular importance in network analysis that the nodes (i.e. variables) selected are the minimum required to assess the network, and that nodes measure a single construct that is distinct from the other nodes [292]. It is inappropriate in network analysis (as in other analyses) for nodes to measure the same specific underlying construct, however, validated questionnaires often use items that are essentially synonymous [254]. Using validated questionnaires without careful consideration of the actual items they contain would be detrimental, therefore there was careful consideration of the items included, and I aimed to reduce items into one latent measure where they clearly intended to measure the same specific construct and there was sufficient overlap in the data. This method has been used in similar studies [e.g. 293].

6.3.2.1. Measure of Apathy

The Apathy Evaluation Scale [AES; 44] assesses the behavioural, emotional and cognitive dimensions of apathy. The systematic review described in 3.7 found that the informant version of the AES had good reliability and validity in NCD. There are 18 items in total, which relate to behavioural, emotional, social or cognitive domains, and general motivation. Each item is rated by a family member, friend or carer on a 4-point Likert scale, based on the last four weeks. Typically, the AES is scored by summing all items together, with total scores ranging from 18 to 72, with higher scores indicating worse apathy, and it is recommended that scores of 41.5 or above indicate clinically relevant apathy [27]. All items from the AES were selected for inclusion in the analysis.

6.3.2.2. Measure of Depression

The Hospital Anxiety and Depression Scale (HADS) assesses both anxiety and depression using separate subscales. The HADS depression subscale (HADS-D) is a self-rated scale of non-somatic symptoms of depression that occurred in the last week, traditionally using sum-scoring to produce a total score between 0 to 21, with a higher score indicating worse depression [294].

HADS-D has been shown to have good reliability and validity in people living in the community [295] and in people with Parkinson's disease [296]. However, it has been primarily designed around the symptom of anhedonia, for which there are five (of seven) items assessing this [294]. Items relating to (loss of) pleasure are: "I still enjoy the things I used to enjoy"; "laugh and see the funny side"; "feel cheerful"; "look forward with enjoyment"; "enjoy book, radio or TV". Therefore, it was planned that these items would be combined to create a general measure of anhedonia, providing there was sufficient agreement between items, indicated by correlations and internal validity assessed by Cronbach's alpha. The remaining items "lost interest in appearance" and "feel as if I am slowed down" were included as standalone nodes.

As the HADS-D was largely developed to measure anhedonia, and specifically avoided the inclusion of somatic or extreme symptoms of depression [294], items measuring depression was included from other scales in the PrAISED2 study to achieve the most comprehensive measure of depression. A content analysis of measures of depression has identified 52 symptoms of depression [297]. An item was only included from another scale if it clearly assessed one of these symptoms.

The DEMQOL was developed to assess quality of life in dementia, including wellbeing, and factor analyses of the DEMQOL and DEMQOL proxy have showed that some items loaded onto factors related to emotion [298]. Therefore it was proposed that these items could be used to assess depression. DEMQOL proxy wellbeing items that loaded onto an emotion factor included: "cheerful"; "worried or anxious"; "frustrated"; "full of energy"; "sad"; "content"; "distressed"; "lively"; "irritable"; "fed-up" and "has things to look forward to". DEMQOL self-report contains the additional wellbeing items of "enjoying life" and "confident" which loaded onto positive emotion factor, and "lonely" which loaded onto the negative emotion factor. Some items could be perceived to map onto one of the 52 symptoms identified by Fried [297]. However, given that the DEMQOL was not developed to measure depression, I took a conservative approach in my selection of items, and did not include items where it was not absolutely clear that it corresponded to a listed symptom, for example, "has things to look forward to" could interpreted as 'enjoyment' or the inverse measure of 'pessimism', but this was not included due to lack of absolute equivalence. There are potential issues with using self-rated items alongside proxy-rated items, as this may artificially increase their separation [275,278], so items were taken from DEMQOL-proxy rather than DEMQOL-self report where they were present in both measures, to be consistent with the proxy-rated method of the AES in this study. Therefore, the items of the DEMQOL-proxy "worried or anxious", "irritable", "sad", "full of energy", and the DEMQOL-self report item "lonely" were included in the analysis as measures of depression.

6.3.2.3. Measure of ADL

The Disability Assessment in Dementia [DAD; 97] is an informant rated measure assessing basic and instrumental activities of daily living in people with dementia. It contains 40 items that are each answered yes, no, or not applicable. In addition to assessing various areas of disability (hygiene, dressing and undressing, continence, eating, cooking, using the telephone, going out, finance, medication and leisure and housework) the DAD also assesses different processes involved in ADL participation, i.e. impairments in: initiation, planning and organisation, and performance. The DAD score is calculated by taking the number of 'yes' responses and dividing this by the total number of items for which the answer was not 'not applicable', then multiplying the result by 100. The score is calculated as a percentage so that not applicable items are excluded from the total, and total scores range from 0 to 100 regardless of number of non-applicable items. Lower scores indicate worse ADL. Subscale scores for 'Initiation', 'Planning and organisation' and 'Effective Performance' can also be obtained in this manner, so these subscales were included as separate nodes in the initial part of the third phase of analysis, which aimed to assess the relationship between external variables with indicators of apathy and depression.

6.3.2.4. Measures of Executive Function

The 'category' or 'semantic' verbal fluency task is a long-standing measure of executive function, involving initiation, attention, and recall [299]. The animal naming version used in this study asks participants to name as many animals they can think of in 60 seconds, with repetitions being ignored.

The Montreal Cognitive Assessment [MoCA; 300] is a brief cognitive screening instrument, composed of 8 domains of cognition and executive function. The test is scored 0 to 30, with lower scores indicating worse cognitive impairment. An additional point is added for participants with 12 years or less of education.

MoCA and verbal fluency were included as separate measures of global cognition and executive function in the initial part of the third phase of analysis.

6.3.2.5. Physical measures

The SHARE frailty instrument [SHARE-FI; 301] assesses frailty on the factors of weakness (assessed by handgrip strength), exhaustion, weight loss, slowness, and low activity. The SHARE-FI uses a complex scoring method (calculators for which are available as supplementary material in the development article [301]), based on these items and participant gender. There is no maximum score for the SHARE-FI, though a score of over 6 in

females and over 7 in males indicates frailty [301]. The SHARE-FI has been found to have good validity in an older adult sample [301].

The Berg Balance Scale (BBS) was developed for an older adult population to assess balance, by asking participants to perform 14 different balance tasks, from sitting unsupported to standing on one leg, for prescribed lengths of time [302]. Participants' performance on each task is rated on a four-point scale, depending on the time for which the participant could perform the task, the level of difficulty achieved, or amount of support required. The measure is scored as the sum of scores on each item, with a maximum possible score of 56, where higher scores represent better performance.

The Timed Up and Go (TUG) [303] assesses mobility by asking participants to rise from an arm chair, walk 3 metres, turn and walk back, and return to sitting on the chair. The task is timed and the score is the number of seconds taken to complete it, so that higher scores reflect worse performance. The TUG has good test-retest reliability in people with dementia [304].

The SHARE-FI, BBS and TUG were included as separate nodes in the initial part of the third phase of analysis.

6.3.3. Developments affecting the network analysis study

Since attending the network analysis course in January 2019, the field of network analysis methods have also undergone some developments, which meant that my analysis method had to be altered somewhat. Initially, I planned to analyse the data using Gaussian Graphical Models (GGMs) with polychoric correlations, despite the ordinal level data (in the apathy and depression items). This is because at the time of planning my analysis, polychoric and polyserial correlations as input in a GGM were thought to be robust to ordinal and non-normally distributed data, and was a standard approach for estimating psychological networks [270]. However, later on in my PhD, further evidence suggested that this may not be the case [269,305], though this is still uncertain, and may be improved by the use of Spearman correlations as input (instead of polychoric or polyserial correlations) [306]. Mixed Graphical Models (MGMs) on the other hand enable ordinal (and binary and continuous) data to be treated as such, in the same model, using neighbourhood multinomial regression in which a multinomial regression model is applied to each node in turn, so that the edge weights reflect the regression coefficients. However, MGMs may not completely resolve the issue of ordinal data [307]. Furthermore, when I tried estimating the network using MGM, I found that assessing the stability of this method was not feasible, as stability analyses of MGMs require intensive computations. These were taking weeks at a time to run, which was beyond the scope of this multiphase network analysis which formed just one part of the mixed methods study. As a result, I chose to perform network analysis using GGM, with Spearman

correlations, perform stability analyses, and then check the sensitivity of this network by repeating the network analysis using MGM.

6.3.4. Network analysis

Data were organised and managed, and descriptive and some preliminary analyses were conducted in IBM SPSS Statistics Version 28.0.0.0(190) for Windows.

A series of GGM network analyses with three main phases (described later) were conducted, with two parts each. First, these were conducted without reducing items or variables (part A). Then these were re-run following reduction of variables or items with topological overlap (part B), as described below. To check the sensitivity of the models to the method used, this series of analyses was also re-performed using MGM analysis, as described below. Code for these analyses were written in and conducted via the statistical software R Studio (version 2022.07.1 Build554) using the packages: *networktools* (v1.5.0), *mgm* (v1.2-13), *bootnet* (v1.5), *qgraph* (v1.9.2), *EGAnet* (v.1.2.0), *mice* (v3.14.0).

6.3.4.1. Preliminary analysis

6.3.4.1.1. Inspecting accuracy of data

Accuracy checks on the data were performed as part of the data cleaning within the main PrAISED2 study, described in section ‘6.5.4.2. Identifying errors in the quantitative data’.

6.3.4.1.2. Normal distribution, linear relationships, transformations and internal consistency

Network analyses based on continuous data assume that data is normally distributed and have linear relationships with one another [270]. The distribution of each node was checked using histograms, rather than statistical techniques to assess skewness, as these are not recommended for large sample sizes [308]. Linearity of the pairwise relationships of continuous variables was checked using scatterplot matrices.

Non-normally distributed continuous data can be transformed if the data is truly continuous [270]. Transformations were performed for the continuous variables, however, the item-level data was ordinal, so transformations were not possible. No further action was deemed necessary for this ordinal data, as the method used is thought to be robust to skewed ordinal data [307]. Following transformations, histograms and scatterplots were rechecked.

Internal consistency of sum-score measures was checked using Cronbach’s alpha.

6.3.4.1.3.Reversal of data

Where measures contained reverse scored items, the data were entered in reverse, so no reversal after data input was required to obtain the overall score. However, the planned analyses use data from different measures. On one measure, a high score may indicate better performance or functioning, whilst on another, it may indicate the reverse. To avoid difficulties interpretating the network plots, some data was reversed. All AES and HADS-D items were reversed in SPSS using the 'recode' function, so that a low score on the AES indicated greater apathy, and similarly, a low score on the HADS-D indicated greater depression. The TUG was reversed as part of its transformation to make this normally distributed and the SHARE-FI was reversed using a reflect transformation. This meant that for all variables included in the network (except age, which was considered inappropriate to reverse), a low score indicated a worse performance or outcome. See Appendix 7 for details of each node and its meaning.

6.3.4.1.4.Missing data

6.3.4.1.4.1. Identifying missing data

The amount and patterns of missing data were first inspected in SPSS, and then in R Studio using the `md.pattern` function in *mice* package and `summary` function in *base* package. To explore potential reasons for missing data, histograms were created on missing data, split by MoCA and then age, as possible determinants of missing data. This is because it was hypothesised that lower cognition may affect ability to answer questions and understand tasks, and increased age may affect ability to undertake physical performance tasks.

6.3.4.1.4.2. Handling missing data in GGM analysis

Where item-level data was missing, and the items were the focus of analysis, the missing item was excluded pairwise (i.e. the participant with the missing item was excluded from analysis of that item but not the analysis of other non-missing items), to avoid unnecessarily reducing the usable sample size. None of the sum-score measures used for this study had missing data guidance included in the manual. Therefore sum-score data was imputed using person-mean substitution (i.e. that participant's average item score across the measure replaced the missing value), providing there were less than 20% of items with missing data for that participant, as advised elsewhere [309]. Where single performance-based data was missing, e.g. TUG, no imputation occurred, as there were no other within person data on this measure from which to impute, so the missing data was excluded pairwise from the analysis.

6.3.4.1.4.3. Handling missing data in MGM analysis

Unlike with the GGM, MGM does not allow for exclusion of missing data pairwise. Data was imputed where possible, as described above. Where this was not possible, missing data was excluded listwise at each phase (i.e. the participant with missing data was excluded from the whole analysis).

6.3.4.1.5. Assessing Topological Overlap

Network analysis also assumes each node is measuring a unique trait; inclusion of nodes that measure the same thing are problematic [251,254,310]. To check for this, nodes were assessed for potential redundancy by inspecting the visual network, performing Unique Variable Analysis (UVA), and, importantly, by considering whether there was substantial theoretical overlap between the items.

The UVA function of the *EGAnet* package [311], was developed to identify and merge redundant items in the network [312]. This estimates a network and uses a weighted topographical overlap method to assess the overlap of connectedness of the nodes in the network; i.e. it takes into account how similar two nodes' weighted connections are with other nodes in the network, in addition to the strength of the nodes' connections to each other [312,313]. Spearman correlation was selected as this was consistent with the network estimation method chosen in this (GGM) study, and the recommended threshold of .25 was selected [311]. This analysis was performed again at each phase.

Where it was decided that nodes should be reduced, the nodes were combined using a reflective latent trait method, as recommended by Christensen, Golino and Silvia [313], which is integrated into the UVA function of the *EGAnet* package. Where new nodes were created in this way, these were integrated into the dataset and the nodes from which they were produced were removed. The new variables were checked for normal distribution and transformed where necessary and possible. Following this process at each phase, the analysis was re-run. It was expected that some items of the apathy and depression measures may be redundant, in particular the anhedonia items of the HADS, and was also expected that the subscales of the ADL may be redundant.

6.3.4.1.6.Additional checks for MGM

MGM resampling requires that a (unspecified) reasonable portion of all response options are selected, and evidence suggests that collapsing categories (or treating it as continuous) is superior to using ordinal data in which a small number of respondents have chosen this category [314]. Ordinal data was checked for the proportion of respondents who selected

those response options, and where there was less than 4% of a response category selected, the response option was collapsed into the next category.

6.3.4.2. Gaussian Graphical Model analysis

Gaussian Graphical Model (GGM) networks were constructed in three phases, with the next phase expanding on the previous phase. To aid visualisation, different colours were assigned to each node in the network plot based on their 'group', which depended on the phase of analysis, as described below.

Phase 1: The 18 items of the AES were imported from an SPSS document into R Studio. Groups were allocated based on the apathy domain that the item was proposed to reflect: 'Cognitive'; 'Behaviour'; 'Emotional'; 'Other' [44].

Phase 2: Depression items from the HADS-D and DEMQOL(/proxy) were then integrated into the final network produced from phase 1. Groups were allocated based on whether the item was proposed to reflect apathy (AES items) or depression.

Phase 3: External variables were then added to the final network produced from phase 2. These were: 3 types of ADL (DAD); executive function (verbal fluency; MoCA); physical ability (BBS; TUG); and frailty (SHARE-FI). Groups were allocated based on whether the item measured apathy, depression, executive function, ADL or physical ability.

6.3.4.2.1.GGM Model Estimation and Selection

As described above, this study used ordinal data, so this had to be taken into consideration when deciding what type of (partial) correlation to use. Polychoric correlations were previously recommended for ordinal data [270], however, this method is now thought to be problematic with ordinal data [307]. Instead, Spearman correlations were used as the input, as this is proposed to be suitable for both ordinal and skewed continuous data [307].

As in other statistical methods, the best model, that describes the most data whilst also maintaining as much simplicity as possible, should be selected. Methods of model selection for network analysis include thresholding or pruning, where a fixed criterion (such as a *p* value) is used to determine whether an edge should be included, and regularisation, where a penalty is applied the more complicated a model is, to enable selection of the best model [306]. Thresholding and pruning makes the level of error that is accepted explicit, for example, alpha is typically set at .05, and the *p* value must be equal to or below this for the null hypothesis to be rejected [315]. However, this can be too conservative in smaller samples (i.e. not sensitive enough), and if a false edge is included, it may be prominent [306]. Graphical 'Least Absolute Shrinkage and Selection Operator' is a regularisation method commonly used in psychopathology networks, and this applies a penalty based on the edge weights, shrinking the smallest edge weights to 0 (rather

than only selecting edges above a certain significance level), so no thresholding is required [306]. This has been used in fields such as genomics, where the number of observations (participants) is smaller than the number of nodes (genes) [254]. This method has gained popularity in psychology and psychopathology, as it enables the analysis of a larger number of variables. Though more false positives may be included, these are faded in the network, so the visual interpretation is more reliable [306]. Therefore graphical Least Absolute Shrinkage and Selection Operator was used to estimate the GGM. The strength at which this method shrinks edges to 0 is controlled by a (lambda) parameter. This is selected automatically, and can be done so via Extended Bayesian Information Criterion (EBIC), which selects the best model from N iterations (default N =100). EBIC chooses the model (of the 100 estimated models) with the lowest lambda parameter whilst retaining the best model. The tuning hyperparameter (gamma) of this EBIC was set to 0, as this has been recommended to increase sensitivity, allowing detecting of smaller edges, whilst maintaining a conservative estimate [254]. EBICglasso regularisation method is preferrable in GGM for smaller sample sizes such as in the present study, providing conclusions are not being drawn regarding individual edges, and instead the analysis is focused on the structure and centrality [306,307]. Furthermore, this method has been used in similar studies that have estimated networks containing nodes of symptoms and external variables [288,e.g. 316].

The GGM was estimated using the `estimateNetwork` function from the package *bootnet* [317].

6.3.4.2.2.Plotting the GGM network

Networks were plotted using the `qgraph` function, in the package *qgraph*, [318]. A modified version of the Fruchtermanreingold algorithm [319] was used to determine node placement in the network, which results in more strongly connected nodes being pulled closer together. No cut was used; a minimum argument was applied, so that the smallest edges were not visualised, to enhance the readability of the network, and the maximum edge was the maximum edge weight included in the graph (as is the default). Nodes were labelled with a short description (Appendix 7 includes a list of all item labels and their descriptions, before item reduction).

6.3.4.2.3.Centrality assessment in the GGM

6.3.4.2.3.1. Strength / Expected Influence

At the end of each phase, strength or expected influence centrality was assessed, as these offer the most important and interpretable information within a psychopathology network [254]. Where all nodes in the network are coded in the same direction, so that lower score always equates to poor

performance, ‘expected influence’ is recommended, as this takes into account the sign of the relationships, producing an estimate of overall effect in a single direction, in contrast to strength, which can merge the opposing effects of negative and positive edges [320]. In the first two phases, all nodes were coded so that lower scores (e.g. on AES) reflected worse outcomes (e.g. more apathy), so expected influence was assessed in these networks. In contrast, strength was used to assess centrality in the third phase, as not all measures indicated the same outcome: it was considered inappropriate to reverse score ‘age’, as lower age is naturally reflective of better, rather than worse, outcome in this population. The use of strength centrality therefore prevented any associations with ‘age’ from diluting the centrality values of the nodes associated with it. Raw centrality values and plots were obtained using the functions `centralityTable` and `centralityPlot` from *qgraph* package, as recommended by Isvoranu and Epskamp [321].

6.3.4.2.3.2. Bridge Centrality

Bridge centrality provides a centrality estimate, however, in bridge centrality, only the edges that connect a node to nodes from different (theoretically, researcher-assigned) groups, are considered. Bridge Expected Influence was assessed in phase 2 to determine which nodes most strongly connected apathy and depression indicators with one another, and Bridge Strength was assessed in phase 3 to determine which nodes were most strongly connected to nodes outside their proposed group. Bridge centrality estimates were obtained using the function `bridge` in the *networktools* package [273].

6.3.4.2.4.GGM stability checking (bootstrapping)

Following the construction of a network at the end of each phase, the accuracy of the network was estimated, using post-hoc stability analyses, as recommended by Epskamp, Borsboom and Fried [317].

First, the stability of edge weights was assessed through bootstrapped confidence intervals. Non-parametric bootstrapping using 1000 iterations was conducted to obtain bootstrapped confidence intervals for the edge weights, using the `bootnet` function of the *bootnet* package [317]. These were plotted using the `plot` function of *bootnet*.

Next, the stability of the centrality indices was assessed using the case-dropping subset bootstrap technique. This technique assesses the impact of dropping a random selection of participants from the network, and produces a Correlation Stability (CS) coefficient which indicates what proportion of the sample can be dropped so that a correlation of at least .70 remains with the original centrality estimates in 95% of cases. It is recommended that this CS-coefficient should be above .25 for centrality estimates to be interpretable, and ideally above .50 [270]. A plot was also created to display the CS-coefficient

with the increasing drop in sample, as recommended by Fried and colleagues [322], using the `plot` function of *bootnet*.

6.3.4.3. Mixed Graphical Model analysis

To assess the sensitivity of the model to methods used, the network analyses were re-performed using the Mixed Graphical Model (MGM) method.

6.3.4.3.1.Estimating the network and model selection (MGM)

MGM networks were constructed as described in the GGM analysis: i.e. in three phases, with the next phase expanding on the previous phase. The MGM analysis required the additional actions of inputting whether data was Gaussian or categorical, and the number of levels if categorical.

The networks were estimated using the `mgm` function of *mgm* package [323] to estimate pairwise interactions via multinomial regression. Regularisation was performed using a Cross Validation (CV) lambda selection method, as recommended for use in MGM with smaller sample sizes [306,307]. Cross validation is another method of selecting the penalty used by Least Absolute Shrinkage and Selection Operator, based on splitting the data into a number of sub-samples and selecting the best penalty from these sub-samples [324]. As in the GGM, no threshold was applied.

6.3.4.3.2.Plotting the network

In contrast to the GGM network described above, where all edges were colorized to reflect a negative or positive relationship, edges between nodes of categorical data will appear grey, as colorizing edges of nodes with more than two categories (response options on an item) is not possible in an MGM [323]. In addition to the methods used to plot the GGM network, a pie chart was displayed around each node, indicating predictability.

6.3.4.3.3.Predictability

Predictability refers to how well the node is predicted by the nodes immediately connected to it [325]. With continuous data, this is the percentage of variance explained and is strongly linked to edge weight and centrality measures such as strength. However, with categorical data (e.g. item level data with four response options), edge weights are less linked to predictability, as predictability is also affected by the proportion of participants selecting the different response options on the item [325]. If a large proportion (e.g. 90%) of participants select the same response option, the model will predict this as the score, and the predictability will be highly accurate (90% correctly classified), before including other variables in the model; this is referred to as ‘marginal’ correct classification. This only leaves 10% of the remaining classifications to be predicted by the other variables (i.e. beyond the marginals) [325].

Predictability was obtained using function `predict` from the package *mgm*. For continuous nodes, r^2 was selected. For ordinal data, the correct classification marginals were obtained, and the 'correct classification beyond marginals' (i.e. the additional predictability provided by the other variables) were calculated and added to the network plot to form a predictability pie chart around each node, as recommended by Haslbeck and Waldorp [325].

6.4. Qualitative Sub-study Methods

The methods specific to the qualitative sub-study will now be described.

6.4.1. Participants and recruitment

Participants were a sub-sample from the main PrAISED2 study in Lincolnshire, Nottinghamshire, Derbyshire, and Bath and North-East Somerset. Recruitment and interviews took place between September 2019 and December 2019. In addition to the eligibility criteria of the main study, inclusion criteria for the qualitative interview were: prior agreement to be contacted about related studies; willing to consent to take part in this additional interview and had the capacity to do so.

I decided to only interview participants who were allocated to the control group, or who had received a maximum of three sessions of the PrAISED2 therapy programme before an interview, as this was equal to the number of visits the control group could receive. This was to try to keep the focus of the interviews on the participants' experiences with daily life, rather than their experience or changes that may have been brought about by the intervention. However, in practice it was not possible to ascertain the number of visits the participant would have received before the interview date, as therapist visit time schedules were varied and communication of these to the research team were post-hoc. In lieu of information about the number of visits the participants had received, I aimed to interview participants in the intervention that had been randomised within a four-week time frame. This was an arbitrary number of days chosen to reflect that therapists had up to four weeks to schedule their first visit with participants.

Maximal variation sampling was used to recruit a diverse group of participants with respect to location, gender, ethnicity, carer relationship, and cognitive abilities within the pool of possible participants. Participants enrolled in the PrAISED2 study who met the eligibility criteria were invited to take part in an interview to explore their interests, activities, and feelings. Participants were both patients and carers. They were offered the opportunity to be interviewed either separately or together, as there are both benefits and limitations to dyadic and individual interviews, and participants who are not given the choice may be reluctant to take part either with or without each other [326]. All participants were also offered the choice of being interviewed at home or at a

neutral location (such as a private room in a local arts centre). Initially, it was intended that participants with an AES score indicating clinical apathy would be interviewed. However, following discussions with experienced qualitative researchers it was decided that participants with a variety of AES scores would be recruited, to explore a range of experiences and views. Interviewing a diverse group of people in this way can enable better triangulation, through comparing and contrasting during analysis [327].

6.4.2. Data collection

In-depth semi-structured interviews were conducted with individual participants or dyads. Focus groups with people with early dementia are regularly conducted, however personal experience have indicated that the voices of people with dementia can become lost in a focus group that also has cognitively intact people. This is echoed by Morgan and colleagues [328], who note that focus groups can be overwhelming for people with dementia, whereas dyadic interviews enable better communication, whilst still allowing for expanding of ideas that may not otherwise have been remembered by the individual. However, in contrast to individual interviews, dyadic interviews may prevent openness, as participants may not wish to discuss certain topics or be honest in front of each other [326]. Therefore, I chose to provide participants with the option of an individual or dyadic interviews, with the aim of obtaining a mixture of both methods, to benefit from the strengths of each, as recommended by Pratt [329].

Interviews were guided by an interview schedule. Critical realism emphasises the importance of building on existing theories [263], so the initial interview schedule was developed based on current definitions of apathy (i.e. domains of behaviour, interests, emotions, and social life). The interview schedule was amended prior to the start of the study based on feedback from PPI members (for further details, see section ‘6.4.4.2. Patient and Public Involvement’), and was revised iteratively throughout the study. These alterations were prompted by points of interest and lines of enquiry based on the findings of interviews. For example, putting things off, or saying “I’ll do it tomorrow” was an unexpected but frequently reported occurrence for many of the participants.

Field notes were made after every participant interview. Interviews were audio-recorded using a Dictaphone, and transcribed verbatim into Microsoft Word. I transcribed the first interview and the remaining interviews were transcribed by a professional transcription service, though were checked and edited by me. See Appendix 8 for details of transcription notations and their meanings. Transcripts were then imported into NVivo.

6.4.3. Data Analysis

Data were analysed using thematic analysis, which involves an iterative process of six phases: familiarisation with the data; code generation; constructing themes; reviewing themes; defining and naming themes; and producing a report [330]. Specifically, I chose to use reflexive thematic analysis, which focuses on constructing themes based on rich meaning, rather than simply summarising what participants have reported [331]. Analysis was also guided by Fletcher's [287] example of critical realist thematic analysis, which combines deductive and inductive approaches. This enables the code generation phase of analysis to be constructed from the data but also informed by theory and prior research. The six phases of thematic analysis are further elaborated below.

1. I began by listening to the audio-recording of each interview and re-read the notes I made after the interviews. I then undertook multiple readings of the transcripts whilst listening to the audio recordings as a familiarisation technique. This also served as a formatting and checking process, as I used this to check and correct for inaccuracies in transcription, then format the transcripts appropriately, and then ensure identifiable information was removed from the transcripts. During this familiarisation phase, I recorded notes on points of interest and notable observations in an 'analytic memos' document, which has previously been recommended [332], and noted down quotes that felt poignant or important in another document.

2. Complete coding was an iterative process, tho ADDIN ZOTERO_ITEM CSL_CITATION {"citationID":"6lQQdThJ","properties":{"formattedCitation":"[333]","plainCitation":"[333]","noteIndex":0},"citationItems":[{"id":650,"uris":["http://zotero.org/users/2637212/items/3ZY9S558"],"itemData":{"id":650,"type":"chapter","container-title":"Analysing qualitative data in psychology","edition":"2nd ed.","event-place":"Los Angeles","page":"84-103","publisher":"Los Angeles : SAGE","publisher-place":"Los Angeles","title":"Thematic Analysis","container-author":[{"family":"Lyons","given":"Evanthia"},{"family":"Coyle","given":"Adrian"}],"author":[{"family":"Clarke","given":"Victoria"},{"family":"Braun","given":"Virginia"}],"issued":{"date-parts":[["2016"]]}}}],"schema":"https://github.com/citation-style-language/schema/raw/master/csl-citation.json"} [333]ugh was largely split into two main cycles. This two cycle approach served to avoid 'coding drift', in which later established codes are not applied to earlier coded transcripts [333], but also enabled separate inductive and deductive rounds of coding. The first cycle of coding was inductive, in which both semantic and latent codes were applied based on the data.

Whilst analysis is never truly inductive, I attempted to maintain a data-driven approach in the first cycle by not conducting a review of the literature on mechanisms of apathy until the first round of (inductive) coding was completed (though not the nosology of apathy as this was used to inform the interview schedule). Two experienced PPI members, one of whom had received formal training in qualitative analysis, were given two transcripts each to informally 'code'. A member of the supervisory team was also given two (different) transcripts to code, so observations and insights could be compared and built upon (see sections '6.4.4.1. Collaboration' and '6.4.4.2. Patient and Public Involvement' and for a more detailed description). The emerging codebook was also discussed at supervisory meetings. This allowed further identification of new codes, which were applied during the remaining coding. A second cycle coding phase was conducted which had a deductive focus, though any inductive semantic or latent codes deemed relevant were also applied at this second stage. Second cycle coding was conducted after the feedback from the PPI members and supervisors and after reviewing the literature on mechanisms and models of apathy. Codes were reviewed throughout to ensure they represented the data appropriately, without relying on the data it contained to make sense (as advised by [333]) and refined where necessary. The first two transcripts were coded manually on paper using highlighters and notes in the margins, and these were later transferred to NVivo, which was used to organise all transcripts and codes. The remaining codes were applied using NVivo 'nodes'.

3. Codes were organised into folders within NVivo based on their overlap and shared meaning, reflecting codes that could later be collapsed or might develop into a subtheme or theme. Codes were also organised using a mind map, with connections and clusters indicating tentative themes and subthemes, and possible alternatives. These organised codes were further categorised, and points of weaknesses, inconsistencies, similarities and differences were identified and discussed within the supervisory team throughout. Following data collection from all participants and the two cycles of coding of all interviews, codes were further refined and re-categorised, until candidate themes were generated. The research questions were placed in front of the researcher throughout tentative theme construction, to encourage a greater focus on themes that were relevant to the research questions.

4. The tentative themes were formally presented to the supervisory team via Microsoft PowerPoint and regularly discussed at supervisory meetings. Tentative themes were reviewed for their relevance to the research questions and whether they were representing a coherent

pattern of meaning. Themes were refined, with some codes being moved, expanded and collapsed or removed.

5. The final themes were organised into a word document containing a short summary, of one to three sentences, for each theme and subtheme, and the associated codes with descriptions and one or two example quotes. These final themes were also named. These final themes were presented again along with their codes and summary descriptions to the supervisory team, and the names and appropriateness of each theme was discussed.
6. Final themes were written up as a draft thesis chapter with integrated discussion, and presented at two conferences and submitted for publication (see Publications and Dissemination section for details). This led to further theme development and refinement, the development and refining of some themes, reflecting the iterative nature of qualitative research, and highlighting that writing is part of the analytic process. Quotes were selected as illustrative or analytical demonstrations of the relevant argument, and were chosen based on their relevance but also to maximise the variety of respondents being presented. The discussion was later separated from the results, as it was deemed preferable for this mixed methods study to have an integrated discussion of findings from both qualitative and network analysis studies.

I was able to collect data in parallel with listening back to interviews and memos, and checking, correcting and re-reading transcripts as soon as possible after each interview, which enabled me to think critically about my approach to each interview, develop my skills, adapt the interviews, but also enhance familiarity with the data, schedule and make notes about tentative ideas and points of interest throughout the data collection process (i.e. phase one of thematic analysis).

6.4.4. Enhancing quality

6.4.4.1. Collaboration

In line with Braun and Clarke's [334] view of qualitative research, a coding reliability approach, in which two researchers code the same data separately and inter-rater reliability is assessed, was not considered appropriate. The value of the qualitative method was to enable deeper understanding and offer the best possible explanations, so varied interpretations from different perspectives were encouraged, and therefore agreement between researchers was not the goal. Similarly, member checking was not deemed appropriate, which is further discussed in section '6.5.4.3. Qualitative analysis data checking'. Nevertheless, collaboration can be beneficial as though no

view is considered ‘correct’, collaboration enables multiple perspectives to inform the interpretations. Furthermore, collaboration enables greater reflexivity, as the interpretations of others help us notice our own perspective [335].

To obtain the benefits of collaboration within qualitative research, whilst also maintaining independence in the research I:

- Sought feedback from an experienced qualitative research fellow (CD) who listened to two audio recordings. CD provided constructive feedback on the interview approach, enabling me to adapt my approach in future interviews to enable greater focus and encourage a balance between patient and carer contributions.
- Discussed my methods, analysis and interpretations with my PhD supervisors (RH, a geriatrician, SG, a nurse, and VvdW a researcher with a background in psychology), throughout the study, and VvdW conducted complete coding of two of the early transcripts. This multi-disciplinary collaboration enabled greater reflexivity, as it highlighted my perspectives when contrasted with others, and greater trustworthiness as it encouraged me to demonstrate and justify my coding and interpretations, keeping the analysis grounded in the data, as recommended by Cornish and colleagues [335]. Furthermore, this enabled multiple perspectives to be included in the analysis.
- Sought the involvement of key stake holders, as outlined in the next section: ‘Patient and Public Involvement’.

6.4.4.2. Patient and Public Involvement

The proposed methods were discussed with the ‘dementia, frail older people, and palliative care PPI’ group at the University of Nottingham on 15th February 2019. There were six PPI members, who were people with experience of caring for someone with dementia, and older adults, in attendance at this meeting. The PPI members were supportive of the semi-structured interview method and felt it was important that the concept of apathy is understood so that people with dementia and carers can be educated. PPI members advised that one hour was an appropriate interview time for someone with dementia, and that a ‘neutral’ location was very important. They felt strongly that carers should be included in interviews, and that participants should be given the opportunity to be interviewed separately so that both carers and people with dementia alike had the opportunity to be more open. PPI members did not like the word ‘apathy’ and felt that this has negative connotations. Instead, they recommended using language of ‘experiences’, ‘activities’ and daily life. Therefore, the word apathy was not mentioned when inviting the participant to an interview or at the beginning of the interview. I asked the initial few participants how they felt about this at the end of the interview. Some

participants reported that they were not sure what to expect, having not taken part in a qualitative interview previously, but felt that this was a good description of what the interview was about and were satisfied with how it was explained. Some participants had not heard the word ‘apathy’ or did not know what it meant, so this gave me reassurance that not using the term apathy to introduce the subject was an inclusive decision.

The first draft interview schedule was sent to two PPI members (MD and MG), who read this and provided comments and feedback. This was used to amend the interview schedule. Two anonymised transcripts were sent to two PPI members (MD and MW) who read these and made notes and comments on what they felt was interesting and their interpretations of the interviews. One PPI member chose to make comments in the document margin (MD) whilst another preferred to send detailed feedback listed in an email (MW). This feedback led to the development of further codes which were applied in the second coding cycle.

6.4.4.3. Reflexivity

Critical realism recognises that there is no ‘objective’ position, so researchers’ own values and beliefs must be recognised and acknowledged to understand how these might influence the methods and findings [336]. A reflexive journal was kept throughout, from the conception of the study until the write up phase, in an attempt to enhance awareness of my own perspectives and interpretations and where these might come from [337]. A record of key decisions and justifications was also maintained.

6.4.4.3.1.Reflexive notes

Throughout my career I have met with many people with cognitive impairment and carers in my professional roles as a research assistant, research psychologist and Alzheimer’s Society volunteer. This likely facilitated my rapport and understanding of the experiences of people with dementia and their carers. However, it may also have meant that I became less affected by participant’s accounts, or less likely to notice something as unusual or unique to people with dementia. My personal experiences of my father who was diagnosed with AD in 2015, and my mother who cares for him, will have led me to have certain assumptions about the everyday lives of participants with NCD and their carers. My father was diagnosed with young onset dementia and was very physically healthy at the time I was conducting the interviews and analysis, so I held prior expectations that the participants would be different from him due to their age. Furthermore, I am relatively young and healthy. This contrast between the age and health of my father’s and I, and that of the participants, may have led me to over emphasise its importance in my analysis and interviews. However, the involvement of others in the analysis was used to check whether this was a legitimate focus from the data.

Furthermore, this contrast may have helped highlight mechanisms that were otherwise implicit.

I decided not to disclose my personal experiences in any of the interviews as I felt that presenting as someone who was naïve to their experiences elicited a telling nature in participants. I also felt that as I am not a carer for my father meant that my perspective was still that of an outsider, and should be presented as such. Disclosing my own experience could have encouraged participants to be more open due to a perceived shared experience. However, this implied shared understanding may have prevented participants from sharing and elaborating on their experiences further.

I started this project with limited experience with qualitative research, and a Psychology degree with a heavy emphasis on the quantitative paradigm. I attended various courses on qualitative interviewing and analysis to increase my knowledge of qualitative research methods but also my confidence in conducting the research. This was beneficial, however, the main advice I received was that the best way to improve my understanding and approach to qualitative research was through practice. I had not allowed for any pilot interviews within my plan. Continued reading, seeking advice and inevitable experience gained as my data collection and analysis progressed meant that my later interviews and analysis were more focused and confident than the earlier ones. This may have led to earlier interview participants views and experiences not being as thoroughly represented. On the other hand, because I lacked confidence at these earlier stages, I also may have been more deliberate in my interviewing and dedicated more time and iterations to these interviews in analysis. Furthermore, the iterative nature of analysis meant that these order effects would have been minimised.

6.5. Ethics

This next section will discuss the ethics of both studies, including the parts of the overall PrAISED2 study relevant to this PhD.

Ethical approval for this set of studies were granted by the Yorkshire and Humber, Bradford and Leeds research ethics committee (18/YH/0059). The research was carried out in accordance with Good Clinical Practice.

6.5.1. Informed consent process

Both patients and carers provided written informed consent for the PrAISED2 study. Participants who took part in the interview study signed an additional consent form after reading a separate participant information sheet for this. Given that apathy is a contested concept, in that people may have their own differing opinions on what it means, and some of these may be particularly negative or have stigmatised connotations, it was decided that apathy would

not be referred to directly in the information sheet. To ensure that informed consent was still given, the information sheet detailed the aspects of apathy that would be covered by the interview: activities, interests and feelings.

Both information sheets were provided at least 24 hours prior to consent. A short simplified information sheet for the main PrAISED2 study was used to supplement the standard version in order to facilitate understanding. Details of the study were explained and participants were given the opportunity to ask any questions. It was made clear to participants that the research studies were optional, their care would not be affected, and that they could withdraw from either study at any point.

Participants had cognitive impairment but all were required to have capacity to give informed consent to take part in the study. Capacity was assessed prior to taking consent for each study in accordance with the requirements of the Mental Capacity Act. Those who lost capacity throughout the PrAISED2 study remained in the study if a personal consultee agreed. Potential personal consultees were identified with the participant at the time of recruitment.

6.5.2. Privacy and confidentiality

For the main PrAISED2 study, potential participants were contacted via the Join Dementia Research register, memory assessment services, GPs, and dementia support groups. Those registered with Join Dementia Research agreed to their details being accessed by researchers for screening purposes, and to be contacted directly by researchers. Potential participants identified via clinical services had their healthcare records checked by a member of their clinical team to see if they match the entry criteria prior to being contacted. Clinicians and support group leaders ensured each person agreed to be contacted about PrAISED2 and for their details to be passed on, prior to passing on their information to the researchers.

For the interview study, participants were only invited to take part if they agreed to be contacted about related studies. This was an optional statement on the PrAISED2 consent form. Pseudonyms were used instead of participants' real names, and identifiable information was removed from transcripts.

Participants were allocated an identification number, and their anonymised data was stored separately from identifiable data. Participants were informed that their information would be kept confidential unless there was a risk of significant harm to them or to others. The researchers conducting quantitative research visits, including myself, had all completed good clinical practice training.

6.5.3. Burden and Distress

Participants all had a diagnosis of dementia or MCI, which will likely have been distressing for both the person living with it and their family and friends. However, these participants should not be denied the opportunity to take part in research. As described in section '6.5.1. Informed consent process', participants were only recruited if they had mental capacity to take part.

There was a risk that the PrAISED2 baseline and follow-up visits would be burdensome to participants. However, the length of the visits (around 2.5 hours) was made clear to any potential participants, and breaks were taken and assessments were shared over two visits when needed.

A variety of sensitive topics were discussed in the quantitative and qualitative studies, such as including changes in memory and abilities and low mood. I strove to adopt a sensitive and listening approach and participants were reminded they were able to take a break or withdraw at any time, consistent with Pratt's [329] recommendations for interviews with people with dementia. There were also provisions for signposting participants to the Alzheimer Society (including the national dementia helpline) or their GP if needed.

6.5.4. Credibility / trustworthiness

6.5.4.1. Researcher training

Every researcher conducting the quantitative assessments received the same training, supported by a research manual and regular teleconference meetings. Researchers were trained on the overall research project aims, the recruitment and consent process, and delivery of assessments. Researcher training was developed and delivered by the core research team working in the randomised controlled trial work package of PrAISED2 (which included myself).

I also undertook additional formal training to prepare for my mixed methods study, including attending a 2-day 'mixed methods in health research' course at the University of Nottingham, a 5-day network analysis course at the University of Amsterdam, a 1-day qualitative interviewing course provided by the Social Research Association, and a qualitative coding course at the University of Nottingham.

6.5.4.2. Identifying errors in the quantitative data

As described in section 6.2.3, accuracy and completeness of data was checked as part of data management in the wider PrAISED2 programme. Data entry accuracy was checked during monitoring visits (conducted by the central research team, including myself) throughout the study, and an additional 5% of data was checked against the source data at study close out. The core research team (including myself) were tasked with directly checking

outliers, identified by the clinical trials unit, or liaising with sites to check these. Where outliers were due to a data entry error, the error was corrected. Where outliers were correct, the data were not changed and remained in the dataset.

6.5.4.3. Qualitative analysis data checking

'Member checking', 'member validation', 'respondent validation' or 'reflexive elaboration' are methods of allowing the participants involved in the study to check or comment on the analysis, that aim to enhance the credibility of qualitative research [338,339]. However, it was not deemed appropriate or possible for this study. Member checking assumes that participants' have greater insight than the researcher and therefore should be empowered to provide a dominant perspective over the analysis [338]. However, it is argued that participants do not necessarily have "epistemological privilege" [339], i.e. they may not hold a greater insight into the issues under study. Without the assumption of a more epistemological privilege, the issue of how to combine perspectives if there is disagreement between the patient, carer and researcher is unresolved. Member checking also requires a high level of engagement, understanding and insight from participants. This may be challenging for people with dementia as they may not remember the interview or even taking part in the study. Member checking solely by carers could result in an inappropriate imbalance in whose interpretations and understandings are represented, as it is recognised that participants have their own agenda and motivations [327]. Member checking could also revert the higher-level interpretations and analysis to a lower-level descriptive perspective of their own individual interview.

6.5.4.4. Triangulation

Triangulation, in which more than one method or data source is used, was performed as part of the mixed methods approach [256]. Triangulation in this study was used for completeness, to combine different perspectives, and to facilitate retroduction, i.e. the post-hoc explanation of phenomena [256].

Triangulation was also used for the purposes of confirmation within the network analysis sub-study, as the analysis was re-performed using a different method (see section 6.3.3 for details).

6.6. Integration and construction phase

Integration occurred once all data had been collected. Initially, a joint display table was created in which the results were compared side by side against each objective. The network analysis results were written in the table first, with qualitative results used to explain these where available, or point out sources of contradiction. This was then used to develop a narrative discussion of both qualitative and network analysis results.

Chapter 7. Network Analysis Study Findings

7.1. Introduction

The aims of this thesis were to characterise apathy and its boundaries, and to understand the possible mechanisms and impact of apathy. The network analysis sub-study sought to: (1) Determine how apathy indicators might influence each other, whether there are distinct domains of apathy, and which indicators may be most important; (2) Assess whether apathy and depression form distinct clusters and which indicators of apathy and depression relate to each other; (3) Determine how external factors relate to apathy indicators.

This chapter will present the results of the network analysis sub-study, including participant characteristics, and the resulting networks from the three phases of analysis outlined in the methods section above. Limitations and conclusions of this sub-study will be noted, with further discussion presented following the integration of the network analysis and qualitative results in Chapter 9.

7.2. Network Analysis Results

7.2.1. Recruitment

Recruitment to the PrAISED2 study and associated baseline visits were conducted by researchers (including myself) across the five sites between October 2018 and June 2021. 1540 potential participants were pre-screened, 475 were screened (formally assessed for eligibility, typically with a face-to-face visit), and 365 participants were recruited. 303 participants were recruited from memory assessment services, 22 from the Join Dementia Research register, and 40 from GP practices.

7.2.2. Descriptive Statistics

7.2.2.1. Sample Characteristics

Participants' mean age was 79.9 (*SD*=6.6). Mean Share-FI score was 1.9 (*SD*= 1.7) indicating an average 'pre-frail' status [301], and the mean HADS-D score was 4.5 out of 21 (*SD*= 2.8), which indicates the absence of clinical depression [340]. Mean AES score was 41.9 out of 72 (*SD*= 12.2), and 51% of participants (with complete AES scores) scored >41.5 on the AES, which has previously been argued to be indicative of 'clinical apathy' [27]. These participants had worse cognition (p<.001, Cohen's d=.37, indicating a small effect size), and were older, but this was not significant (p=.06, Cohen's d=-.20, indicating a small effect size). Participants with and without clinical apathy did not significantly differ in terms of depression (p= .08, r=.09, indicating a very small effect size) or gender (p=.45, *Phi*=.04, indicating a very small effect

size). The remaining participant characteristics are displayed in Table 8, and carer characteristics are displayed in Table 9.

Table 8. *Patient Characteristics*

Characteristic	Distribution	% missing
Gender	42.5% female; 57.5% male	0.0%
Marital Status	2.5% single; 1.1.% partnered; 67.7% married; 0.6% separated; 4.7% divorced; 23.0% widowed; 0.3% other;	0.3%
Ethnicity	98.1% White; 0.6% Black 0.3% South Asian 1.1% Other: (self-described as: *Afro Caribbean; Asian; Indian; Lebanese New Zealand*)	0.0%
Highest level of education	2.2% primary school or less; 46.3% secondary education; 20.6% vocational training; 31.0% college or university degree	0.0%
Living situation	25.2% lived alone; 74.8% lived with someone	0.0%
Diagnosis	19.2% MCI; 38.9% AD; 19.5% VaD; 0.8% FtD; 16.2% Mixed; 0.8% Other type of dementia 2.7% Dementia of unknown type	1.4% 'don't know' 0.8% missing

AES, Apathy Evaluation Scale; HADS-D, Hospital Anxiety and Depression Scale – Depression; DAD, Disability Assessment for Dementia; SHARE-FI, SHARE frailty index; MoCA, Montreal Cognitive Assessment; MCI, Mild Cognitive Impairment; AD, Alzheimer's Disease; VaD, Vascular Dementia; FtD, Frontotemporal Dementia.

Table 9. *Carer Characteristics*

Characteristic	Distribution	Missing
Carer Gender	72.3% Female; 27.4% Male	0.3%
Carer Relationship	64.7% Spouse 0.8% Sibling 28.8% Son/ Daughter 2.2% Friend 1.1% Paid carer 2.2% Other relative: 0.8% Brother/Sister; 0.6% Brother/Sister-in-law; 1.1% Son/Daughter-in-law; 0.3% Niece/ Nephew; 0.3% Grandchild	0.3%

7.2.2.2. Node characteristics

7.2.2.2.1.Phase 1 Node characteristics.

The apathy indicator with the highest mean score (indicating worse apathy) was ‘S/he is interested in learning new things’ (*M*= 2.9 out of 4, *SD*= 1.0), and the apathy indicator with the lowest mean score (indicating less apathy) was ‘S/he is interested in things’ (*M*= 2.0 out of 4, *SD*= 0.9). Apathy items were generally not normally distributed, and transformations were not appropriate given the four-point ordinal scale.

Table 10 provides further details of descriptive data for each node.

7.2.2.2.2.Phase 2 Node characteristics.

The HADS-D item with the highest mean score was ‘I feel as if I am slowed down now’ (*M*= 1.2 out of 3, *SD*= 0.9) and the item with the lowest mean score was ‘I can enjoy a good book or radio or TV programme’ (*M*= 0.3 out of 3, *SD*= 0.6); where high scores indicated worse depression. The DEMQOL(/proxy) item with the highest mean score was ‘irritable’ (*M*= 3.2 out of 4, *SD*= 0.8), whilst ‘sad’ was the lowest mean score (*M*= 1.8 out of 4, *SD*= 0.8), where low scores indicated worse depression.

The proposed HADS anhedonia subscale had low internal consistency, (α= .59), calling into question whether these items should be combined into one node as planned. ‘EnjoyMedia’ was the least correlated, yet excluding this item from the proposed sum-score did not increase internal consistency sufficiently (α= .64) to warrant the inclusion of these items as a single node in the network. Therefore, the HADS ‘anhedonia’ items were included in the network separately. As the items no longer formed an overall measure of anhedonia, the HADS ‘cheerful’ item was replaced with the DEMQOL-proxy ‘cheerful’ item, as this was consistent with the proxy assessment method of the majority of items in the network. Depression items were generally not normally distributed, and as with apathy items, transformations were not appropriate.

7.2.2.2.3.Phase 3 Node characteristics

The DAD subscale with the highest mean score was DAD-initiation (*M*= 78.9 out of 100, *SD*= 20.9) and the lowest was DAD-planning and organization (*M*= 71.4 out of 100, *SD*= 25.6). The characteristics of the remaining nodes, prior to item reduction, entered in this phase are reported in Table 10.

The DAD subscales, BBS, and TUG had relatively skewed distributions, whilst the remaining external variables were normally distributed. All correlations between the continuous variables were linear, as identified by scatterplots (available in Appendix 9).

Table 10. *Node characteristics prior to item reduction*

Item	Mean	SD	Min	Max	Missing
AES 1	2.0	0.9	1	4	0.6%
AES 2	2.3	0.9	1	4	0.6%
AES 3	2.2	1.1	1	4	1.1%
AES 4	2.7	1.0	1	4	0.8%
AES 5	2.9	1.0	1	4	0.8%
AES 6	2.1	1.0	1	4	1.1%
AES 7	2.8	1.0	1	4	0.8%
AES 8	2.3	1.2	1	4	0.8%
AES 9	2.1	1.0	1	4	0.8%
AES 10	2.2	1.1	1	4	0.8%
AES 11	2.2	1.0	1	4	1.1%
AES 12	2.2	1.0	1	4	0.6%
AES 13	2.3	1.1	1	4	1.6%
AES 14	2.2	1.0	1	4	0.6%
AES 15	2.1	1.0	1	4	0.6%
AES 16	2.4	1.1	1	4	0.6%
AES 17	2.5	1.1	1	4	0.6%
AES 18	2.5	1.0	1	4	0.6%
HADS-D 2	1.2	0.9	0	3	0.3%
HADS-D 6	0.6	0.8	0	3	0.0%
HADS-D 3	0.7	0.8	0	3	0.3%
HADS-D 7	0.3	0.6	0	3	0.3%
HADS-D 10	0.5	0.7	0	3	0.3%
HADS-D 14	0.3	0.6	0	3	0.3%
DEMQOL 8	3.6	0.7	1	4	0.0%
DEMQOL-P 1	2.7	0.8	1	4	1.4%
DEMQOL-P 2	2.9	0.8	1	4	1.1%
DEMQOL-P 4	1.8	0.8	1	4	1.1%
DEMQOL-P 5	3.2	0.8	1	4	1.6%
DEMQOL-P 9	3.0	0.9	1	4	0.8%
DAD initiation	78.9	20.9	7.7	100	1.9%
DAD planning & organisation	71.4	25.6	0	100	1.4%
DAD effective performance	78.7	21.3	6.3	100	3.0%
MoCA	19.9	3.2	13	26	2.7%
Verbal Fluency	12.1	4.6	0	25	0.3%
BBS	46.4	9.4	4	56	5.2%
TUG	15.4	8.3	5	84	3.8%
SHARE - FI	1.9	1.7	-2.1	5.9	2.2%
Participant's Age	79.9	6.6	65	95	0.0%

AES, Apathy Evaluation Scale; HADS-D, Hospital Anxiety and Depression Scale – Depression; DEMQOL; Dementia Quality of Life measure; DEMQOL-P, DEMQOL- Proxy measure; DAD, Disability Assessment for Dementia; MoCA, Montreal Cognitive Assessment; BBS, Berg Balance Scale, TUG, Timed Up and Go; SHARE- FI, SHARE - frailty index.

7.2.2.3. Preliminary analysis

7.2.2.3.1.Data transformation

Transformations were performed on the TUG and BBS to produce a normal distribution. Data transformations on the DAD were unsuccessful, however the distribution was skewed and did not reflect any bimodal distribution or non-linear relationships with other variables, so it was deemed acceptable to treat this as continuous. A reflect transformation was performed on the SHARE-FI for the purposes of reversing the data only.

7.2.2.3.2.Missing data and imputation

Percentage missing data for each variable included in the network is reported in Table 10. MoCA scores did not appear to explain missing data, however missing depression scores and missing SHARE-FI did appear to be higher in older participants (see Appendix 10).

All missing MoCA scores (*N*= 10) were due to missing the item 'years of education' that is used to adjust the MoCA score based on education of the participant. This data was imputed based on the participants' response to another question in the dataset related to highest level of education: those with secondary or less were scored '1' (i.e. assumed to have 12 or less years of education) and those who reported a higher level of education (such as college or university) were assumed to have over 12 years education, and not given the additional point. MoCA scores were recalculated on this assumption, so all participants had available MoCA scores. BBS and DAD score was imputed with person mean substitution. Three participants had DAD scores with >20% items missing and two participants had >20% missing BBS items, so data could not be imputed for these participants. Data transformation was reapplied to the BBS dataset with imputed data. No other imputation occurred.

7.2.2.3.3.Outliers

Two outliers were identified in the TUG, however these were reduced after transformation, so no data was removed or altered.

7.2.2.3.4.Collapsing item responses for MGM

All AES items had over 4% of responses falling in each response category, whilst seven depression items had less than 4% in at least one response category. HADS items 3, 6, 7 and 10 and DEMQOL items 4 and 8 had their number of item levels reduced from 4 to 3, whilst HADS item 14 had its levels reduced from 4 to 2 for MGM methods.

7.2.2.3.5.Internal consistency

Internal consistency of the sum-score variables are reported in Table 11. Though internal consistency of the MoCA was low, this was deemed acceptable given that it is intended to measure global cognition [300].

Table 11.*Internal consistency of sum-score measures*

Measure	Cronbach's α
DAD – Initiation	.80
DAD – Planning and Organization	.82
DAD – Effective Performance	.85
MoCA	.56
SHARE-FI	Unable to calculate
BBS	.90

Note: Cronbach's α ≥.70 indicates sufficient internal consistency [168].

7.2.2.3.6.Multicollinearity

Possible multicollinearity was identified amongst the DAD subscales, but as their correlation did not reach >.90, they were included in the analysis separately at this stage, until unique variable analysis was performed.

7.2.2.3.7.Unique Variable Analysis (UVA)

In phase 1, UVA identified 3 possible sets of variables for merging: "S/he is interested in having new experiences" and "S/he is interested in learning new things"; "S/he has friends" and "Getting together with friends is important to her/him"; and "S/he has initiative" and "S/he has motivation". The first two pairs of items were merged into two new variables (named "Novelty" and "Social" respectively), as these items were thought to have sufficient theoretical overlap. Motivation and initiative were considered theoretically distinct so were not merged. The effect of not merging these items was checked, by combining these items and re-running the phase 1 analysis to check for differences. Though differences in edge weights occurred, the overall structure generally was replicated, and the centrality of the new node 'motivation-initiative' remained consistent with the two nodes when they were separate in the model (see Appendix 11), so they remained separate for the subsequent analyses.

In phase 2, UVA identified one additional possible set of variables for merging: DEMQOL items related to being "worried or anxious" and "sad". These were considered substantially theoretically distinct and not merged, so no further item reduction occurred in this phase.

In phase 3, UVA identified three additional sets of variables suitable for merging: the three DAD subscales; Verbal fluency and MoCA; and BBS and TUG. All three sets were considered to have sufficient theoretical overlap, so were merged, resulting in the creation of three new latent variables named "ADL", "EF" (executive function), and "Mobility" respectively.

The next section presents the results of the network analyses of each phase after this item reduction was performed. The results of the network analyses prior to UVA are not presented here, as the findings were largely similar, except for the revelation of some of the effects of the measures that were merged, which is expected given the known effects of topological overlap on networks [251,254]. For example, merging the three ADL subscales into one ADL measure resulted in greater centrality indices for ADL. Appendix 12 shows the networks and their centrality estimates before and after item reduction.

The labels for each node in the (post-UVA) networks, and what these represent, are outlined in Table 12.

Table 12. *List of nodes in the three phases (after item reduction) and their descriptions*

Node label	Measure / item	Item description	Low score indicative of
Interest	AES 1	S/he is interested in things.	More apathy
GetDone	AES 2	S/he gets things done during the day	More apathy
Started	AES 3	Getting things started on his/her own is important to him/her	More apathy
Novelty	AES 4 & 5	'S/he is interested in having new experiences' and 'S/he is interested in learning new things' combined after UVA	More apathy
Effort	AES 6	S/he puts little effort into anything*	More apathy
Intensity	AES 7	S/he approaches life with intensity	More apathy
Completion	AES 8	Seeing a job through to the end is important to her/him	More apathy
TimeInterest	AES 9	S/he spends time doing things that interest her/him	More apathy
WhatToDo	AES 10	Someone has to tell her/him what to do each day*	More apathy
SelfConcern	AES 11	S/he is less concerned about her/his problems than s/he should be*	More apathy
Social	AES 12 & 13	'S/he has friends' and 'Getting together with friends is important to her/him' combined after UVA	More apathy
Excited	AES 14	When something good happens, s/he gets excited	More apathy
Insight	AES 15	S/he has an accurate understanding of her/his problems	More apathy
ImportDone	AES 16	Getting things done during the day is important to her/him	More apathy
Initiative	AES 17	S/he has initiative.	More apathy
Motivation	AES 18	S/he has motivation.	More apathy
NotSlowed	HADS 2	I feel as if I am slowed down*	More depression
InterestinAppear	HADS 6	I have lost interest in my appearance*	More depression
StillEnjoy	HADS 3	I still enjoy the things I used to enjoy	More depression
Laugh	HADS 7	I can laugh and see the funny side of things	More depression
LookForward	HADS 10	I look forward with enjoyment to things	More depression
EnjoyMedia	HADS 14	I can enjoy a good book or radio or TV programme	More depression
Cheerful	DEMQOL proxy 1	Cheerful*	Less cheerful
NotWorried	DEMQOL proxy 2	Worried or Anxious	More worry

Node label	Measure / item	Item description	Low score indicative of
Energy	DEMQOL proxy 4	Full of Energy*	Less energy
NotSad	DEMQOL proxy 5	Sad	More sadness
NotIrritable	DEMQOL proxy 9	Irritable	More irritability
NotLonely	DEMQOL 8	Lonely	More loneliness
ADL	DAD subscales	DAD subscales (initiation, planning & organization, effective performance) combined after UVA	Worse ADL
EF	MoCA & Verbal Fluency	MoCA scored as per manual & Verbal fluence (Number of animals correctly named), combined after UVA	Worse executive function
Mobility	BBS & TUG	BBS score (sum of item scores) and TUG score (time to walk in seconds), combined after UVA analysis	Worse mobility
Frailty	SHARE-FI	SHARE frailty index scored as per manual	More frail
Age	-	Participant Age in years	Lower age

*reverse scored as part of data entry. Some measures were reversed again so that low scores equated to worse outcomes, as stated in the final column.

Abbreviations: AES, Apathy Evaluation Scale [44]; HADS, Hospital Anxiety and Depression Scale [294]; DEMQOL, Dementia Quality of Life measure [298]; MoCA, Montreal Cognitive Assessment [300]; EF, Executive Function; BBS, Berg Balance Scale [302]; TUG, Time Up and Go [303]; SHARE-FI, Survey of Health, Ageing and Retirement in Europe - Frailty Instrument [301].

7.2.3. Gaussian Graphical Model

7.2.3.1. Phase 1. Results of apathy network

A network of 16 apathy indicators with 84 edges was produced (a table of edge weights is provided in Appendix 13). The edges reflect partial correlations, which were generally weak (mean absolute edge weight= .06): the strongest absolute edge was between ‘Motivation’ and ‘Initiative’ (r_S= .51), the weakest absolute edges were <.01, and 30% of possible edges were set to 0 (see Appendix 13 for edge weight matrix).

Figure **3** shows the resultant network plot of apathy indicators. Due to the large number of weak edges included in the network, a minimum value was applied to the network plot to enable easier interpretation. Initially, this was set at .05, however following stability checks, this was increased to .08 (further explanation for this is provided below). Visual inspection of the network indicated that symptoms did not form distinct clusters. 'Initiative', closely followed by 'Motivation' had the greatest expected influence, and 'Effort' and 'Self-concern' had the least expected influence, as shown in Figure 4. The majority of included edges were positive, and six were negative.

The sample edge weight estimates were similar to the bootstrapped means (see Appendix 13 for edge weight stability plot). Thirty-two of the edges included in the model were present in ≥95% of the bootstrapped samples. Of the edges that did not meet these criteria (*N*= 52), all but one had absolute edge weights of <.08, thus were not visualised in the network when the minimum argument of .08 was applied. This was with the exception of the negative edge between 'Social' and 'Completion' (r_S= -.12), so particular caution should be taken when inferring from this edge in the plot. Centrality indices showed good stability, as the CS-coefficient reached the highest level tested (.75). Figure 5 shows centrality stability plot.

Figure 3. Phase 1 Network plot.

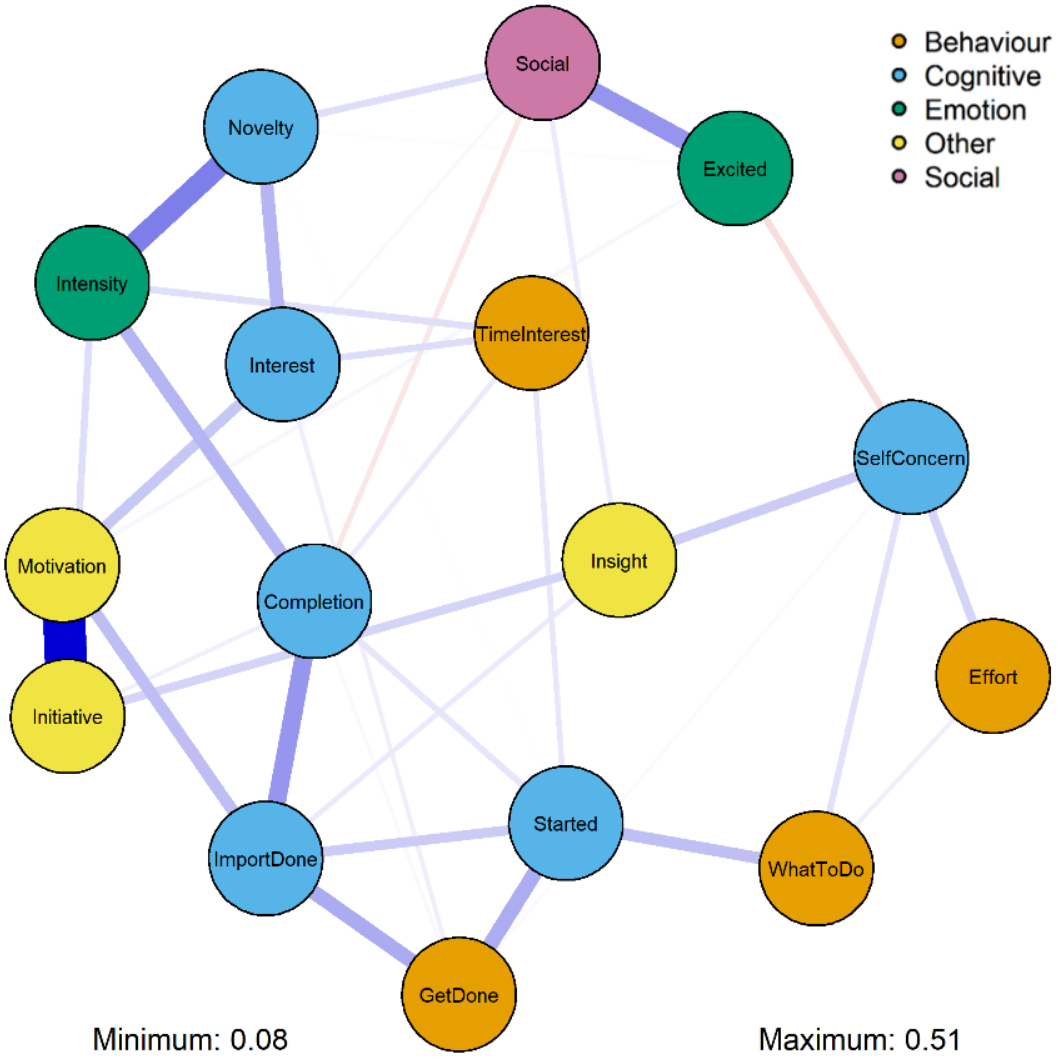


Figure shows apathy indicators, after item reduction via Unique Variable Analysis, estimated with a Gaussian Graphical Model. Blue edges (lines) indicate a positive association, and negative edges indicate a negative association. The thickness of the line corresponds to the relative edge weight (strength of association), with edges <.08 not visible. Node colours indicate the proposed domain of apathy that the item measures, as indicated by the figure legend.

Figure 4. Phase 1 Centrality Plot

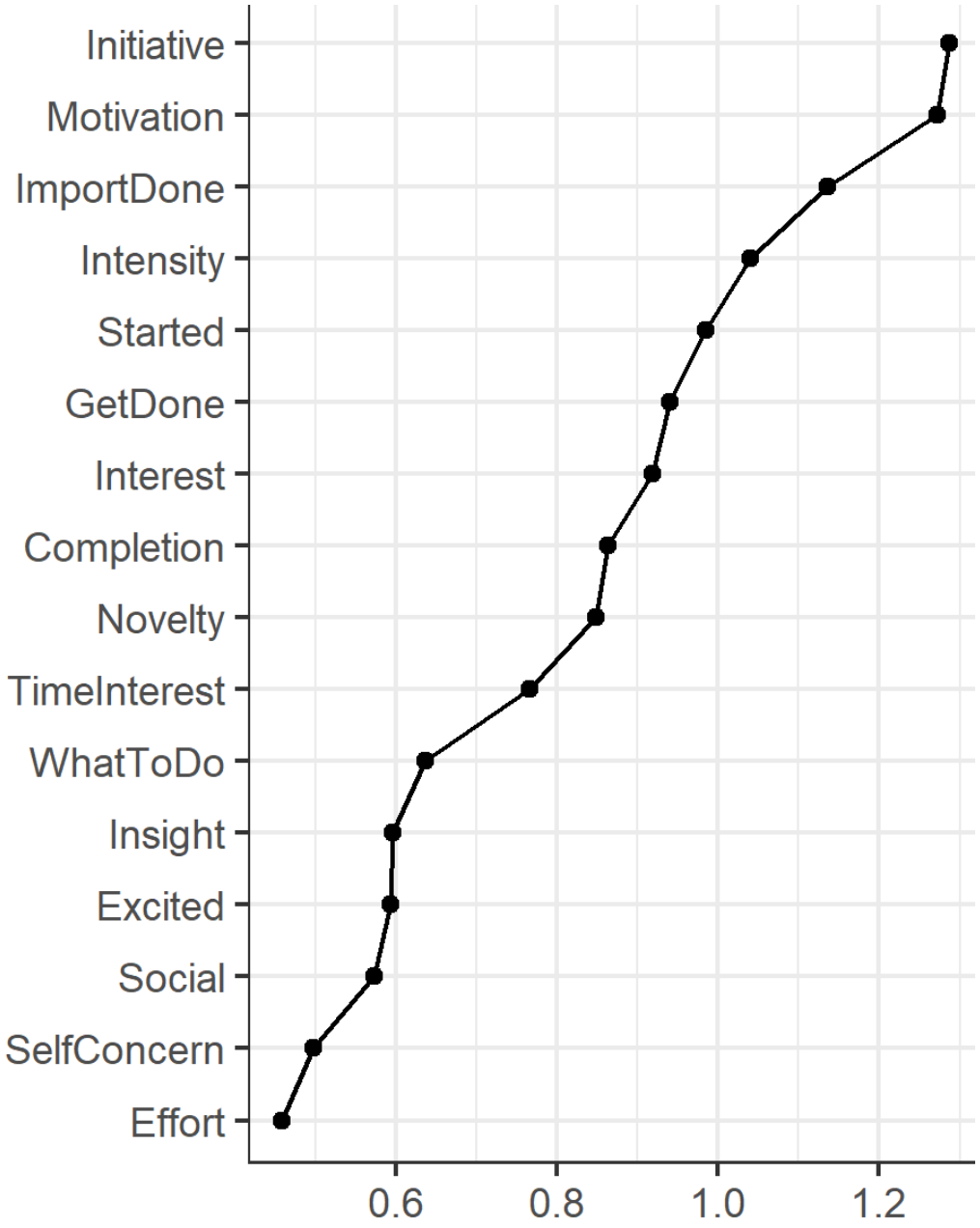


Figure shows centrality plot of apathy indicators after item reduction (y-axis), ordered by expected influence (x-axis), with higher values indicating greater centrality.

Figure 5. Phase 1 Centrality Stability Plot

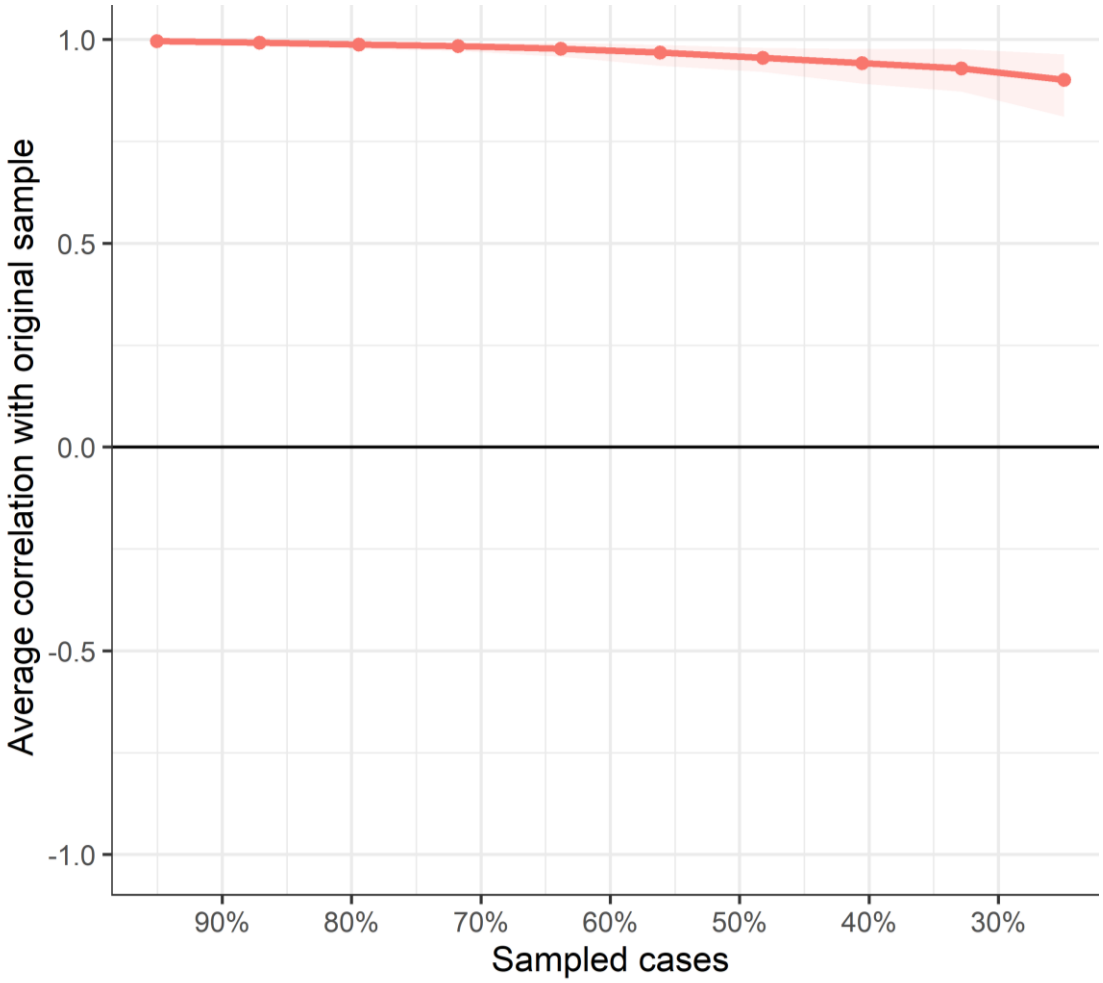


Figure shows correlation between bootstrapped average expected influence and original sample expected influence with increasing dropped cases. The line indicates the mean, and the shaded area indicates the 2.5th to 97.5th quantile.

7.2.3.2. Phase 2. Results of apathy and depression network

A network of 16 apathy and 12 depression indicators with 141 edges was produced (see Appendix 14 for table of edge weights). The edges reflect partial correlations, which were generally weak (mean absolute edge weight= .03): the strongest absolute edges were r_S= .44 (connecting 'Motivation' and 'Initiative', as in phase 1) and r_S= .42 (connecting 'NotWorried' and 'NotSad'), the weakest were <.01, and 63% were set to zero. As in the previous phase, a minimum value of .08 was applied to the network plot to enable easier visual interpretation.

The resultant network of apathy and depression indicators is displayed in Figure 6. From visual inspection, indicators did appear to cluster depending on whether they were intended to measure apathy or depression, with the apathy indicators forming one cluster, and depression clustering separately. Depression could be seen to contain two clusters, one characterised by mood: 'Cheerful', 'NotSad', 'NotIrritable', 'NotWorried'; and another perhaps characterised by outlook: 'Still Enjoy', 'LookForward', 'Laugh' and to a lesser extent: 'NotSlowed' and 'InterestInAppear'.

Figure 7 presents the centrality (unstandardized expected influence) of each node. The nodes with the largest expected influence were 'Initiative' and 'Motivation', as in the first phase. The depression indicator with the highest expected influence was 'Cheerful', whilst 'EnjoyMedia' and 'NotLonely' had the least overall expected influence. As in phase 1, stability of the centrality indices was sufficient, as the CS-coefficient reached the highest level tested (.75). Centrality stability is displayed in Figure 8.

Stability analysis showed that the sample estimated edge weights were similar to bootstrapped means (see Appendix 14 for edge weight stability plot). Forty-nine of the 141 edges included in the model were present in ≥95% of the bootstrapped estimates, and all of these were positive. Edges that did not meet these criteria (N= 92) all had an absolute edge weight of ≤.08, so their visualization in the network plot was minimal.

Bridge centrality (i.e. expected influence of each node only on the domain it did not belong to) is displayed in Figure 9. 'Energy' had the highest bridge centrality, followed by 'Cheerful'. Taking only the most stable edges (i.e. present in ≥95% of the bootstrapped estimates) of these bridges: 'Energy' was connected to depression items of 'Cheerful' (r_S= .19) and 'NotSlowed' (r_S= .08), and apathy items of 'GetDone' (r_S= .11) and 'Novelty' (r_S= .09); and 'Cheerful' was connected to the depression items 'NotSad' (r_S= .25) and 'NotIrritable' (r_S= .08) and the apathy item 'Interest' (r_S= .09) in addition to 'Energy'.

Figure 6. Phase 2 Network Plot

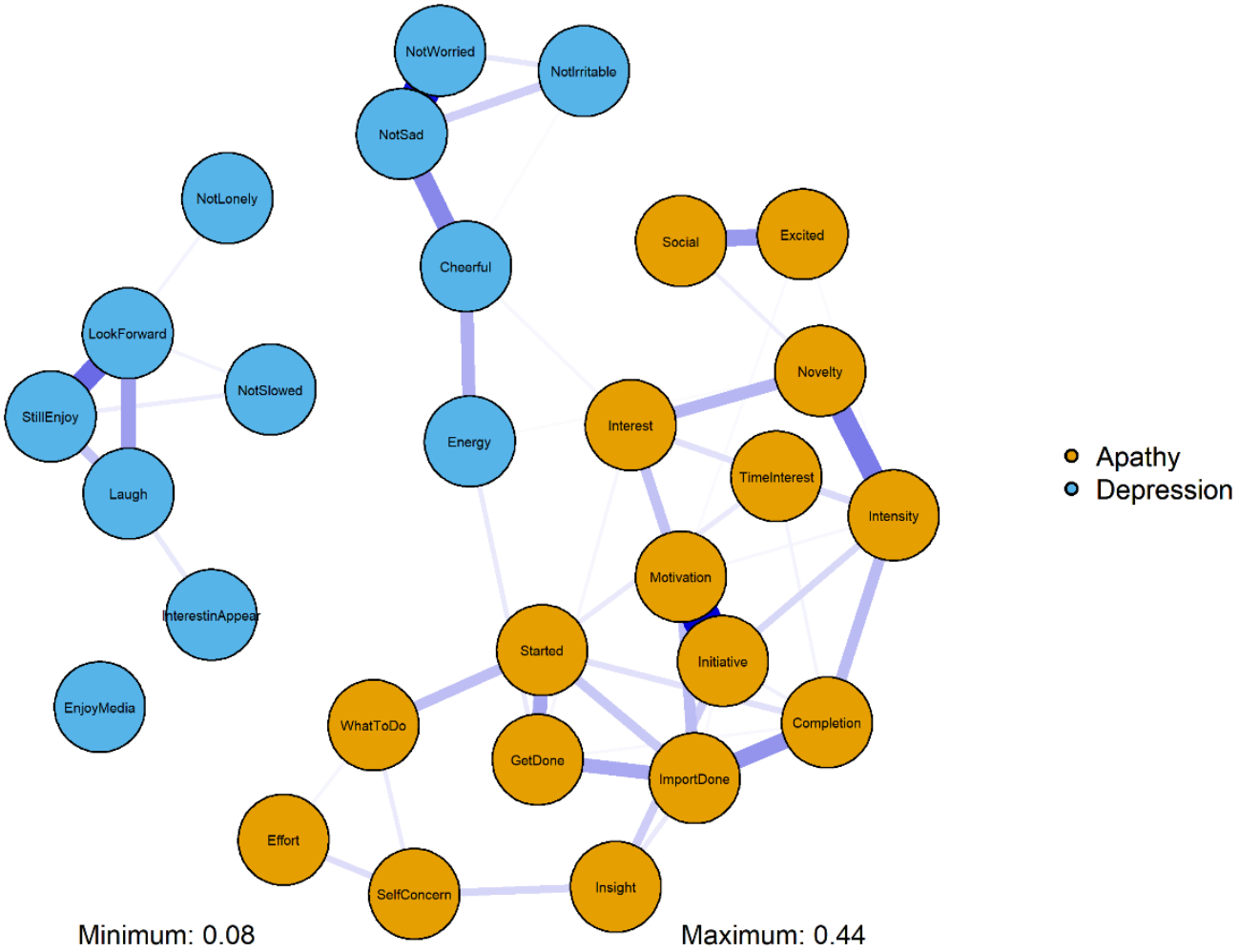


Figure shows network plot of apathy and depression indicators, after item reduction Unique Variable Analysis, estimated with a Gaussian Graphical Model. Blue lines indicate a positive association. The thickness of the line corresponds to the relative edge weight (strength of association), with edges <.08 not visible. The colour of the node reflects the proposed construct it measures, as indicated by the figure legend.

Figure 7. Phase 2 Centrality Plot

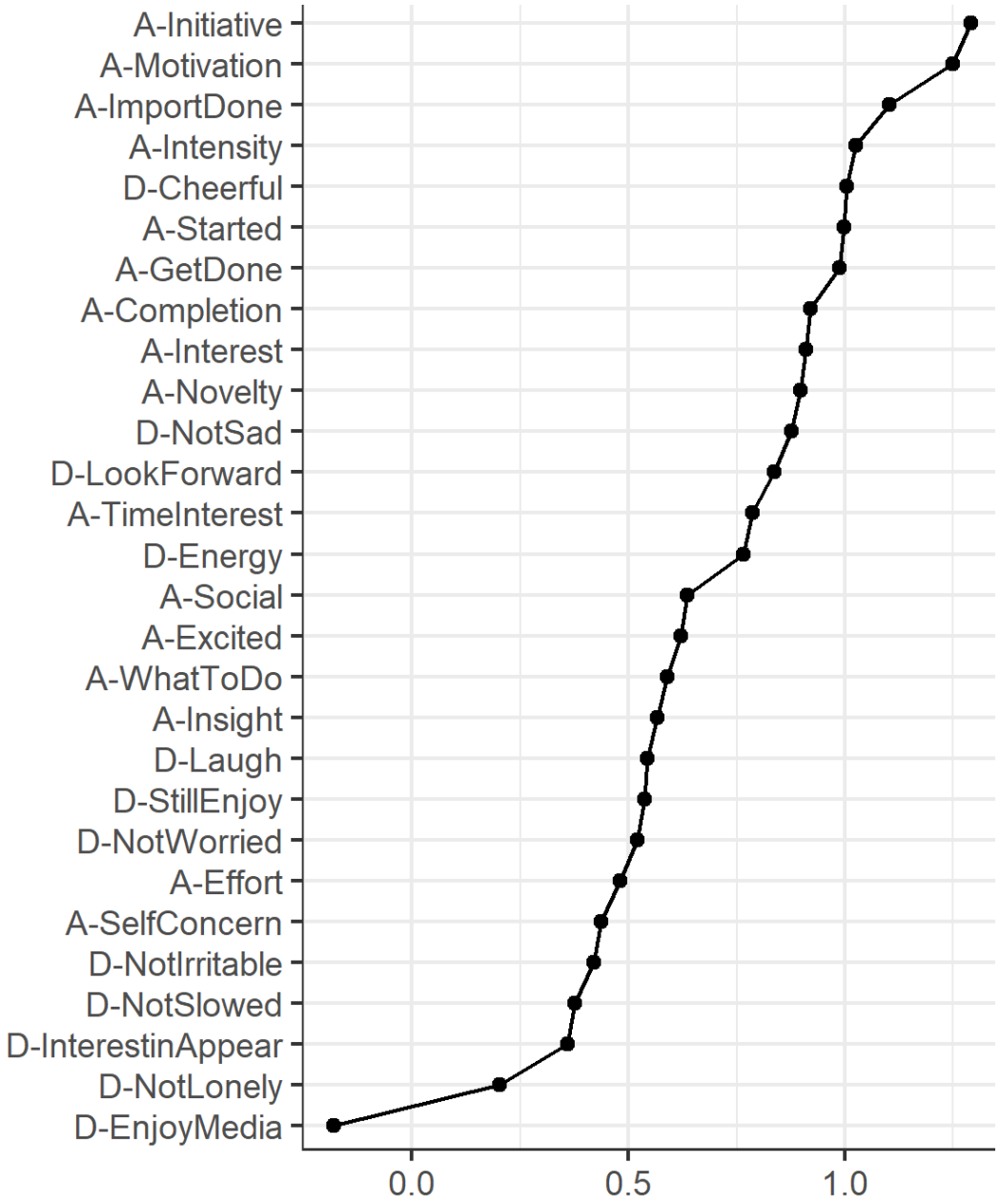


Figure shows centrality plot for apathy indicators (with prefix ‘A-’) and depression indicators (with prefix ‘D-’) after item reduction (y-axis), ordered by expected influence (x-axis), with higher values indicating greater centrality.

Figure 8. Phase 2 Centrality Stability Plot

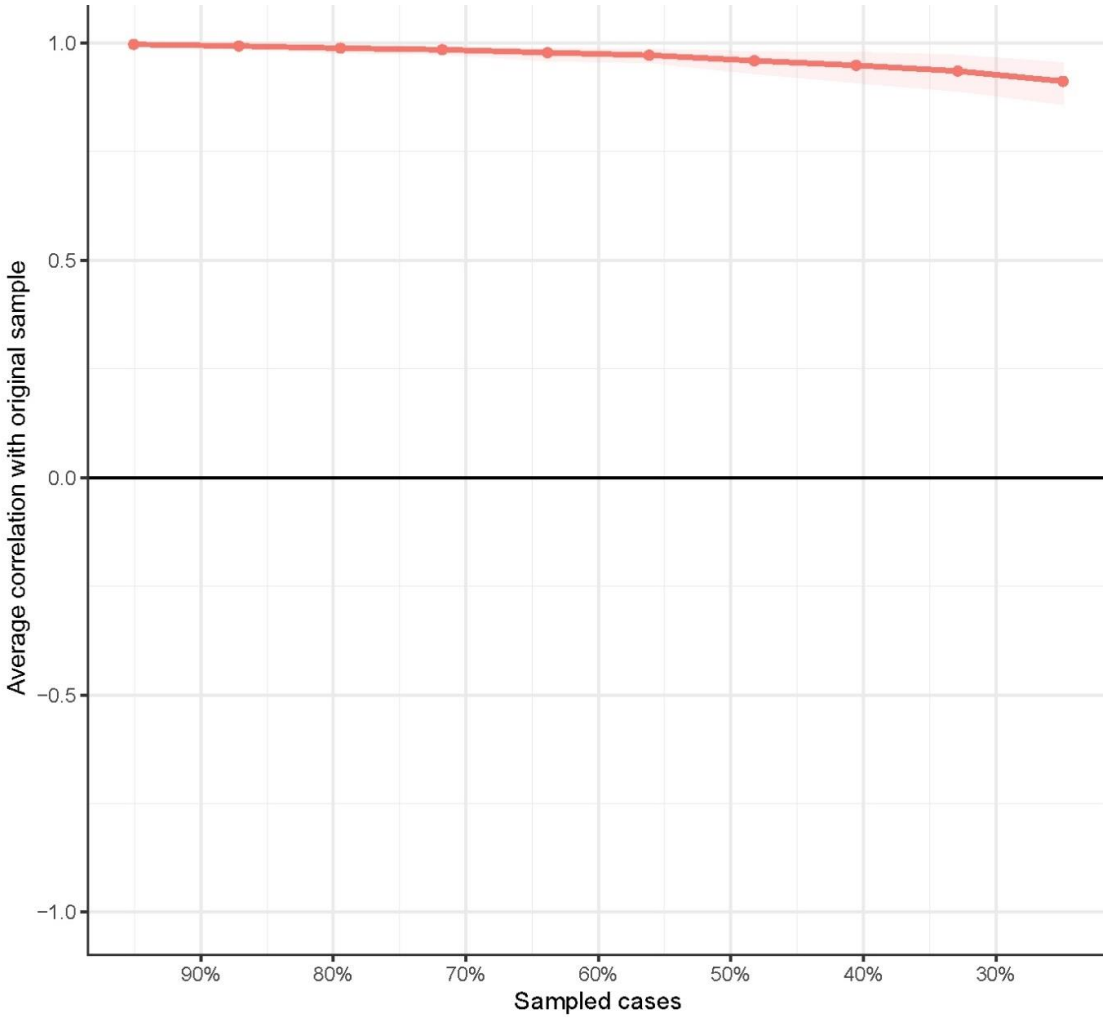
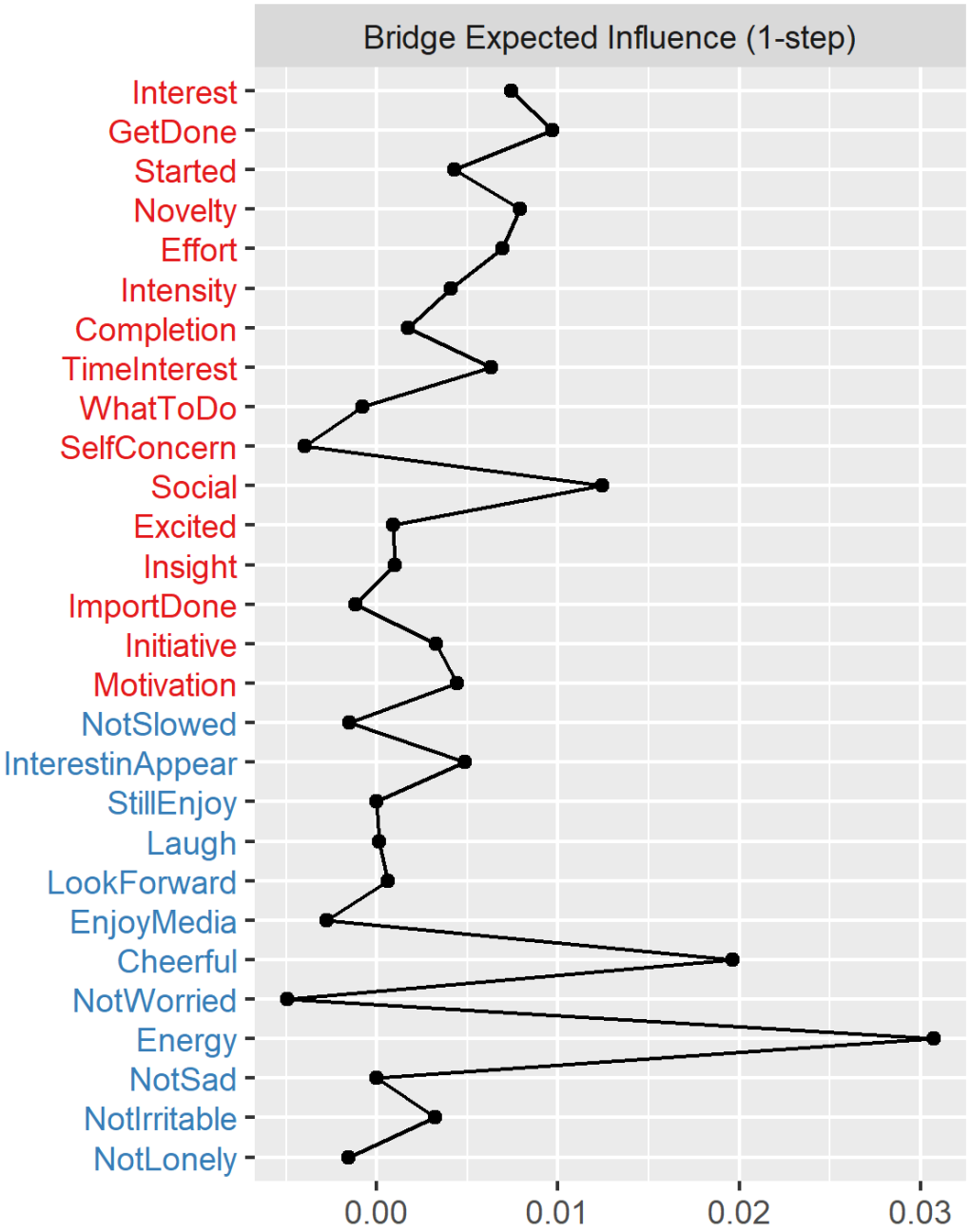


Figure shows correlation between bootstrapped average expected influence and original sample average expected influence with increasing dropped cases. The line indicates the mean, and the shaded area indicates the 2.5th to 97.5th quantile.

Figure 9. Phase 2 Bridge Centrality Plot



Bridge centrality plot for apathy indicators (red) and depression indicators (blue) after item reduction (y-axis), with higher values along the x-axis indicating greater bridge expected influence (i.e. expected influence with nodes not from the same construct).

7.2.3.3. Phase 3. Results of overall network

A network of 16 apathy indicators, 12 depression indicators, and 5 ‘external’ nodes, with 185 edges was produced. The edges were further weakened compared to the previous phases (mean absolute edge weight =.02), with 65% of possible edges set to zero. The two strongest edges remained the same as in phase 2. A table of edge weights is provided in Appendix 15. As in the previous phases, a minimum value was applied to the network plot to enable easier visual interpretation.

Figure 10 shows the resultant network plot of apathy indicators, depression indicators and external variables of ADL, Frailty, Mobility, Executive Function, and age. Visual inspection of the plot shows that apathy and depression remained relatively distinct, were less internally connected, and connected to each other through external variables and the depression item ‘Cheerful’ which connected with apathy item ‘Interest’. Figure 11 shows the centrality (unstandardized strength) of the nodes included in the network. ‘Initiative’, ‘Motivation’ and ‘ADL’ had the greatest strength centrality, and, as in the previous phase, ‘Enjoy Media’ and ‘Not Lonely’ had the lowest centrality.

Fifty-seven of the edges included in the model were present in ≥95% of the bootstrapped samples. All these edges were positive, except those connecting Age with Mobility and Frailty, reflecting that increasing age was associated with worse Mobility and Frailty. All edges estimated as non-zero in <95% of bootstrapped samples (*N*=128) had an absolute edge weight of <.08, meaning they were not visible in the network plot. As in the previous phases, stability of the centrality indices was sufficient, as the CS-coefficient reached the highest level tested. A centrality stability plot is displayed in Figure 12. Centrality was only weakly negatively correlated with SD (r= -.13), and this was not significant (t= -0.70, df= 31, p= .49).

ADL, Energy and Frailty had the highest bridge strength respectively (as shown in Figure 13. Of the stable edges (i.e. present in ≥95% of the bootstrapped estimates) of these bridge nodes: ADL connected to executive function (r_S= .20) and Frailty (r_S= .12), apathy items ‘What To Do’ (r_S= .25), ‘Get Done’ (r_S= .13), and ‘Initiative’ (r_S= .11), and no depression items; ‘Energy’ connected to the depression item ‘Cheerful’ (r_S= .19), the apathy item ‘Novelty’ (r_S= .08) and the physical measure ‘Frailty’ (r_S= .24); and ‘Frailty’ connected to Age (r_S= -.11) and ‘Mobility’ (r_S= .31) in addition to ADL and ‘Energy’. Mobility was connected to Frailty (as above), Age (r_S= -.15), depression item ‘Not Slowed’ (r_S= .16) and Executive Function (r_S= .11).

Figure 10. Phase 3 Network Plot

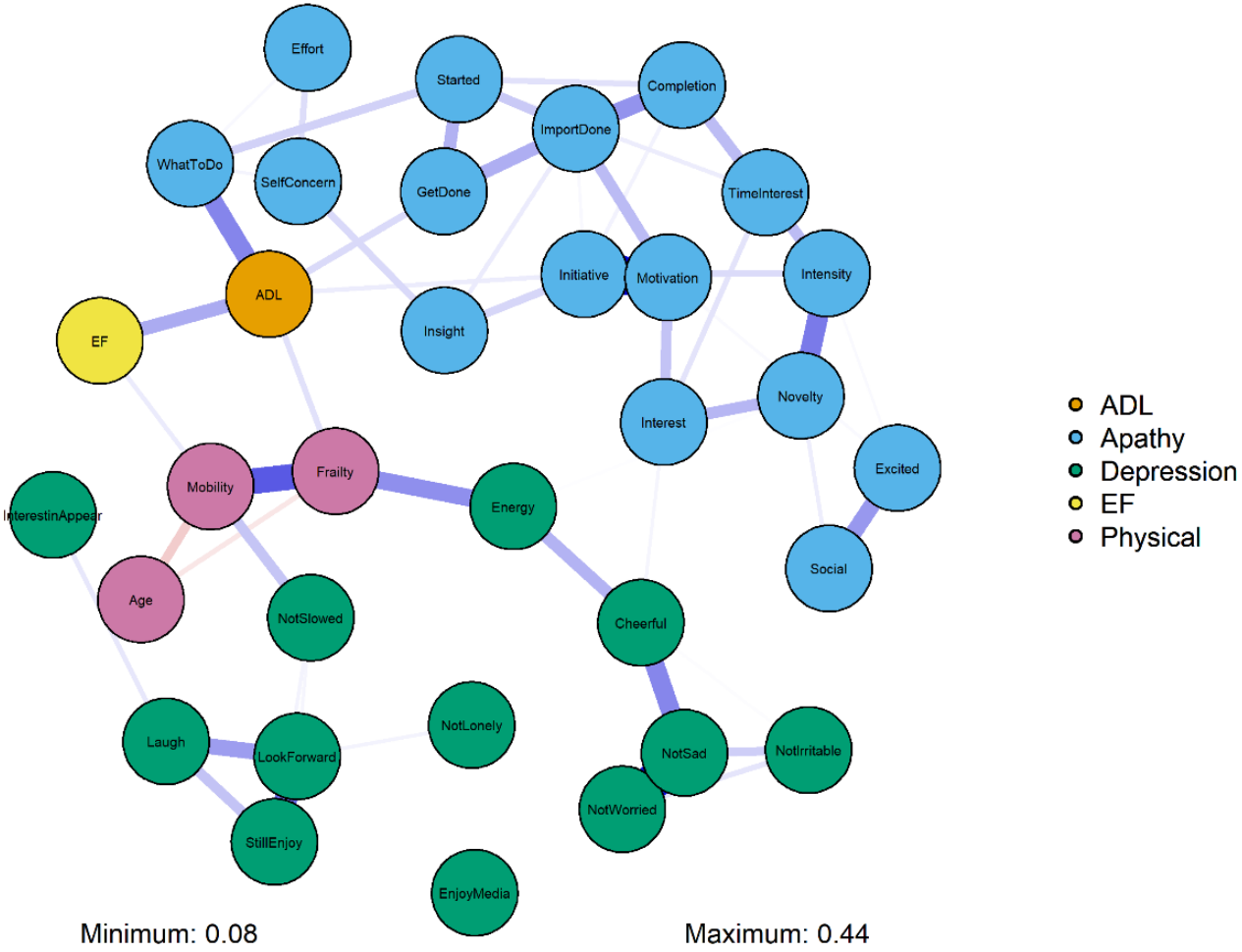


Figure shows network plot of apathy, depression and external variables, after item reduction Unique Variable Analysis, estimated with a Gaussian Graphical Model. Blue lines indicate a positive association, and red lines indicate a negative association. The thickness of the line corresponds to the relative edge weight (strength of association), with edges <.08 not visible. Node colours indicate the proposed construct measured, as indicated by the figure legend. Abbreviations: ADL, Activities of Daily Living; EF, Executive Function.

Figure 11. Phase 3 Centrality Plot

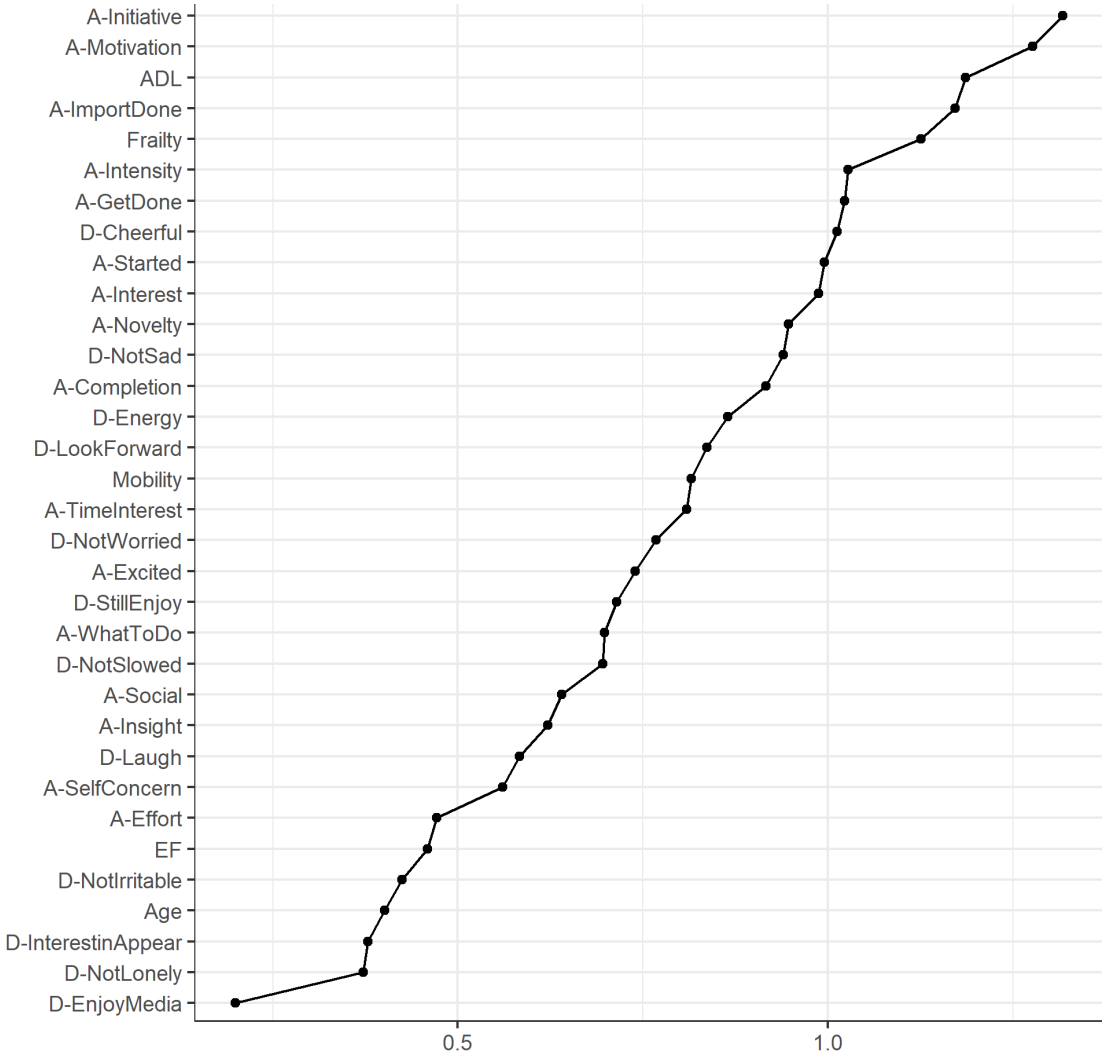


Figure shows centrality plot for apathy indicators (with prefix 'A-'), depression indicators (with prefix 'D-'), and external variables, after item reduction (y-axis), ordered by expected influence (x-axis), with higher values indicating greater centrality. Abbreviations: ADL, Activities of Daily Living; EF, Executive Function.

Figure 12. Phase 3 Centrality stability plot

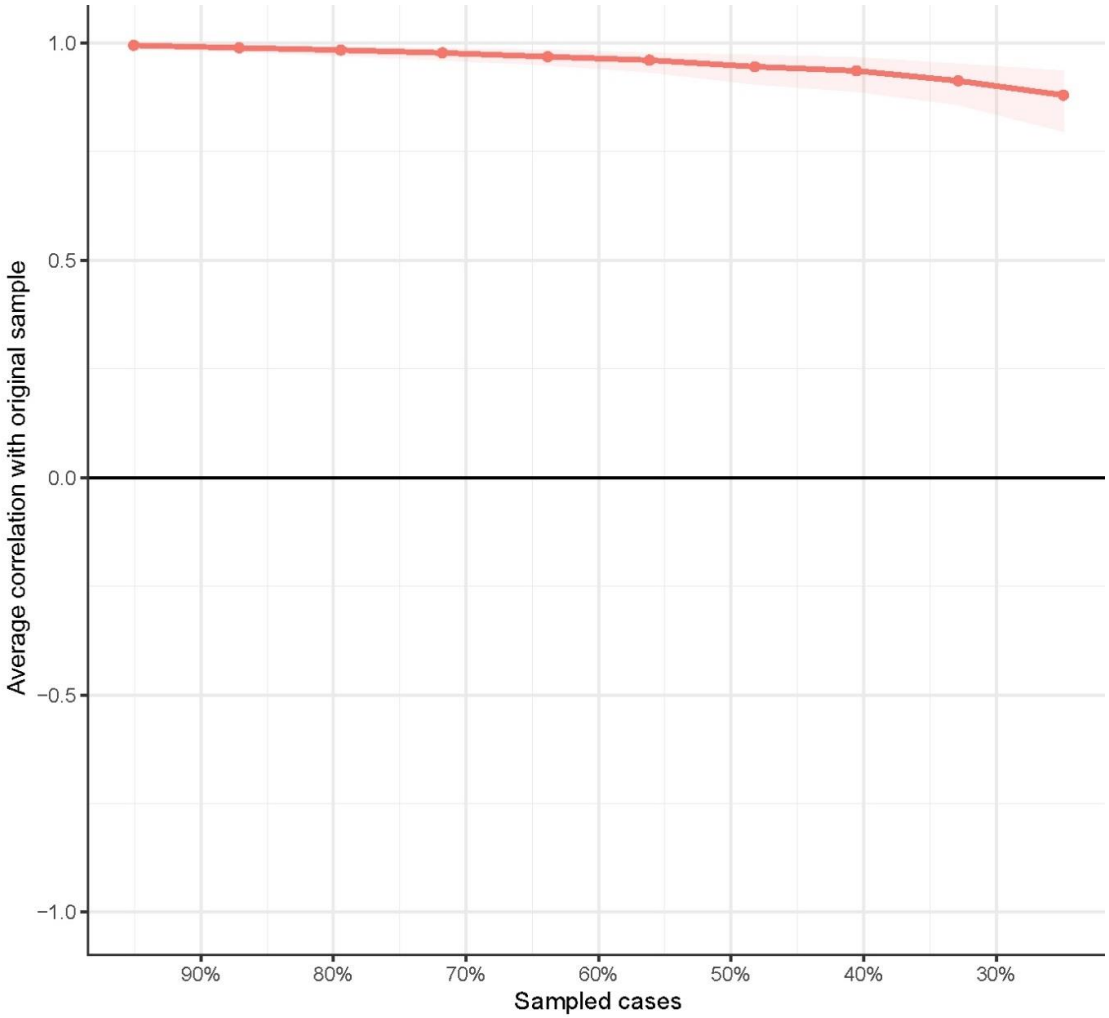


Figure shows correlation between bootstrapped average strength and original sample average strength with increasing dropped cases. The line indicates the mean, and the shaded area indicates the 2.5th to 97.5th quantile.

Figure 13. Phase 3 Bridge Centrality Plot

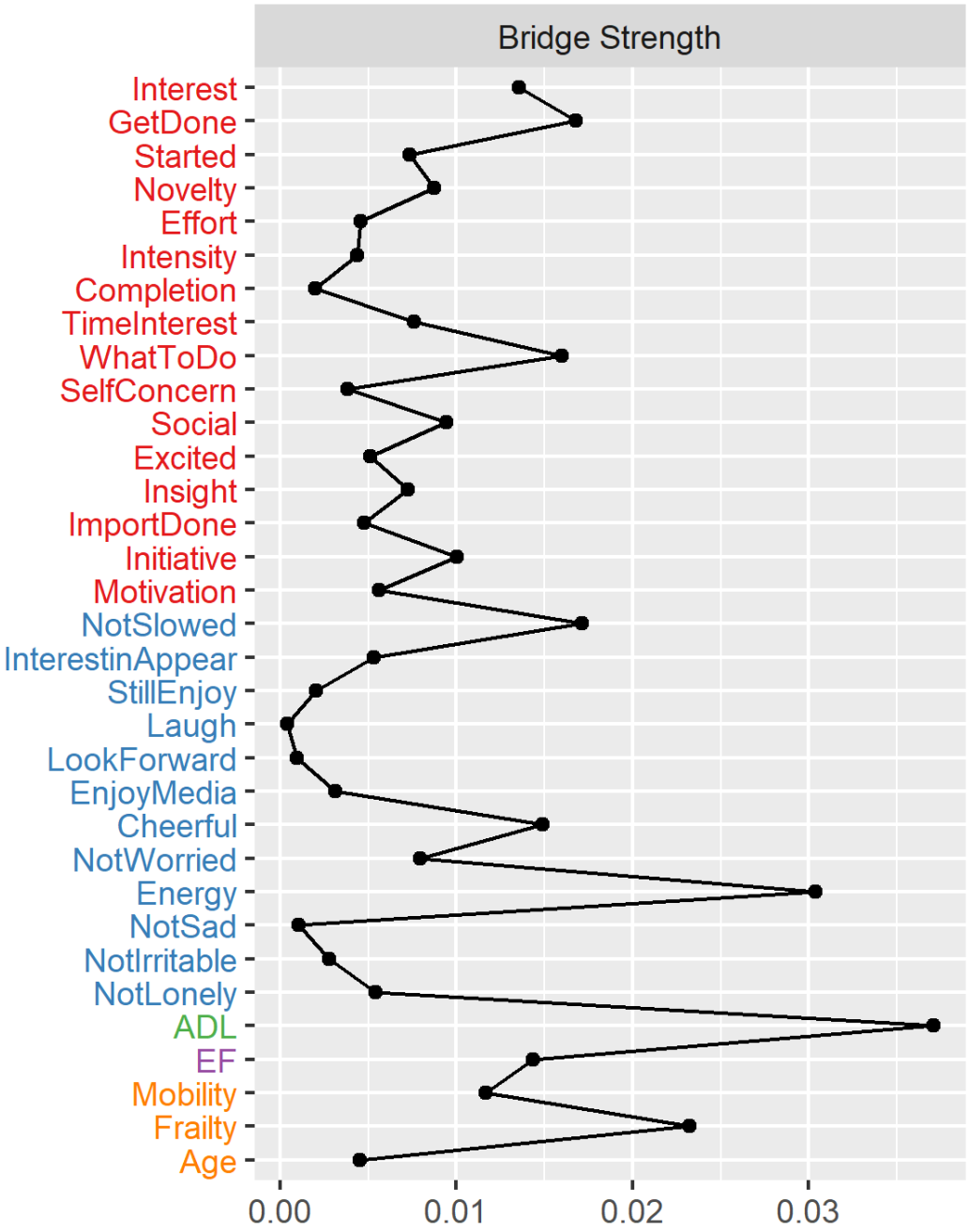


Figure shows bridge centrality plot for apathy indicators (red), depression indicators (blue), ADL (Activities of Daily Living; green), EF (executive function; purple), and physical characteristics (orange) after item reduction (y-axis), with higher values along the x-axis indicating greater bridge expected influence (i.e. expected influence with nodes not from the same construct).

7.2.4. Mixed Graphical Model

7.2.4.1. Sensitivity analysis

Mixed Graphical Model (MGM) analyses yielded mostly similar results in all three phases, though as expected, some differences occurred. MGM

estimated more edges than GGM (111, 193 and 235 edges estimated in phases 1, 2 and 3 respectively). In all phases, few edges were found in the GGM that were not also included in the MGM, and those that were had absolute edge weights <.08. The MGM did include some larger edges that were not included in the GGM, mostly in phase 1, where the largest edge not included in GGM was .14, ('Get Done'– 'Intensity'), and others between .12 and .09 ('GetDone'–'TimeInterest'; 'Effort' –'TimeInterest'; 'GetDone'–'Motivation'; 'Started'–'Motivation'). In phase 2, all edges found in the MGM not included in the GGM were ≤.08 except the largest edge found in MGM phase 1 ('Get Done'–'Intensity'), which was also present here and again in phase 3. Phase 3 also resulted in another edge >.08 in the MGM, that was not present in GGM, this was .10 ('EnjoyMedia'–'NotWorried'). Due to the differences in methods used to obtain edges, absolute edge weights are not comparable between methods. Side by side comparisons of plots are provided in Appendix 16.

7.2.4.2. Predictability

Predictability (proportion of variance explained) for most nodes was high, suggesting that the networks included many important predictors. However, for some nodes, a high proportion of variance was explained by the measure itself, i.e. the marginal correct classification, for example, in the node 'EnjoyMedia' this was .80, reflecting the poor distribution in the item response. Predictability is displayed in MGM plots in Appendix 16, and values are presented in Appendix 17.

7.3. Discussion

In this next section, the overall findings of the network analysis sub-study will be briefly summarised, along with the limitations relevant to this method. A more detailed discussion of the results will be provided in Chapter 9, where the findings will be integrated with the findings of the qualitative sub-study, and their interpretations will be discussed together.

7.3.1. Summary of findings

Three networks were produced using the GGM method, with the subsequent network expanding on the previous network. These networks showed that apathy items did not cluster into distinct domains, and motivation and initiative reflected the most central indicators of apathy. Apathy and depression clustered separately, with lack of energy and feeling cheerful acting as shared indicators of apathy and depression. Furthermore, ADL was associated with apathy indicators, but not depression indicators, except via physical impairment, and physical and cognitive impairment was associated with apathy indicators largely through ADL. These networks were generally replicated using the MGM method, with the strongest edges (associations)

being the most reliable (i.e. stable according to boot-strapped estimates in GGM and replicated in MGM). However, the networks did not perfectly replicate, with MGM largely providing a more conservative estimate, which is expected given the additional parameters required to estimate multi-level MGMs, and the lower sample size due to the list-wise deletion that was required to estimate the MGM.

7.3.2. Limitations of network analysis study

The (regularisation) methods used in the present study assume a sparse network, however little guidance is available to determine what constitutes a sparse network, except that this is relative to the number of possible edges [269]. It has been argued it is unlikely that networks are sparse where all items are designed to measure the same concept [341], in this case, the first phase network of apathy items. However, the items were designed to assess different domains of apathy [44], and similar network analysis studies have been conducted in which items intended to measure one construct are explored [e.g. 277].

Unlike in traditional thresholding methods in which an arbitrary threshold (e.g. p <.05) is selected, the methods used in this study mean that the level of error is unknown. Interpretation of individual edge weights is limited, as Least Absolute Shrinkage and Selection Operator regularisation was used to enhance sensitivity, at the cost of specificity. Furthermore, the bootstrapped confidence intervals in each phase were relatively wide, indicating that comparisons should not be made between individual edge weights. Nevertheless, EBICglasso in GGMs and CV selection in MGMs is recommended, though this is providing conclusions are not being drawn regarding individual edges [306]. Similarly, it is important to note that whilst centrality indicates important nodes in a network, this does not necessarily equate to causally important symptoms [342]. High centrality may be the result of a strong association with an underlying latent trait [310], and the unknown direction of effects in the network means that it is impossible to know when high centrality is the result of a nodes' causal effects on other symptoms, or due to being highly affected by other symptoms [342]. This means conclusions that can be made from this analysis are limited. However, associations should never be understood to equal causation [260], and this was intended as an exploratory study, not designed to make generalisable conclusions nor identify causation.

Inclusion of nodes that essentially measure the same particular concept is highly problematic for network analysis, and this can affect the network structure and centrality estimates [254,292]. I attempted to combat this in the present study by checking for topological overlap using UVA, and reducing nodes identified by UVA if there was sufficient theoretical overlap. Though the

two networks (before and after UVA) were largely similar, the reduction of nodes via UVA was successful in revealing some important edges that were previously hidden by the presence of overlapping nodes, for example, it was only by combining the three ADL subscales, that the importance of ADL in the third phase network was revealed. However, UVA did indicate overlap between some nodes which were not reduced in this analysis, as they were deemed to be theoretically distinct, such as ‘Motivation’ and ‘Initiative’. Whilst it could be argued this is what resulted in their high centrality, the analysis was re-run in phase one with these items combined, and the relative centrality was unchanged (i.e. this new node also had the greatest centrality).

Centrality measures can be biased by different variance between nodes [254]. In the present study, some items had poor variation in response, for example, 'EnjoyMedia', which could explain its low centrality, and items assessed on a sum-score approach had greater variance, which could have inflated the centrality of ADL and frailty for example. However, the correlation between standard deviation and centrality was checked in this final network phase (as recommended by McNally [254]) and found to be minimal.

Network analysis in the field of psychopathology typically assesses symptoms using single items taken from a measure designed for a latent-trait model [251,310]. Whilst there are limitations to this latent trait sum-score approach to measuring complex phenomena such as apathy, their strength is that they are designed in this way to handle measurement error [239]. Methods are available to combine a latent and network approach, with each symptom assessed by a few items, and each node modelled as a latent trait within the network, with measurement error included [271], however this may have required developing a new apathy measure, which was beyond the scope of this PhD.

Following the results of the stability tests, the minimum value of the visualized networks was set to .08, so that only the strongest edges (associations) were visualised. This was done primarily to aid visual interpretation, due to the large number of small edges estimated. This highlights two potential issues. Firstly, the large number of small edges could be suggestive of a dense network, suggesting the regularisation method used may not have been appropriate. Secondly, whilst the minimum value was set specifically at .08 as almost all edges present in <95% of bootstrapped estimates were ≤.08 or less, it is important to recognize that the visualized network plots are not specifically showing the most stable edges.

This study used a cross-sectional ‘macro’ approach, examining symptoms from a group of participants, experienced over a relatively long time period (one to four weeks). It is important to recognise that this type of group-level analysis does not necessarily predict individual networks [264,274], though this was not the aim here. Time series within-person data, where information

collected from one person at regular intervals, a short space apart (e.g. every few hours over a few days), enables the production of an individual-level psychopathology network, indicating which intensity or presence of symptoms at the first time point predict intensity or presence of symptoms at subsequent time points. This has been argued to enable better theorisation of the true causal interactions between symptoms, as it allows for generalisation to individual symptom networks [264]. However, collecting data over multiple time series could be burdensome, particularly for people with dementia. The viability of collecting time series data in this population should be investigated.

As discussed previously, the nodes within a network analysis must not be synonymous, and one should not directly dictate the other [254,292,310]. Whilst every effort was made to ensure the nodes included were indeed distinct, including the use of UVA to identify topological overlap, there are some potential issues with conflation across the measures used. SHARE-FI assesses frailty, and included a question regarding the 'energy' of the participant, and a separate node related to 'energy' was included in the network, which may have driven the association between 'Energy' and 'Frailty' in phase 3. The SHARE-FI also contains an assessment of hand-grip strength, which may also be conflated with apathy. As discussed in section 3.3.2, studies have shown that people with greater apathy are less willing to exert effort on grip strength tasks, and this is thought to be due to reward discounting or over-estimation of effort [64,127]. However, it has been shown that once the decision to act has been made, participants with apathy exert just as much effort on hand-grip strength tests as people without apathy [122].

Network analysis also requires that no important concepts are missing from the network, as they could act as latent traits, resulting in associations between nodes that are not explained by the nodes in the network [310]. Whilst the present study included numerous constructs relevant to apathy, some may have been missed. For example, personality factors such as extroversion may be important, and could explain the relationship between some nodes, such as 'Social', 'Novelty' and 'Excited'. This is discussed further in section 9.6.

7.3.3. Conclusion

This analysis offers a starting point for future research to explore, rather than conclusive statements regarding the interaction of symptoms with one another and external variables. The results of this analysis will be discussed further and integrated with the results from the qualitative study in Chapter 9.

Chapter 8. Qualitative study findings

8.1. Introduction

The aims of this thesis were to characterise apathy and its boundaries, and to understand the possible mechanisms and impact of apathy. The qualitative sub-study in particular sought to: (1) explore participants' understandings and beliefs about apathy; (2) explore experiences of and explanations for apathy indicators; (3) explore how apathy impacts the everyday lives of people living with NCD. This chapter presents the results of the qualitative study. Limitations and conclusions of this sub-study will be noted, with further discussion presented following the integration of the network analysis and qualitative results in Chapter 9.

8.2. Participant characteristics

Twenty-eight participant dyads were approached to take part in the qualitative interview study between September and November 2019. For clarity, here I will refer to people with dementia or MCI as 'people living with NCD', family members or friends who took part as 'carers' and the two collectively or interchangeably as 'participants'. Thirty participants were recruited: two people living with NCD were interviewed alone, and fourteen were interviewed together with a carer. No carers wished to be interviewed separately. Thirteen people living with NCD lived with their spouses, whilst the remaining three lived alone. Five people living with NCD were female, all but one was white, and seven were educated to a college or university level, whilst the remaining had secondary education. The mean age of people living with NCD (at the time of the qualitative interview) was 75.6 years (range=66-91), mean AES score was 46.9 (range=24-66), out of a possible 72, where 41.5 or above indicates clinically significant apathy [27], and the mean MoCA score was 20 (range=16-24), out of a possible 30, with scores of 25 or below indicative of cognitive impairment [300]. Four people living with NCD were participating in the intervention group of PrAISED2 research, and had been allocated this group an average of 24.8 days (range=21-28) prior to their interview, whilst the remaining participants were in the control group. Of the carers that were interviewed, all were spouses of the person living with NCD, except for one who was a daughter, and four were male. The mean age of carers was 72.3 years. Individual participant characteristics are detailed in Table 13. All interviews took place at participants' homes, across four geographical areas: Nottinghamshire (*N*=4), Derbyshire (*N*=5), Lincolnshire (*N*=4) and Bath and North East Somerset (*N*=3). Interviews lasted between 38 and 115 minutes.

Table 13. *Qualitative interview participant characteristics*

PATIENT CHARACTERISTICS							CARER CHARACTERISTICS		
Pseudonym	**Age**	**AES**	**MoCA**	**Highest level of education**	**Diagnosis**	**Ethnicity**	**Pseudonym**	**Age**	**Relationship**
Mary	66	51	18	College / University	AD	White	Charles	70	spouse
David	69	47	23	Secondary education	AD	White	Patricia	69	spouse
John	87	50	24	College / University	VaD	White	Betty	85	spouse
Rahul	76	24	20	College / University	Mixed	South Asian	Ruby	71	spouse
Peter	73	51	22	Secondary education	MCI	White	Diane	69	spouse
Robert	66	50	20	Secondary education	AD	White	Sandra	61	spouse
Paul	72	41	23	Secondary education	VaD	White	Karen	50	child
Adrian	73	42	21	Secondary education	VaD	White	Susan	71	spouse
Chris	91	41	20	College / University	AD	White	Judith	89	spouse
Anne	87	47	16	College / University	AD	White	William	87	spouse
Stephen	70	26	16	Secondary education	MCI	White	-	-	-
Richard	77	63	21	College / University	AD	White	Martha	73	spouse
Helen	66	29	24	Secondary education	MCI	White	-	-	-
Nick	85	59	19	Secondary education	VaD	White	Cynthia	76	spouse
Linda	67	64	17	Secondary education	AD	White	Donald	67	spouse
Jean	85	66	16	College / University	AD	White	Joseph	74	spouse

Abbreviations: AD, Alzheimer's Disease; Mixed, Mixed dementia; MCI, Mild Cognitive Impairment; VaD, Vascular Dementia.
- Indicates that no carer took part alongside this participant in the qualitative interview.

8.3. Themes (results)

Four themes were constructed: one functioned as a domain summary theme 'Apathy is Poorly Understood', and three were interpretive themes: '"*Too much trouble*": Mediating Effort and Outcome'; 'Preserving Identity in the Face of Loss of Capability and Autonomy'; 'Opportunity and Exclusion'.

8.3.1. 'Apathy is Poorly Understood'

This domain summary theme was closely related to the research question 'what constitutes apathy?'. It captures participants' understanding of apathy and instances where people living with NCD displayed what would be considered a symptom of apathy based on proposed diagnostic criteria.

Participants' interpretations and understandings of apathy varied. For some, it was an unfamiliar term, not used in everyday conversation.

> Nick: *Never heard it before. Apathy?*
> Interviewer: *Apathy or apathetic.*
> Nick: *What do you mean, feeling sorry for yourself? Is that what you mean?*
> Interviewer: *So, I'm being difficult and asking you what you think it means[...]*
> Nick: *Oh I don't know what it means.*
> Cynthia: *No, I don't.*
> Nick: *Apathy… no.*
> Cynthia: *I have to say it's not one of our words, let's put it that way.*
>
> (Nick, 85 year old with VaD, and Cynthia, 76-year old carer)

Where 'apathy' was a familiar term, it was interpreted in different ways by participants, but often described as someone who "can't be bothered", suggestive of an intentional lack of effort.

> Interviewer: *I wonder what you thought of that word, what does it mean to you?*
> Paul: *It means you can't be bothered. I think that's all it means isn't it? No, no, I'm not apathy, I'm not that.*
>
> (Cynthia, 72-year-old with VaD)

Apathy was typically viewed as a negative term, and therefore not surprisingly this label was rejected by many participants.

> Charles: *[Apathy]'s a bit of a negative word really.*
> Mary: *It is a bit of a negative word. […] I think it is a bit negative. And I prefer to think about … what's happening rather than-*

> Charles: *I think motivation is the – reasonably more accurate. Certainly in her case, rather than apathy, it's not. It's not, she's … <sighs> lacking in … well you know it's subtle isn't it.*
>
> (Mary, 66-year-old with AD, and Charles, 70-year-old carer)

The above indicate that participants sometimes lacked awareness of apathy in the context of NCD. Despite this, most people living with NCD experienced what is typically described in the literature as the components of apathy, i.e. lack of or reduced: motivation, initiation, activity, going out, interest, enjoyment, and social engagement.

> *But you don't… you don't get motivated to do anything, do you… particularly? You know in the morning, he doesn't get up "ooh, shall we go for a walk or- ?" You know, that motivation has gone, hasn't it, a bit?*
>
> (Martha, 73-year-old carer)

Occasionally it was difficult for participants to explain or understand these experiences. It was sometimes felt that people living with NCD lacked some form of drive, and participants often expressed this apparent apathy in alternate terms.

> *[I] Need to just stop being a lazy cow! It's like my friend, she finishes work at half past 12 on a Tuesday, and probably quarter to 12 I might get out of bed and have a shower and be all ready for when she comes <laughs>. And there's no need for it is there?*
>
> (Helen, 66-year-old with MCI)

> *[referring to how she feels] You might have the desire to do things, but, you don't have the get up and go.*
>
> (Anne, 87-year-old with AD)

In summary, participants' understanding of apathy varied. For some, apathy was an unfamiliar term, and for others it had negative connotations. Generally, though participants described indicators of apathy, they did not view apathy as applicable to them or part of NCD. Overwhelmingly however, patients and carers' explanations for these experiences was that things were more difficult, which formed the basis of the next theme.

8.3.2.'*Too much trouble*': Mediating Effort and Outcome

This theme captures how participants' accounts of their experience were characterised by struggle: they faced difficulties, failures and setbacks in everyday life, and things seemed to require more effort (subtheme 1: Daily Struggle). Participants had to reduce this effort, and wanted to avoid negative consequences, which often meant withdrawing from activities, interests, and social interaction. Things could be considered not worth the effort or risk

unless they were particularly necessary or purposeful (subtheme 2: Impact of Consequence and Purpose).

8.3.2.1. Daily Struggle

People living with NCD described experiences of impaired attention, forgetting to do things, finding it difficult to make decisions, difficulties in navigating the nuances of social interaction, and failing to initiate something due to memory problems. Cognitive impairment meant that everyday life was more difficult, and previously simple tasks were more complicated now.

> *He can't …. He doesn't - you don't think as clear do you? Things don't come so easy to you do they? [...] He's always been very hands-on, always very handy. And it's a big prob – it's a big headache really for us now because if anything wants doing now we've got to get somebody in to do it. He can't do it anymore, because he struggles. Well, he struggles to get his head around it don't you? […] Even just little things you struggle don't you- just to get it straight.*
>
> (Susan, 71-year-old carer)

Many people living with NCD also experienced physical impairment associated with age-related changes or comorbidities, such as issues with mobility, fatigue and ill health. This could make things less enjoyable and prevent interest and activity.

> *I used to love knitting but I've got - had an operation for carpel tunnel in that hand and I've got it in this hand, but it's not as bad but I can't -. I find my wrists ache.*
>
> (Helen, 66-year-old with MCI)

> William: *[to Anne:] -as soon as you play [piano] just for a, you know, you feel tired and exhausted.*
> Anne: *Yes, we've both got older you see.*
>
> (Anne, 87-year-old with AD, and William, 87-year-old carer)

Participants often experienced failure, and negative experiences that were remembered as significant and sometimes dangerous.

> Robert: *Because, I used to do woodwork, but I can't use the machinery anymore. You know what I mean?*
> Interviewer: *What's difficult about using the machine?*
> Robert: *Well, because it's a handsaw, in case I chop my fingers off. Because what I did once, I put my hand in the hot oven and picked a pan up, because I didn't… use it like, so I don't…*
> Sandra: *You didn't pick the oven glove up, did you?*
> Robert: *No, and I just put my hand in.*
>
> (Robert, 66-year-old with AD, and Sandra 61-year-old carer)

Impairments could cause people living with NCD to have to start again, increasing the efforts and time required to do something.

> *There are many a time I'm going down the allotment, I've got halfway there and I've thought "I haven't got a key to get in". So I've had to come all the way back and get my key.*
>
> (Robert, 66-year-old with AD)

> *Often I shall have to write more than one cheque […] and I've had them sent back because I haven't got them right, and things like that*
>
> (Chris, 91-year-old with AD).

Difficulties meant that people living with NCD were often unable to do things, at least to a previous standard, without help, or were unwilling to do things alone, creating a dependency and reliance on others.

> *I'd like to start walking properly… but I have to have somebody with me all the time… frightened I'm going to fall over.*
>
> (Nick, 85-year-old with VaD)

Both cognitive and physical impairments meant that every-day life was more effortful for people living with NCD. This feeling of struggle and additional effort was beyond specific instances and situations, but reflected a general experience that *"life is a bit harder"* (Chris, 91-year-old with AD).

> *He can't walk like he used to. And he can't think like he used to, because he was so clever, so bright, and put me into the shade, you know, he did. But now everything is an effort, isn't it?*
>
> (Martha, 73-year-old carer)

> *I couldn't get up [a ramp in the garden centre]. In the end the [wheel]chair came back down the ramp. I said "you'll just have to go with a trolley Nick, I cannot push you". And that's the last time we went to that sort of garden centre.*
>
> (Cynthia, 76-year-old carer)

> *Now, I avoid doing anything I don't have to do. And I say that seriously.* […] *Just the effort of doing it is difficult.*
>
> (Chris, 91-year-old with AD)

8.3.2.2. Impact of Consequence and Purpose

8.3.2.2.1.Negative experience and increased vulnerability

Difficulties experienced by people living with NCD, described in the first subtheme, caused anxiety and stress. People living with NCD could become frustrated at the decline in abilities that they experienced.

Sandra: *Because with Alzheimer's he can remember things from years ago immaculate; ask him what he did yesterday, gone.*
Robert: *Gone.*
Sandra: *Then he'll get-, he can get extremely frustrated and angry, and it's-, I have to be a bit cute and distract and comfort-*
Robert: *Yeah, it's terrible.*

(Robert, 66-year-old with AD, and Sandra, 61-year-old carer)

At the same time, people living with NCD were less resilient, and could become easily overwhelmed.

Well, I've noticed something, because Rahul can't cope with stress like he used to, can't cope with any sort of stress now.
[…] Our youngest son […] can be very challenging. He's got certain issues and he can be quite a challenge, but Rahul can't seem to cope with it as well as he used to be able to, and I'm sure that's part of the dementia.

(Ruby, 71-year-old carer)

The combination of being less able to cope, increased experience of difficulty, including fatigue, and being required to exert more effort to achieve the same result sometimes meant that people living with NCD preferred or had to adopt a slower pace, both physically and mentally.

Charles: *[when there's] Four boisterous grandchildren charging around […] its kind of intense and loud, so it's a bit, its tiring for anybody. And then it's trying to follow what's going on because you know everything's… "oh we're gonna do this and that and then you do that and then that", and then Mary can't keep up with that. Umm, so, she can be a bit bewildered […] I've found Mary's less emotionally engaged with the grandchildren now I think. […] I think there's a subtle difference there. She still loves them but it's I don't think […she's] quite so engaged with them.*
Mary: *Oh, yes I mean we used to have them a lot didn't we?– initially.*
Charles: *I think it's partly a tiredness thing as well. It's a little goes a long way, particularly with - the Canadians have got loud American shouty voices so everything's "pssh!" up there somewhere, so I think that's part of it. And that it's just too much like hard work to be there. Be with them and get engaged with them too long. Cause they're more full on aren't they, those two particularly <laughs>. […]*
Mary: *And yet [Granddaughter 1], I mean she just sits at the dining room table*
Charles: Yeah I mean well that's the difference, I think that's the problem.

(Mary, 66-year-old with AD, and Charles, 70-year-old carer)

As such, participants described doing what they were able, and opted for convenience and ease where possible.

> *Anything I've dropped it's because it's been not convenient really*
>
> (John, 87-year-old with VaD)

> *Oh yes, we get out a bit. I mean I trot over to the shops, it's no big deal that. It's just across the road, which is why we're here. And occasionally we go into town, but, again we have to make it easier to-, you know, have a taxi.*
>
> (Chris, 91-year-old with AD).

8.3.2.2.2.Reaction to difficulty and consequences

The inability, increased difficultly, and additional effort were attributable to loss of enjoyment, loss of interest, avoidance, and giving up activities.

> *[to David:] You've lost interest in say dramas because you're not following the storyline.*
>
> (Patricia, 69-year-old carer)

In particular, people living with NCD feared and avoided situations that could lead to negative experiences.

> Sandra: *Robert mixing with new people is not great.*
> Robert: … *yes. Because it seems like I'm frightened, because I don't know what I'm going to say and it's hard to get conversation. You know what I mean?*
>
> (Robert, 66-year-old with AD, and Sandra, 61-year-old carer)

> *If David started something and then can't remember where he's up to, he'll get frustrated and then you don't bother, don't want to pick it up again.*
>
> (Patricia, 69-year-old carer)

Where people living with NCD were still competent, they generally retained their interests and remained active.

> Linda: *[I] Do the pots, yeah, plant the pots out.*
> Donald: *That's right, you're good at that.*
> Linda: *There's still quite a few things I can remember with the garden. I know what plants are, what they're called, and things like that can't I?*
>
> (Linda, 67-year-old with AD and Donald, 67-year-old carer)

People living with NCD would often try or intend to take action or make an effort, but the additional effort required to engage was sometimes seen as too much.

> *On the whole it's because it's too much effort. It's easier not to do it.*

(Jean, 85-year-old with AD)

Jigsaws are a bit different. Depending on what size it is. If it's like that <indicates size of large piece> it's OK. But little ones, I can't be doing with them, or anything too hard I can't do.

(Linda, 67-year-old with AD)

Underlying these depictions was the tacit assumption that reducing activity, interest and engagement was unsurprising and reasonable given the circumstances. The additional effort required along with the risk of negative consequences meant that engagement was sometimes seen as not worth the effort or risk, resulting in loss of interest, disengagement, inactivity and giving up. However, some participants did report occasionally being surprised that the effort or consequence experienced was not as bad as expected.

Betty: *Before we started, [John said] "I'm not going to - I shan't be doing with this, blah-blah-blah." I said "oh give it a chance, you've not been". And as soon as he'd been once yes he was quite happy to go.*
John: *I quite enjoyed it.*
Interviewer: *What made you think that it wasn't going to be for you?*
John: *Well, I don't know really. It sounded a bit uh, dry and dusty.*

(John, 87-year-old with VaD, and Betty, 85-year-old carer)

8.3.2.2.3.Role of purpose, and necessity

People living with NCD frequently did not engage where for them, there was no purpose or need, but would often continue doing what they deemed necessary, and engage in purposeful activity.

You don't see the point, or the need – if I want – like today I'll do the housework, because I'm getting on and doing things. So then if I ask him it's no different to him, it's just another day and it's "oh I'll do it tomorrow".

(Patricia, 69-year-old carer)

I think that's what keeps me going, you know, because there's so much every day that I'm supposed to be doing so it just keeps me going. Motivating isn't it?

(Helen, 66-year-old with MCI)

Therefore a sense of purpose and necessity in activity may encourage engagement and reduce apathy. Whether something was viewed as necessary was linked to participants' sense of roles and responsibilities.

William: *One of the examples I gave was "she's a brilliant cook but not anymore". And now you really can't remember where the utensils are can you dear?*

Anne: *No, why don't I? Because I don't have to!*
William: *No.*[…] *She stays where she is now.*
Anne: *I do the things that I do, which is the-*
William: *So physically-*
Anne: *-bathrooms and…*
William: *-I'd say after the fall and the hip bone business for a time when the carers were coming in she really couldn't do much at all. Then you got going. And uh, so now you do make the bed every day, because I'm too lazy, and you clean the bathroom and the cloakroom every day. Those are the physical things she does.*

(Anne, 87-year-old with AD, and William, 87-year old carer)

The support provided by the carer could mean that the person living with NCD felt certain things were not their responsibility, and not necessary for them to do, as the carer would do it.

Martha: *Um, but one thing he would never do and that was the shower room or the bathroom. "I wouldn't do that". But yeah, you did. You used to, didn't you?*
Richard: *Yeah.*
Martha: *And now, I don't know. Could it be we've just got in a rut, where I do it, and he knows I will do it?*

(Richard, 77-year-old with AD, and Martha, 73-year-old carer)

Conversely, a sense of purpose could particularly be achieved where the patient could adopt a caring role themselves. For example, Chris's wife Judith had health problems and Chris explained:

Chris: *I do the washing up to relieve Judith the-.*
Judith: *You have done that for some years, since he's retired he's done it.*
Chris: *I try to do what I can to help, you know […] I mean particularly if it's got a purpose. I mean the washing up for me isn't to keep me active, it's to make sure Judith doesn't have to do it. You see. But I'm not going to take the cooking on-*
Judith: *You used to do.*
Chris: *- or the washing.*

(Chris, 91-year-old with AD, and Judith, 89-year-old carer)

Participants' personal beliefs about necessity and purpose were important. Those that believed that it was important to stay engaged; to do things, to get out, and that activity was beneficial, often made a conscious effort to keep active and involved.

I've got to do something because I've got to keep fit. And my swimming helps me a hell of a lot. I go twice a week and it really helps me, I really enjoy swimming.

(Robert, 66-year-old with AD)

Carers' beliefs were also influential, as carers who believed keeping active was important encouraged this within their relative. However, where views were not consistent, a tension existed between what the carer encouraged and what the person living with NCD wanted.

> Cynthia: *But you don't come home and do any of the exercises that [a Pilates instructor] show you do you? Like getting on and off the chair, which you, you should.*
> Nick: *Well I have enough don't I?*
> Cynthia: *Yeah, but an hour a week Nick, is not-.*
> Nick: *Yeah I know that much, but it's still knackering at the time.*
> Cynthia: *I'm afraid I have no sympathy because he comes out, "oh my legs", and I just go "that's because you moved a muscle"*
>
> (Nick, 85-year-old with VaD and Cynthia, 76-year-old carer)

Unsurprisingly, people living with NCD did the things they enjoyed, and avoided the things they did not. Behaviour was understood to be dictated by, and even part of, an individual's personality and identity. Participants sometimes explained lack of activity, interest or emotion as who they were or how they had always been.

> *But he doesn't show a lot of feelings. You know, I mean, if the family suddenly descended, oh he'd be pleased. But he wouldn't be effusive. But then again he never was; I suppose that's always been my side of the partnership.*
>
> (Betty, 85-year-old carer)

> Peter*: I find it very hard to start a conversation, especially with people that you only know them when they're on the bowls rink. You know it's… Uuh. I don't know why I am-*
> Diane*: You've always been quiet haven't you?[…]*
> Peter*: Yeah, […]*
> Diane*: Um. <sighs> He doesn't chat to [the grandchildren]. Our youngest grandson at the moment, he-, […] I think he's slightly autistic. […] And his granddad can say things, and I'll know they're going to upset him. You can't criticise him, do you know what I mean? So Peter can soon upset him […] but you don't very often open a conversation with [another grandson] either do you?*
> Peter*: I don't very often open a conversation with anyone really.*
> Diane*: No, that's true, even our son you don't.*
> Peter*: No, you know, I'll answer when I've sort of got to.*
> Diane*: Yeah, I mean our daughter's noticed how quiet he's got.*
> Peter*: I'm not a conversationalist, yeah but look at [Daughter 1], she's non-stop rabbit isn't she?*

Diane*: Yeah, but she has noticed that you don't talk much.*
Peter*: Mm. But it's the same wherever I go, I don't talk much.*
Diane*: Yeah, so there you go.*

(Peter, 73-year-old with MCI, and Diane, 69-year-old carer)

In the above example, Peter is described as having "always been quiet", but later Diane states that this lack of social engagement is new, or at least worsening, and this has been observed by their daughter, who "noticed how quiet he's got". This seems to indicate that there is post-hoc rationalisation taking place, in which Peter's identity is reconstructed as someone who is naturally quiet, which relates to the next theme. Of note is the discussion that takes place in which Diane describes how Peter can make errors in conversations with the grandson, which supports this theme's finding that it can be 'too much trouble' to become engaged.

This theme captured how impaired cognitive and physical capabilities meant people living with NCD experienced greater difficulties and negative consequences and loss of enjoyment. It was considered reasonable to be less active and interested in the face of the additional effort and time required to participate, though purpose and necessity could provide additional motivation.

8.3.3. Preserving Identity in the Face of Loss of Capability and Autonomy

This theme describes how people living with NCD experienced threats to their sense of competency, autonomy and overall self (subtheme 1: 'Threatened Identity'), and made attempts to combat this, which could occur through withdrawal and avoidance (subtheme 2: 'Preserving the Competent and Autonomous Self').

8.3.3.1. Threatened identity

Cognitive and physical difficulties meant that people living with NCD experienced frequent threats to their sense of self, particularly their identity as a competent, able and independent person.

8.3.3.1.1. Threatened Competency

The difficulties and negative experiences described by the first theme resulted in loss of confidence and generalised feelings of incompetency.

That's regular that things go wrong for me

(Robert, 66-year-old with AD)

Stephen: *Yes, that's one of the reasons I don't like going out because I get anxious, you know? I get a bit irate. I do control it, to a certain degree.*

Interviewer: *What kind of things do you get anxious about or irate about?*
Stephen: *Crowded spaces; I don't like people in shops.*
Interviewer: *Do you mind me asking what it is about that?*
Stephen: *Uhh. I don't feel confident anymore. If that makes any sense. Yeah.*
Interviewer: *In what sorts of things?*
Stephen: *Shopping, you know. "Have I forgotten something? Have I forgot something?" you know, and when I get back I have forgotten something. And then I get stressed with myself then.*

(Stephen, 70-year-old with MCI)

A discrepancy occurred between the present capabilities of the person living with NCD and their own expectations as well as expectations of others, often based on their previous abilities.

I used to get angry with myself, because I knew I could do it, but somehow I couldn't, if you know what I mean.

(Linda, 67-year-old with AD

Even now people want me to do jobs, but obviously I can't do them.

(Adrian, 73-year-old with VaD)

8.3.3.1.2.Threatened Autonomy

Participants described a lack of control over situations, their abilities and themselves, which threatened their sense of autonomy.

Jean: *Frustrated. Fed up.*
Interviewer: *At anything in particular?*
Jean: *Um, oh that I can't do things that I used to do and want to do. I feel like standing up and screaming, but that doesn't do any good, so I don't <laughs>.*
Interviewer: *What kind of things can't you do now that you used to be able to do and would like to do?*
Jean: *Well, go shopping on my own, I'd like to be able to go out on my own obviously. Mainly that. I think it's a loss of independence that I feel most. And I don't like it, but I have to admit-, you know I have to give in to it to a certain extent, because I know there are certain things that I cannot do now that I used to be able to do. So it gets frustrating for me.*

(Jean, 85-year-old with AD)

The necessary dependency on others, noted in the previous theme, resulted in a power imbalance, forcing people living with NCD to adopt a lower status in their relationships. For example, Peter said: *"I do whatever I'm asked, or told"* by his wife to do. This change in relationship was not always welcome.

Susan: *He'll flare up. At me, not at anybody else, you know, just…*
Adrian: *She's the one telling me off all day. I'm doing things wrong. […]*
Susan: *But unfortunately I'm the one who's got to tell him when he forgets to do something or, you know, I'm here so I'm the one who's got to…*
Interviewer: *You're the reminder.*
Susan: *I've got to remind him about things. And it gets on his nerves, it gets on my nerves, but it gets on his nerves as well, don't it? But it's frustrating at times for both of us.*

(Adrian, 73-year-old *with VaD*, and Susan, 71-year-old carer)

Carers and others worried about the safety of the person living with NCD, which could lead to direct restrictions being placed on them, further contributing to their loss of autonomy.

Robert: *And if I want owt doing like the raised beds, [son-in-law]'s done them for me.*
Interviewer: *Do you ever work on things together with him, with him using the saw?*
Robert: *No. No. No. He wouldn't let me use any of the tools anyway, because he knows what might happen.*

(Robert, 66-year-old with AD)

[when discussing cooking:]
Paul: *they don't trust me with a knife.*
Karen: *No, well, there's a good reason for that isn't there? [referring to an incident where the patient accidentally cut himself]*

(Paul, 72-year-old with VaD and Karen, 50-year-old carer)

Cognitive and physical decline was seen as unavoidable. People living with NCD expressed feeling powerless to change their situation, even when this was not a result of dementia. In particular, it was often assumed to be inevitable that ageing causes problems, which in turn restricts activities and interests.

I would like to get fit but I don't think it will happen, you know. Never in a million years.

(Stephen, 70-year-old with MCI)

Well I think your interests change as you get older, because you can't do the things that you used to do.

(Anne, 87-year-old with AD)

8.3.3.1.3.Threatened Self

No longer being able to do things was not just a matter of giving up specific activities, but was seen as giving up a part of their self, or even becoming a different person.

> *I've become a bit of a hermit really. Yeah. I used to. I used to be out a lot but, I just can't be bothered with it. Bit sad really isn't it?*
>
> (Stephen, 70-year-old with MCI)

> *Well, I would love it if she was the way she was. You know, it grieves me that she can't garden because it was so important, she loved the garden.*
>
> (William, 87-year-old carer)

> *I can sit here all day looking out of that window without even thinking about it. And that does worry me, because it's not me at all.*
>
> (Chris, 91-year-old with AD).

At the same time as experiencing an internal change, people living with NCD had to navigate a changing world that was sometimes seen as inherently different.

> John: *I was a photographer as well.*
> Betty: *Yeah, he lectured in photography.*
> John: *When photography was photography. […]*
> Interviewer: *How do you feel about those things now then?*
> John: *<sighs>. Well, photography I don't get involved because it's just not the same game.*
>
> (John, 87-year-old with VaD, and Betty, 85-year-old carer)

People living with NCD sometimes found it difficult to handle challenges to their sense of self; they found changes frustrating and hard to accept.

[when asked how they were feeling:]

> *Annoyed… with myself. You know? Yeah, when I think what I used to be like and what I'm like now.*
>
> (Stephen, 70-year-old with MCI)

8.3.3.2. Preserving the Capable and Autonomous Self

These threats to sense of self were unsurprisingly not welcome. Many people living with NCD did not wish to be seen as incompetent, were embarrassed or proud, and resisted the identity of someone of a lower status, less autonomous and less competent. This often meant avoiding situations where this would be required.

> *[when discussing reluctance to join an organised group:] I don't really like that, after being top dog and sort of organising everything, I might*

find it a bit difficult not to. I would want to interfere […] I would want to take control I feel <laughs>, even if I haven't got the capacity or capability to do it, I'd still feel I would want to if I was there.

(Jean, 85-year-old with AD)

[when discussing not going to family parties:] I think Linda's just resided herself, doesn't like people seeing the way she is I think it is. I think that's what in the back of her mind.

(Donald, 67-year-old carer)

In particular, people living with NCD could be fearful of the unknown and unfamiliar, which could threaten their sense of autonomous and capable self, preferring to stay within their comfort zone.

Interviewer: *You mentioned that if you were to plan a holiday in the days leading up to a holiday Jean, you would be a bit unsure about going.*
Jean: *Yes. I'd be a bit iffy.*
Interviewer: *Yeah, anything in particular about the idea of going away that makes you feel like that?*
Jean: *Um, no, I suppose I'll be moving out of my comfort zone, to put it simply. Yeah.*
Interviewer: What kind of things in particular? … So out of your comfort zone in terms of location or knowing where things are…?
Jean: *Well I don't know, I sort of -it's a kind of feeling of safety I guess. Here, I feel safe and secure and things like that.*
Joseph: *I've wondered about that, is it that here is predictable?*
Jean: *Yes.*
Joseph: *You still feel you've got some control over it, whereas if you go somewhere else-.*
Jean: *Yeah, that's right yeah.*
Joseph: *Yeah, I can see that.*
Jean: *Yes.*
Interviewer: *Whereas if you were to go away on holiday, then you're-.*
Jean: *I wouldn't have the same amount of control.*
Interviewer: *Mm. No. Don't know what to expect in a new place?*
Jean: *Yes.*

(Jean, 85-year-old with AD, and Joseph, 74-year-old carer)

People living with NCD also avoided experiencing failure and situations in which they did not feel competent, in particular if they had previously experienced a failure or set back, as described in the first theme. The need to avoid challenges to competency could mean giving up or not doing something rather than adapting it to declining abilities.

Patricia: *if you don't think you can do something, you don't want to try it. Oh no not, if you don't think you'll be good at it, you're not going to do it. Like the choir. Because [friend 1] said to you, we've got a community choir now and I'm in and [friend 1] and [friend 2] are in it. And he said, we were listening to the 60s music or something one time and you were singing away and he said "you've got a voice, you should come along and join the choir". And he just blanked at [friend 1].*
David: […] *It sort of didn't interest me enough, let me put it that way, to do it. There have been one or two things where, just think, I'll just sing low in the background. There's no way I would ever-*
Patricia: *You weren't singing low in the background when we went to that Kinks revival thing. Belted those out.*
David: *It's because I know all the songs.*

(David, 69-year-old with AD and Patricia, 69-year-old carer)

In the above example, David explains not wishing to join the choir as it is not of interest to him, however he later states that he sang along to the Kinks because the songs were familiar, suggesting that he may not have wished to join the choir as to avoid a threat to his feeling of competency, rather than not wishing to sing.

Some people living with NCD fought against their declining autonomy by resisting encouragement, leading to further decreased engagement:

Joseph: *Jean wouldn't go, basically. Everybody, doctors, the re-ablement team, physiotherapists, all telling you to do stuff, but.*
Jean: *Once I made up my mind. I can be stubborn I must admit.*

(Jean, 85-year-old with AD, and Joseph, 74-year-old carer)

Others did not outwardly resist requests from others, but frequently put things off. They appeared to intend to do something, but this did not occur despite encouragement.

Judith: *I'm afraid things come like post and things, and he can't be bothered with them, and he just piles it up until it looks awful. It's piled up, and I keep saying "you must do it". "Yes, I'll do it this afternoon", but he doesn't.*
Chris: *- He doesn't. Whereas if it had been in the office-*
Judith: *He would have done it straightaway.*
Chris: *Well either that or made sure I'd got somebody else to do it <laughs>.*
Judith: *So it can be like that to the point where I said "you really will have to do something, because there must -", I try and keep tabs on bills and things, I mean he does deal with them.*
Chris: *I deal with all, there's nothing-*
Judith: *You do, but they get put on there.*

Chris: *Yes, but we never owe any money.*
Judith: *No, but it's things like that. No. I mean a week ago, decided under the stairs that go upstairs to the flat, is the glory hole, well I won't show you in it because it's just…*
Chris: *It's the neatest glory hole you can-*
Judith: *And he said a week ago "we'll empty all that and get rid". "Right". And we set a date. He didn't feel like doing it.*
Chris: *Well I wasn't well enough to do it. It wasn't that I didn't feel like it.*
Judith: *You didn't feel like it.*

(Chris, 91-year-old with AD, and Judith, 89-year-old carer)

It was important to people living with NCD and carers that people living with NCD retained their independence where possible. People living with NCD enjoyed and engaged with activities with variety and choice, and those in which they could take on a role with higher status and have greater control, supporting their sense of capable and autonomous self.

I think also he likes to be his own person. Because when he goes to these clubs, he's his own person, he's not protected by me, or I protect him, and I wouldn't dream-. Because somebody said why don't you join it? I said "no, no, no that is Nick's time, he must get on and enjoy his self".

(Nick, 85 year old with VaD, and Cynthia, 76-year old carer)

90% of the people that are here couldn't do anything anyway. If they bent down they wouldn't be able to get up again. Um, so I'd jump in for that, without a doubt.

(Peter, 73-year-old with MCI)

However, carers sometimes found it difficult to balance people living with NCD's need for independence with their need for support and supervision.

I must admit I feel it's hard sometimes me taking over so much 'cause I just feel like I'm control - I'm trying not to be controlling and it's hard to get that balance, because sometimes you just have to step in and sort something out... I find it difficult as well ... giving her a bit of free reign without <laughs> I don't mean that, that sounds horrible doesn't it? 'free reign' but you know what I mean? […] Again, you don't want to take over all the roles either. I don't wanna sort of say right every night I'm gonna do that then we're gonna do this cause it just … devalues Mary's contribution and makes her more dependent still. So it's trying to get that … happy balance. Which doesn't always work but we try to make it work. We try and share it out. So there's still a fair bit of responsibility still with Mary, but within safe limits really.

(Charles, 70-year-old carer)

In contrast to avoidance or resistance, some participants felt there was nothing to be done. They had no choice but to accept the changes they experienced; they had to "cope" (Stephen) and "manage" (Adrian) with how things were. This acceptance allowed participants to let go of what they could no longer do, and move forward.

> Anne: *Well, it's one of those things. If you can't do it you can't can you? And um, I don't rail against things that I can't do, I just get on and do the things I can do, which isn't very much. […] you can't do the things that you used to do. So you… continue to do the things that you can do.*
> William: *I mean in a way you miss your independence. But you don't miss it because you – as you say if it can't be done [...]so you don't let it worry you.*
> Anne: *No.*
> William: *But I mean you enjoyed being able to just do things.*
> Anne: *Yes I did. But, uh, if you can't you can't.*
>
> (Anne, 87-year-old with AD, William, 87-year-old carer)

Some participants preferred to focus on what the patient could "still" do, helping them retain their sense of self. For example, Linda, who spoke about feeling angry at herself due to not being able to do what she expected, said:

> Linda: *It was frustrating.[…] Yeah, got used to it all now.*
> Interviewer: *So how are you feeling now?*
> Linda: *Yeah, I'm fine actually yeah. Quite happy in myself and OK, can still do things that I've always done since we've been married, like cooking, cleaning, whatever.*
>
> (Linda, 67-year-old with AD)

Similarly, some did not see apathy as a problem, and were content with how things were. This acceptance could still mean giving up activities and losing interest in previously enjoyed activities, and activities were sometimes substituted for something more passive that required less engagement.

> *He's not as motivated and he's not as keen as he was to do things and, like I say, to get up and get on and do things as he used to be. But I don't think that bothers him like it would have done either. I think he's more, he probably is more relaxed.*
>
> (Susan, 71-year-old carer)

> Martha: *So he does tend to sit at home too much. More so in the last four months or so really, six months, we haven't been out much for lunch or anything like that.*
> Richard: *No.*
> Martha: *Where, we used to go on a Friday with our friends or we'd go to [a café] in town, but we don't seem to do it so much now, do we?*

Richard: *No.*
Martha: It seems harder work.
Interviewer: *What do you think about that?*
Richard: *It doesn't worry me too much. Um, because, as Martha says, I'm not a very sociable person.*
(Richard,77-year-old with AD, and Martha 73-year-old carer)

In summary, declining capabilities and increasing dependency meant that people living with NCD experienced a threat to their sense of self. This could result in withdrawal and dissociation from previous interests, activities and social interaction, in an attempt to preserve their sense of competent and autonomous self. However, sometimes this change was accepted, which enabled them to move on.

8.3.4. Opportunity and Exclusion

This theme captures how the engagement of people living with NCD in activities, interests and social interaction was dependent upon appropriate opportunities provided to them, and their social support networks, which were seen as particularly important.

Opportunities in the environment, prompting and activities being set by others enabled participation and engagement. For example, Ruby, a 71-year-old carer, described the various support groups that Rahul attended: "*we didn't go looking for things, it was just it turned up, so we decided to go*". This contrasts with the experience of other people living with NCD who were less active, such as Nick, an 85-year-old with VaD, who said: *"I don't worry about it, just… sit here. That's all you can do. She says I don't do nothing, I'm sat here all day. But what else can I do?"*. Thus, without such opportunities being readily provided, people living with NCD felt there were no options for meaningful activity, leading to decreased engagement.

Reducing the difficulty and effort of everyday life was an important need that had to be met to support participation. Accessibility was important to support independence and activity. Activities had to be consciously altered to make them more manageable for people living with NCD.

Charles: *We used to have great meals. Um so it's only very recently where [Mary] really struggles with just following recipes - that's the main thing I think isn't it?*
Mary: *Yeah recipes.*
Charles: *And she'll get so far and then panic because she can't remember where she is so … we try and do some of that together. But there is a […] great shop […] and [the shopkeeper] … makes up packs. So she weighs all the stuff up to a recipe and the instructions and so you just go and buy cakes, brownies or whatever it is and you've got*

the ingredients there so you haven't got any waste. Um and so Mary still enjoys doing that.
Mary: *And they're all in bags. One, two, three, four, five, six, seven.*
Charles: *yeah they're all numbered. It's nice.*
Mary: *They're all numbered. Absolutely fantastic. So you put one in. Stir it around. Next one in. Stir it around. <laughs> Fantastic.*

(Mary 66-year-old with AD and Charles 70-year-old carer)

We get off the bus and it's right opposite and they've got a nice coffee shop. And um, it sounds as though we're always drinking coffee, but we always use it for a sit down. You know what I mean?

(Chris, 91-year-old with AD)

Familiar environments provided an important sense of security, and unfamiliarity could be distressing and difficult. Similarly, plans, routines, and scheduled activity facilitated participation.

If I haven't got anything arranged I'm terrible. Like I say, I'll just stay in bed in the morning, instead of getting up and doing something.

(Helen, 66-year-old with MCI)

Diane: *The days Peter plays bowls, because he knows what he's doing, are the days he's better. You know, it's structured, "right, I'm getting up, I'm doing this, I'm getting ready, I'm going to bowls". And I have noticed that he's better on those days than not having a-, "right we're going to do this, this, this and this".*
Peter: *Probably is, because I know where I'm going, I know what I'm doing, I know who's going to be there.*

(Peter, 73-year old with MCI, and Diane, 69-year-old carer)

People living with NCD found that the world they lived in was not always inclusive of someone with cognitive or physical impairment. Difficulties were not always catered for, and sometimes others, including carers, lacked understanding and failed to provide extra support for dementia symptoms, resulting in exclusion.

Cynthia: *But it's like, he said the other day "I must get that puzzle down", and I sat here and I thought "yeah, well you go up and get it and bring it down", because there's no reason for him not to go up and get it. He just thought about it-*
Nick: *I don't know where they are.*
Cynthia: *The back of the wardrobe where they usually are.*

(Nick, 85-year-old with VaD, and Cynthia 76-year-old carer)

In the above example, Cynthia is unable to understand why Nick does not get the puzzle himself, as for her, the location is obvious, as it has not changed.

Similarly, Rahul described activities that he had to give up due to difficulties following the instructions.

> Rahul*: I did speak to him and I said I'm finding it difficult to follow you, can you give me notes? He said no, I won't give you notes. That's what he said. […] and he gave notes at the end but then, the notes didn't make any sense to me.*
> Ruby: *They perhaps assumed a level of knowledge that you haven't got.*
> Rahul: *Yeah. My biggest problem, as I said, is the short term memory's gone. I'll acknowledge things, but I can't remember it.*
> (Rahul, 76-year-old with MD, and Ruby, 71-year-old carer)

Of note is that Rahul remained very active in dementia friendly activities offered to him when he received his diagnosis, suggesting that catering for cognitive impairment enabled his continued participation. This also may be explained by a sense of belonging that dementia friendly organised groups offer, which is discussed later in this theme.

8.3.4.1. Importance of social support and social networks

Social support networks, including carers, friends, other relatives, and formal support staff or group facilitators were important providers of opportunities. Whilst new people could be difficult to engage with, established relationships helped provide a sense of security and confidence, facilitating participation.

> *I see one friend on a Tuesday; I see another one on a Wednesday. Oh, I go to my other friend's on a Monday night for tea. Thursday, I go to bingo with another friend. So it's all the time I've got friends - because I've lived round this area since I was 21. So I've got friends that I've known 40-odd years. So yes, they're keeping an eye on me*
> (Helen, 66-year-old with MCI)

Others frequently tried to motivate the patient and were required to facilitate activities to enable participation. For example, Robert and Sandra described how an organised activity group providing set activities helped facilitate conversation, enabling Robert to participate.

> Sandra: *Robert was in the other room [at a charity 1 group] with the others chatting and laughing and playing games, weren't you?*
> Robert: *Yeah.[…] We had a ball game and we answered questions and we could have a go at getting it in the net and things like that. You know what I mean? […]*
> Sandra: *Then the contract changed, didn't it?*
> Robert: *Yeah and then we got [charity 2]. I didn't enjoy [charity 2].*
> Interviewer: *What was the thing that you didn't enjoy about that compared to the other?*

Robert: *It was something that… they put on to us, what we wanted to do. Now, to me I-, I didn't feel like that. I wanted something to get my brain going and it's something that we wanted to do and things like that. You know what I mean? […]*
Sandra: *So, [charity 1] brought things to us, you know, they'll perhaps do a quiz or a game or something like that and it just starts the conversation off.*
Robert: *Yeah and we all have a conversation and it gets me going, and I find that easier with Alzheimer's.*

(Robert, 66-year-old with AD, and Sandra, 61-year-old carer)

Some carers adopted a role of proactive organiser, to ensure people living with NCD had opportunities for engagement.

I think dad gets frustrated sometimes when we make him do stuff <laughs>, rather than us do it for him. […] It's like shopping. My sister said "well do an online shop", which, yes, that would be easier for me, but, it's getting dad out to go and do his shopping and he can choose his own food. So, I'm conscious that that's important

(Karen, 50-year-old carer)

As well as creating opportunities for the person living with NCD, carers also had to take on more responsibilities as the participant's abilities declined. For example, Robert described making mistakes with his money and worrying about handling his finances, so his wife explained how this led her to take over this responsibility.

He was whittling about money, I went right, I said "I'm not messing about, I said why don't you have pocket money then", that you have a certain amount in your wallet. So he's not got large amounts. Do you see what I mean?

(Sandra, 61-year-old carer)

In the example above, Sandra describes giving Robert "pocket money", suggesting a shift in the dynamic of their relationship. Though carers' support is undoubtedly necessary and positive for the patient, this could be seen to reduce the participant's feelings of autonomy and may reflect a change in the dynamic and power balance of their relationship, which may further result in a change in identity, described in the previous theme.

Furthermore, there was tension between the needs of the person living with NCD and with the carers' own needs. Carers efforts to motivate and support the person living with NCD could lead to the carer having less freedom and feelings of burden.

At the moment we can't fit anything else in in the morning. In fact I was moaning the other day that I'm missing out some of the things I used to

do because of going to the dementia things, like I used to play tennis, quite a lot, and now I've had to cut down on that

(Ruby, 71-year-old carer)

I think he just needs motivation to get out there. I feel like kicking him up the backside sometimes, because it gets you like that then, you know? You think oh right, no. And I do go off and leave him for an hour or so if I go swimming, because I need the exercise. But he's got his […] alarm that if anything should happen, so he's never really on his own as such. And if I'm out I keep thinking my phone's going to ring, or I'm gonna-, so I don't really relax because he won't let me relax if you know what I mean.

(Cynthia, 76-year-old carer)

In contrast to taking on a proactive organiser role, some carers struggled to navigate civic society and services, were not sure what was available and were not provided with the same opportunities for engagement.

Diane: *a social gathering I think would do him good […] But I don't know how much is the cognitive impairment, you know, we don't get any help with that, nobody explains anything to us. You know, I think he needs to be reassessed, but, who do you talk-, I find there's a real lack of help. A real lack of help.*
Peter: *Yeah, but I've not asked for help.*
Diane: *No we haven't.*

(Peter, 73-year-old with MCI, and Diane, 69-year-old carer)

Socialising was seen as inherently positive, though some people living with NCD struggled with social engagement, and familiar support networks provided important sense of security, and confidence which in turn supported engagement.

Interviewer: *What kind of support does [Daughter 1] give you then?*
Jean: *She gives me confidence, the fact that she's around and that I can rely on her. And she's very caring isn't she Joseph? Not sort of over caring, she doesn't let me get away with a lot, but she is very caring in the end.*

(Jean, 85-year-old with AD)

Stephen: *I've started to go to a senior citizens keep fit with my neighbour, the neighbour from hell [said affectionately]. I started that two or three weeks ago.*
Interviewer: *Wonderful. What got you into that then?*
Stephen: *She did. Yes. Yeah.*
Interviewer: *And what made you say yes?*
Stephen: *She's a good friend. And if I said no she'd give me stress. 24/7! 24/7!*

Interviewer: *And how do you feel when you go to those keep fit things?*
Stephen: *Uh. All right really, all right yeah. Yeah. Because I know [neighbour], upstairs, she'll keep an eye on me.*

(Stephen, 70-year-old with MCI)

Some participants experienced a loss of this support however, due to others' having their own busy lives and problems for example, Joseph said: *"we've tried keeping in touch with friends, but they have issues of their own".* However, sometimes people living with NCD felt they were avoided by others due to their condition. Betty, an 85-year-old carer observed that "*some friends [of John] have sort of dropped by the wayside that little bit because they probably find it disconcerting, the spectre at the damn feast*".

I always see my friends at Christmas. I don't see them though-, they used to-, my mates used to come, but since the Alzheimer's, they haven't bothered [...] I think they feel embarrassed, or they're frightened, or what have you, so they don't come.

(Robert, 66-year-old with AD)

Belonging to a community or being a 'member' of something supported engagement. Similarly, feelings of not belonging led to rejection or avoidance of activities.

Well, I found getting sort of, I wouldn't say membership, but getting into - like the [local dementia support group] group, I've found that useful.

(Rahul, 76-year-old with MD)

You can do walking football yes. I know about it yes […] I've just not considered it. Just doesn't really interest me. It's all these old blokes.

(David, 69-year-old with AD)

The above example also demonstrates the importance of identity described in the previous theme, as David does not see himself as an "old bloke". Identity dictates whether a sense of belonging can be achieved, and thus whether particular communities and groups are deemed appropriate.

Carers could only do so much to encourage motivation, and sometimes it was easier to exclude the person living with NCD, and take over doing something, than try the difficult task of involving and encouraging them.

I mean in the end I just gave up quite frankly. You could only spend-, it is hard Jean, to see you doing this thing to yourself, and to know that it's basically, after a time, it's just a waste to have a row with you to get you physically wound up enough to make you do something. You can't go on doing it, it's not good for my blood pressure apart from anything else.

(Joseph, 74-year-old carer)

I perhaps ask him less often now to do things than before. Unless I know he's really going to get involved in it and enjoy doing it.

(Patricia, 69-year-old carer)

This suggests that insufficient social support may be one mechanism via which people living with NCD appear to display apathy.

In summary, this theme demonstrated how impairments and increased needs meant that people living with NCD required additional support to remain engaged and participate in a meaningful way. Effort could be reduced and consequences could be improved by opportunities and help from others. However, others were not always able to meet their additional needs, and people living with NCD sometimes found themselves in a world designed for abled people, from which they faced exclusion.

8.4. Discussion

A fully integrated discussion of findings is presented in Chapter 9, so only a brief summary of findings and limitations relevant to this particular sub-study are presented below.

8.4.1. Summary of findings

Apathy is experienced as an understandable response to the everyday difficulties people with NCD face. It can be understood to be a coping mechanism to preserve identity in the face of declining physical and cognitive abilities and associated threats to competency and autonomy. This is exacerbated by lack of opportunities and social support.

8.4.2. Limitations relevant to the qualitative sub-study

Thematic analysis was chosen for this study as it offers a flexible method of qualitative analysis that is consistent with various epistemological perspectives, data types, and research questions [333]. However, using thematic analysis may have meant that insights into individual participants' overall 'stories' were lost. Interpretive phenomenological analysis may have enabled a more thorough exploration of the lived experiences of apathy in people with NCD and their carers, as it is particularly suited to exploring how participants make sense of their world. However, interpretive phenomenological analysis may also have failed to situate the participants in the context of the social environment [338]. In contrast, the approach used in the present study enabled the role of the social environment to become a key finding. Grounded theory is another approach that could have been chosen for this study, as this is an analytical method that aims to produce theory from the data [343], and so this may have enabled a focused production of a novel theory of apathy. However, as a grounded theory approach only uses inductive methods and ignores prior theory, this could be considered

inconsistent with a critical realist approach, which emphasizes using theory from the start [287]. In contrast, thematic analysis allowed for the application of both inductive and deductive coding, enabling the consideration of prior theory within the analysis itself.

In this qualitative study, participants frequently spoke of reduced interest, cognition, activity and social engagement, however, 'emotion' was less salient. It could be argued that this was due to participants' reluctance to discuss emotions with an unknown researcher or in the presence of each other (as many interviewers were dyadic), or the interview questions failed to elicit responses related to the emotional experience of the participants. However, many participants discussed various negative (and therefore perhaps more difficult to discuss) emotional experiences of frustration, irritation, and anger. Furthermore, in a similar study of apathy in six people with AD, there was little reference to emotion presented in their findings [284]. This suggests that rather than being a limitation of the interview approach used in the present study, reduced emotions or lack of emotional expression were not as prevalent or as important to the participants in these studies.

The sample of participants were varied in their experiences of apathy. Five participants did not reach the 'clinical cut off' for apathy according to suggested criteria [27], and only four participants felt the term apathy applied to them, with others believing they did not experience it or did not understand the term enough to comment. This may explain why findings suggest apathy is understood to be and experienced as a reasonable reaction to problems experienced by people living with NCD, rather than experienced as a disorder or symptom of dementia (see Chapter 9 for further discussion of this finding). If clinical cut-off criteria were used or if the term apathy had been used in the information sheets for the study, this may have allowed selection of participants who felt the issue was particularly relevant for them, and could have led to different findings. However, this would have also missed the perspectives of those who had not heard the term before and may have limited the sample to more educated and informed participants and carers. Furthermore, these findings echo other studies of people with Parkinson's disease who had exceeded proposed clinical cut-off scores on an apathy measure [283] or were identified as having clinically relevant apathy by a clinician [146]. These findings are discussed further in Chapter 9.

As discussed in section '6.4.2. Data collection', I aimed to conduct a mixture of dyadic and individual interviews, as recommended by Pratt [329]. Whilst this was achieved in the present study, only two people living with NCD chose to be interviewed with their carer, and no carers took part in an interview on their own. This may have restricted how open and honest people living with NCD and carers felt they could be in the presence of one another [326]. However, it also may have allowed people living with NCD to be more supported in their

communication [329], and may have provided an insight into the dyads' shared experiences rather than two separate stories [326]. In the present study, the back-and-forth discussion between patient and carer enabled useful insights that may not otherwise have been realized.

Culture influences how dementia is experienced [344], and it is important to recognise that the findings should be understood within their context. Participants were mostly of white ethnicity, and all were educated to at least secondary level. Furthermore, all people living with NCD in this study were aged 65 or over, which may explain the focus on comorbidities and physical impairment in the second theme. Nevertheless, this study adds to previous qualitative studies of apathy thus far, as the views of non-spousal carers, and people with MCI, VaD and mixed dementia from a variety of geographical locations within England were included. Furthermore, it is important to note that it is not the aim of this, or other critical realist research to produce 'representative' findings nor make generalisations, but instead, understand mechanisms at work within their context [336].

Chapter 9. Integration & Discussion

9.1. Introduction

This thesis aimed to 1) Determine what measures of apathy are available and their quality for use with people living with NCD; 2) Determine how apathy should be characterised and what its boundaries are in people living with NCD; 3) Understand the possible mechanisms and impact of apathy in people living with NCD. The first aim has been addressed and discussed in the systematic review presented in Chapter 4. This chapter presents an integrated narrative discussion of the key findings of the systematic review, network analysis and qualitative sub-studies relevant to the final two aims. These findings and the literature discussed throughout this thesis are then combined in a tentative model of apathy for people living with NCD. The challenges to and changes made throughout this study and the strengths and limitations of this mixed methods research are then discussed. Finally, some recommendations for future directions are made.

9.2. Nosology of apathy

9.2.1. Clustering of apathy indicators

The network analysis sub-study found that within apathy itself, indicators did not cluster in their proposed domains of behavioural, cognitive and emotional apathy. This is relatively consistent with previous factor analytic studies of the AES (assessed as part of the systematic review presented in Chapter 4), which have found that in general, items loaded into one main general apathy factor [44,46,47], though smaller second and third factors were present, including those reflecting ‘interest’ [46], ‘completion’ [189], ‘insight’ [44,48], ‘novelty’ [44,48], and ‘friendship’ or ‘social’ factors [48,50]. The present study expands these previous studies, as it demonstrates that specific indicators of apathy from the same domain did not form strong connections with one another, nor weaker connections with indicators from a different domain. For example, in the first phase analysis, the two items proposed to measure the emotional domain (‘Excited’ and ‘Intensity’) were only weaky connected, and were connected to items from other domains more strongly. This lack of clustering of apathy indicators in their proposed domains calls to question the various proposed apathy criteria which require a number of symptoms are present from multiple domains [33,36,39].

In the present study, two sets of apathy indicators were merged into nodes that represented ‘Social’ and ‘Novelty’ indicators of apathy, perhaps suggesting consistency with the previous studies that found these were separate factors [44,48,50]. However, in the present study, ‘Novelty’ was

strongly independently associated with ‘Intensity’ though this was not included in the ‘novelty’ factors found in previously reported factor analyses [44,48].

9.2.1.1. Possibility of a social domain

In section 2.3.3.1, it was argued that particular attention should be paid to how indicators of social apathy relate to other apathy indicators, and whether there is a distinct social domain of apathy. Like many apathy measures identified in the systematic review, the AES did not contain items intending to measure a social domain of apathy, however a ‘Social’ node in the present network analysis sub-study was created from two items intended to measure apathy in the behavioural (‘HasFriends’) and cognitive (‘GetTogether’) domains, due to their strong relationship and similar relatedness to other nodes. This ‘Social’ node had low expected influence, indicating it was not well connected to other apathy indicators. This may suggest that there is a distinct social domain of apathy. This is of particular importance, as though a distinct social domain was previously included in criteria for apathy [36], it was recently removed due to lack of evidence (but importantly, not evidence to suggest it was not distinct) [39]. The idea that apathy may be distinct for social activity and interest may be supported by the qualitative sub-study, as this found that people were more engaged with activity with a purpose, particularly caring for others, and that social interaction was frequently seen as inherently beneficial, which may amplify its sense of purpose.

9.2.2. Motivation and initiative

The network analysis sub-study found that the most central indicators of apathy were ‘Initiative’ and ‘Motivation’, tentatively suggesting they might be the key features of apathy. Apathy has previously been defined as a loss of motivation [38], highlighting its importance in the concept. Whilst it could be argued that the high centrality of motivation and initiative could be explained as the result of their strong connection to each other, without which, their centrality in the network may be less apparent, combining these made little difference. Furthermore, in a previous factor analysis study of the AES, initiative and motivation had the highest squared multiple correlation and corrected item-total correlation of all the items [191]. These findings may suggest the importance of motivation and initiative in apathy, and perhaps their causal influence over other indicators of apathy. Whilst the network analysis sub-study was not designed to be confirmatory, this highlights the importance of not excluding motivation from the concept of apathy, which most recently proposed criteria and definitions for apathy have attempted to do [36,39]

Participants in the qualitative study described experiencing what are typically described in the literature as indicators of apathy, including reduced motivation and initiation. A loss of motivation and loss of interest was similarly

reported by Baber and colleagues in their qualitative study of apathy in people with AD [284]. This further emphasizes the role of reduced motivation and initiation in apathy.

9.2.3. People living with NCD and carers understanding of apathy

Participants sometimes found it difficult to understand experiences of apathy, describing them in different terms, such as lacking ‘get up and go’, and being ‘lazy’, as they knew they would enjoy something once started. This is echoed by findings of previous qualitative interviews, that revealed people with dementia found it difficult to do things that they knew they would enjoy once they started [345]. These qualitative findings provide an interesting contrast with reward-based understandings of apathy discussed in section 3.3.2, which proposed that apathy may reflect a deficit in anticipatory reward. This would suggest that people living with NCD would not report knowing they would enjoy something but not feeling able to get started. Instead, this may lend support to the explanation of apathy as an impairment in initiation as an executive function, i.e. a higher order cognitive process, which may explain why participants found this difficult to explain. However, the remaining qualitative findings emphasize that apathy is more complex than this, as it appears to be affected by cognitive and physical impairment, activity limitations and participation restrictions. This is discussed in section 9.3.

Some participants felt ‘apathy’ was different from terms used to define it in the literature, or were not familiar with the term. This was in contrast to a previous qualitative study of people with AD, that reported that participants’ understanding of apathy was consistent with current definitions [284]. This difference could be explained by the difference in recruitment method, as I invited potential participants to talk about activity, interests, and emotions, instead of using the term apathy initially. This would have allowed the inclusion of participants who were less aware of the concept of apathy in NCD. This explanation is supported by a qualitative study of experiences of people with dementia, which reported that carers found apathy difficult to understand [345].

9.2.4. Apathy and depression

One of the objectives of this thesis, related to characterising what apathy is, was to assess how depression and apathy indicators relate to each other and cluster. This was largely investigated through the network analysis sub-study, so these findings will now be the focus of this next section of discussion.

9.2.4.1. Clustering of apathy and depression indicators

The network analysis sub-study found that indicators proposed to assess apathy and depression did appear to cluster separately. This contrasts with previous network analysis studies that found that apathy and depression do not form distinct clusters, as apathy indicators were highly connected to depression indicators, and only sometimes highly connected to each other [275–277]. These studies used a single measure of depression (GDS-15), which includes items that have since been claimed to assess apathy (GDS-3a). Therefore their origins as a measure of depression may explain the finding that these items clustered together in these previous network analysis studies.

9.2.4.2. Shared indicators of apathy and depression

'Energy' and 'Cheerful' had the greatest bridge expected influence in phase 2, suggesting they may be the key determinants of the relationship between apathy and depression. This is consistent with previous network analysis studies which, though they did not assess bridge centrality specifically, found that nodes related to happiness and energy were connected to both apathy and depression indicators [275–277].

9.2.4.2.1.Energy

In phase 3, 'Energy' remained high in bridge centrality (i.e. connectedness to nodes outside of depression), supporting the assertion that it is a shared indicator of depression and apathy. This is consistent with previous network analysis studies of depression that found lack of energy was one of the most common bridge symptoms between depression and other disorders (i.e. had the greatest connections with indicators of other disorders), emphasizing the importance of energy in various comorbidities [274,346]. Energy is also one of the most commonly reported symptoms of depression, particularly in people living with NCD [42,103] and is also important in physical impairment (in particular, frailty) [71]. The qualitative sub-study found that fatigue added to the overwhelming sense of difficulty and effort that people living with NCD experienced, which could make them less willing to engage in activity (further discussed in section 9.3.1.3). These findings suggest that energy may be causally important to apathy, and that it could help explain the relationship of apathy with depression and physical impairment.

9.2.4.2.2.Lack of positive affect

In the network analysis sub-study, the node 'Cheerful', and a similar node in aforementioned network studies ('Happy') [275–277] are the reverse measure of lack of positive affect, i.e. emotional blunting, suggesting these findings support the argument (discussed in Section 2.5.5) that emotional blunting is a shared symptom of apathy and depression. This is important as it contrasts

with assertions that apathy and depression are distinguishable on the basis of emotional blunting in apathy versus negative emotions in depression [30,71]. Furthermore, though previous network analyses of depression alone have found varied results regarding the relations between different symptoms, they have consistently found that 'cheerfulness' or lack of positive affect is a central symptom of depression [274]. Though this is perhaps unsurprising given that it is a core symptom of major depression in the DSM, it further highlights the important role of lack of positive affect in the construct of depression.

9.2.4.3. Mood and outlook

The network analysis sub-study appeared to show that depression was formed of two clusters, one characterised by mood and another characterised by general outlook. This was a novel finding not reported in other network analyses of apathy and depression.

9.2.4.3.1.Outlook cluster

This cluster characterised by general outlook referred to indicators of still enjoying things, looking forward to things, being able to laugh, not feeling slowed, and interest in appearance. The outlook cluster was only linked to apathy via the item 'Energy' (in phase 2), and 'Mobility' (in phase 3), through its item 'NotSlowed'. This may be explained as the item 'NotSlowed' failing to separate the negative bias of feeling slowed (related to outlook) and being physically slowed, highlighting the difficulty of assessing distinct single constructs. The present qualitative study and the other qualitative literature (discussed in section 9.3.3) found that people with dementia sometimes took an approach of acceptance and moving on [345,347–349], which could be seen to be related to outlook. This did not necessarily equate to being more engaged or less apathetic in the present study and in others [347,349], so the lack of direct relationship of outlook with apathy is perhaps not surprising.

9.2.4.3.2.Mood cluster

It has been argued that negative or dysphoric mood should not be present in apathy [71]. The present study found a cluster of indicators of depression related to mood (cheerfulness, sadness, irritability, and worry) and these were associated with apathy only though the indicator 'cheerful' i.e. lack of positive affect, as described above. This could be seen to support the assertion that depression is characterised by negative mood, whilst apathy is not.

In the qualitative sub-study, participants reported varied emotional experiences including negative mood, in particular frustration, anxiety and stress. They were less resilient and able to cope than previously, and could become easily overwhelmed which could lead to avoidance behaviour, loss of interest and adopting a slower pace. This was similarly found in a qualitative interview study of people with AD, which similarly reported that patients were

more easily overwhelmed [284]. These findings support Massimo's progressive lowered stress threshold framework of apathy, which proposes that personal, carer and environmental factors result in apathy through increased vulnerability to stressors [153].

A review of network analysis studies of depression did not find separate clusters of mood and outlook symptoms, but highlighted how symptoms, such as worry, can exist on a continuum of micro-level 'affective states' i.e. in the moment feelings, usually measured over hours, to macro level symptoms i.e. continued state usually measured over days or weeks [274]. The network analysis study measured symptoms over the last one to four weeks, indicating a macro focus. In contrast, participants in the qualitative study largely discussed how their momentary experiences led to specific outcomes, for example Patricia recounts that when David forgets things, "*he'll get frustrated and then [doesn't] bother*". This could explain the difference in findings, and highlights the benefits of qualitative approaches, which enable a deeper look at possible mechanisms rather than averaged associations.

9.2.4.4. Avoidance

Marin argued that another way apathy and depression could be distinguished was by the active avoidance of activity by people with depression but the passive compliance of people with apathy when prompted [30]. In contrast, the present qualitative study found that people living with NCD both actively avoided situations as well as put things off despite frequent prompting. This was understood to be a mechanism of preserving their identity as someone competent and able. This may suggest that apathy and depression cannot be distinguished based on the passiveness of their resistance to activity, however, it is recognised that the participants in this study were not specifically selected for apathy.

9.2.4.5. Differential association with ADL

In the network analysis sub-study, ADL was independently associated with apathy indicators, but not depression indicators. This is in line with previous literature that reports that ADL is strongly linked to single measures of apathy but not depression [16,77,96,141]. Furthermore, this is consistent with a previous network analysis study which found that apathy (assessed by a single node) was associated with bADL and iADL but depression indicators (measured by multiple nodes) were not [278]. Another network analysis study has found that ADL was associated with all three indicators of apathy, but also with the depression indicator 'helpless' [277]. However, the symptom of helpless/ hopelessness was not included in the present study, highlighting that the measure of depression included in this study was not comprehensive, a limitation which is further discussed in section 9.6.1.

9.2.5. Conclusions on the nosology of apathy

In contrast to previous network analysis studies, the present study found that apathy and depression do not cluster together, and have differential associations with ADL. This supports the argument that depression and apathy are relatively distinct phenomena, even in people living with NCD. The present study also found that lack of positive affect was an important shared indicator of apathy and depression, suggesting that they should not be distinguished on the basis of this symptom. Lack of positive affect and energy connected indicators of apathy with depression, and as such warrant further exploration as they could help explain the overlap between and offer useful treatment targets for both constructs in people living with NCD.

Apathy may be best understood as reduced motivation, characterised by many indicators, as proposed by Marin [38], however, in contrast to the nosology proposed by previous understandings of apathy, these may not consist of distinct behavioural, cognitive and emotional domains. The social domain of apathy was recently removed from criteria [39], however this study highlights that this may reflect a different domain of apathy that should not be dismissed.

Definitions and proposed ‘diagnostic criteria’ for apathy in NCD have consistently excluded behaviours that are the direct result of cognitive or physical impairment and environmental context [36,39,350]. However, the present study indicated that problems due to cognitive and physical impairment, lack of opportunity or social support, and exclusion are important in apathy in people living with NCD. These findings are discussed further in the next section, however they are highlighted here as this finding informs our understanding of and current criteria for apathy. Whilst the role of neurobiological changes in apathy is not contested here, the role of social factors and the wider context must not be underplayed or ignored [351]. The process appears to be more complex and nuanced than diagnostic criteria allow for.

9.3. Mechanisms and impact of apathy

9.3.1. Impairments and Activity Limitations and Participation Restrictions

As outlined in section 3.3.4, the ICF proposes that disability (i.e. deficit in functioning) is determined by impairments, restrictions on or difficulties in activity and participation, and environmental and individual contextual factors [139].

9.3.1.1. Executive functioning

Within the ICF framework, dysfunction in executive function can be understood as an 'impairment' i.e. an underlying deficit in the body function that can result in (but does not dictate) overall impairment in functioning [139]. It has been proposed that executive dysfunction may directly underlie apathy, as many of these executive functions are required for producing goal-directed behaviour [111]. As outlined in section 2.5.4.3.1, the investigation of the relationship between executive function and apathy in NCD thus far has resulted in mixed findings, with some suggesting no association, whilst others suggesting an association between apathy, in particular the behavioural domain of apathy, and executive function.

Specifically, it has been proposed that apathy may be the result of impaired option generation and initiation, and thus expected that verbal fluency would be associated with apathy [120]. The present study found that participants with 'clinical' apathy were significantly more likely to have worse MoCA scores. However, the network analysis found that executive function (including verbal fluency before its merging with the MoCA) was not directly associated with apathy indicators, which could indicate that apathy is not a result of deficit in option generation or initiation. Instead, executive function was indirectly associated with a variety of indicators of apathy, through their joint association with ADL.

9.3.1.2. Age

In the network analysis sub-study (phase 3), age was largely only associated with other nodes via frailty and physical impairment, which is perhaps not surprising, and consistent with some previous research that suggests age is related to apathy only through physical functioning [125]. This contrasts with the views that participants expressed in the qualitative study, which suggested that some elements of apathy are expected "as you get older", indicative of assumptions about inevitable decline in old age. Interventions challenging these assumptions and promoting good physical health could support people in maintaining motivation in later life.

9.3.1.3. Frailty and Mobility

Previous literature found that apathy was associated with increased frailty independent of cognition, age [151], physical function and mood [152]. The present study found that this was also the case. It also expands on this, as it identifies which indicators of apathy may be driving this relationship: taking only the strongest, most stable edges (associations), frailty was connected with apathy indicators through ADL, and 'Energy'. Similarly, mobility was not independently associated with apathy indicators, but was directly associated with frailty, and to a lesser extent, executive function. Mobility's relationship with executive function may be explained by the effect of cognitive impairment

on physical ability which is proposed to underly the increase in frailty and falls seen in people with dementia [352]. However, it could also be explained by a possible confounding effect on the measurement of mobility, as the walk speed and balance tasks required participants to follow instructions.

Semprini and colleagues' [353] propose that the behavioural indicator of apathy, 'inactivity', results in frailty. This may be supported by the network analysis sub-study, as ADL can be interpreted as a measure of inactivity, though it is recognised that the network analysis sub-study was not designed to be confirmatory. Additionally, the direct relationship between frailty and energy, the latter of which acted as a bridge (linked other nodes to) apathy indicators, could also suggest that apathy results in frailty through loss of energy, in addition to inactivity. However, in the qualitative study, physical impairments such as mobility, fatigue (which encompasses loss of energy) and illnesses or comorbidities (which can be considered to reflect the concept of frailty in cumulative deficit models) were seen to make things less enjoyable and prevent activity and interest. Frailty has previously been found, through qualitative interviews, to result in various losses, including loss of social connection and reduced ability to participate in activity [150], which are described as part of the apathy construct. This highlights the complex and multidirectional causal interaction of apathy, frailty, and functional ability.

9.3.1.4. Activity limitation

In the network analysis sub-study, ADL had one of the strongest centrality measures, and was independently associated with various apathy indicators suggesting its importance in apathy, though not necessarily revealing a causal link or direction. ADL also had the strongest bridge centrality (in phase 3), directly connecting apathy indicators with other nodes (largely frailty and executive function). The qualitative sub-study can enable a deeper insight into these findings. The sub-theme 'Daily Struggle' described how everyday life is more difficult and effortful for the people with NCD, who were often unable to do things without additional help and experienced setbacks and failures due to physical and cognitive impairment. This subtheme shares some similarities with Baber and colleague's [284] theme of "hindered by invisible obstacles", in which participants experienced disruptions to their everyday activities. Motivation theories indicate that people must feel they can successfully complete tasks to be motivated to do them [115,354], which could suggest that impaired ADL are a mechanism to apathy.

9.3.1.5. Conclusions regarding Impairments and Activity Limitations and Participation Restrictions

These findings support the well-known process in which impairments (both cognitive and physical) can result in activity limitations [139]. Additionally, this highlights how impairment may not directly result in apathy, but may do so via

activity limitations and participation restrictions. The next section will expand on how activity limitations can make things “too much trouble”, resulting in apathy.

9.3.2. Effort and decision-making

9.3.2.1. Qualitative findings regarding effort

The qualitative study suggested that willingness to exert effort was reduced by experiences of difficulties, additional effort required, and risk of negative consequences, resulting in loss of interest, disengagement, inactivity and giving up, as captured in the theme ‘“Too Much Trouble”: Mediating Effort and Outcome’. Similar findings have been reported from qualitative interview studies with people with schizophrenia and negative symptoms, who experienced difficulties with concentration and loss of motivation, viewing things as too much effort [155].

As discussed in section 3.3.2, it has been proposed that apathy is the result of an impairment in the reward-based decision-making process, in which people with apathy are more likely to inaccurately estimate task effort and consequences, making them less likely to exert effort for potential reward [28,121], in particular when the reward is perceived as low level [127]. Whilst the participants in this study did indeed seem to be making judgements about potential effort and consequences, these appeared to be appropriate estimates based on previous experience of difficulty, effort and negative consequences, rather than an overestimation of effort, underestimation of or insensitivity to reward. This is supported by other qualitative studies that have found people with AD disengage due to experiencing struggle and stigma [284] and to avoid negative consequences such as failure and stress [348,355]. Similarly, it has been concluded that people with Parkinson’s disease and apathy make realistic changes to activities and interests in response to their impairment [146].

In this way, apathy may be understood to be a narrowing of interests and behaviours, to focus overstretched efforts on activities that provide the lowest risk of negative consequences.

It has been argued, however, that people with apathy are reluctant to act initially, but sometimes enjoy something once started, as they hold pessimistic views which are only sometimes based on actual experience [284]. In the present qualitative sub-study, some participants did also report occasionally being surprised that the outcome was not as bad as expected. It is possible that their prior negative experiences of struggle, effort, failure and difficult emotional consequences may make people with dementia or MCI hypersensitive to negative consequences, and generalise this to other scenarios, perpetuating the reluctance to act.

This does however contradict an experimental study in people with Huntington's Disease that found that apathy was associated with a lack of sensitivity to negative consequences [132]. The process of subjective experience of and learning from negative consequences requires further exploration in apathy in NCD.

It is also worth noting that in a similar study, whilst it was found that people with dementia are "hindered by invisible obstacles" [284, p.4], the authors also argued that participants purposefully externalise reasons for non-participation to preserve their self. The need to preserve self was also found in the present study, and is discussed below (in section 9.3.3). This argument highlights that the theme in this study '"Too much trouble": mediating effort and outcome' may not be a direct indication of real phenomena. Though for critical realism, qualitative interviews with participants can provide a window to real phenomena for interpretation, data may also be reflective of constructions rather than mechanisms, and all data will be viewed through a lens, rather than present direct access to phenomena [336]. Nevertheless, our theme echoes findings of other qualitative studies which conclude that people with dementia disengage due to loss of ability and to manage the stress associated with task failure [348].

Furthermore, these studies suggest that greater interest and motivation is required for something to be considered worth the effort. In other words, the threshold for action becomes higher. This can be seen as consistent with the proposals that apathy can occur from a deficit in reward sensitivity. For example, the conclusion that low effort, everyday tasks do not offer sufficient reward to motivate people with apathy to perform [122]. However rather than interpret this as exclusively a neuropathological mechanism, this may be learnt overtime. This has been recently acknowledged by Riehle and colleagues, who state that apathy occurs when "the assumed costs [of a behaviour] seem to outweigh the assumed benefits too often" [356, p.2], and that this may be the result of deficits in executive function processes that underlie decision-making, including impaired feeling or anticipation of reward, or can result from learned 'demotivating beliefs' from repeated experiences of failure. However, the latter mechanism was underemphasized and based on literature in the field of schizophrenia. The present study supports this proposed mechanism in people living with NCD.

9.3.2.2. Network analysis findings regarding effort

In the network analysis sub-study (phase 1), 'Effort' was largely not independently associated with other apathy indicators. This is consistent with findings of other studies investigating the structure of the AES, in which 'Effort' had low correlation with the remaining AES items [50,191]. Though on the surface the low expected influence of 'Effort' may appear surprising, given the importance of effort found in the qualitative study, the qualitative study may

further help us understand this (lack of) relationship. The qualitative study found that *"even just the little things"* are effortful for participants, and suggested great effort was often exerted for minimal everyday tasks. In effort-based decision-making tasks, discussed in section 3.3.2, though participants were less willing to exert effort for reward, they exerted the same level of effort once the decision had been made [122,131], which may be supported by the present network analysis findings regarding the item related to effort. Alternatively, this could be explained by potentially confusing double negative created by the combination of the item "s/he puts little effort into anything" with the given response options (from "not at all" to "a lot"), resulting in inconsistent responses from participants (discussed further in section 9.3.3.3).

9.3.2.3. Conclusions on effort and decision-making

These findings suggest that apathy may occur during unimpaired process of decision-making, as well as in impaired processing. Participants who experience increased need to exert effort, and reduced likelihood of a positive outcome and increased likelihood of negative consequences may suitably learn that many activities are not worth the effort. Participants may also become hypersensitive to these negative outcomes, and perhaps over-generalise this cautionary approach. This highlights the need for an integration of neurocognitive models of apathy in which various underlying impairments are proposed to impair the decision-making process, with a more behavioural outlook in which participants appropriately use these mechanisms. This can be expanded on further by the qualitative findings that participants act to preserve their identity, in ways that may produce apathy type behaviours, which will now be discussed.

9.3.3. Preserving Identity

The qualitative sub-study found that the struggle which participants experienced and the dependency that this could create meant that they experienced threats to their sense of competency and autonomy. No longer being able to do things was not just a matter of giving up specific activities, but was seen as giving up a part of their self, or even becoming a different person. A similar theme was found by Baber and colleagues [284], who reported the theme of "losing one's sense of self" with the subthemes: "Juxtaposition of the past and present self"; "Loss of interest in hobbies and activities"; "loss of motivation"; "loss of confidence". Loss of self in the context of illness has long been reported [357], and may be exacerbated in dementia where the person with dementia not only sees their situation as not improving, but worsening [358]. Though this assumes intact insight, which is in contrast negatively associated with apathy. Nevertheless, meta-syntheses of qualitative studies have demonstrated that changes experienced in dementia, in particular, memory problems, and the diagnostic label itself, threaten individuals' overall

identity, and sense of competency and autonomy, which in turn further threatens identity [347–349]. In summary, the findings of the present qualitative sub-study, that people living with NCD experience a change and threat to their identity, are well recognised. Whether this results in apathy, however, has not been as thoroughly explored, and will now be discussed.

9.3.3.1. Competency

The qualitative sub-study sub-theme 'preserving the capable and autonomous self' suggests that people living with NCD withdrew from activities and interests to avoid challenges to their sense of autonomy and competency, preserving their identity as someone capable and independent. This sub-theme shares some commonalities with Baber and colleagues' [284] 'feeling like a burden' theme, in which participants feared failure, were pessimistic, struggled against changes, and were avoidant. However, participants in the present study rarely discussed feelings of burden. This may be because they were interviewed together with a carer, and did not wish to speak openly in front of them. Alternatively, interviewing people living with NCD together with the person who cared for them may have helped reveal patients' avoidance strategies and thus present this as self-preservation as opposed to burden. Certainly, many instances were revealed only through the back-and-forth discussion between people living with NCD and carers about the patient's behaviours.

The proposal that people living with NCD withdraw to retain their identity as a competent individual is supported by various qualitative research studies into experiences of apathy and similar constructs such as negative symptoms, as well as general qualitative work about experiences of people with dementia. A qualitative study of people with psychosis, found that negative symptoms, similar to apathy, were explained as due to avoidance of negative experiences (in particular, stigma and embarrassment) [359]. Similarly, based on qualitative interviews about apathy in people with Parkinson's disease, it was theorised that withdrawal sometimes occurs due to embarrassment and loss of confidence [146]. Meta-syntheses have described that a common theme in qualitative studies of people with dementia is that avoidance occurs, and this can be due to threats to identity, and to avoid being 'exposed' [347], and that studies have reported experiences of difficulties that lead to feelings of incompetency, resulting in avoidance and withdrawal [348].

Competency is considered an inherent universal basic psychological need within Self-Determination Theory, a motivational theory which posits that fulfilment of basic psychological needs enables individuals to seek interesting and enjoyable activities and engagement with others [354]. Self-Determination Theory proposes that when these needs are not satisfied, individuals can develop maladaptive strategies to cope, such as withdrawal, resulting in negative outcomes such as loss of motivation and psychological distress

[360]. In particular, competency is proposed to be necessary for all motivation, and experiencing failure undermines this need [354]. Alongside competency, autonomy is also considered a basic psychological need within Self-Determination Theory, and is proposed to be required for intrinsic motivation, in which individuals are self-motivated [354].

9.3.3.2. Autonomy

The present study also found that people living with NCD experienced threats to their sense of autonomy and this may lead to withdrawal, which is supported by other qualitative work. For example, it was reported in a study of carers and people with young onset dementia that loss of autonomy prevents feelings of usefulness, and the changes and losses they experienced could lead to withdrawal [345]. A meta-synthesis found that in the post-diagnostic phase, individuals experience lack of control, as well as lack of competency, leading to reduced engagement, and that individuals use strategies to retain their original identity that can include withdrawal [349].

Whilst people living with NCD in our study often used these avoidance strategies, they also sometimes felt that they had no choice but to accept their impairments and changes this brought. This acceptance has previously been characterised as another, more ‘adaptive’ coping [345] strategy to preserve identity, in which, as in the present study, people with dementia adjusted to the changes they experienced through reluctant acceptance or making active decision to move on, focusing on what they could still do, rather than what they were no longer able to do [347–349].

Participants in the qualitative interview study sometimes felt that they had no choice but to move on, and this is echoed in other studies in which participants experienced ‘resigned acceptance’ of what they could no longer do [347] may suggest that feelings of lack of autonomy (i.e. that nothing can be done) is not necessarily detrimental, as it can lead to acceptance. However, the apparently ‘adaptive’ coping strategies of ‘moving on’ could still result in less engagement and more limited interests and activities, both in the present study and others [347,349]. For example, Steeman and colleagues [349] characterise these as self-protective, in which the individual strives to maintain life as normal, but is met with challenges to this and adopts an avoidance approach to manage this, and self-adjustment strategies, in which the individual accepts the disorder and moves on, sometimes by giving up activities they can no longer do.

9.3.3.3. Lack of insight

It has been argued that this ‘adaptive’ moving on described above may be facilitated by insight into problems, enabling acceptance, whereas lack of insight may result in withdrawal and avoidance of situations where participants

are unexpectedly faced with challenge [348]. This may explain the proposed link between apathy and lack of insight (outlined in section 2.3.2.1).

The present study found that people experienced that their expectations did not always match their present capabilities, which could result in distress. It is possible that unachievable, or a mismatch in, expectation is particularly stressful, as opposed to lack of ability on its own, as the 'surprise' of discovering that something is no longer achievable may lead to additional feelings of disappointment and dissatisfaction. This relates to both Self-Determination Theory, which posits that thwarted feelings of competency will result in amotivation [354], and (as concluded in section 3.3.2) decision-making models, which emphasize the need to learn from outcomes to maintain motivation for goal-directed activity [121]. It has additionally been theorised that lack of insight could lead to worse adaptations to impaired ADL, which could in turn lead to apathy, which is similarly relevant to decision-making models and the importance of learning [61]. However, to my knowledge neither of these models of motivation have integrated impaired insight.

In the network analysis sub study, the node 'SelfConcern' ("S/he is less concerned about problems than s/he should be"), could be seen to assess insight. This had low expected influence, perhaps suggesting its lack of importance to the construct of apathy, or perhaps that it is a distinct domain. However, the item 'SelfConcern' also results in a double negative when combined with the response options, and has previously been found to have a low correlation with the other AES items [50]. Issues with the double-negatives created in the AES and their possible effect on factor loadings have previously been noted [47,142]. This highlights the importance of developing comprehensible questions in context of the response options, as highlighted by the COSMIN criteria, and discussed in the systematic review (Chapter 4). Nevertheless, the item 'Insight' also had relatively low expected influence, further emphasizing its separation from other indicators of apathy.

9.3.3.4. Conclusion on preserving identity

Withdrawal, avoidance, and reduced interests and activities seen in apathy may be understood as responses to threats to basic psychological needs of autonomy and competency, in an attempt to retain identity. The role of identity and lack of insight has been ignored by models of apathy in people living with NCD. The importance of identity is also similarly highlighted by findings regarding participants' beliefs about what is important and necessary, that will now be discussed.

9.3.4. Role of purpose and necessity and beliefs

In the network analysis sub-study, the belief that 'getting things done during the day is important' ('ImportDone') was the third most central node in the first network phase, and was independently associated with other nodes about belief, including whether getting things started on their own was important ('Started'), and seeing jobs through to the end was important ('Completion'); but also general motivation, and whether they got things done during the day ('GetDone'). The belief that getting things done is important suggests a sense of duty, need or purpose. This could be explained by findings of the qualitative sub-study, in which people living with NCD described retaining more interest and activity where this fulfilled a sense of purpose or where these were deemed necessary. Perceived roles and responsibilities were also important within this, which has also been found elsewhere. Baber and colleagues [284] similarly found that participants were 'kept going' by a desire to help others, and people with schizophrenia have been found to gain a sense of purpose from caring for others, which could help reduce negative symptoms [155]. A meta-synthesis of qualitative studies of coping in dementia found that some participants wished to be useful and helpful to others, and in doing so, retain meaning in their activities [347]. In caring for others, people with dementia may retain a sense of purpose. A loss of purpose has been described as resulting in a reduced desire to do things in adolescents with depression related anhedonia [73], further suggesting purpose is important in motivation.

It is also possible that caring for others enables participants to assume a role of higher status, as the qualitative study also found that participants avoided activities in which they felt or were seen as of a lower status. Similar to the present qualitative sub-study, another qualitative study regarding apathy described a participant who engaged in day centre activities, but perceived themselves as a teacher rather than an attendee to a day centre, and interpretated this as a way of retaining his "sense of identity as someone who is in charge" [285, p.6]. Being able to assume a higher status may enable a sense of competency and autonomy which support motivation and participation.

These studies indicate that a sense of purpose and necessity in activity, in particular, caring for others, may encourage engagement and reduce apathy. However, it is worth noting that the reverse cause and effect could be true: if people lose the willingness to do something, they may no longer feel it is their role, useful or necessary. This relates to the aforementioned issue that though qualitative research enables a closer look at mechanisms, the participants do not necessarily have direct insight into the causal processes. Nevertheless, others have similarly argued that feeling useful is important in activity participation, at least for people in the earlier stages of dementia [345], and loss of control or purpose was a proposed personal factor that could act as a

stressor and lead to apathy in Massimo and colleagues framework for apathy [153]. Furthermore, personal beliefs and experiences have previously been acknowledged to inform the decision-making process, within Levy's explanation of apathy as a 'dysfunction in the valuation system' [107], though this was only recently included in this largely neurocognitive explanation of apathy, was underemphasized and lacked elaboration.

9.3.4.1. Difficulty in maintaining purpose in NCD

The role of purpose and necessity deserve special consideration within a model of apathy for people living with NCD, as retaining roles, purpose and a sense of necessity may become increasingly difficult in people living with NCD. Though people living with NCD wish to and do participate in meaningful activity, this becomes increasingly challenging to achieve as dementia progresses [347]. Meaningful engagement may need to be reconsidered towards the later stages, where vicarious activity may suffice [361]. Pool has emphasised the importance of matching abilities ('activity levels') with what is offered at all severity levels in dementia [362].

Furthermore, the present qualitative sub-study found that the carer support that the person living with NCD received could result in a change in their sense of responsibility, believing some things were no longer necessary or relevant for them. This has been echoed elsewhere, in descriptions of carers 'over-helping' which is deemed to risks causing 'excess disability' in people with dementia [363].

9.3.5. Opportunity and Exclusion

9.3.5.1. Importance of social support and social networks

The importance of social support and social networks was emphasized by the qualitative sub-study (see section 8.3.4.1), with maintaining social contact seen as inherently positive. This is echoed by other qualitative research that has shown that social support and feelings of security are important for managing negative symptoms (which share similarities with apathy) in schizophrenia [155] and engagement in people with dementia [345,364]. Furthermore, 'relatedness' is considered the third basic psychological need in Self-Determination Theory (alongside competency and autonomy), and is proposed to be important for intrinsic motivation [354], suggesting it is important for engagement and activity.

The present study found that carers sometimes took on an important organizer and facilitator role, proactively engaging participants, and could be vital to motivation, engagement and activity. This finding expands the previous qualitative interview study with people with dementia, as though it was reported that people with AD acknowledged their need for their spouses to motivate them [284], the importance of and ways in which carers did this was

not discussed. This may have been due to participants being interviewed alone, so they may have lacked insight into the ways in which their carers facilitated their activities. The importance of the role of carers in keeping participants engaged was found in a qualitative interview study of carers of people with AD [285], supporting this assertion. Furthermore, carers have previously been reported to be a source of support in managing various behavioural and psychological symptoms of dementia, including apathy [283,363], and supporting 'volition' of people with dementia in care homes [361].

9.3.5.2. Importance of opportunities, understanding and education

The present qualitative sub-study found that opportunities in the environment, and carer or others support through prompting and organizing activities, and reducing their difficulty, encouraged participation and engagement, which supports the proposal that apathy can result from unmet needs such as lack of stimulation [154]. People living with NCD had increased needs and required additional support to remain engaged and participate in a meaningful way. This is consistent with Kitwood's notion that increased 'work' was required of carers to enable people living with dementia to be included, occupied and related to [351]. Sometimes, this additional work was too much, and carers had to take over or not include the participant. Sometimes carers or others failed to provide extra support for cognitive symptoms, because they assumed a greater capability or lacked understanding of the symptoms, resulting in exclusion. Lack of education of carers regarding dementia symptoms has previously been identified as a hypothesized mechanism for apathy [153]. This is also supported by the findings from a qualitative study in which people with Parkinson's disease were said to experience 'psycho-emotional disablism' [146]. This theory highlights how people with impairments experience not just structural disablism, in which people experience discrimination and inaccessibility, but also the effect this has on their psychological and emotional well-being, for example direct and structural discrimination can cause individuals to be reminded of their 'other' status, and be negatively emotionally impacted by others' and their own internalized stigma [365]. In addition to dementia symptoms, understanding of apathy was mixed, as discussed in 9.2.3. This highlights the need to support and educate others, in particular, informal carers about NCD and apathy to enable them to effectively support people with NCD experiencing apathy. Similar calls for more information and education for carers regarding apathy have recently been made [282,345].

9.3.5.3. Familiarity and routine

In the qualitative sub-study, it was found that people avoided threats to their sense of competency and autonomy, and the unknown and unfamiliar, in particular, threatened these and could cause distress. As a result, individuals

preferred to stay within their ‘comfort zones’. A supportive environment offering familiar activity and routines was important for engagement. This is supported by other qualitative research which has similarly found that familiarity supports a sense of self and supports participation and engagement of people with dementia [348,364]. Routine was also identified as a sub-theme in Baber and colleagues theme of ‘what keeps me going’, which they argue reduced cognitive load, taking away the need for decision-making [284], which complements existing neurocognitive models of apathy that theorize apathy may be the result of a deficit in the decision-making process [107,120,121]. These findings are also consistent with Self-Determination Theory, as familiarity and routine could be seen to facilitate feelings of competency, important for engagement [354].

9.3.5.4. Conclusion regarding opportunity and exclusion

The finding that apathy may be understood in part to be a response to exclusion and lack of opportunity is consistent with the unmet needs model of neuropsychiatric symptoms outlined in section 3.5.1. These findings highlight that interventions to reduce apathy by providing opportunities for activity, routine, and understanding and familiar environments may be successful in reducing apathy.

9.4. Conclusions on mechanisms of apathy

9.4.1. Neurocognitive models

Neurocognitive models can be criticised for their bottom-up, impairment focused explanation of apathy, with evidence relying on neurobiological associations and experimental studies applying a closed system approach. However, the present thesis does not dismiss the role of these mechanisms and in fact supports some of these proposed mechanisms. Neurocognitive decision-making models conclude that due to various disrupted processes, apathy is the result of reduced willingness to exert effort [127]. The network analysis sub-study found that ADL was associated with indicators of apathy, and the qualitative study overwhelmingly found that these activity limitations meant that things were seen as not worth the effort, or ‘too much trouble’. However, rather than being a disordered process, this appeared a legitimate estimation, though there was some suggestion of a hypersensitivity to negative consequences which is also hypothesised by neurocognitive models. As a result of difficulty in everyday life, including increased effort and likelihood of negative consequences, it could be understood that individuals’ natural threshold for actions are heightened. However, it is important to recognise that the findings of the present study were largely not understood to be the result of impaired processing, and indeed executive function was not independently associated with apathy indicators, but was associated through

ADL, i.e. through activity limitation. This highlights how neurocognitive models need to be integrated with other ICF components that affect activity, participation and disability, to enable a comprehensive understanding of apathy.

9.4.2. Massimo's conceptual framework for apathy

Although one attempt has been made to produce a biopsychosocial conceptual framework for apathy [153] which is consistent with the ICF's understanding of disability, this lacked a clear integration with neurocognitive models.

The present study found support for some aspects of Massimo's model. Carer factors of mismatch of expectations and lack of education about dementia, and environmental factors of lack of activity, structure and routine were supported by the qualitative study, which emphasised the role of opportunity and exclusion. Personal factors of unmet needs (which include loss of control and purpose), and changes in ability to interact with the environment could be seen to be supported by the network analysis sub-study, in the finding of the relationship with ADL and apathy indicators related to beliefs and by the qualitative study, in the finding of the importance of declining ability and sense of purpose and necessity.

More integration of mechanisms is required, and there is much potential for overlap of Massimo's model and decision-making models. Whilst a sense of purpose is understood as an unmet need in Massimo's model, it can be additionally understood to be part of the decision-making process, with whether an activity is necessary or meaningful forming a key part of the valuation process, which has previously been briefly recognised [107].

The present study also expands our understanding of the potential mechanisms of apathy, not currently recognised by the models discussed. Massimo recognised the role of mismatch in expectations of the carer, but not of the individual. The present study found that a mismatch in expectations of the self could result in threats to identity. Consistent with Self-Determination Theory, competency and autonomy were important and threats to these were experienced as a threat to identity, so individuals showed withdrawal and avoidance or a narrowed focus of activities to preserve their identity as a capable and autonomous individual. Current models of apathy fail to recognise this, but this could be integrated as part of the normal learning process, with individuals learning to avoid scenarios that lead to negative outcomes, including these threats to self. A carer factor that could be included in Massimo's model of apathy could be 'overhelping' as this was found in the present qualitative study and in others. Additionally, a greater emphasis could be placed on the role of social support within the environmental factors of

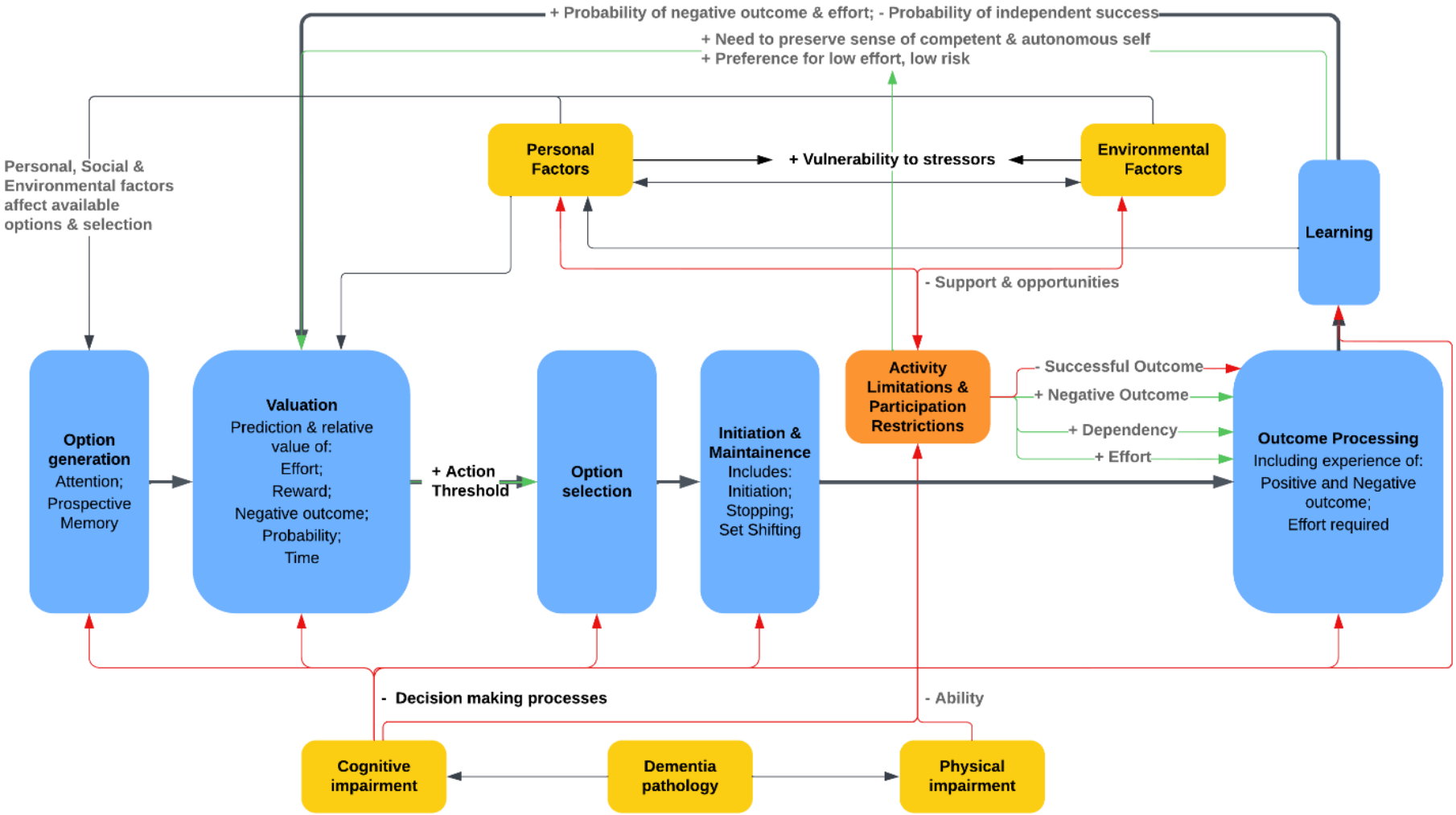
Massimo's framework for apathy, as this was key to participation and engagement in the present study.

9.4.3. An updated conceptual model of apathy

The proposed additions to a conceptual model of apathy in people living with NCD are summarised in Figure 14. This model integrates neurocognitive models (drawing on figures previously provided by decision-making models [120,121]) with the previously proposed conceptual framework for apathy [153,157] into a framework consistent with the ICF. This model also extends the current models in numerous ways. The proposed carer factors are incorporated in a more inclusive single 'environmental' factor, which includes not just carer factors but wider social influences [238]. Descriptions of environmental factors and personal factors are not included due to space limitation, though it is noted that these may include the factors outlined previously. In particular, emphasis should be placed on personal factors of personal beliefs regarding the value of activities, particularly in regard to their perception of purpose and necessity, individuals' outlook and personality, and the extent to which needs for competency and autonomy are (un)met. Environmental factors that warrant emphasis are the absence of routine, familiarity, prompting, and other support for cognitive and physical impairment, a lack of education of others regarding symptoms of NCD, and a mismatch in expectations from others. Furthermore, the present model recognises the complex and interactive nature of apathy. This model expands neurocognitive models as in addition to impairment underlying potential disruptions in the decision-making process, activity and participation are recognised as a different level of phenomena, in which different interactions can occur. In line with the ICF, in the present model, in addition to physical and cognitive impairment, personal and environmental, including social, factors, are also proposed to influence activity limitations and participation restrictions. For example, carers can reduce the effort required in activity and individuals may use strategies such as establishing routines to support participation. It is proposed that these activity limitations and participation restrictions can result in increased experiences of increased effort and negative outcomes, and decrease in successful outcomes, which are integrated into the decision-making process itself. These experiences inform probability estimations in future valuations, leading to increased thresholds for action. This model also proposes that environmental and personal factors not only act as stressors, as proposed by Massimo and colleagues [153] in their conceptual framework for apathy, but also are informed by ability and activity of people living with NCD. For example, people who experience activity limitations due to cognitive and physical impairment may also experience stigma, which in turn may reduce support from others, and this additionally may affect individuals' own feelings of self-worth and place in society, termed psycho-emotional disablism

[146,365]. A need to preserve the self was also included in the model, as this was found to be key in the present study and in previous qualitative literature exploring the experience of people living with NCD [349]. It is proposed that this may inform the decision-making process, as individuals withdraw from scenarios which they feel may challenge their identity as a capable and autonomous individual.

Figure 14. Updated Conceptual Model for Apathy in People living with NCD



Black text reflects existing theories whilst grey text indicates an additional or more specific proposed mechanism proposed in this thesis. However, it should be recognised that there will be some cross-over. '-' Denotes decrease (also indicated by red lines) whilst '+' denotes an increase (also indicated by green lines). Blue boxes and thick lines reflect the existing decision-making models of apathy, in particular those proposed by [28,121]. Yellow and orange boxes reflect theories and proposed mechanisms consistent with the ICF model [139,238], with yellow boxes additionally indicating consistency with the previous conceptual framework for apathy [153,157].

9.5. Challenges and changes to the study

9.5.1. Impact of COVID-19

The Coronavirus Disease 2019 (COVID-19) was declared a pandemic by the World Health Organization on 11th March 2020 [366]. On 16th March 2020, the UK prime minister advised that all non-essential contact and travel should cease, and that would be particularly important for vulnerable people, including the over 70s [367]. The Director of Science, Research and Evidence at the Department of Health and Social Care and co-lead for the National Institute for Health and Care Research (NIHR) additionally issued a statement advising that many NIHR-funded studies would need to be paused unless doing so would be detrimental to participants [368].

9.5.1.1. Impact of COVID-19 on the wider PrAISED2 study

All recruitment and face-to-face follow-ups of all participants to the main PrAISED2 study was suspended on 17th March 2020. At this time, 300 participants had been recruited to the trial, with 63 participants having completed follow-up assessments. Face-to-face follow-up assessments and recruitment resumed on 1st September 2020. There was an additional pause in recruitment between 15th December 2020 and 15th February 2021 due to waiting for extension. Working from home and being unable to conduct visits in person for an unknown amount of time, and subsequent risk of COVID-19 after in-person work was enabled meant that the numerous and complex changes were made to the PrAISED2 study.

In the initial months of COVID-19 lockdown, the following changes were made:

- An amendment to the protocol to enable immediate remote collection of data of those who had been in the trial for 10.5-12 months. Therefore, some participants were followed-up up to 6 weeks early.
- An additional remote ‘interim’ data collection time point of those who had been in the trial for 3-10.5 months. This was intended to capture a snapshot of participants’ status’, as much as possible before continued COVID-19 restrictions potentially altered this.
- An amendment to the protocol to enable participants who had not completed their initial therapist assessment (both in the control and intervention groups, *N*=26) to be ‘paused’ so that they did not continue to receive therapy until face-to-face visits could recommence.
- Remote data collection was conducted across all sites. This meant that some measures could not be conducted, such as those requiring physical

assessments, for example the TUG, BBS, SHARE-FI, or the researcher to administer the test, such as cognitive assessments.

- Remote data collection involved new procedures: Researchers first called participants to inform them of the changes, then posted the case report forms to participants. Participants returned this case report form postally, then were followed-up with a telephone call to support this where needed, or to ask for any missing or unclear responses. This sometimes meant asking all questions over the telephone. Follow-ups were sometimes late due to delayed return of remotely collected data.

In August 2020, the government was encouraging the return to work and usual face-to-face contact was allowed, providing precautions were taken, such as the use of personal protective equipment. As a result, the following changes were made to the protocol and submitted as a non-substantial amendment:

- Face-to-face visits were resumed, however, many participants did not wish for a visit in person, so a hybrid method of data collection was used, meaning that some participants data continued to be collected remotely.
- A new face-to-face (baseline and follow-up) case report form was produced, which did not include some measures that were originally included prior to the pandemic, due to their infection control risk.

Recruitment was paused again between mid-December 2020 and mid-February 2021, as the initial approved time spent in active recruitment had elapsed. A substantial amendment to the protocol was submitted on 5th January 2021:

- To enable extension of the recruitment period, from the initial 21 months, to 27 months, enabling recruitment to continue until July 2021 (and therefore follow-ups to continue until July 2022).
- To enable extension of paused participants' time in the study, so that their follow-up was not due until 12 months after they completed their assessment visit once face-to-face therapy resumed. This required an additional consent process. This meant that some participants were followed-up up to 21 months after their baseline data was collected.

9.5.1.2. Impact of COVID-19 on the network analysis study

Initially, I had planned a fourth phase of network analysis, in which I aimed to assess the stability of the network overtime, by comparing the network at baseline (presented in this thesis) with a network produced from the follow-up data. The availability of the follow-up data was deferred due to the delays and extensions to the recruitment period due to COVID-19. Additionally, there were issues with the incompleteness of data, and the variability in follow-up

time and method (remote vs face-to-face), which would have meant that checking stability would have been inherently undermined by the variability in data collection.

This also meant that I had to change the measures used in the third phase of the analysis, as whilst baseline data collection was always conducted face-to-face, some measures were dropped upon return to face-to-face data collection after the pause in recruitment. These were an accelerometer-based measure of activity, a personality questionnaire, and a computer based cognitive assessment. The response rate for these measures was already poor due to technical issues, and some acceptability issues (i.e. were not tolerated by all participants), so this led to the decision to abandon these assessments when face-to-face visits resumed, as their benefits were deemed to be insufficient comparative to the infection risk associated.

9.5.1.2.1.Changes to planned measures

I initially planned to use two of the measures that were subsequently dropped from the baseline data collection. These were a measure of executive functioning, and a measure of activity:

- The 'CAmbridge Neuropsychological Test Automated Battery' (CANTAB®) [369] assesses executive functioning, and is administered directly to participants on a computer tablet device. This measure would have been included instead of the MoCA to assess executive function, as CANTAB is argued to be more sensitive, as it is able to score precise reaction times, and prevents ceiling and floor effects by enabling trials to become progressively difficult depending on the participants' success with previous trials.
- The 'Misfit Shine 2' is a commercial-grade accelerometer that records steps taken per day, and minutes of high, moderate and low-level activity. It can be worn on the wrist and its water resistance enables it to be worn all day, and participants in the PrAISED2 study were initially (before COVID) asked to wear this for one week, twice during the study. This measure would have been included as an objective measure of activity, which has been previously encouraged for apathy research [236].

Instead, I used the MoCA and verbal fluency score as an additional measure of global cognition/ executive function, and the BBS and TUG were included in analysis as physical measures instead of the accelerometer data. Though BBS and TUG were not measures of activity, these were included as evidence from the qualitative interviews was emerging that suggested that physical mobility was a barrier to goal-directed activity and participation and could even affect interests.

9.5.1.3. Impact of COVID-19 on the qualitative study

I initially aimed to complete transcription and transcript checking for each interview before the next interview was conducted, and analyse some interviews whilst continuing to collect data. This was to allow for lessons to be learned, and for interesting points and emerging ideas from earlier interviews to be explored in later interviews, as well as developing theories to be tested in subsequent interviews [336]. However, I was unable to do this, as transcription return took longer than expected, and I had not accounted for this when arranging the initial local interviews. Furthermore, some interviews required travelling large distances, so I chose to cluster them together to reduce expenses, rather than wait to conduct analysis between each interview. I was able to collect data in parallel with listening back to interviews and memos, and checking, correcting and re-reading transcripts as soon as possible after each interview, which still enabled me to enhance familiarity with the data, think critically about my approach to each interview, develop my skills, adapt the interview schedule and make notes about tentative ideas and points of interest throughout the data collection process. I was not able to formally analyse the data through coding (phase 2 and beyond of thematic analysis) whilst also conducting the interviews. As a result, I decided to pause my qualitative data collection and conduct a full iteration of reflexive thematic analysis prior to continuing data collection. My intention was to allow this initial analysis phase to inform a re-development of the interview schedule that would allow me to test my tentative theories on the final four participants. Whilst this could have been done by re-interviewing the recruited participants, building on what was discussed previously, this was deemed inappropriate for people with NCD who would have difficulties recalling discussions from the previous interview. However, following government restrictions regarding visiting people in their homes, due to COVID-19, I was unable to collect further qualitative data using face-to-face semi-structured interviews. This is not necessarily problematic, as when coding occurs is flexible [338], so this was not deemed detrimental to the project. Therefore, data collection did not recommence after the pause due to COVID-19.

9.5.2. Impact of personal circumstances

When face-to-face visits were resumed in September 2020, I was entering my third trimester of pregnancy, at which time the advice was that pregnant women should work from home from 28 weeks' gestation [370]. During the birth of my child in February 2021, I sustained an injury which required major surgery and left me with a long-term medical condition which was emotionally and physically disruptive. I required major surgery again in April 2022. As a result, I did not personally conduct any further face-to-face baseline or follow-up visits with participants after March 2020 when visits were paused due to COVID-19, though, I continued to support remote data collection.

9.6. Strengths and Limitations

The limitations specific to the systematic review, network analysis and qualitative sub-studies are outlined in sections 4.5.3, 7.3.2 and 8.4.2 respectively. The strengths and limitations of the overall mixed methods study will now be discussed.

This study was conducted in a general NCD population, rather than those with 'clinical' or self-identified apathy. With regards to the network analysis sub-study, the network structure of apathy may appear different in those exhibiting severe symptoms, as indicators of apathy may causally influence each other differently when the indicators are experienced more severely [251]. This problem is not limited to network analysis, and the issue has been described by McNally [254], who used the example of predictors of success in basketball: whilst in the general population, height would be a strong predictor of basketball performance, if only professional basketball players (with much taller than average height) were included in the analysis, height may no longer be a predictor for success. Though the average apathy score for the population in the present study was suggestive of 'clinical apathy', these results may not replicate in a population with self-identified apathy or those meeting latest proposed criteria for apathy [39]. It is also worth noting that exploring apathy in this way assumes that apathy is a homogenous construct [251]. If there are different types of apathy, as has been previously proposed [24,32,60], sampling from general cognitively impaired participants, as in this study, or those meeting general criteria for 'clinical apathy', may result in a dilution of the interactions between indicators of apathy. This may explain why apathy did not cluster in the proposed domains. Future research should therefore conduct network analyses of apathy indicators in different populations, selected for general apathy and different types of apathy. Nevertheless, these limitations are not so problematic for the qualitative sub-study, which sought to contrast participants experiences rather than be limited to assessing the average experience. The qualitative sub-study similarly included participants who would be considered to have clinical apathy and those that would not, however this facilitated triangulation, as it enabled contrasts and comparisons to be made, which is a key for interpretation of qualitative data [327].

Patient and public involvement representatives were consulted throughout this study where possible. This helped ensure the methods were sensitive, appropriate, and relevant to people living with NCD and their carers [371]. This also further enhanced reflexivity in the qualitative study, as this provided another perspective from which to interpret the findings, highlighting my own assumptions and perspectives [335].

A convergent parallel-databases mixed methods design was chosen due to time constraints. An explanatory sequential method could have been used so

that network analysis findings could have been explored and explained through qualitative interviews, which may have enabled better exploration of possible mechanisms for apathy identified in the network analysis study, and this could have better supported integration of findings. However, it was not feasible to wait until all quantitative data was collected and analysed before conducting qualitative interviews. Furthermore, this enabled the qualitative research to focus on different phenomena, such as the importance of social support and opportunity, which may not have been discovered if the goal was to explain the network analysis findings.

The present findings and literature discussed have largely focused on apathy as a reduction in goal-directed behaviour. This has meant that emotional indicators, such as reduced emotional expression, are less understood and expanded on in this study. Though the qualitative study aimed to understand the emotional component of apathy, this was less discussed, as detailed in section 8.4.2.

The present findings appeared to suggest the presence of a distinct social domain of apathy, however, it is also worth noting however that ‘Social’ was most related to ‘Excited’ and ‘Novelty’, and the clustering of these nodes could possibly be attributed to a common cause, latent trait, separate from apathy, not included in the present model: extroverted personality. Network analysis assumes that there is no underlying latent trait that is not included in the model that is causing relationships between nodes, however it has been noted that this might not be plausible in psychopathology research [271]. Similarly, as all nodes were developed to measure a singular construct (apathy), and reversed to ensure that a low score on each node reflected worse apathy, prominent negative edges (associations) were not expected. The finding of the stable negative edge between ‘SelfConcern’ and ‘Excited’ was therefore surprising. However, ‘Excited’ has been previously found to have a negative factor loading in the AES [50]. It is possible that ‘Excited’ is also influenced by personality or attitude, with increased excitement reflecting a more positive attitude that is negatively associated with self-concern, which could be seen to reflect a negative attitude. This highlights the difficulty in developing measures to assess apathy independent of overlapping constructs, such as underlying personality. Future research should consider personality and its relationship to apathy.

Critical realism recognises that causation is context dependent, and so these theorised mechanisms are not expected to generalise outside this context [336]. For example, people living with NCD in this study were all aged 65 or over, so the focus on comorbidities and physical impairment found in the qualitative study, and their influence in the network analysis, may be more pronounced than in people with young onset NCD. This study also focused on people with mild to moderate NCD, so findings are likely to differ in people

with severe dementia, many of whom may live in residential care or nursing homes and may have their liberty restricted.

9.6.1. Limitations regarding measurements

The network analysis sub-study adds to previous network analysis studies as it is the first study that used dedicated apathy and depression measures assessing multiple indicators of both constructs. This is in contrast to previous network analysis studies that have taken items from a measure of depression [276,277] or used a single item to assess apathy [278]. However, this meant that the measures used had different assessment methods, time frames, and response options, which may have artificially separated the nodes belonging to the different measures. For example, the HADS-D items were self-rated and asked participants to rate their symptoms in the last week, whilst the AES was proxy-assessed and asked the respondents to rate their symptoms over the last four weeks. This is particularly relevant as there are known differences in carer proxy and self-ratings from people living with NCD regarding apathy and depression [372,373], and weaker relationships have previously been found between proxy and self-rated apathy and depression [275]. However, DEMQOL-proxy items (which are also carer-rated, assessing symptoms over the last four weeks) were included to assess depression where possible to minimise this effect. This could somewhat explain the increased bridge centrality of ‘Energy’ and ‘Cheerful’, as both these items were proxy-rated, which may have increased their association with apathy items. However, the remaining (proxy and self-rated) depression items had similar bridge centrality to one another, suggesting energy and cheerfulness are likely shared indicators of apathy and depression. Nevertheless, carer-rated proxy measures have been criticised for their bias, as they can be influenced by carer factors such as burden [153].

The HADS-D was largely designed to assess anhedonia, and notably does not include some of the most extreme symptoms of depression, such as suicidal ideation [294]. Furthermore, neither the DEMQOL(-proxy) nor HADS-D includes somatic symptoms of depression. Therefore, key components of depression were missing from the present study. However, extreme symptoms are problematic for network analysis, as they do not have an underlying normal distribution, i.e. by their nature, they are only experienced by a small percentage of the population [270], and somatic symptoms are problematic to assess in older people with NCD, as they are often the same symptoms caused by comorbidities experienced by this population [294,340]. Further work is needed to understand how best to assess these symptoms in future network analyses in this population.

ADL was associated with apathy indicators, including one related to needing to be prompted, and executive function was found to be associated with a

variety of apathy indicators only through its association with ADL. Though the ADL measure used in this study attempted to separate ‘initiation’ from ‘performance’ through its different subscales, it required that participants performed tasks ‘without prompting’. This may have meant that the measure failed to distinguish between problems with initiation and task performance which may have inflated the relationship between ADL (performance) and executive function (initiation) and the apathy indicator related to prompting.

As described above, the MoCA was used to replace a measure of executive function (the CANTAB) that could not be completed due to COVID restrictions. The MoCA was designed to measure global cognition, rather than executive function per se. Different executive functions may have different relations to different apathy indicators, so the summation of various cognitive processes into one measure may result in missing some important associations. This is indicated by its poor internal consistency. Nevertheless, verbal fluency was included as a separate measure of executive function to assess initiation, and was only merged with the MoCA score following item reduction, which indicated the two assessments had sufficient shared associations with the other variables in the network.

9.7. Recommendations for future directions

Network analysis is a novel method in this field, and relatively few studies have assessed apathy using this method. A number of recommendations can be made for future network analysis studies. Future studies could assess participants only with ‘clinical’ apathy though the identification of this using a measure or criteria would itself influence the network structure, so findings would not be expected to replicate. Whether the collection of time-series data is feasible with people with living with NCD should be investigated, as this could provide another method via which to assess possible interaction of symptoms. Further research should also include a depression scale validated for people with living with NCD that assesses all aspects of depression. Furthermore, an alternative measure of apathy could be used that more suitably represents the varied proposed dimensions of apathy, for example, the dimensional apathy scale, which showed promising psychometric properties in the systematic review. Measures of apathy need to be developed with people living with NCD and their carers to ensure they are comprehensible, relevant and comprehensive. These measures should not be refined through factor loadings, to enable indicators of apathy to be assessed individually in network analysis. Wichers and colleagues [374] have similarly suggested that patients and carers could be involved in identifying distinct symptoms for inclusion in future network analysis studies. Future research could conduct network analyses using additional concepts from other models of apathy and motivation, for example the basic psychological needs of competency, autonomy and relatedness that are essential for intrinsic

motivation according to Self-Determination Theory. Furthermore, the concepts of personality or outlook could also be investigated, as it is possible these concepts were affecting relationships in the network analysis.

The present study aimed to conduct the emergent theory building phase of critical realist mixed methods research. The next step might be to test some of these assertions. Further qualitative studies could present a theorised model of apathy in a realist interviewing method, to gain perspectives and insights from those with lived experience [375]. A systematic review and meta-analysis could be conducted to investigate the evidence for each proposed process in people living with NCD, as there is much overlap from other fields, such qualitative studies on the experience of dementia, and experimental studies investigating the process of decision-making, and research in the field of motivation.

The participants within this mixed methods study were mostly white, and all were over 65. As noted previously, culture influences how cognitive impairment is experienced [344] and age may have influenced the association between physical ability, functioning and apathy across the two sub-studies. Further research should aim to understand the mechanisms of apathy in people from Black, Asian and minority ethnic backgrounds, and people with young onset dementia.

The qualitative study suggested that future research and therapies targeting apathy may benefit from describing the components of apathy and perhaps avoiding the term, given the varied and negative interpretations of apathy. Future research with carers and people with dementia or MCI may benefit from describing the components of apathy, rather than using the term apathy directly, to ensure the views and experiences of those who are not as well-informed about apathy are also included.

Whilst it is emphasized that this thesis presents exploratory work, the findings, along with the literature discussed, and proposed model for apathy provided, can be used to inform the development of future therapeutic interventions. Taken together, these studies suggest that apathy could be managed at least in part by environmental alterations, including carer support. This highlights importance of schemes to promote inclusion and participation in wider society, such as ‘dementia friends’, and ‘dementia friendly spaces’, as well as ‘age-friendly’ schemes, as some people living with NCD faced barriers to participation due to physical health in addition to cognitive difficulties. Provision of education programmes and information for carers could be important in dispelling misconceptions about and providing information on the occurrence of apathy in NCD. Management of apathy could involve supporting individuals’ sense of competency and autonomy and reducing effort and risk of negative consequences by the use of prompting, establishing routines, maintaining familiarity, adaptations to reduce effort of activity, engaging

people living with NCD with established interests and social support groups, and providing opportunity for purposeful or necessary activity.

9.8. Conclusions

The current study found that motivation may be the key feature of apathy, and apathy may not be composed of distinct domains, except perhaps a social domain. Apathy may be understood as a loss of motivation in reaction to challenges to individuals' sense of competent and autonomous self, and lack of appropriate opportunities and support. This questions current criteria for apathy which have removed the role of motivation, and required that indicators are present from different supposedly distinct dimensions. The exclusion of cognitive or physical impairment or lack of opportunity from the diagnostic criteria for apathy [39] is problematic, given the importance of cognitive and physical impairment, social support and opportunity in apathy demonstrated in this study. Apathy may be improved by inclusive opportunities for people living with NCD and carer education regarding NCD and apathy.

References

[1] Bajwa RK, Goldberg SE, Van der Wardt V, Burgon C, Di Lorito C, Godfrey M, et al. A randomised controlled trial of an exercise intervention promoting activity, independence and stability in older adults with mild cognitive impairment and early dementia (PrAISED) - A Protocol. Trials 2019;20:815. https://doi.org/10.1186/s13063-019-3871-9.

[2] Oxford Dictionary of English. apathy. Oxford Dictionary of English 2010.

[3] Oxford English Dictionary. apathy. OED Online n.d.

[4] Cambridge Advanced Learner's Dictionary & Thesaurus. Cambridge Advanced Learner's Dictionary & Thesaurus n.d.

[5] Chase TN. Apathy in neuropsychiatric disease: diagnosis, pathophysiology, and treatment. Neurotox Res 2011;19:266–78. https://doi.org/10.1007/s12640-010-9196-9.

[6] American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-5®). Washington, UNITED STATES: American Psychiatric Publishing; 2013.

[7] World Health Organization. International Classification of Diseases, Eleventh Revision (ICD-11). 2021.

[8] Gauthier S, Reisberg B, Zaudig M, Petersen RC, Ritchie K, Broich K, et al. Mild cognitive impairment. The Lancet 2006;367:1262–70. https://doi.org/10.1016/S0140-6736(06)68542-5.

[9] Wittenberg R, Bo Hu, Luis Barraza-Araiza, Amritpal Rehill. Projections of older people with dementia and costs of dementia care in the United Kingdom, 2019–2040. London School of Economics and Political Science; 2019.

[10] Finkel SI, Silva JC e, Cohen G, Miller S, Sartorius N. Behavioral and Psychological Signs and Symptoms of Dementia: A Consensus Statement on Current Knowledge and Implications for Research and Treatment. International Psychogeriatrics 1997;8:497–500. https://doi.org/10.1017/S1041610297003943.

[11] Bergh S, Selbæk G. The prevalence and the course of neuropsychiatric symptoms in patients with dementia. 1 2012;22. https://doi.org/10.5324/nje.v22i2.1570.

[12] Cipriani G, Lucetti C, Danti S, Nuti A. Apathy and dementia. Nosology, assessment and management. J Nerv Ment Dis 2014;202:718–24. https://doi.org/10.1097/NMD.0000000000000190.

[13] van der Linde RM, Dening T, Stephan BCM, Prina AM, Evans E, Brayne C. Longitudinal course of behavioural and psychological symptoms of dementia: systematic review. The British Journal of Psychiatry 2016;209:366–77. https://doi.org/10.1192/bjp.bp.114.148403.

[14] Leung DKY, Chan WC, Spector A, Wong GHY. Prevalence of depression, anxiety, and apathy symptoms across dementia stages: A systematic review and meta-analysis. Int J Geriatr Psychiatry 2021;36:1330–44. https://doi.org/10.1002/gps.5556.

[15] Sherman C, Liu CS, Herrmann N, Lanctôt KL. Prevalence, neurobiology, and treatments for apathy in prodromal dementia. International Psychogeriatrics 2018;30:177–84. https://doi.org/10.1017/S1041610217000527.

[16] Benoit M, Andrieu S, Lechowski L, Gillette-Guyonnet S, Robert PH, Vellas B, et al. Apathy and depression in Alzheimer's disease are associated with functional deficit and psychotropic prescription. Int J Geriatr Psychiatry 2008;23:409–14. https://doi.org/10.1002/gps.1895.

[17] Mukherjee A, Biswas A, Roy A, Biswas S, Gangopadhyay G, Das SK. Behavioural and Psychological Symptoms of Dementia: Correlates and

Impact on Caregiver Distress. DEE 2017;7:354–65. https://doi.org/10.1159/000481568.
[18] Rattinger GB, Sanders CL, Vernon E, Schwartz S, Behrens S, Lyketsos CG, et al. Neuropsychiatric symptoms in patients with dementia and the longitudinal costs of informal care in the Cache County population. Alzheimers Dement (N Y) 2019;5:81–8. https://doi.org/10.1016/j.trci.2019.01.002.
[19] Volicer L, Frijters DHM, van der Steen JT. Apathy and weight loss in nursing home residents: longitudinal study. J Am Med Dir Assoc 2013;14:417–20. https://doi.org/10.1016/j.jamda.2012.12.004.
[20] Willem van Dalen J, van Wanrooij LL, Moll van Charante EP, Brayne C, van Gool WA, Richard E. Association of Apathy With Risk of Incident Dementia: A Systematic Review and Meta-analysis. JAMA Psychiatry 2018;75:1012–21. https://doi.org/10.1001/jamapsychiatry.2018.1877.
[21] Yeager CA, Hyer L. Apathy in dementia: relations with depression, functional competence, and quality of life. Psychol Rep 2008;102:718–22. https://doi.org/10.2466/pr0.102.3.718-722.
[22] Landes AM, Sperry SD, Strauss ME, Geldmacher DS. Apathy in Alzheimer's disease. J Am Geriatr Soc 2001;49:1700–7.
[23] Sultzer DL. Why Apathy in Alzheimer's Matters. Am J Psychiatry 2018;175:99–100. https://doi.org/10.1176/appi.ajp.2017.17111258.
[24] Stuss DT, Van Reekum R, Murphy KJ. Differentiation of states and causes of apathy. The neuropsychology of emotion, New York, NY, US: Oxford University Press; 2000, p. 340–63.
[25] Ruthirakuhan MT, Herrmann N, Abraham EH, Chan S, Lanctôt KL. Pharmacological interventions for apathy in Alzheimer's disease. Cochrane Database of Systematic Reviews 2018. https://doi.org/10.1002/14651858.CD012197.pub2.
[26] Pickett J, Bird C, Ballard C, Banerjee S, Brayne C, Cowan K, et al. A roadmap to advance dementia research in prevention, diagnosis, intervention, and care by 2025. International Journal of Geriatric Psychiatry 2018;33:900–6. https://doi.org/10.1002/gps.4868.
[27] Cummings JL, Friedman JH, Garibaldi G, Jones M, Macfadden W, Marsh L, et al. Apathy in Neurodegenerative Diseases: Recommendations on the Design of Clinical Trials. J Geriatr Psychiatry Neurol 2015;28:159–73. https://doi.org/10.1177/0891988715573534.
[28] Nobis L, Husain M. Apathy in Alzheimer's disease. Current Opinion in Behavioral Sciences 2018;22:7–13. https://doi.org/10.1016/j.cobeha.2017.12.007.
[29] Starkstein SE, Leentjens AFG. The nosological position of apathy in clinical practice. J Neurol Neurosurg Psychiatry 2008;79:1088–92. https://doi.org/10.1136/jnnp.2007.136895.
[30] Marin RS. Differential diagnosis and classification of apathy. AJP 1990;147:22–30. https://doi.org/10.1176/ajp.147.1.22.
[31] Gilmore-Bykovskyi A, Block L, Johnson R, Goris ED. Symptoms of apathy and passivity in dementia: A simultaneous concept analysis. Journal of Clinical Nursing 2019;28:410–9. https://doi.org/10.1111/jocn.14663.
[32] Levy R, Dubois B. Apathy and the functional anatomy of the prefrontal cortex-basal ganglia circuits. Cereb Cortex 2006;16:916–28. https://doi.org/10.1093/cercor/bhj043.
[33] Robert PH, Onyike CU, Leentjens AFG, Dujardin K, Aalten P, Starkstein S, et al. Proposed diagnostic criteria for apathy in Alzheimer's disease and other neuropsychiatric disorders. Eur Psychiatry 2009;24:98–104. https://doi.org/10.1016/j.eurpsy.2008.09.001.
[34] Mulin E, Leone E, Dujardin K, Delliaux M, Leentjens AFG, Nobili F, et al. Diagnostic criteria for apathy in clinical practice. International Journal of Geriatric Psychiatry 2011;26:158–65. https://doi.org/10.1002/gps.2508.

[35] Ducharme S, Price BH, Dickerson BC. Apathy: a neurocircuitry model based on frontotemporal dementia. J Neurol Neurosurg Psychiatry 2018;89:389–96. https://doi.org/10.1136/jnnp-2017-316277.
[36] Robert PH, Lanctôt KL, Agüera-Ortiz L, Aalten P, Bremond F, Defrancesco M, et al. Is it time to revise the diagnostic criteria for apathy in brain disorders? The 2018 international consensus group. European Psychiatry 2018;54:71–6. https://doi.org/10.1016/j.eurpsy.2018.07.008.
[37] World Health Organization. The ICD-10 Classification of Mental and Behavioural Disorders: Clinical Descriptions and Diagnostic Guidelines. Geneva, Switzerland: World Health Organization; 1992.
[38] Marin RS. Apathy: a neuropsychiatric syndrome. JNP 1991;3:243–54. https://doi.org/10.1176/jnp.3.3.243.
[39] Miller DS, Robert P, Ereshefsky L, Adler L, Bateman D, Cummings J, et al. Diagnostic criteria for apathy in neurocognitive disorders. Alzheimer's & Dementia 2021;17:1892–904. https://doi.org/10.1002/alz.12358.
[40] Dickson SS, Husain M. Are there distinct dimensions of apathy? The argument for reappraisal. Cortex 2022;149:246–56. https://doi.org/10.1016/j.cortex.2022.01.001.
[41] Manera V, Abrahams S, Agüera-Ortiz L, Bremond F, David R, Fairchild K, et al. Recommendations for the non-pharmacological treatment of apathy in brain disorders. The American Journal of Geriatric Psychiatry 2019. https://doi.org/10.1016/j.jagp.2019.07.014.
[42] Benoit M, Berrut G, Doussaint J, Bakchine S, Bonin-Guillaume S, Frémont P, et al. Apathy and depression in mild Alzheimer's disease: a cross-sectional study using diagnostic criteria. J Alzheimers Dis 2012;31:325–34. https://doi.org/10.3233/JAD-2012-112003.
[43] Robert PH, Berr C, Volteau M, Bertogliati-Fileau C, Benoit M, Guerin O, et al. Importance of lack of interest in patients with mild cognitive impairment. Am J Geriatr Psychiatry 2008;16:770–6. https://doi.org/10.1097/JGP.0b013e31817e73db.
[44] Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. Psychiatry Res 1991;38:143–62.
[45] Radakovic R, Starr JM, Abrahams S. A Novel Assessment and Profiling of Multidimensional Apathy in Alzheimer's Disease. J Alzheimers Dis 2017;60:57–67. https://doi.org/10.3233/JAD-170292.
[46] Clarke DE, Reekum R van, Simard M, Streiner DL, Freedman M, Conn D. Apathy in dementia: an examination of the psychometric properties of the apathy evaluation scale. J Neuropsychiatry Clin Neurosci 2007;19:57–64. https://doi.org/10.1176/jnp.2007.19.1.57.
[47] Johansson M, Johansson P, Stomrud E, Hagell P, Hansson O. Psychometric testing of a Swedish version of the Apathy Evaluation Scale. Nord J Psychiatry 2017;71:477–84. https://doi.org/10.1080/08039488.2017.1334820.
[48] Lee S-H, Wen M-C, Chao C-C, Chen YJ, Yen C-F. Apathy in late-life depression among Taiwanese patients. Int Psychogeriatr 2008;20:328–37. https://doi.org/10.1017/S1041610207005698.
[49] Pedersen KF, Alves G, Larsen JP, Tysnes O-B, Møller SG, Brønnick K. Psychometric properties of the Starkstein Apathy Scale in patients with early untreated Parkinson disease. Am J Geriatr Psychiatry 2012;20:142–8. https://doi.org/10.1097/JGP.0b013e31823038f2.
[50] Lueken U, Evens R, Balzer-Geldsetzer M, Baudrexel S, Dodel R, Gräber-Sultan S, et al. Psychometric properties of the apathy evaluation scale in patients with Parkinson's disease. Int J Methods Psychiatr Res 2017;26. https://doi.org/10.1002/mpr.1564.
[51] van der Eijk C, Rose J. Risky Business: Factor Analysis of Survey Data – Assessing the Probability of Incorrect Dimensionalisation. PLOS ONE 2015;10:e0118900. https://doi.org/10.1371/journal.pone.0118900.

[52] Holgado–Tello FP, Chacón–Moscoso S, Barbero–García I, Vila–Abad E. Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. Qual Quant 2008;44:153. https://doi.org/10.1007/s11135-008-9190-y.
[53] Starkstein SE, Petracca G, Chemerinski E, Kremer J. Syndromic validity of apathy in Alzheimer's disease. Am J Psychiatry 2001;158:872–7. https://doi.org/10.1176/appi.ajp.158.6.872.
[54] Treusch Y, Page J, Niemann-Mirmehdi M, Gutzmann H, Heinz A, Rapp M. Apathy and its nonpharmacological treatment in dementia. GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry 2011;24:155–64. https://doi.org/10.1024/1662-9647/a000041.
[55] Marin RS, Wilkosz PA. Disorders of diminished motivation. J Head Trauma Rehabil 2005;20:377–88.
[56] Starkstein SE, Sabe L, Chemerinski E, Jason L, Leiguarda R. Two domains of anosognosia in Alzheimer's disease. J Neurol Neurosurg Psychiatry 1996;61:485–90. https://doi.org/10.1136/jnnp.61.5.485.
[57] Mograbi DC, Morris RG. On the Relation among Mood, Apathy, and Anosognosia in Alzheimer's Disease. Journal of the International Neuropsychological Society 2014;20:2–7. https://doi.org/10.1017/S1355617713001276.
[58] Mak E, Chin R, Ng LT, Yeo D, Hameed S. Clinical associations of anosognosia in mild cognitive impairment and Alzheimer's disease. Int J Geriatr Psychiatry 2015;30:1207–14. https://doi.org/10.1002/gps.4275.
[59] Sockeel P, Dujardin K, Devos D, Denève C, Destée A, Defebvre L. The Lille apathy rating scale (LARS), a new instrument for detecting and quantifying apathy: validation in Parkinson's disease. J Neurol Neurosurg Psychiatry 2006;77:579–84. https://doi.org/10.1136/jnnp.2005.075929.
[60] Radakovic R, Abrahams S. Multidimensional apathy: evidence from neurodegenerative disease. Current Opinion in Behavioral Sciences 2018;22:42–9. https://doi.org/10.1016/j.cobeha.2017.12.022.
[61] Starkstein SE, Brockman S, Bruce D, Petracca G. Anosognosia Is a Significant Predictor of Apathy in Alzheimer's Disease. JNP 2010;22:378–83. https://doi.org/10.1176/jnp.2010.22.4.378.
[62] Apps MAJ, Rushworth MFS, Chang SWC. The Anterior Cingulate Gyrus and Social Cognition: Tracking the Motivation of Others. Neuron 2016;90:692–707. https://doi.org/10.1016/j.neuron.2016.04.018.
[63] Chang SWC, Fagan NA, Toda K, Utevsky AV, Pearson JM, Platt ML. Neural mechanisms of social decision-making in the primate amygdala. PNAS 2015;112:16012–7. https://doi.org/10.1073/pnas.1514761112.
[64] Lockwood PL, Hamonet M, Zhang SH, Ratnavel A, Salmony FU, Husain M, et al. Prosocial apathy for helping others when effort is required. Nature Human Behaviour 2017;1:0131. https://doi.org/10.1038/s41562-017-0131.
[65] Ang Y-S, Lockwood P, Apps MAJ, Muhammed K, Husain M. Distinct Subtypes of Apathy Revealed by the Apathy Motivation Index. PLOS ONE 2017;12:1–15. https://doi.org/10.1371/journal.pone.0169938.
[66] M'Barek L, Radakovic R, Noquet M, Laurent A, Allain P. Different aspects of emotional processes in apathy: Application of the French translated dimensional apathy scale. Curr Psychol 2020;39:564–70. https://doi.org/10.1007/s12144-017-9775-5.
[67] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equation Modeling: A Multidisciplinary Journal 1999;6:1–55. https://doi.org/10.1080/10705519909540118.
[68] Santangelo G, Barone P, Cuoco S, Raimo S, Pezzella D, Picillo M, et al. Apathy in untreated, de novo patients with Parkinson's disease: validation

study of Apathy Evaluation Scale. J Neurol 2014;261:2319–28. https://doi.org/10.1007/s00415-014-7498-1.
[69] Esposito F, Rochat L, Juillerat Van der Linden A-C, Lekeu F, Charnallet A, Van der Linden M. Apathy in aging: Are lack of interest and lack of initiative dissociable? Archives of Gerontology and Geriatrics 2014;58:43–50. https://doi.org/10.1016/j.archger.2013.09.002.
[70] Robert P, Manera V. Definition of Apathy and Differential Diagnosis. Apathy, Oxford University Press; 2021, p. 1–18. https://doi.org/10.1093/med/9780198841807.003.0001.
[71] Lanctôt K, Aleman A. Apathy: Clinical and Neuroscientific Perspectives from Neurology and Psychiatry. 1st ed. Oxford: Oxford University Press; 2021.
[72] Mortby ME, Maercker A, Forstmeier S. Apathy: a separate syndrome from depression in dementia? A critical review. Aging Clin Exp Res 2012;24:305–16. https://doi.org/10.3275/8105.
[73] Watson R, Harvey K, McCabe C, Reynolds S. Understanding anhedonia: a qualitative study exploring loss of interest and pleasure in adolescent depression. Eur Child Adolesc Psychiatry 2020;29:489–99. https://doi.org/10.1007/s00787-019-01364-y.
[74] Groeneweg-Koolhoven I, Comijs HC, Naarding P, de Waal MW, van der Mast RC. Presence and correlates of apathy in non-demented depressed and non-depressed older persons. The European Journal of Psychiatry 2015;29:119–30. https://doi.org/10.4321/S0213-61632015000200003.
[75] Landes AM, Sperry SD, Strauss ME. Prevalence of Apathy, Dysphoria, and Depression in Relation to Dementia Severity in Alzheimer's Disease. JNP 2005;17:342–9. https://doi.org/10.1176/jnp.17.3.342.
[76] Starkstein SE, Jorge R, Mizrahi R, Robinson RG. A prospective longitudinal study of apathy in Alzheimer's disease. J Neurol Neurosurg Psychiatr 2006;77:8–11. https://doi.org/10.1136/jnnp.2005.069575.
[77] Zahodne LB, Tremont G. Unique effects of apathy and depression signs on cognition and function in amnestic mild cognitive impairment. International Journal of Geriatric Psychiatry 2013;28:50–6. https://doi.org/10.1002/gps.3789.
[78] Starkstein SE, Migliorelli R, Manes F, Tesón A, Petracca G, Chemerinski E, et al. The prevalence and clinical correlates of apathy and irritability in Alzheimer's disease. Eur J Neurol 1995;2:540–6. https://doi.org/10.1111/j.1468-1331.1995.tb00171.x.
[79] Siafarikas N, Selbaek G, Fladby T, Benth JŠ, Auning E, Aarsland D. Frequency and subgroups of neuropsychiatric symptoms in mild cognitive impairment and different stages of dementia in Alzheimer's disease. International Psychogeriatrics 2018;30:103–13. https://doi.org/10.1017/S1041610217001879.
[80] Tagariello P, Girardi P, Amore M. Depression and apathy in dementia: Same syndrome or different constructs? A critical review. Archives of Gerontology and Geriatrics 2009;49:246–9. https://doi.org/10.1016/j.archger.2008.09.002.
[81] Dulawa SC, Janowsky DS. Cholinergic Regulation of Mood: From Basic and Clinical Studies to Emerging Therapeutics. Mol Psychiatry 2019;24:694–709. https://doi.org/10.1038/s41380-018-0219-x.
[82] Mineur YS, Obayemi A, Wigestrand MB, Fote GM, Calarco CA, Li AM, et al. Cholinergic signaling in the hippocampus regulates social stress resilience and anxiety- and depression-like behavior. Proceedings of the National Academy of Sciences 2013;110:3573–8. https://doi.org/10.1073/pnas.1219731110.
[83] Levy R. Apathy: A pathology of goal-directed behaviour. A new concept of the clinic and pathophysiology of apathy. Revue Neurologique 2012;168:585–97. https://doi.org/10.1016/j.neurol.2012.05.003.

[84] Kaiser S, Schlagenhauf F. Brain Reward Systems and Apathy. In: Lanctot K, Aleman A, editors. Apathy: Clinical and Neuroscientific Perspectives from Neurology and Psychiatry, Oxford University Press; 2021, p. 0. https://doi.org/10.1093/med/9780198841807.003.0011.
[85] Kolanowski A, Boltz M, Galik E, Gitlin LN, Kales HC, Resnick B, et al. Determinants of behavioral and psychological symptoms of dementia: A scoping review of the evidence. Nursing Outlook 2017;65:515–29. https://doi.org/10.1016/j.outlook.2017.06.006.
[86] Borsje P, Wetzels RB, Lucassen PL, Pot AM, Koopmans RT. The course of neuropsychiatric symptoms in community-dwelling patients with dementia: a systematic review. Int Psychogeriatr 2015;27:385–405. https://doi.org/10.1017/S1041610214002282.
[87] Palmer K, Di Iulio F, Varsi AE, Gianni W, Sancesario G, Caltagirone C, et al. Neuropsychiatric predictors of progression from amnestic-mild cognitive impairment to Alzheimer's disease: the role of depression and apathy. J Alzheimers Dis 2010;20:175–83. https://doi.org/10.3233/JAD-2010-1352.
[88] Mourao RJ, Mansur G, Malloy-Diniz LF, Costa EC, Diniz BS. Depressive symptoms increase the risk of progression to dementia in subjects with mild cognitive impairment: systematic review and meta-analysis. International Journal of Geriatric Psychiatry 2016;31:905–11. https://doi.org/10.1002/gps.4406.
[89] Vicini Chilovi B, Conti M, Zanetti M, Mazzù I, Rozzini L, Padovani A. Differential impact of apathy and depression in the development of dementia in mild cognitive impairment patients. Dement Geriatr Cogn Disord 2009;27:390–8. https://doi.org/10.1159/000210045.
[90] Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology 1994;44:2308–14.
[91] Suchy Y. Executive functioning: overview, assessment, and research issues for non-neuropsychologists. Ann Behav Med 2009;37:106–16. https://doi.org/10.1007/s12160-009-9097-4.
[92] Brown RG, Pluck G. Negative symptoms: the 'pathology' of motivation and goal-directed behaviour. Trends in Neurosciences 2000;23:412–7. https://doi.org/10.1016/S0166-2236(00)01626-X.
[93] Levy ML, Cummings JL, Fairbanks LA, Masterman D, Miller BL, Craig AH, et al. Apathy Is Not Depression. JNP 1998;10:314–9. https://doi.org/10.1176/jnp.10.3.314.
[94] Starkstein SE, Ingram L, Garau ML, Mizrahi R. On the overlap between apathy and depression in dementia. J Neurol Neurosurg Psychiatr 2005;76:1070–4. https://doi.org/10.1136/jnnp.2004.052795.
[95] Lü W, Duan J, Zhang W, Yang W, Yu W. Relationship between neuropsychiatric symptoms and cognitive functions in patients with cognitive impairment. Psychogeriatrics 2021;21:773–82. https://doi.org/10.1111/psyg.12738.
[96] You SC, Walsh CM, Chiodo LA, Ketelle R, Miller BL, Kramer JH. Neuropsychiatric Symptoms Predict Functional Status in Alzheimer's Disease. J Alzheimers Dis 2015;48:863–9. https://doi.org/10.3233/JAD-150018.
[97] Gélinas I, Gauthier L, McIntyre M, Gauthier S. Development of a functional measure for persons with Alzheimer's disease: the disability assessment for dementia. Am J Occup Ther 1999;53:471–81.
[98] Lam LCW, Tam CWC, Chiu HFK, Lui VWC. Depression and apathy affect functioning in community active subjects with questionable dementia and mild Alzheimer's disease. Int J Geriatr Psychiatry 2007;22:431–7. https://doi.org/10.1002/gps.1694.
[99] Olin JT, Schneider LS, Katz IR, Meyers BS, Alexopoulos GS, Breitner JC, et al. Provisional diagnostic criteria for depression of Alzheimer's disease:

description and review. Expert Review of Neurotherapeutics 2003;3:99–106. https://doi.org/10.1586/14737175.3.1.99.
[100] Gallo JJ, Rabins PV. Depression without sadness: alternative presentations of depression in late life. Am Fam Physician 1999;60:820–6.
[101] Paradiso S, Duff K, Vaidya JG, Hoth A, Mold JW. Cognitive and daily functioning in older adults with vegetative symptoms of depression. Int J Geriatr Psychiatry 2010;25:569–77. https://doi.org/10.1002/gps.2376.
[102] Hybels CF, Landerman LR, Blazer DG. Age differences in symptom expression in patients with major depression. International Journal of Geriatric Psychiatry 2012;27:601–11. https://doi.org/10.1002/gps.2759.
[103] Forsell Y, Jorm AF, Fratiglioni L, Grut M, Winblad B. Application of DSM-III—R criteria for major depressive episode to elderly subjects with and without dementia. The American Journal of Psychiatry 1993;150:1199–202. https://doi.org/10.1176/ajp.150.8.1199.
[104] Forsell Y, Winblad B. Major Depression in a Population of Demented and Nondemented Older People: Prevalence and Correlates. Journal of the American Geriatrics Society 1998;46:27–30. https://doi.org/10.1111/j.1532-5415.1998.tb01009.x.
[105] Cheng S-T, Chan ACM. Withdrawal, apathy and lack of vigor in late life depression: factorial validity and relationship to diagnosis. Aging Ment Health 2007;11:532–7. https://doi.org/10.1080/13607860601086447.
[106] Lee JR, Suh SW, Han JW, Byun S, Kwon SJ, Lee KH, et al. Anhedonia and Dysphoria Are Differentially Associated with the Risk of Dementia in the Cognitively Normal Elderly Individuals: A Prospective Cohort Study. Psychiatry Investig 2019;16:575–80. https://doi.org/10.30773/pi.2019.06.07.
[107] Levy R. Apathy: A Pathology of Goal-Directed Behaviour and Prefrontal Cortex–Basal Ganglia Circuits. Apathy, Oxford University Press; 2021, p. 156–73. https://doi.org/10.1093/med/9780198841807.003.0009.
[108] Geda YE, Roberts RO, Knopman DS, Petersen RC, Christianson TJH, Pankratz VS, et al. Prevalence of neuropsychiatric symptoms in mild cognitive impairment and normal cognitive aging: population-based study. Archives of General Psychiatry 2008;65:1193–8. https://doi.org/10.1001/archpsyc.65.10.1193.
[109] Peters ME, Rosenberg PB, Steinberg M, Tschanz JT, Norton MC, Welsh-Bohmer KA, et al. Prevalence of Neuropsychiatric Symptoms in CIND and Its Subtypes: The Cache County Study. The American Journal of Geriatric Psychiatry 2012;20:416–24. https://doi.org/10.1097/JGP.0b013e318211057d.
[110] Sheikh F, Ismail Z, Mortby ME, Barber P, Cieslak A, Fischer K, et al. Prevalence of mild behavioral impairment in mild cognitive impairment and subjective cognitive decline, and its association with caregiver burden. International Psychogeriatrics 2018;30:233–44. https://doi.org/10.1017/S104161021700151X.
[111] Aleman A. Neural Basis of Apathy: Functional Imaging Studies. In: Lanctot K, Aleman A, editors. Apathy: Clinical and Neuroscientific Perspectives from Neurology and Psychiatry, Oxford University Press; 2021, p. 0. https://doi.org/10.1093/med/9780198841807.003.0012.
[112] Agartz I, Mørch-Johnsen L. Neural Basis of Apathy: Structural Imaging Studies. In: Lanctot K, Aleman A, editors. Apathy: Clinical and Neuroscientific Perspectives from Neurology and Psychiatry, Oxford University Press; 2021, p. 0. https://doi.org/10.1093/med/9780198841807.003.0010.
[113] Le Heron C, Apps. MAJ, Husain M. The anatomy of apathy: A neurocognitive framework for amotivated behaviour. Neuropsychologia 2018;118:54–67. https://doi.org/10.1016/j.neuropsychologia.2017.07.003.

[114] Stella F, Radanovic M, Aprahamian I, Canineu PR, de Andrade LP, Forlenza OV. Neurobiological correlates of apathy in Alzheimer's disease and mild cognitive impairment: a critical review. J Alzheimers Dis 2014;39:633–48. https://doi.org/10.3233/JAD-131385.
[115] Forstmeier S, Maercker A. Motivational processes in mild cognitive impairment and Alzheimer's disease: results from the Motivational Reserve in Alzheimer's (MoReA) study. BMC Psychiatry 2015;15:293. https://doi.org/10.1186/s12888-015-0666-8.
[116] Forstmeier S, Maercker A. Motivational reserve: lifetime motivational abilities contribute to cognitive and emotional health in old age. Psychol Aging 2008;23:886–99. https://doi.org/10.1037/a0013602.
[117] Starkstein SE. Anosognosia in Alzheimer's disease: Diagnosis, frequency, mechanism and clinical correlates. Cortex 2014;61:64–73. https://doi.org/10.1016/j.cortex.2014.07.019.
[118] Perri R, Turchetta CS, Caruso G, Fadda L, Caltagirone C, Carlesimo GA. Neuropsychological correlates of cognitive, emotional-affective and auto-activation apathy in Alzheimer's disease. Neuropsychologia 2018;118:12–21. https://doi.org/10.1016/j.neuropsychologia.2018.01.039.
[119] Radakovic R, Abrahams S. Developing a new apathy measurement scale: Dimensional Apathy Scale. Psychiatry Research 2014;219:658–63. https://doi.org/10.1016/j.psychres.2014.06.010.
[120] Husain M, Roiser JP. Neuroscience of apathy and anhedonia: a transdiagnostic approach. Nature Reviews Neuroscience 2018;19:470–84. https://doi.org/10.1038/s41583-018-0029-9.
[121] Sinha N, Manohar S, Husain M. Impulsivity and apathy in Parkinson's disease. J Neuropsychol 2013;7:255–83. https://doi.org/10.1111/jnp.12013.
[122] Le Heron C, Plant O, Manohar S, Ang Y-S, Jackson M, Lennox G, et al. Distinct effects of apathy and dopamine on effort-based decision-making in Parkinson's disease. Brain: A Journal of Neurology 2018;141:1455–69. https://doi.org/10.1093/brain/awy110.
[123] Gonçalves S de AB, Caramelli P, Mariano LI, Guimarães HC, Gambogi LB, Resende E de PF, et al. Apathy in frontotemporal dementia is related to medial prefrontal atrophy and is independent of executive dysfunction. Brain Research 2020;1737:146799. https://doi.org/10.1016/j.brainres.2020.146799.
[124] Hartmann MN, Hager OM, Reimann AV, Chumbley JR, Kirschner M, Seifritz E, et al. Apathy But Not Diminished Expression in Schizophrenia Is Associated With Discounting of Monetary Rewards by Physical Effort. Schizophr Bull 2015;41:503–12. https://doi.org/10.1093/schbul/sbu102.
[125] Raffard S, Bortolon C, Burca M, Gely-Nargeot M-C, Capdevielle D. Multidimensional model of apathy in older adults using partial least squares--path modeling. Age (Dordr) 2016;38:55. https://doi.org/10.1007/s11357-016-9916-z.
[126] Esposito F, Gendolla GHE, Van der Linden M. Are self-efficacy beliefs and subjective task demand related to apathy in aging? Aging Ment Health 2014;18:521–30. https://doi.org/10.1080/13607863.2013.856865.
[127] Bonnelle V, Veromann K-R, Burnett Heyes S, Lo Sterzo E, Manohar S, Husain M. Characterization of reward and effort mechanisms in apathy. J Physiol Paris 2015;109:16–26. https://doi.org/10.1016/j.jphysparis.2014.04.002.
[128] Le Heron C, Manohar S, Plant O, Muhammed K, Griffanti L, Nemeth A, et al. Dysfunctional effort-based decision-making underlies apathy in genetic cerebral small vessel disease. Brain 2018;141:3193–210. https://doi.org/10.1093/brain/awy257.
[129] Bayard S, Jacus J-P, Raffard S, Gely-Nargeot M-C. Apathy and Emotion-Based Decision-Making in Amnesic Mild Cognitive Impairment and

Alzheimer's Disease. Behavioural Neurology 2014;2014:e231469. https://doi.org/10.1155/2014/231469.
[130] Martínez-Horta S, Riba J, Bobadilla RF de, Pagonabarraga J, Pascual-Sedano B, Antonijoan RM, et al. Apathy in Parkinson's Disease: Neurophysiological Evidence of Impaired Incentive Processing. J Neurosci 2014;34:5918–26. https://doi.org/10.1523/JNEUROSCI.0251-14.2014.
[131] Muhammed K, Manohar S, Ben Yehuda M, Chong TT-J, Tofaris G, Lennox G, et al. Reward sensitivity deficits modulated by dopamine are associated with apathy in Parkinson's disease. Brain 2016;139:2706–21. https://doi.org/10.1093/brain/aww188.
[132] McLauchlan DJ, Lancaster T, Craufurd D, Linden DEJ, Rosser AE. Insensitivity to loss predicts apathy in huntington's disease. Mov Disord 2019;34:1381–91. https://doi.org/10.1002/mds.27787.
[133] McPherson S, Fairbanks L, Tiken S, Cummings JL, Back-Madruga C. Apathy and executive function in Alzheimer's disease. J Int Neuropsychol Soc 2002;8:373–81.
[134] Drijgers RL, Verhey FRJ, Leentjens AFG, Köhler S, Aalten P. Neuropsychological correlates of apathy in mild cognitive impairment and Alzheimer's disease: the role of executive functioning. Int Psychogeriatr 2011;23:1327–33. https://doi.org/10.1017/S1041610211001037.
[135] Chau SA, Chung J, Herrmann N, Eizenman M, Lanctôt KL. Apathy and Attentional Biases in Alzheimer's Disease. Journal of Alzheimer's Disease 2016;51:837–46. https://doi.org/10.3233/JAD-151026.
[136] Esposito F, Rochat L, Van der Linden A-CJ, Lekeu F, Quittre A, Charnallet A, et al. Apathy and executive dysfunction in Alzheimer disease. Alzheimer Dis Assoc Disord 2010;24:131–7. https://doi.org/10.1097/WAD.0b013e3181c9c168.
[137] Giebel CM, Sutcliffe C, Challis D. Hierarchical Decline of the Initiative and Performance of Complex Activities of Daily Living in Dementia. J Geriatr Psychiatry Neurol 2017;30:96–103. https://doi.org/10.1177/0891988716686835.
[138] Esposito F, Rochat L, Linden A-CJV der, Linden MV der. Apathy and Prospective Memory in Aging. DEE 2012;2:456–67. https://doi.org/10.1159/000345037.
[139] World Health Organization. Towards a Common Language for Functioning, Disability and Health: ICF The International Classification of Functioning, Disability and Health. Geneva, Switzerland: 2002.
[140] Mlinac ME, Feng MC. Assessment of Activities of Daily Living, Self-Care, and Independence. Arch Clin Neuropsychol 2016;31:506–16. https://doi.org/10.1093/arclin/acw049.
[141] Boyle PA, Malloy PF, Salloway S, Cahn-Weiner DA, Cohen R, Cummings JL. Executive dysfunction and apathy predict functional impairment in Alzheimer disease. Am J Geriatr Psychiatry 2003;11:214–21.
[142] Clarke DE, van Reekum R, Simard M, Streiner DL, Conn D, Cohen T, et al. Apathy in dementia: clinical and sociodemographic correlates. J Neuropsychiatry Clin Neurosci 2008;20:337–47. https://doi.org/10.1176/jnp.2008.20.3.337.
[143] Norton LE, Malloy PF, Salloway S. The impact of behavioral symptoms on activities of daily living in patients with dementia. Am J Geriatr Psychiatry 2001;9:41–8.
[144] Strauss ME, Sperry SD. An informant-based assessment of apathy in Alzheimer disease. Neuropsychiatry Neuropsychol Behav Neurol 2002;15:176–83.
[145] Brodaty H, Altendorf A, Withall A, Sachdev P. Do people become more apathetic as they grow older? A longitudinal study in healthy individuals. Int Psychogeriatr 2010;22:426–36. https://doi.org/10.1017/S1041610209991335.

[146] Simpson J, McMillan H, Leroi I, Murray CD. Experiences of apathy in people with Parkinson's disease: a qualitative exploration. Disabil Rehabil 2015;37:611–9. https://doi.org/10.3109/09638288.2014.939771.
[147] Cook C, Fay S, Rockwood K. Decreased initiation of usual activities in people with mild-to-moderate Alzheimer's disease: a descriptive analysis from the VISTA clinical trial. Int Psychogeriatr 2008;20:952–63. https://doi.org/10.1017/S1041610208007230.
[148] Lechowski L, Benoit M, Chassagne P, Vedel I, Tortrat D, Teillet L, et al. Persistent apathy in Alzheimer's disease as an independent factor of rapid functional decline: the REAL longitudinal cohort study. Int J Geriatr Psychiatry 2009;24:341–6. https://doi.org/10.1002/gps.2125.
[149] Kiely KM, Mortby ME, Anstey KJ. Differential associations between sensory loss and neuropsychiatric symptoms in adults with and without a neurocognitive disorder. International Psychogeriatrics 2018;30:261–72. https://doi.org/10.1017/S1041610217001120.
[150] Nicholson C, Meyer J, Flatley M, Holman C. The experience of living at home with frailty in old age: A psychosocial qualitative study. International Journal of Nursing Studies 2013;50:1172–9. https://doi.org/10.1016/j.ijnurstu.2012.01.006.
[151] Parrotta I, Maltais M, Rolland Y, Spampinato DA, Robert P, de Souto Barreto P, et al. The association between apathy and frailty in older adults: a new investigation using data from the Mapt study. Aging & Mental Health 2020;24:1985–9. https://doi.org/10.1080/13607863.2019.1650890.
[152] Maruta M, Shimokihara S, Makizako H, Ikeda Y, Han G, Akasaki Y, et al. Associations between apathy and comprehensive frailty as assessed by the Kihon Checklist among community-dwelling Japanese older adults. Psychogeriatrics 2022;22:651–8. https://doi.org/10.1111/psyg.12867.
[153] Massimo L, Kales HC, Kolanowski A. State of the Science: Apathy As a Model for Investigating Behavioral and Psychological Symptoms in Dementia. Journal of the American Geriatrics Society 2018;66:S4–12. https://doi.org/10.1111/jgs.15343.
[154] Cohen-Mansfield J. Theoretical Frameworks for Behavioral Problems in Dementia. Alzheimer's Care Today 2000;1:8–21.
[155] Butcher I, Berry K, Haddock G. Understanding individuals' subjective experiences of negative symptoms of schizophrenia: A qualitative study. British Journal of Clinical Psychology 2020;59:319–34. https://doi.org/10.1111/bjc.12248.
[156] Trahan MA, Kuo J, Carlson MC, Gitlin LN. A Systematic Review of Strategies to Foster Activity Engagement in Persons With Dementia. Health Educ Behav 2014;41:70S-83S. https://doi.org/10.1177/1090198114531782.
[157] Kales HC, Gitlin LN, Lyketsos CG. Assessment and management of behavioral and psychological symptoms of dementia. BMJ 2015;350. https://doi.org/10.1136/bmj.h369.
[158] Smith M, Gerdner LA, Hall GR, Buckwalter KC. History, Development, and Future of the Progressively Lowered Stress Threshold: A Conceptual Model for Dementia Care. Journal of the American Geriatrics Society 2004;52:1755–60. https://doi.org/10.1111/j.1532-5415.2004.52473.x.
[159] Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R, Marin RS. Are the available apathy measures reliable and valid? A review of the psychometric evidence. Journal of Psychosomatic Research 2011;70:73–97.
[160] Burgon C, Goldberg SE, Van Der Wardt V, Brewin C, Harwood RH. Apathy Measures in Older Adults and People with Dementia: A Systematic Review of Measurement Properties Using the COSMIN Methodology. Dementia and Geriatric Cognitive Disorders 2021;50:111–23.
[161] Radakovic R, Harley C, Abrahams S, Starr JM. A systematic review of the validity and reliability of apathy scales in neurodegenerative conditions. Int

Psychogeriatr 2015;27:903–23. https://doi.org/10.1017/S1041610214002221.
[162] Weiser M, Garibaldi G. Quantifying motivational deficits and apathy: A review of the literature. European Neuropsychopharmacology 2015;25:1060–81. https://doi.org/10.1016/j.euroneuro.2014.08.018.
[163] Mohammad D, Ellis C, Rau A, Rosenberg PB, Mintzer J, Ruthirakuhan M, et al. Psychometric Properties of Apathy Scales in Dementia: A Systematic Review. J Alzheimers Dis 2018;66:1065–82. https://doi.org/10.3233/JAD-180485.
[164] Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Medical Research Methodology 2003;3:25. https://doi.org/10.1186/1471-2288-3-25.
[165] de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: A Practical Guide. Cambridge: Cambridge University Press; 2011. https://doi.org/10.1017/CBO9780511996214.
[166] Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research 2018. https://doi.org/10.1007/s11136-018-1798-3.
[167] Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation 2018. https://doi.org/10.1007/s11136-018-1829-0.
[168] Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. Trials 2016;17:449. https://doi.org/10.1186/s13063-016-1555-2.
[169] Burgon C, Goldberg S, van der Wardt V, Harwood RH. A systematic review of measures of apathy for older adults: validity, reliability and conceptualisation of apathy. PrAISED Discussion Paper Series 2019.
[170] Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of Life Research 2009;18:1115–23. https://doi.org/10.1007/s11136-009-9528-5.
[171] Bramer W, Bain P. Updating search strategies for systematic reviews using EndNote. J Med Libr Assoc 2017;105:285–9. https://doi.org/10.5195/jmla.2017.183.
[172] Boland A, Cherry MG, Dickson R. Doing a systematic review : a student's guide. 2nd ed. 2017.
[173] Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation 2017. https://doi.org/10.1007/s11136-017-1765-4.
[174] Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. Journal of Clinical Epidemiology 2007;60:34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012.
[175] Tappen RM, Williams CL. Development and Testing of the Alzheimer's Disease and Related Dementias Mood Scale. Nurs Res 2008;57:426–35. https://doi.org/10.1097/NNR.0b013e31818c3dcc.
[176] Tappen RM, Williams C. Attribution of emotion in advanced Alzheimer's disease: Family and caregiver perspectives. American Journal of Alzheimer's Disease 1998;13:257–64. https://doi.org/10.1177/153331759801300506.

[177] Greene JG, Smith R, Gardiner M, Timbury GC. Measuring behavioural disturbance of elderly demented patients in the community and its effects on relatives: a factor analytic study. Age Ageing 1982;11:121–6. https://doi.org/10.1093/ageing/11.2.121.
[178] Devanand DP, Brockington CD, Moody BJ, Brown RP, Mayeux R, Endicott J, et al. Behavioral syndromes in Alzheimer's disease. Int Psychogeriatr 1992;4 Suppl 2:161–84.
[179] Burgess PW, Alderman N, Wilson BA, Evans JJ, Emslie H. The Dysexecutive Questionnaire. Behavioural assessment of the dysexecutive syndrome., Bury St Edmunds, UK: Thames Valley Test Company; 1996.
[180] Grace J, Malloy PF. Frontal Systems Behavior Scale (FrSBe): Professional manual. Lutz, FL: Psychological Assessment Resources; 2001.
[181] Brink TL, Yesavage JA, Lum O, Heersema PH, Adey M, Rose TL. Screening tests for geriatric depression. Clinical Gerontologist: The Journal of Aging and Mental Health 1982;1:37–43. https://doi.org/10.1300/J018v01n01_06.
[182] Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: A preliminary report. Journal of Psychiatric Research 1982;17:37–49. https://doi.org/10.1016/0022-3956(82)90033-4.
[183] Verstraten P, Van Eekelen C. Handleiding voor de GIP: Gedragsobservatieschaal voor de Intramurale Psychogeriatrie. Deventer: Van Loghum Slaterus; 1987.
[184] Cucinotta D, Ambrosoli L, Poli A, Martorelli M, Savorani GC, Anzivino F. Clinical assessment of mental decline in elderly people: A proposal for a new quantitative index. Aging Clin Exp Res 1995;7:29–34. https://doi.org/10.1007/BF03324289.
[185] Kolitz BP, Vanderploeg RD, Curtiss G. Development of the Key Behaviors Change Inventory: a traumatic brain injury behavioral outcome assessment instrument. Arch Phys Med Rehabil 2003;84:277–84. https://doi.org/10.1053/apmr.2003.50100.
[186] Fahn S, Elton RL, Members of the UPDRS Development Committee. Unified Parkinson's Disease Rating Scale. Recent Developments in Parkinson's Disease Volume II, Macmillan Healthcare Information; 1987, p. 153–63.
[187] Robert PH, Clairet S, Benoit M, Koutaich J, Bertogliati C, Tible O, et al. The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. Int J Geriatr Psychiatry 2002;17:1099–105. https://doi.org/10.1002/gps.755.
[188] Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R, Robinson RG. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. J Neuropsychiatry Clin Neurosci 1992;4:134–9. https://doi.org/10.1176/jnp.4.2.134.
[189] Guercio BJ, Donovan NJ, Munro CE, Aghjayan SL, Wigman SE, Locascio JJ, et al. The Apathy Evaluation Scale: A Comparison of Subject, Informant, and Clinician Report in Cognitively Normal Elderly and Mild Cognitive Impairment. J Alzheimers Dis 2015;47:421–32. https://doi.org/10.3233/JAD-150146.
[190] Köller L, Knebel M, Haberstroh J, Krause K, Sahlender S, Jakob M, et al. Apathy in dementia care: An investigation in community-dwelling persons using the German informant version of the Apathy Evaluation Scale. GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry 2016;29:129–39. https://doi.org/10.1024/1662-9647/a000150.
[191] Stankevich Y, Lueken U, Balzer-Geldsetzer M, Dodel R, Gräber-Sultan S, Berg D, et al. Psychometric Properties of an Abbreviated Version of the Apathy Evaluation Scale for Parkinson Disease (AES-12PD). The American Journal of Geriatric Psychiatry 2018;26:1079–90. https://doi.org/10.1016/j.jagp.2018.06.012.

[192] Valentino V, Iavarone A, Amboni M, Moschiano F, Picillo M, Petretta V, et al. Apathy in Parkinson's disease: differences between caregiver's report and self-evaluation. Funct Neurol 2018;33:31–5. https://doi.org/10.11138/fneur/2018.33.1.031.
[193] Stella F, Andrade LP de, Garuffi M, Vital TM, Hernández SSS, Ruocco M, et al. Validation of the Brazilian version of the Apathy Inventory. International Journal of Geriatric Psychiatry 2013;28:979–86. https://doi.org/10.1002/gps.3917.
[194] Leone E, Lafont V, Filleau C, Baudu C, Benoit M, Deudon A, et al. Validation de la version soignant de l'Inventaire Apathie. Revue Geriatr 2008;33:777–85.
[195] Ang Y-S, Lockwood PL, Kienast A, Plant O, Drew D, Slavkova E, et al. Differential impact of behavioral, social, and emotional apathy on Parkinson's disease. Annals of Clinical and Translational Neurology 2018;5:1286–91. https://doi.org/10.1002/acn3.626.
[196] Kay DB, Kirsch-Darrow L, Zahodne LB, Okun MS, Bowers D. Dimensions of apathy in Parkinson's disease: exploratory factor analysis of the apathy scale. J Parkinsons Dis 2012;2:161–6. https://doi.org/10.3233/JPD-2012-12085.
[197] Guimarães HC, Fialho PPA, Carvalho VA, Santos EL dos, Caramelli P, Guimarães HC, et al. Brazilian caregiver version of the Apathy Scale. Dementia & Neuropsychologia 2009;3:321–6. https://doi.org/10.1590/S1980-57642009DN30400010.
[198] Morita H, Kannari K. Reliability and validity assessment of an apathy scale for home-care patients with Parkinson's disease: a structural equation modeling analysis. J Phys Ther Sci 2016;28:1724–7. https://doi.org/10.1589/jpts.28.1724.
[199] Serrano-Dueñas M, Martínez-Martín P, Merchán T, Bravo R, Serrano M. Properties of the Apathy Scale (AS) for use on Parkinson's patients. Advances in Parkinson's Disease 2013;2:53–7. https://doi.org/10.4236/apd.2013.22010.
[200] Lopez FV, Eglit GML, Schiehser DM, Pirogovsky-Turk E, Litvan I, Lessig S, et al. Factor Analysis of the Apathy Scale in Parkinson's Disease. Movement Disorders Clinical Practice 2019;6:379–86. https://doi.org/10.1002/mdc3.12767.
[201] Wetmore JB, Arbelo JM, Catalán MJ, Valldeoriola F, Rodriguez-Blazquez C, Martinez-Martin P. Psychometric Properties of the Apathy Scale in Advanced Parkinson's Disease. Parkinson’s Disease 2019;2019:e1965394. https://doi.org/10.1155/2019/1965394.
[202] Radakovic R, Davenport R, Starr JM, Abrahams S. Apathy dimensions in Parkinson's disease. Int J Geriatr Psychiatry 2018;33:151–8. https://doi.org/10.1002/gps.4697.
[203] Santangelo G, D'Iorio A, Piscopo F, Cuoco S, Longo K, Amboni M, et al. Assessment of apathy minimising the effect of motor dysfunctions in Parkinson's disease: a validation study of the dimensional apathy scale. Qual Life Res 2017;26:2533–40. https://doi.org/10.1007/s11136-017-1569-6.
[204] Radakovic R, McGrory S, Chandran S, Swingler R, Pal S, Stephenson L, et al. The brief Dimensional Apathy Scale: A short clinical assessment of apathy. Clin Neuropsychol 2020;34:423–35. https://doi.org/10.1080/13854046.2019.1621382.
[205] Radakovic R, Gray D, Dudley K, Mioshi E, Dick D, Melchiorre G, et al. Reliability and validity of the brief dimensional apathy scale. Arch Clin Neuropsychol 2020;35:539–44. https://doi.org/10.1093/arclin/acaa002.
[206] Shinagawa Y, Nakaaki S, Hongo J, Murata Y, Sato J, Matsui T, et al. Reliability and validity of the Japanese version of the Dysexecutive Questionnaire (DEX) in Alzheimer's disease: validation of a behavioral

rating scale to assess dysexecutive symptoms in Japanese patients with Alzheimer's disease. Int J Geriatr Psychiatry 2007;22:951–6. https://doi.org/10.1002/gps.1768.
[207] Carvalho JO. Revision of the frontal systems behavior scale (FrSBe). Doctoral Dissertations Available from Proquest 2011:1–107.
[208] Midden AJ, Mast BT. IRT Analysis of the Frontal Systems Behavior Scale: Identifying the Best Items for Use with Geriatric Patients in Primary Care. Clinical Gerontologist 2017;40:295–306. https://doi.org/10.1080/07317115.2016.1276115.
[209] Davis TE. The Effectiveness of the Geriatric Depression Scale to Distinguish Apathy From Depression in Alzheimer's Disease and Related Dementias. 2008.
[210] Bertens AS, Moonen JEF, de Waal MWM, Foster-Dingley JC, de Ruijter W, Gussekloo J, et al. Validity of the three apathy items of the Geriatric Depression Scale (GDS-3A) in measuring apathy in older persons. Int J Geriatr Psychiatry 2017;32:421–8. https://doi.org/10.1002/gps.4484.
[211] Diesfeldt HFA. [Interpreting change scores of the Behavioural Rating Scale for Geriatric Inpatients (GIP)]. Tijdschr Gerontol Geriatr 2013;44:166–74. https://doi.org/10.1007/s12439-013-0035-3.
[212] Kolitz BP. Development of the Key Behaviors Change Inventory (KBCI), a TBI behavioral outcome assessment instrument. University of South Florida, 2000.
[213] Belanger HG, Brown LM, Crowell TA, Vanderploeg RD, Curtiss G. The Key Behaviors Change Inventory and executive functioning in an elderly clinic sample. Clin Neuropsychol 2002;16:251–7. https://doi.org/10.1076/clin.16.3.251.13848.
[214] Dujardin K, Sockeel P, Delliaux M, Destée A, Defebvre L. The Lille Apathy Rating Scale: validation of a caregiver-based version. Mov Disord 2008;23:845–9. https://doi.org/10.1002/mds.21968.
[215] Fernández-Matarrubia M, Matías-Guiu JA, Moreno-Ramos T, Valles-Salgado M, Marcos-Dolado A, García-Ramos R, et al. Validation of the Lille's Apathy Rating Scale in Very Mild to Moderate Dementia. The American Journal of Geriatric Psychiatry 2016;24:517–27. https://doi.org/10.1016/j.jagp.2015.09.004.
[216] García-Ramos R, Villanueva Iza C, Catalán MJ, Reig-Ferrer A, Matías-Guíu J. Validation of a Spanish Version of the Lille Apathy Rating Scale for Parkinson's Disease. The Scientific World Journal 2014;2014:e849834. https://doi.org/10.1155/2014/849834.
[217] Camozzato AL, Kochhann R, Simeoni C, Konrath CA, Franz AP, Carvalho A, et al. Reliability of the Brazilian Portuguese version of the Neuropsychiatric Inventory (NPI) for patients with Alzheimer's disease and their caregivers. International Psychogeriatrics 2008;20:383–93. https://doi.org/10.1017/S1041610207006254.
[218] Choi SH, Na DL, Kwon HM, Yoon SJ, Jeong JH, Ha CK. The Korean version of the neuropsychiatric inventory: a scoring tool for neuropsychiatric disturbance in dementia patients. J Korean Med Sci 2000;15:609–15.
[219] Davidsdottir SR, Snaedal J, Karlsdottir G, Atladottir I, Hannesdottir K. Validation of the Icelandic version of the Neuropsychiatric Inventory with Caregiver Distress (NPI-D). Nord J Psychiatry 2012;66:26–32. https://doi.org/10.3109/08039488.2011.593100.
[220] de Medeiros K, Robert P, Gauthier S, Stella F, Politis A, Leoutsakos J, et al. The Neuropsychiatric Inventory-Clinician rating scale (NPI-C): reliability and validity of a revised assessment of neuropsychiatric symptoms in dementia. Int Psychogeriatr 2010;22:984–94. https://doi.org/10.1017/S1041610210000876.
[221] Gallo JL, Schmidt KS, Libon DJ. An itemized approach to assessing behavioral and psychological symptoms in dementia. Am J Alzheimers Dis

Other Demen 2009;24:163–8.
https://doi.org/10.1177/1533317508329480.
[222] Malakouti SK, Panaghi L, Foroughan M, Salehi M, Zandi T. Farsi version of the Neuropsychiatric Inventory: validity and reliability study among Iranian elderly with dementia. Int Psychogeriatr 2012;24:223–30. https://doi.org/10.1017/S1041610211001797.
[223] Politis AM, Mayer LS, Passa M, Maillis A, Lyketsos CG. Validity and reliability of the newly translated Hellenic Neuropsychiatric Inventory (H-NPI) applied to Greek outpatients with Alzheimer's disease: a study of disturbing behaviors among referrals to a memory clinic. Int J Geriatr Psychiatry 2004;19:203–8. https://doi.org/10.1002/gps.1045.
[224] Stella F, Forlenza OV, Laks J, de Andrade LP, Ljubetic Avendaño MA, Gasparetto Sé EV, et al. The Brazilian version of the Neuropsychiatric Inventory-Clinician rating (NPI-C): Reliability and validity in dementia. Int Psychogeriatr 2013;25:1503–11. https://doi.org/10.1017/S1041610213000811.
[225] Vilalta-Franch J, Lozano-Gallego M, Hernández-Ferrándiz M, Llinàs-Reglà J, López-Pousa S, López OL. [The Neuropsychiatric Inventory. Psychometric properties of its adaptation into Spanish]. Rev Neurol 1999;29:15–9.
[226] Kat MG, de Jonghe JFM, Aalten P, Kalisvaart CJ, Dröes RM, Verhey FRJ. [Neuropsychiatric symptoms of dementia: psychometric aspects of the Dutch Neuropsychiatric Inventory (NPI)]. Tijdschr Gerontol Geriatr 2002;33:150–5.
[227] Leung VP, Lam LC, Chiu HF, Cummings JL, Chen QL. Validation study of the Chinese version of the neuropsychiatric inventory (CNPI). Int J Geriatr Psychiatry 2001;16:789–93. https://doi.org/10.1002/gps.427.
[228] Weintraut R, Karádi K, Lucza T, Kovács M, Makkos A, Janszky J, et al. Lille Apathy Rating Scale and MDS-UPDRS for Screening Apathy in Parkinson's Disease. J Parkinsons Dis 2016;6:257–65. https://doi.org/10.3233/JPD-150726.
[229] Kirsch-Darrow L, Zahodne LB, Hass C, Mikos A, Okun MS, Fernandez HH, et al. How Cautious Should We Be When Assessing Apathy with the Unified Parkinson's Disease Rating Scale? Mov Disord 2009;24:684–8. https://doi.org/10.1002/mds.22437.
[230] Starkstein SE, Merello M. The Unified Parkinson's Disease Rating Scale: Validation study of the mentation, behavior, and mood section. Movement Disorders 2007;22:2156–61. https://doi.org/10.1002/mds.21521.
[231] Pedersen KF, Larsen JP, Aarsland D. Validation of the Unified Parkinson's Disease Rating Scale (UPDRS) section I as a screening and diagnostic instrument for apathy in patients with Parkinson's disease. Parkinsonism & Related Disorders 2008;14:183–6. https://doi.org/10.1016/j.parkreldis.2007.07.015.
[232] Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. Mov Disord 2007;22:41–7. https://doi.org/10.1002/mds.21198.
[233] Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord 2008;23:2129–70. https://doi.org/10.1002/mds.22340.
[234] de Jonghe JF, Ooms ME, Ribbe MW. [Short version of the Dutch Behavioral Rating Scale for Psychogeriatric Inpatients (GIP-28)]. Tijdschr Gerontol Geriatr 1997;28:119–23.
[235] Fitts W, Massimo L, Lim N, Grossman M, Dahodwala N. Computerized assessment of goal-directed behavior in Parkinson's disease. J Clin Exp

Neuropsychol 2016;38:1015–25. https://doi.org/10.1080/13803395.2016.1184232.
[236] König A, Aalten P, Verhey F, Bensadoun G, Petit P-D, Robert PH, et al. A review of current information and communication technologies: can they be used to assess apathy? Int J Geriatr Psychiatry 2014;29:345–58. https://doi.org/10.1002/gps.4017.
[237] Massimo L, Evans LK. Differentiating subtypes of apathy to improve person-centered care in frontotemporal degeneration. J Gerontol Nurs 2014;40:58–65. https://doi.org/10.3928/00989134-20140827-01.
[238] World Health Organization. How to use the ICF: A practical manual for using the International Classification of Functioning, Disability and Health (ICF). Exposure draft for comment. Geneva: : WHO; 2013.
[239] Fried EI, Nesse RM. Depression sum-scores don't add up: why analyzing specific depression symptoms is essential. BMC Medicine 2015;13:72. https://doi.org/10.1186/s12916-015-0325-4.
[240] Carter SM, Little M. Justifying knowledge, justifying method, taking action: epistemologies, methodologies, and methods in qualitative research. Qualitative Health Research 2007;17:1316–28. https://doi.org/10.1177/1049732307306927.
[241] Collier A. Critical Realism: An Introduction to Roy Bhaskar's Philosophy. London: Verso; 1994.
[242] Scotland J. Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. English Language Teaching 2012;5:9.
[243] Maxwell JA, Mittapalli K. Realism as a Stance for Mixed Methods Research. In: Tashakkori A, Teddlie C, editors. SAGE handbook of mixed methods in social & Bheavioral Research. 2nd ed., Thousand Oaks, Calif. ; London: Thousand Oaks, Calif. ; London : SAGE; 2010, p. 145–67.
[244] Chalmers AF. What Is This Thing Called Science? : An Assessment of the Nature and Status of Science and its Methods. 2nd ed. Milton Keynes: Open University Press; 1982.
[245] Clark AM. The qualitative-quantitative debate: moving from positivism and confrontation to post-positivism and reconciliation. Journal of Advanced Nursing 1998;27:1242–9.
[246] Rosenberg A. Philosophy of social science. 4th ed. Boulder, Colo.: Boulder, Colo. : Westview Press; 2012.
[247] Pilgrim D. Some implications of critical realism for mental health research. Soc Theory Health 2014;12:1–21. https://doi.org/10.1057/sth.2013.17.
[248] Borsboom D, Cramer AOJ. Network analysis: an integrative approach to the structure of psychopathology. Annu Rev Clin Psychol 2013;9:91–121. https://doi.org/10.1146/annurev-clinpsy-050212-185608.
[249] Bhaskar R. The Possibility of Naturalism: a philosophical critique of the contemporary human sciences. 4th ed. London: London : Routledge, Taylor & Francis Group; 2015.
[250] Bhaskar R. A realist theory of science. [Rev. ed.]. London: Routledge; 2008.
[251] Fried EI, Cramer AOJ. Moving Forward: Challenges and Directions for Psychopathological Network Theory and Methodology. Perspect Psychol Sci 2017;12:999–1020. https://doi.org/10.1177/1745691617705892.
[252] Borsboom D. A network theory of mental disorders. World Psychiatry 2017;16:5–13. https://doi.org/10.1002/wps.20375.
[253] McNally RJ. Can network analysis transform psychopathology? Behaviour Research and Therapy 2016;86:95–104. https://doi.org/10.1016/j.brat.2016.06.006.

[254] McNally RJ. Network Analysis of Psychopathology: Controversies and Challenges. Annu Rev Clin Psychol 2021;17:31–53. https://doi.org/10.1146/annurev-clinpsy-081219-092850.
[255] Michell J. The Quantitative Imperative: Positivism, Naive Realism and the Place of Qualitative Methods in Psychology. Theory & Psychology 2003;13:5–31. https://doi.org/10.1177/0959354303013001758.
[256] McEvoy P, Richards D. A critical realist rationale for using a combination of quantitative and qualitative methods. Journal of Research in Nursing 2006;11:66–78. https://doi.org/10.1177/1744987106060192.
[257] Williams M, May T. Philosophical Issues in the Process of Social Research. Introduction to the Philosophy of Social Research, London: Routledge; 1996, p. 135–54.
[258] Olsen W, Morgan J. A Critical Epistemology of Analytical Statistics: Addressing the Sceptical Realist. Journal for the Theory of Social Behaviour 2005;35:255–84. https://doi.org/10.1111/j.1468-5914.2005.00279.x.
[259] Sayer A. Realism and Social Science. London: SAGE Publications; 2000.
[260] Sayer A. System, Lifeworld and Gender: Associational Versus Counterfactual Thinking. Sociology 2000;34:707–25. https://doi.org/10.1177/S0038038500000432.
[261] Clark AM, Lissel SL, Davis C. Complex critical realism: tenets and application in nursing research. ANS Adv Nurs Sci 2008;31:E67-79. https://doi.org/10.1097/01.ANS.0000341421.34457.2a.
[262] Sayer A. Quantitative Methods in Social Science. Method in social science : a realist approach. 2nd ed., London: London; 1992.
[263] Eastwood JG, Jalaludin BB, Kemp LA. Realist explanatory theory building method for social epidemiology: a protocol for a mixed method multilevel study of neighbourhood context and postnatal depression. SpringerPlus 2014;3:12. https://doi.org/10.1186/2193-1801-3-12.
[264] Robinaugh DJ, Hoekstra RHA, Toner ER, Borsboom D. The network approach to psychopathology: a review of the literature 2008–2018 and an agenda for future research. Psychological Medicine 2020;50:353–66. https://doi.org/10.1017/S0033291719003404.
[265] Borsboom D, Cramer AOJ, Fried EI, Isvoranu A-M, Robinaugh DJ, Dalege J, et al. Chapter 1. Network Perspectives. Network psychometrics with R: a guide for behavioral and social scientists, Routledge, Taylor & Francis Group; 2022.
[266] Constantini G, Perugini M. Network Analysis for Psychological Situations. The Oxford Handbook of Psychological Situations 2017. https://doi.org/10.1093/oxfordhb/9780190263348.013.16.
[267] Epskamp S, Haslbeck JMB, Isvoranu A-M, van Borkulo CD. Chapter 6. Pairwise Markov Random Fields. Network Psychometrics with R: A guide for behavioural and social scientists, Routledge, Taylor & Francis Group; 2022.
[268] Waldorp LJ, Borsboom D, Epskamp S. Chapter 5. Association and Conditional Independence. Network Psychometrics with R: A guide for behavioural and social scientists, Routledge, Taylor & Francis Group; 2022.
[269] Borsboom D, Deserno MK, Rhemtulla M, Epskamp S, Fried EI, McNally RJ, et al. Network analysis of multivariate data in psychological science. Nat Rev Methods Primers 2021;1:1–18. https://doi.org/10.1038/s43586-021-00055-w.
[270] Epskamp S, Fried EI. A tutorial on regularized partial correlation networks. Psychological Methods 2018;23:617–34. https://doi.org/10.1037/met0000167.
[271] Epskamp S, Rhemtulla M, Borsboom D. Generalized Network Psychometrics: Combining Network and Latent Variable Models. Psychometrika 2017;82:904–27. https://doi.org/10.1007/s11336-017-9557-x.

[272] Epskamp S, Maris GKJ, Waldorp LJ, Borsboom D. Network Psychometrics 2018.
[273] Jones PJ, Ma R, McNally RJ. Bridge Centrality: A Network Approach to Understanding Comorbidity. Multivariate Behavioral Research 2021;56:353–67. https://doi.org/10.1080/00273171.2019.1614898.
[274] Wichers M, Riese H, Hodges TM, Snippe E, Bos FM. A Narrative Review of Network Studies in Depression: What Different Methodological Approaches Tell Us About Depression. Frontiers in Psychiatry 2021;12.
[275] Saari T, Hallikainen I, Hintsa T, Koivisto AM. Network structures and temporal stability of self- and informant-rated affective symptoms in Alzheimer's disease. Journal of Affective Disorders 2020;276:1084–92. https://doi.org/10.1016/j.jad.2020.07.100.
[276] Kim KM, Kim D, Chung US, Lee JJ. Identification of Central Symptoms in Depression of Older Adults With the Geriatric Depression Scale Using Network Analysis and Item Response Theory. Psychiatry Investig 2021;18:1068–75. https://doi.org/10.30773/pi.2021.0453.
[277] van Wanrooij LL, Borsboom D, Charante EPM van, Richard E, Gool WA van. A network approach on the relation between apathy and depression symptoms with dementia and functional disability. International Psychogeriatrics 2019;31:1655–63. https://doi.org/10.1017/S1041610218002387.
[278] Saari T, Hallikainen I, Hintsa T, Koivisto AM. Neuropsychiatric symptoms and activities of daily living in Alzheimer's disease: ALSOVA 5-year follow-up study. International Psychogeriatrics 2020;32:741–51. http://dx.doi.org/10.1017/S1041610219001571.
[279] Creswell JW, Plano Clark VL. Designing and conducting mixed methods research. 3rd ed. London: SAGE Publications, Inc.; 2017.
[280] Petty S, Dening T, Coleston DM, Griffiths A. Dementia: beyond disorders of mood. Aging & Mental Health 2019;23:525–8. https://doi.org/10.1080/13607863.2018.1430742.
[281] Gilmore-Bykovskyi A. Commentary on Apathy as a Model for Investigating Behavioral and Psychological Symptoms in Dementia. Journal of the American Geriatrics Society 2018;66:S13–6. https://doi.org/10.1111/jgs.15325.
[282] Dening T, Baber W, Chang M, Yates J. The struggle of apathy in dementia. Aging & Mental Health 2021;0:1–3. https://doi.org/10.1080/13607863.2021.2008309.
[283] Mele B, Goodarzi Z, Hanson HM, Holroyd-Leduc J. Barriers and facilitators to diagnosing and managing apathy in Parkinson's disease: a qualitative study. BMC Neurology 2019;19:101. https://doi.org/10.1186/s12883-019-1329-z.
[284] Baber W, Chang CYM, Yates J, Dening T. The Experience of Apathy in Dementia: A Qualitative Study. Int J Environ Res Public Health 2021;18. https://doi.org/10.3390/ijerph18063325.
[285] Chang CYM, Baber W, Dening T, Yates J. "He Just Doesn't Want to Get Out of the Chair and Do It": The Impact of Apathy in People with Dementia on Their Carers. International Journal of Environmental Research and Public Health 2021;18:6317. https://doi.org/10.3390/ijerph18126317.
[286] Yin RK. Mixed Methods Research: are the methods genuinely integrated or merely parallel? Research in the Schools 2006;13:41–7.
[287] Fletcher AJ. Applying critical realism in qualitative research: methodology meets method. International Journal of Social Research Methodology 2017;20:181–94. https://doi.org/10.1080/13645579.2016.1144401.
[288] Hasson-Ohayon I, Goldzweig G, Lavi-Rotenberg A, Luther L, Lysaker PH. The centrality of cognitive symptoms and metacognition within the interacting network of symptoms, neurocognition, social cognition and

metacognition in schizophrenia. Schizophr Res 2018;202:260–6. https://doi.org/10.1016/j.schres.2018.07.007.
[289] Bernstein EE, Heeren A, McNally RJ. Unpacking Rumination and Executive Control: A Network Perspective. Clinical Psychological Science 2017;5:816–26. https://doi.org/10.1177/2167702617702717.
[290] Santos HP, Kossakowski JJ, Schwartz TA, Beeber L, Fried EI. Longitudinal network structure of depression symptoms and self-efficacy in low-income mothers. PLoS ONE 2018;13:e0191675. https://doi.org/10.1371/journal.pone.0191675.
[291] Ou CHK, Hall WA, Thorne SE. Can nursing epistemology embrace p-values? Nurs Philos 2017;18. https://doi.org/10.1111/nup.12173.
[292] Bringmann LF, Albers C, Bockting C, Borsboom D, Ceulemans E, Cramer A, et al. Psychopathological networks: Theory, methods and practice. Behaviour Research and Therapy 2022;149:104011. https://doi.org/10.1016/j.brat.2021.104011.
[293] Triolo F, Belvederi Murri M, Calderón-Larrañaga A, Vetrano DL, Sjöberg L, Fratiglioni L, et al. Bridging late-life depression and chronic somatic diseases: a network analysis. Transl Psychiatry 2021;11:1–8. https://doi.org/10.1038/s41398-021-01686-z.
[294] Snaith RP. The Hospital Anxiety And Depression Scale. Health and Quality of Life Outcomes 2003;1:29. https://doi.org/10.1186/1477-7525-1-29.
[295] Bjelland I, Dahl AA, Haug TT, Neckelmann D. The validity of the Hospital Anxiety and Depression Scale. An updated literature review. J Psychosom Res 2002;52:69–77.
[296] Marinus J, Leentjens AFG, Visser M, Stiggelbout AM, van Hilten JJ. Evaluation of the hospital anxiety and depression scale in patients with Parkinson's disease. Clin Neuropharmacol 2002;25:318–24.
[297] Fried EI. The 52 symptoms of major depression: Lack of content overlap among seven common depression scales. J Affect Disord 2017;208:191–7. https://doi.org/10.1016/j.jad.2016.10.019.
[298] Smith SC, Lamping DL, Banerjee S, Harwood R, Foley B, Smith P, et al. Measurement of health-related quality of life for people with dementia: development of a new instrument (DEMQOL) and an evaluation of current methodology. Health Technol Assess 2005;9:1–93, iii–iv. https://doi.org/10.3310/hta9100.
[299] Patterson J. Verbal Fluency. In: Kreutzer JS, DeLuca J, Caplan B, editors. Encyclopedia of Clinical Neuropsychology, New York, NY: Springer; 2011, p. 2603–6. https://doi.org/10.1007/978-0-387-79948-3_1423.
[300] Nasreddine ZS, Phillips NA, Bédirian V, Charbonneau S, Whitehead V, Collin I, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. J Am Geriatr Soc 2005;53:695–9. https://doi.org/10.1111/j.1532-5415.2005.53221.x.
[301] Romero-Ortuno R, Walsh CD, Lawlor BA, Kenny RA. A Frailty Instrument for primary care: findings from the Survey of Health, Ageing and Retirement in Europe (SHARE). BMC Geriatrics 2010;10:57. https://doi.org/10.1186/1471-2318-10-57.
[302] Berg K, Wood-Dauphine S, Williams J i., Gayton D. Measuring balance in the elderly: preliminary development of an instrument. Physiotherapy Canada 1989;41:304–11. https://doi.org/10.3138/ptc.41.6.304.
[303] Podsiadlo D, Richardson S. The Timed "Up & Go": A Test of Basic Functional Mobility for Frail Elderly Persons. Journal of the American Geriatrics Society 1991;39:142–8. https://doi.org/10.1111/j.1532-5415.1991.tb01616.x.
[304] Ries JD, Echternach JL, Nof L, Gagnon Blodgett M. Test-Retest Reliability and Minimal Detectable Change Scores for the Timed "Up & Go" Test, the Six-Minute Walk Test, and Gait Speed in People With Alzheimer Disease. Physical Therapy 2009;89:569–79. https://doi.org/10.2522/ptj.20080258.

[305] Johal SK, Rhemtulla M. Comparing estimation methods for psychometric networks with ordinal data. Psychological Methods 2021:No Pagination Specified-No Pagination Specified. https://doi.org/10.1037/met0000449.
[306] Isvoranu A-M, Epskamp S, Waldorp LJ, Borsboom D, editors. Network psychometrics with R: a guide for behavioral and social scientists. Abingdon: New York, NY; 2022. https://doi.org/10.4324/9781003111238.
[307] Isvoranu A-M, Epskamp S. Which estimation method to choose in network psychometrics? Deriving guidelines for applied researchers. Psychological Methods 2021. https://doi.org/10.1037/met0000439.
[308] Tabachnick B, Fidell L. Using Multivariate Statistics : Pearson New International Edition. Harlow, UNITED KINGDOM: Pearson Education, Limited; 2013.
[309] Downey RG, King CV. Missing Data in Likert Ratings: A Comparison of Replacement Methods. The Journal of General Psychology 1998;125:175–91. https://doi.org/10.1080/00221309809595542.
[310] Hallquist MN, Wright AGC, Molenaar PCM. Problems with Centrality Measures in Psychopathology Symptom Networks: Why Network Psychometrics Cannot Escape Psychometric Theory. Multivariate Behavioral Research 2021;56:199–223. https://doi.org/10.1080/00273171.2019.1640103.
[311] Golino H, Christensen A, Moulder R, Garrido LE, Jamison L. EGAnet: Exploratory Graph Analysis – a Framework for Estimating the Number of Dimensions in Multivariate Data using Network Psychometrics 2022.
[312] Christensen AP, Garrido LE, Golino H. Unique variable analysis: A network psychometrics method to detect local dependence. PsyArXiv; 2020. https://doi.org/10.31234/osf.io/4kra2.
[313] Christensen AP, Golino H, Silvia PJ. A Psychometric Network Perspective on the Validity and Validation of Personality Trait Questionnaires. European Journal of Personality 2020;34:1095–108. https://doi.org/10.1002/per.2265.
[314] DiStefano C, Shi D, Morgan GB. Collapsing Categories is Often More Advantageous than Modeling Sparse Data: Investigations in the CFA Framework. Structural Equation Modeling: A Multidisciplinary Journal 2021;28:237–49. https://doi.org/10.1080/10705511.2020.1803073.
[315] Field AP. Discovering statistics using SPSS / Andy Field. 3rd ed. London: London : SAGE; 2009.
[316] Fried EI, von Stockert S, Haslbeck JMB, Lamers F, Schoevers RA, Penninx BWJH. Using network analysis to examine links between individual depressive symptoms, inflammatory markers, and covariates. Psychol Med 2020;50:2682–90. https://doi.org/10.1017/S0033291719002770.
[317] Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. Behav Res Methods 2018;50:195–212. https://doi.org/10.3758/s13428-017-0862-1.
[318] Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D. qgraph: Network Visualizations of Relationships in Psychometric Data. Journal of Statistical Software 2012;48:1–18. https://doi.org/10.18637/jss.v048.i04.
[319] Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. Software: Practice and Experience 1991;21:1129–64. https://doi.org/10.1002/spe.4380211102.
[320] Deserno MK, Isvoranu A-M, Epskamp S, Blanken TF. Chapter 3. Descriptive analyses of network structures. Network psychometrics with R: a guide for behavioral and social scientists, Routledge, Taylor & Francis Group; 2022.
[321] Isvoranu A-M, Epskamp S. Chapter 4. Constructing and drawing networks in qgraph. Network Psychometrics with R: A guide for behavioural and social scientists, Routledge, Taylor & Francis Group; 2022.

[322] E. I. Fried, Epskamp S, Veenman M, van Borkulo CD. Chapter 8. Network stability comparison, and replicability. Network psychometrics with R: a guide for behavioral and social scientists, Routledge, Taylor & Francis Group; 2022, p. 133–53.
[323] Haslbeck JMB, Waldorp LJ. mgm: Estimating Time-Varying Mixed Graphical Models in High-Dimensional Data. Journal of Statistical Software 2020;93:1–46. https://doi.org/10.18637/jss.v093.i08.
[324] Costantini G, Epskamp S, Borsboom D, Perugini M, Mõttus R, Waldorp LJ, et al. State of the aRt personality research: A tutorial on network analysis of personality data in R. Journal of Research in Personality 2015;54:13–29. https://doi.org/10.1016/j.jrp.2014.07.003.
[325] Haslbeck JMB, Waldorp LJ. How well do network models predict observations? On the importance of predictability in network models. Behav Res 2018;50:853–61. https://doi.org/10.3758/s13428-017-0910-x.
[326] Taylor B, de Vocht H. Interviewing separately or as couples? Considerations of authenticity of method. Qual Health Res 2011;21:1576–87. https://doi.org/10.1177/1049732311415288.
[327] Maxwell JA. Using Qualitative Methods for Causal Explanation. Field Methods 2004;16:243–64. https://doi.org/10.1177/1525822X04266831.
[328] Morgan DL, Ataie J, Carder P, Hoffman K. Introducing Dyadic Interviews as a Method for Collecting Qualitative Data. Qual Health Res 2013;23:1276–84. https://doi.org/10.1177/1049732313501889.
[329] Pratt R. 'Nobody's ever asked how I felt'. The Perspectives of People with Dementia: Research Methods and Motivations, Jessica Kingsley Publishers; 2002, p. 165–82.
[330] Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology 2006;3:77–101. https://doi.org/10.1191/1478088706qp063oa.
[331] Braun V, Clarke V, Hayfield N, Terry G. Thematic Analysis. In: Liamputtong P, editor. Handbook of Research Methods in Health Social Sciences, Singapore: Springer; 2018, p. 1–18. https://doi.org/10.1007/978-981-10-2779-6_103-1.
[332] Saldaña J. The coding manual for qualitative researchers. 2nd ed. Los Angeles: SAGE; 2013.
[333] Clarke V, Braun V. Thematic Analysis. Analysing qualitative data in psychology. 2nd ed., Los Angeles: Los Angeles : SAGE; 2016, p. 84–103.
[334] Braun V, Clarke V. Reflecting on reflexive thematic analysis. Qualitative Research in Sport, Exercise and Health 2019;11:589–97. https://doi.org/10.1080/2159676X.2019.1628806.
[335] Cornish F, Gillespie A, Zittoun T. Collaborative Analysis of Qualitative Data. The SAGE Handbook of Qualitative Data Analysis, 1 Oliver's Yard, 55 City Road, London EC1Y 1SP United Kingdom: SAGE Publications Ltd; 2014, p. 79–93. https://doi.org/10.4135/9781446282243.n6.
[336] Maxwell JA. Collecting Qualitative Data: A Realist Approach. The SAGE Handbook of Qualitative Data Collection, 1 Oliver's Yard, 55 City Road London EC1Y 1SP: SAGE Publications Ltd; 2018, p. 19–31. https://doi.org/10.4135/9781526416070.n2.
[337] Finlay L. "Outing" the Researcher: The Provenance, Process, and Practice of Reflexivity. Qual Health Res 2002;12:531–45. https://doi.org/10.1177/104973202129120052.
[338] Braun V, Clarke V. Successful qualitative research : a practical guide for beginners. London: London : SAGE; 2013.
[339] Mason J. Qualitative researching. 3rd edition. Thousand Oaks, CA: SAGE Publications; 2017.
[340] Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand 1983;67:361–70.

[341] Forbes MK, Wright AGC, Markon KE, Krueger RF. Quantifying the Reliability and Replicability of Psychopathology Network Characteristics. Multivariate Behavioral Research 2021;56:224–42. https://doi.org/10.1080/00273171.2019.1616526.
[342] Dablander F, Hinne M. Node centrality measures are a poor substitute for causal inference. Scientific Reports 2019. https://doi.org/10.1038/s41598-019-43033-9.
[343] Silverman D 1943-. Qualitative research / edited by David Silverman. 4th ed. London: London : Sage; 2016.
[344] Gajardo J, Alvarado R, Slachevsky A, Gitlin LN. Self-stigma in people living with dementia in Chile: A qualitative exploratory study. Aging & Mental Health 2021;0:1–8. https://doi.org/10.1080/13607863.2021.1998351.
[345] van Vliet D, Persoon A, Bakker C, Koopmans RTCM, de Vugt ME, Bielderman A, et al. Feeling useful and engaged in daily life: exploring the experiences of people with young-onset dementia. Int Psychogeriatr 2017;29:1889–98. https://doi.org/10.1017/S1041610217001314.
[346] Contreras A, Nieto I, Valiente C, Espinosa R, Vazquez C. The Study of Psychopathology from the Network Analysis Perspective: A Systematic Review. Psychother Psychosom 2019;88:71–83. https://doi.org/10.1159/000497425.
[347] Bjørkløf GH, Helvik A-S, Ibsen TL, Telenius EW, Grov EK, Eriksen S. Balancing the struggle to live with dementia: a systematic meta-synthesis of coping. BMC Geriatrics 2019;19:295. https://doi.org/10.1186/s12877-019-1306-9.
[348] Górska S, Forsyth K, Maciver D. Living With Dementia: A Meta-synthesis of Qualitative Research on the Lived Experience. Gerontologist 2018;58:e180–96. https://doi.org/10.1093/geront/gnw195.
[349] Steeman E, de Casterlé BD, Godderis J, Grypdonck M. Living with early-stage dementia: a review of qualitative studies. J Adv Nurs 2006;54:722–38. https://doi.org/10.1111/j.1365-2648.2006.03874.x.
[350] Marin RS. Apathy: a neuropsychiatric syndrome. The Journal of Neuropsychiatry and Clinical Neurosciences 1991.
[351] Kitwood TM. Dementia reconsidered: the person comes first. Buckingham [England]; Philadelphia: Open University Press; 1997.
[352] Kearney FC, Harwood RH, Gladman JRF, Lincoln N, Masud T. The Relationship between Executive Function and Falls and Gait Abnormalities in Older Adults: A Systematic Review. DEM 2013;36:20–35. https://doi.org/10.1159/000350031.
[353] Semprini R, Lubrano A, Misaggi G, Martorana A. Apathy as Marker of Frail Status. Journal of Aging Research 2012;2012:436251. https://doi.org/10.1155/2012/436251.
[354] Deci EL, Ryan RM. The "What" and "Why" of Goal Pursuits: Human Needs and the Self-Determination of Behavior. Psychological Inquiry 2000;11:227–68. https://doi.org/10.1207/S15327965PLI1104_01.
[355] Gillies BA. A memory like clockwork: Accounts of living through dementia. Aging & Mental Health 2000;4:366–74. https://doi.org/10.1080/713649961.
[356] Riehle M, Kasanova Z, Lincoln TM. Psychosocial Approaches to the Treatment of Apathy. Apathy, Oxford University Press; 2021, p. 242–58. https://doi.org/10.1093/med/9780198841807.003.0014.
[357] Charmaz K. Loss of self: a fundamental form of suffering in the chronically ill. Sociology of Health & Illness 1983;5:168–95. https://doi.org/10.1111/1467-9566.ep10491512.
[358] Harman G, Clare L. Illness Representations and Lived Experience in Early-Stage Dementia. Qual Health Res 2006;16:484–502. https://doi.org/10.1177/1049732306286851.

[359] Gee B, Hodgekins J, Lavis A, Notley C, Birchwood M, Everard L, et al. Lived experiences of negative symptoms in first-episode psychosis: A qualitative secondary analysis. Early Intervention in Psychiatry 2019;13:773–9. https://doi.org/10.1111/eip.12558.
[360] Vansteenkiste M, Ryan RM, Soenens B. Basic psychological need theory: Advancements, critical themes, and future directions. Motiv Emot 2020;44:1–31. https://doi.org/10.1007/s11031-019-09818-1.
[361] Raber C, Teitelman J, Watts J, Kielhofner G. A Phenomenological Study of Volition in Everyday Occupations of Older People with Dementia. British Journal of Occupational Therapy 2010;73:498–506. https://doi.org/10.4276/030802210X12892992239116.
[362] Pool J, Downs M, Downs M. The Pool Activity Level (PAL) Instrument for Occupational Profiling : A Practical Resource for Carers of People with Cognitive Impairment Fourth Edition. London, UNITED KINGDOM: Jessica Kingsley Publishers; 2011.
[363] Polenick CA, Struble LM, Stanislawski B, Turnwald M, Broderick B, Gitlin LN, et al. "I've learned to just go with the flow": Family caregivers' strategies for managing behavioral and psychological symptoms of dementia. Dementia 2020;19:590–605. https://doi.org/10.1177/1471301218780768.
[364] Phinney A, Chaudhury H, O'connor DL. Doing as much as I can do: The meaning of activity for people with dementia. Aging & Mental Health 2007;11:384–93. https://doi.org/10.1080/13607860601086470.
[365] Reeve D. Psycho-Emotional Disablism. Routledge Handbook of Disability Studies, Routledge; 2013. https://doi.org/10.4324/9780203144114.ch7.
[366] World Health Organization. Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. 2020.
[367] Johnson B. Prime Minister's statement on coronavirus (COVID-19): 16 March 2020. 2020.
[368] Wood L. DHSC guidance on the impact of COVID-19 on research funded or supported by NIHR. 2020.
[369] Cambridge Cognition. CANTAB® [Cognitive assessment software] 2022.
[370] Royal College of Obstetricians and Gynaecologists. COVID-19 virus infection and pregnancy: Occupational health advice for employers and pregnant women during the COVID-19 pandemic. 2020.
[371] National Institute for Health and Care Research. Briefing notes for researchers - public involvement in NHS, health and social care research. 2021.
[372] Pfeifer L, Drobetz R, Fankhauser S, Mortby ME, Maercker A, Forstmeier S. Caregiver rating bias in mild cognitive impairment and mild Alzheimer's disease: impact of caregiver burden and depression on dyadic rating discrepancy across domains. International Psychogeriatrics 2013;25:1345–55. https://doi.org/10.1017/S1041610213000562.
[373] Votruba KL, Persad C, Giordani B. Patient Mood and Instrumental Activities of Daily Living in Alzheimer Disease: Relationship Between Patient and Caregiver Reports. J Geriatr Psychiatry Neurol 2015;28:203–9. https://doi.org/10.1177/0891988715588829.
[374] Wichers M, Wigman JTW, Bringmann LF, de Jonge P. Mental disorders as networks: some cautionary reflections on a promising approach. Soc Psychiatry Psychiatr Epidemiol 2017;52:143–5. https://doi.org/10.1007/s00127-016-1335-z.
[375] Mukumbang FC, Marchal B, Van Belle S, van Wyk B. Using the realist interview approach to maintain theoretical awareness in realist studies. Qualitative Research 2020;20:485–515. https://doi.org/10.1177/1468794119881985.

Appendix 1 MEDLINE strategy

MEDLINE search strategy:

	Search terms
1	(instrumentation or methods).sh.
2	(Validation Studies or Comparative Study).pt.
3	exp Psychometrics/
4	psychometr*.ti,ab.
5	(clinimetr* or clinometr*).tw.
6	exp "Outcome Assessment (Health Care)"/
7	outcome assessment.ti,ab.
8	outcome measure*.tw.
9	exp Observer Variation/
10	observer variation.ti,ab.
11	exp Health Status Indicators/
12	exp "Reproducibility of Results"/
13	reproducib*.ti,ab.
14	exp Discriminant Analysis/
15	(reliab* or unreliab* or valid* or coefficient or homogeneity or homogeneous or "internal consistency").ti,ab.
16	(cronbach* and (alpha or alphas)).ti,ab.
17	(item and (correlation* or selection* or reduction*)).ti,ab.
18	(agreement or precision or imprecision or "precise values" or test-retest).ti,ab.
19	(test and retest).ti,ab.
20	(reliab* and (test or retest)).ti,ab.
21	(stability or interrater or inter-rater or intrarater or intra-rater or intertester or inter-tester or intratester or intra-tester or interobserver or inter-observer or intraobserver or intra-observer or intertechnician or inter-technician or intratechnician or intra-technician or interexaminer or inter-examiner or intraexaminer or intra-examiner or interassay or inter-assay or intraassay or intra-assay or interindividual or inter-individual or intraindividual or intra-individual or interparticipant or inter-participant or intraparticipant or intra-participant or kappa or kappa's or kappas or repeatab*).ti,ab.
22	((replicab* or repeated) and (measure or measures or findings or result or results or test or tests)).ti,ab.
23	(generaliza* or generalisa* or concordance).ti,ab.
24	(intraclass and correlation*).ti,ab.
25	(discriminative or "known group" or factor analysis or factor analyses or dimension* or subscale*).ti,ab.
26	(multitrait and scaling and (analysis or analyses)).ti,ab.
27	(item discriminant or interscale correlation* or error or errors or "individual variability").ti,ab.
28	(variability and (analysis or values)).ti,ab.
29	(uncertainty and (measurement or measuring)).ti,ab.
30	("standard error of measurement" or sensitiv* or responsive*).ti,ab.
31	((minimal or minimally or clinical or clinically) and (important or significant or detectable) and (change or difference)).ti,ab.
32	(small* and (real or detectable) and (change or difference)).ti,ab.
33	(meaningful change or "ceiling effect" or "floor effect" or "Item response model" or IRT or Rasch or "Differential item functioning" or DIF or "computer adaptive testing" or "item bank" or "cross-cultural equivalence").ti,ab.
34	1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33
35	exp APATHY/
36	apath*.mp
37	amotivat*.ti,ab.
38	diminished motivation.ti,ab.
39	diminished interest.ti,ab.
40	lack of interest.ti,ab.
41	diminished initiat*.ti,ab.

42	**lack of initiat*.ti,ab.**
43	lack of motivation.ti,ab.
44	emotional* blunt*.ti,ab.
45	abulia.ti,ab.
46	anhedonia.ti,ab.
47	exp Anhedonia /
48	frontal symptom*.ti,ab.
49	emotional responsiv*.ti,ab.
50	asocial*.ti,ab.
51	avolition*.ti,ab.
52	lassitude.ti,ab.
53	35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52
54	34 and 53
55	limit 54 to "all adult (19 plus years)"

Appendix 2 Data Extraction Table Headers

Data Extraction Table Headers:

Identifier	**Number** (for randomisation)
	DOI
	Author
	Year
	Title of study
	Name of measure
Interpretability	**% of missing items**
	% of missing total scores
	Floor/ ceiling effects: % min score; % max score
	Apathy scores & change scores for relevant (sub)groups (e.g. cognitive impairment vs healthy controls) ***Mean (SD), range or similar***
	minimal important change or minimal important difference
	Information on response shift
Study characteristics	**Design**
	Sampling method
	Setting (location, time, context)
	Target population
	eligibility criteria
	N (in each sub-analysis where appropriate)
	Measurement properties assessed (i.e. relevant COSMIN boxes to complete)
	Further description of measure if needed (e.g. changes to original)
	Country from which research was conducted
	Language of measure
Participant Descriptives	**Age**
	Mean (SD), Range or similar
	Ethnicity
	Distribution of sex
	% M/ F
	Disease characteristics (disease status, severity, duration)
	Cognitive status
	Mean (SD), range of MMSE or similar
	Residential status
	Type and distribution e.g. % at home in community
Content validity	**Content validity** *Describe method describe method briefly, e.g. asking stakeholders (e.g. patients, carers, experts) about the relevance, comprehensiveness, comprehensibility of the measure*

Psychometric properties	**Structural validity -** (e.g. Comparative Fit Index or Tucker Lewis Index)
	Internal consistency (e.g. Cronbach's alpha)
	Reliability (e.g. intra-class correlation co-efficient or weighted Kappa)
	Measurement error (Standard error of measurement, Smallest Detectable change, Limits of Agreement, or % agreement)
	Hypothesis testing / Construct validity : comparisons Describe all comparator instruments used
	Hypothesis testing / Construct validity (convergent and discriminative; cross-sectional data only) Describe statistical method, result for each relevant comparator measure)
	Cross cultural validity/ measurement invariance. Differences between group factors
	Responsiveness (Longitudinal data only. Compared to another measure, compared across groups, or before & after intervention)

Appendix 3 COSMIN RoB checklist

Boxes of the COSMIN Risk of Bias Checklist

Category	Boxes of the COSMIN Risk of Bias Checklist
Content Validity	Box 1. PROM development
	Box 2. Content validity
Internal Structure	Box 3. Structural validity
	Box 4. Internal consistency
	Box 5. Cross-cultural validity
Remaining measurement properties	Box 6. Reliability
	Box 7. Measurement error
	Box 8. Criterion validity
	Box 9. Hypothesis testing for construct validity
	Box 10. Responsiveness

Adapted with permission from Mokkink et al. (2017)

Each risk of bias checklist box is to be completed for each study that assesses that measurement property. Box 1 is to be completed for original development studies, whereas box 2 is to be completed for any additional content validity studies, or studies developing an established measure in a different population. Box 8 will not be completed for any study in this systematic review, as no gold standard measure of apathy exists. For details of how risk of bias is assessed for each measurement property, see Mokkink et al (2017).

Full guidelines followed for this review are found in the comprehensive COSMIN user manual version 1.0 dated February 2018 downloaded from: https://www.cosmin.nl/wp-content/uploads/COSMIN-syst-review-for-PROMs-manual_version-1_feb-2018-1.pdf and the content validity user manual version 1.0, downloaded from: https://www.cosmin.nl/wp-content/uploads/COSMIN-methodology-for-content-validity-user-manual-v1.pdf

Appendix 4 COSMIN criteria for good measurement properties

COSMIN criteria for good measurement properties

Measurement property	Rating	Criteria
Structural Validity	+	**CTT:** CFA: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08[a] **IRT/Rasch:** No violation of unidimensionality[b]: CFI or TLI or comparable measure >0.95 OR RMSEA <0.06 OR SRMR <0.08 *AND* no violation of local independence: residual correlations among the items after controlling for the dominant factor < 0.20 OR Q3's < 0.37 *AND* no violation of monotonicity: adequate looking graphs OR item scalability >0.30 *AND* adequate model fit IRT: $\chi 2$ >0.01 Rasch: infit and outfit mean squares ≥ 0.5 and ≤ 1.5 OR Z-standardized values > -2 and <2
	?	CTT: Not all information for '+' reported IRT/Rasch: Model fit not reported
	-	Criteria for '+' not met
Internal Consistency	+	At least low evidence[c] for sufficient structural validity[d] AND Cronbach's alpha(s) ≥ 0.70 for each unidimensional scale or Subscale[e]
	?	Criteria for "At least low evidence4 for sufficient structural Validity[d]" not met
	-	At least low evidence[c] for sufficient structural validity[5] AND Cronbach's alpha(s) < 0.70 for each unidimensional scale or subscale[e]
Reliability	+	ICC or weighted Kappa ≥ 0.70
	?	ICC or weighted Kappa not reported
	-	ICC or weighted Kappa < 0.70
Measurement error	+	SDC or LoA < MIC[d]
	?	MIC not defined
	-	SDC or LoA > MIC[d]
Hypotheses testing for construct validity	+	The result is in accordance with the hypothesis[f]
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis[f]
Cross-cultural validity \ measurement invariance	+	No important differences found between group factors (such as age, gender, language) in multiple group factor analysis OR no important DIF for group factors (McFadden's R^2 < 0.02)
	?	No multiple group factor analysis OR DIF analysis performed
	-	Important differences between group factors OR DIF was found
Criterion validity	+	Correlation with gold standard ≥ 0.70 OR AUC ≥ 0.70
	?	Not all information for '+' reported
	-	Correlation with gold standard < 0.70 OR AUC < 0.70
Responsiveness	+	The result is in accordance with the hypothesis[f] OR AUC ≥ 0.70
	?	No hypothesis defined (by the review team)
	-	The result is not in accordance with the hypothesis[f] OR AUC < 0.70

AUC area under the curve, *CFA* confirmatory factor analysis, *CFI* comparative fit index, *CTT* classical test theory, *DIF* differential item functioning, *ICC* intraclass correlation coefficient, *IRT* item response theory, *LoA* limits of agreement, *MIC* minimal important change, *RMSEA* root mean square error of approximation, *SEM* standard error of measurement, *SDC* smallest detectable change, *SRMR* standardized root mean residuals, *TLI* Tucker–Lewis index

"+" = sufficient, "–" = insufficient, "?" = indeterminate

[a] To rate the quality of the summary score, the factor structures should be equal across studies

[b] unidimensionality refers to a factor analysis per subscale, while structural validity refers to a factor analysis of a (multidimensional) patient-reported outcome measure

[c] As defined by grading the evidence according to the GRADE approach

[d] This evidence may come from different studies

[e] The criteria 'Cronbach alpha < 0.95' was deleted, as this is relevant in the development phase of a PROM and not when evaluating an existing PROM.

[f] The results of all studies should be taken together and it should then be decided if 75% of the results are in accordance with the hypotheses

Table and footnotes reproduced with permission from Prinsen et al. (2018)

In addition criteria listed in this table, criteria previously described by Prinsen et al. (2016, p. 7) for studies using exploratory factor analysis will be applied:
+ "First factor accounts for at least 20% of the variability AND ratio of the variance explained by the first to the second factor greater than 4";
? "Not all information for '+' reported";
- "Criteria for '+' not met".

Appendix 5 Additional decisions for systematic review

COSMIN section	Description of issue / decision to be made	Decision and reason
General / multiple properties		
Box 4 standard 2. Boxes 6 & 7 standards 4 & 5	What classes as continuous and ordinal scores?	Though Likert scales produce ordinal level data, they are commonly treated as continuous by researchers [1]. Given that it is common practice among researchers, it was deemed too harsh to penalise studies that treated data from Likert scales as continuous. Furthermore, it has been demonstrated that parametric tests are robust to violations of assumptions of continuous data [2]. Therefore, it was decided that measures can be treated as continuous data if they contain at least seven points from which participants can choose that contribute to the score in question. For example, a measure made of two items, each rated on a 4-point Likert scale would be considered continuous, but not a measure made of two items each rated on 3 weighted options (e.g. not at all = 0; a little = 2; a lot = 5). A requirement of the scale to have at least 7 rateable points decided following discussions with the review team: if 7 rateable items are considered optimal [3] then the whole scale will require at least 7 rateable points to be considered continuous.
Content validity		
General	What constitutes a development study and what is a content validity study?	Distinguishing between content validity and development studies is sometimes difficult when a measure is based on the adaptation of an existing measure. In these scenarios they were treated as development studies only if the new measure was deemed sufficiently different from the original measure, as indicated by the addition of new items and changes in wording.
Quality criteria 1 (development study rating)	What constitutes 'evidence from concept elicitation, literature or professionals that at least 85% of the items refer to the construct of interest'	Email correspondence with the COSMIN team confirmed that this could be a simple statement to say that items were developed from the literature or developed by professionals.
Quality criteria 1 (reviewer rating)	What items are deemed 'relevant to apathy'	Items may cover: apathy, motivation, goal-directed behaviour (including social), goal-direct cognition or 'interest' (including social), emotion / emotional responsiveness, and insight. These have been previously argued to be a part of the apathy construct. Items that are easily conflated with dysphoria/ depression, cognition, or fatigue were rated as not relevant in order to discourage inclusion of items that are likely to measure another related but different construct.
Quality criteria 4 (reviewer rating)	What constitutes appropriate response options?	Appropriate response options are those with more than 2 options and categorical or Likert response options (rather than an analogue scale in which scores are not defined by named categories). This was decided following consultation with PPI members and literature. It has been shown that respondents prefer more than 2 response options and people with dementia can use 4 response options successfully [4]. It has also been noted that Likert scales require no training, where as visual analogue scales require some additional explanation [5], which is not desirable for a cognitively impaired population.
Quality criteria 5	What is an appropriate recall period?	When the respondent is a person with dementia, the recall period should be 1 to 2 weeks. For other respondents, the recall period should be 1 to less than 4 weeks.

COSMIN section	Description of issue / decision to be made	Decision and reason
(reviewer rating)		Reference to before the onset of disease is not appropriate as it is difficult for respondents to know when the disease started, and diagnosis or disease onset may have been a long time ago. There was little guidance for this in the literature, with researchers commenting what is appropriate is dependent on many factors [6]. Therefore this decision was informed entirely by consultation with PPI members.
Quality criteria 6 (reviewer rating)	What is required for a comprehensive assessment of apathy?	The measure must include measurement of goal-direct behaviour, goal-direct cognition and emotion / emotional responsiveness. This is in line with Marin's [7] definition of apathy, and other conceptualisations [8,9]. This is not however in line with the latest proposed criteria for apathy that stipulates apathy also has a social domain [10]. It was decided not to require a social dimension because this is was easily conflated with the other domains e.g. social interest, initiation of social contact.
RoB Box 2d Standard 23. RoB Box 2e Standard 28.	What constitutes 'professionals from all relevant disciplines'?	All relevant disciplines are: psychiatry, neurology, psychology, and nursing/OT/support work/health care assistant. Neurology is important as apathy is often considered a symptom of physiological brain dysfunction. Psychiatry is important as apathy is considered a psychiatric symptom, and overlaps with depression. Psychology because this may focus on more behavioural and non-clinical symptoms. Nursing and related disciplines have traditionally had a different focus (holistic, person centred) and could offer a different perspective from more medical disciplines. This selection of professionals also represents the diverse range of professionals likely to be using apathy measures in clinical practice.
Overall quality criteria for content validity	How should content validity be rated when the combination of scores from the 3 domains of content validity does not fall into any of the listed scenarios in the COSMIN guide? i.e. when one or more of the 3 content validity domains is rated mixed.	Allow rating of sufficient or insufficient if just one is rated mixed and the other two are consistent with the sufficient/ insufficient rating. This was decided to avoid giving too much weight to more trivial aspects of content validity, such as justifying the recall period and response options. However if two or more content validity domains were rated mixed, then the overall rating was decided to be rated mixed.
Structural validity		
General	Should studies of wider measures be considered evidence for structural validity of an apathy sub-measure?	Studies were only deemed to assess structural validity if they assessed the structure of a measure that only related to apathy. Studies assessing the structure of a measure of different concepts were not classed as evidence for structural validity of an apathy subcomponent, even if the result was that apathy was its own factor, as the aim was not to assess validity of the apathy measure itself, and results may differ if the subcomponent of apathy was investigated on its own.
Quality criteria for exploratory studies	The final COSMIN manual does not provide criteria for structural validity for studies using EFA.	An earlier version of COSMIN provided the following criteria for EFA, agreed in a Delphi study: "Unidimensionality: EFA: First factor accounts for at least 20% of the variability AND ratio of the variance explained by the first to the second factor greater than 4" [11]. This criteria will be applied to EFA studies.
RoB Box 3 Standard 4.	What constitutes an appropriate rotation method?	Orthogonal rotation methods (equamax, orthomax, quartimax, and varimax) should be used when factors in the analysis are uncorrelated. Oblique rotation methods (oblimin & promax) should be used when the factors are correlated. Tabachnick and Fidell [12] recommend researchers start with an oblique

COSMIN section	Description of issue / decision to be made	Decision and reason
		rotation method then assess the correlations between factors, and advise that if correlations exceed .32, then there is sufficient relationship between factors to warrant oblique rotation. Therefore, if oblique rotation was used, but correlation between factors was less than .32, this will be considered inappropriate (inadequate). If no evidence to the contrary, this will be considered appropriate, as it can be expected that factors of an apathy scale will be related. If orthogonal rotation used, but no evidence for whether there was correlation between factors was less than .32, this will be marked as doubtful. If correlation indicates more than .32, this will be considered an inappropriate (inadequate). If correlation indicates less than .32 this will be considered appropriate.
RoB Box 3 Standard 4.	What might indicate a flaw in the design or statistical methods of the study? (KMO)	KMO indicates whether factors contained overlapping correlations between items, in which case factor analysis would be inappropriate. If below .50, then score 'inadequate'; if between .69 and .5, score 'doubtful' as this is considered "barely acceptable" [13]. Do not penalise if not reported.
RoB Box 3 Standard 4.	What might indicate a flaw in the design or statistical methods of the study? (Bartlett test of sphericity)	Bartlett's test of sphericity tests whether there was sufficient relationships between variables for factor analysis to be appropriate. If .05 or above (i.e. not significant), then mark inadequate, as this indicates factor analysis was not appropriate [13]. Do not penalise if not reported.
Internal Consistency		
RoB Box 4. Standard 1	What constitutes evidence that the scale or subscale was unidimensional?	Where studies of structural validity did not have a conclusive result, or where there were none included in the review, internal consistency was not determined for that measure. This is because internal consistency should be assessed for unidimensional scales, and COSMIN guidelines recommend at least low evidence of sufficient structural validity. Whilst this evidence could be taken from studies that did not meet eligibility criteria, this would only give evidence for structural validity for populations different from those that are the focus of this review, so it was decided to only use studies that met the inclusion criteria as validity is population and context dependent.
Reliability		
Quality criteria	COSMIN criteria for good measurement properties only provided criteria for reliability when measured by ICC or weighted kappa	The same criteria was decided to be applicable to results of other statistical methods (such as unweighted Kappa, Pearson's r and Spearman's rho). This is because the studies' methodological quality is already penalised for using these less optimal methods, but the results are still indicative of relative reliability.
RoB Box 6 Standard 2.	What is an appropriate time interval for a test-retest reliability study?	There is no gold standard time interval for test-retest reliability studies [14] though a time interval of two weeks is common [15]. Apathy is a relatively stable, but not an enduring trait, so a time interval that exceeded 28 days or 1 month was considered inappropriate. A time interval of less than 3 days for people with memory problems, and less than 7 days for people without memory problems, was also considered inappropriate as memory of previous answers may inflate reliability estimates. These were arbitrary numbers chosen in lieu of guidance, but was deemed acceptable by PPI members.
RoB Box 6 Standard 2.	What is an appropriate time interval for an interrater reliability study	Guidance regarding appropriate time intervals for interrater reliability is limited. Less than 7 days was chosen as an arbitrary cut off to avoid conflating interrater variability with change over time. Simultaneous assessment was not deemed appropriate to assess interrater reliability (see decision regarding box 6, standard 8)

COSMIN section	Description of issue / decision to be made	Decision and reason
RoB Box 6 Standard 8	"Were there any other important flaws in the design or statistical methods of the study" of interrater reliability?	Interrater reliability studies that conducted assessments in tandem, or based on observation of a pre-recorded interview were judged to contain 'minor methodological flaws' (therefore rated doubtful) as they are not assessing the whole of interrater reliability. When interrater reliability assessments are made at the same time as each other this ignores a key part of the interrater reliability: whether there are differences in the way the scales are administered, as well as scored. Whilst it may be interesting to know what error exists when these differences in administration are controlled for, allowing this would give an inappropriate impression of the evidence for overall reliability in these scenarios.
Hypothesis testing		
General	What should be considered as a test of construct validity?	Studies that assessed the relationship of apathy with another apathy measure, depression, executive functioning, activities of daily living, anxiety, fatigue, and global memory were considered to be studies of construct validity. Studies assessing the relationship of apathy with another construct were not considered construct validity studies. Nor were studies testing a measure against a version of itself. This is in line with COSMIN guidance that states that studies can mislabel their testing of associations incorrectly as construct validity studies.
Quality criteria (a-priori hypotheses)	What are the a-priori hypotheses?	Construct validity ratings require that a-prior hypotheses are developed by the review team. These hypotheses should go beyond significant testing [14]. The following hypotheses were informed by COSMIN suggested generic hypotheses [16] and other researchers' a-priori predictions about the expected relationships between apathy and other constructs [17]. For convergent validity, it was decided that apathy measures should demonstrate a strong correlation with another apathy measure r≥.6. However, it was predicted that there would be slightly less correlation if information was gathered from a different source (e.g. participant rated measure compared to carer reported measure of apathy; r≥.5), and at least a small correlation if the apathy scale was compared to just part of the apathy construct, for example a subscale of an apathy measure (r≥.3). A moderate association with depression was expected, if the depression scale was designed to be inclusive of anhedonia and apathy-like symptoms (r=.3 to.59), and again allowances were made if the information source differed (r=.2 to .59). As initiation is a key component of apathy, and included in some measures of Executive Functioning (EF), EF measures that included initiation components were expected to also have a moderate correlation with apathy measures (r=.3 to .59). Similar, as goal-directed activity is a key component of apathy, Activities of Daily Living (ADL) measures that included instrumental ADL or initiation of ADL were also expected to have a moderate correlation with apathy measures (r=.3 to .59). Anxiety, fatigue, EF and memory without an initiation component, basic ADL, and depression measures that did not include anhedonia or apathy-like symptoms were not expected to correlate with apathy measures (r<.4).
RoB Box 9b Standard 5.	What are the important characteristics of subgroups that should be reported?	Important characteristics of subgroups are: Number; age; gender; disease status (where multiple are present); apathy scores; cognitive status (only where included in study). Rate 'most' are present if up to 2 are missing. Rate 'poor' if more than 2 of these are missing. Age, gender and disease characteristics are suggested by COSMIN guidelines. Number is relevant to the statistical method applied. Cognitive status and disease status are important determinants of apathy scores, and it is good practice to report the actual average apathy scores per group as this is what is being tested.

COSMIN section	Description of issue / decision to be made	Decision and reason
GRADE		
Indirectness	How should studies be penalised for indirectness when the target population is older adults?	For single studies: -1 if there are no healthy adults and -1 if there is evidence that there are under 65s in the study. For multiple studies: -1 if all results have no healthy adults and -1 if all results have evidence that there are under 65s in the study. This decision was informed by PPI feedback: Members felt that people with a disorder could represent older adults generally, as many older adults live with health problems, so it was not necessary for the entire sample to be made of healthy adults, but simply some healthy adults. People of a younger age were deemed not to be representative of older adults however, so the decision was made to penalise studies that included younger adults. Many studies did not report a minimum age however, so if it was unclear whether under 65s were included, the study was not penalised.
Indirectness	How should studies be penalised for indirectness when the target population is people with dementia/ cognitive impairment?	For single studies: -1 if some participants do not have a diagnosis of cognitive impairment (dementia or MCI), or no diagnosis for any participant but average MMSE <27; -2 if all participants have no diagnosis and no cognitive impairment (as defined by average MMSE <27). For multiple studies: 0 if at least one study with all participants with diagnosed cognitive impairment. -1 if at least one study has some participants with diagnosed cognitive impairment or all participants have no diagnosis but average MMSE <27. - 2 if no studies have people diagnosed with cognitive impairment and average MMSE 27 or above. Where MMSE is not reported, assume it is 27 or above. MMSE criteria was chosen as a score of 27 is reported to indicate cognitive impairment. However, studies are still penalised -1 as MMSE is not the best classification of cognitive impairment [18,19]. MMSE was nevertheless chosen due to its common use in research.
Inconsistency	Whether to rate studies as +/- or pool results and downgrade for inconsistency.	Where studies do not meet criteria for sufficient or insufficient ratings, but the discrepancy can be explained for instance by worse quality studies producing a different result, or differences in population studied, they will be rated sufficient or insufficient (depending on which score is consistency with the higher quality studies or similar population etc.) and scored down for inconsistency (-1). When there is no explanation for discrepancies, the studies will be rated mixed, in which case there is no need to score down for inconsistency.
Additional criteria	What to do with hypothesis testing results when there has been no studies of convergent validity	Where there were no studies of convergent validity, downgrade 1. This is because convergent validity is a better indicator of construct validity than divergent and known-group hypothesis testing [20]. Downgrading evidence that does not contain results from convergent validity accounts for this at the pooling stage.

RoB, Risk of Bias; MMSE, Mini Mental State Examination; KMO, Kaiser-Meyer-Olkin measure of sampling adequacy;

Appendix 6 Systematic Review Tables

Overview of Studies Table

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[54]	AES-C; AES-I; AES-S	nr	Structural validity; Hypothesis Testing (convergent & divergent).	Community-dwelling (95.8%) and nursing home residents (4.2%).	N=121	Dementia: AD (55.2%); MD (AD-DLB, 14.3%; AD-VaD, 5.7%); DLB (9.5%); VaD (5.7%), FtD (4.8%); 'other dementia' (4.8%).	73.7 (9.4)	47.1%	nr	nr
[55]	AES-C; AES-I; AES-S	English^	Structural validity; Internal consistency; Hypotheses testing (divergent & known groups).	Outpatient and community sample – confirmed all community dwelling via correspondence with author	N=75	MCI (N=57); Cognitively normal (Ctrl N=18)	MCI: 74.5 (8.6, 53 to 86) Ctrl: 75.4 (6.0, 63 to 84) Total: 74.7 (8.0, 53 to 86)	MCI: 70.2% Ctrl: 22.2% Total: 58.7%	MCI: 27.3 (1.9, 23 to 30) Ctrl: 29.4 (0.8, 28 to 30) Total: 27.8 (1.9, 23 to 30)	AES-C: MCI: 60.9±7.7 (39–72) Ctrl: 68.4±4.3 (55–72) Total: 62.7±7.7 (39–72) AES-I: MCI: 61.1 (8.0, 42 to 72) Ctrl: 68.3 (4.5, 58 to 72) Total: 62.8 (7.9, 42 to 72) AES-S: MCI: 63.3 (8.0, 40 to 72) Ctrl: 67.2 (4.2, 56 to 72) Total: 64.3 (7.4, 40 to 72)
[33]	AES-C; AES-I; AES-S	English^	Development (item elicitation and pilot); Structural validity; Internal consistency; Reliability (interrater &test-retest); Hypothesis testing (divergent & known groups).	Community-dwelling	N=123 (N=40 for pilot) (n/a for item elicitation)	Mixed sample: Healthy controls (Ctrl, N=31); Probable AD (N=21); Major Depression (Dep; N=30); Left Hemisphere Stroke (LHS, N=19); Right Hemisphere Stroke (RHS =22).	Ctrl: 68.3 (5.7,nr) AD: 70.8 (7.6,nr) Dep: 71.6 (5.7,nr) LHS: 66.2 (6.6,nr) RHS: 70.1 (5.0,nr) Total: 69.53 (6.03)* 55 to 85)	Ctrl: 45.16% AD: 47.62% Dep: 10.00% LHS: 57.89% RHS: 54.55% Total: 40.65%	Ctrl: 29.1 (1.1, nr) AD: 19.1 (6.5, nr) Dep: 28.0 (1.7, nr) LHS: 25.0 (4.6, nr) RHS: 26.9 (2.3, nr)	AES-C: reported separately for the 2 clinician ratings: Ctrl: 26 (6.2, nr); 25.8 (5.8, nr) AD: 44.4 (11.1, nr); 45.2 (11.7, nr); Dep: 40.5 (9.7, nr); 36.6 (8.3, nr) LHS: 31.9 (9.6, nr); 32.0 (11.7, nr) RHS: 34.7 (7.3, nr); 35.4 (9.6, nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
										AES-I: Ctrl: 26.3 (7.5, nr) AD: 49.1 (9.9, nr) Dep: 41.7 (15.0, nr) LHS: 28.1 (6.9, nr) RHS: 35.4 (10.9, nr)
										AES-S: Ctrl: 28.1 (6.4, nr) AD: 35.5 (8.1, nr) Dep: 38.7 (9.8, nr) LHS: 32.2 (8.6, nr) RHS: 31.6 (6.7, nr)
[71]	AES-I; AES-I-16	German	Internal consistency; Hypothesis Testing (divergent).	Community-dwelling.	N=100 (AES-I N=80.)	Dementia	83.19 (8.32, 59 to 100, N=99)	29%	16.35 (7.60, 0 to 29, N=65)	AES-I: 31.74 (10.43, 8 to 48)
[56]	AES-I; AES-S	Swedish	Structural validity; Internal consistency; Measurement error; Hypothesis Testing (divergent).	Outpatients and community sample – No confirmation from correspondence whether the outpatients were community-dwelling.	N=511 Complete AES-I N=367. Complete AES-S N=496.	Neurodegenerative disease and cognitive impairment: MCS (N=222. AES-I N=192. AES-S N=209) with subgroups of subjective cognitive decline (SCD, N=97) and MCI (N=125). Parkinson's Symptoms (PS, N=88. AES-I N=76. AES-S N=88), with subgroups of PD (PD, N=71); Parkinson's Disease Dementia or Dementia with Lewy Bodies (PDD-DLB, N=17). Ctrl (N=201. AES-I N=135; AES-S N=199)	MCS: 70 (6) MCI: 71 (6) PD: 67 (9) PDD-DLB: 74 (6) Ctrl: 75 (5) Total: 72 (7)	MCS: 44.3%* MCI: 52%* PD: 56.3%* PDD-DLB: 76.5%* Ctrl: 37.8%* Total: 46.4%*	*median (Q1 to Q3)* MCS: 29 (27 to 29) MCI: 27 (26 to 28) PD: 29 (27 to 30) PDD-DLB: 23 (20 to 24) Ctrl: 29 (28 to 30) Total: 29 (27 to 29)	AES-I MCS: 36.2 (10.6, nr) PS: 52.3 (11.4, nr)) Ctrl: 28.7 (8.2, nr) Total: 36.6 (12.9, nr) AES-S MCS: 32.6 (8.8 , nr) PS: 53.3 (10.6, nr) Ctrl: 28.0 (5.7, nr) Total: 34.2 (11.9, nr)
[89]	AES I; AES-S	Italian	Hypothesis Testing (divergent).	Outpatients – No confirmation from correspondence	N=48	Parkinson's Disease (PD)	72.21 (9.01, nr)	64.58%*	22.83 (4.71,nr)	AES-I: 45.14 (13.09, nr) AES-S: 49.85 (10.37, nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
				whether community-dwelling.						
[57]	AES-S	German^	Structural validity; Internal consistency; Hypothesis testing (convergent & divergent).	Author confirmed all community via correspondence.	N=665	Parkinson's Disease Sub-sample of PD excluding comorbidities of dementia or depression (PDexclDd; N=339)	PD: 67.3 (7.90,nr) PDexclDd: 66.52 (7.96,nr)	PD: 67.9% PDexclDd: 66.52%	PD: 27.94 (2.23) PDexclDd: 28.47 (1.58)	PD: 30.63 (9.49) PDexclDd: 27.96 (7.59)
[78]	AES-12PD	German	Internal consistency; Hypothesis testing (convergent & divergent)	Data taken from a study that has been confirmed community via correspondence.	N=339	Parkinson's Disease. (Sample split for analyses: Sample 1: N=170; Sample 2: N=169) Subsample of PDDd: N=42	Sample 1: 68 (nr, nr) Sample 2: 68 (nr, nr)	Sample 1: 70.00% Sample 2: 70.41%	*median (Q1 to Q3)* Samples 1&2: 29 (nr, nr)	*median (Q1 to Q3)* AES: Samples 1&2: 27.0 (nr) AES-12PD: Sample 1: 17.0 (nr) Sample 2: 18.0 (nr)
[34]	AI	French^	Development (item elicitation)	n/a no participants.	n/a	n/a	n/a	n/a	n/a	n/a
			Internal consistency; Reliability (test-retest & interrater), hypothesis testing (convergent, divergent & known groups)	Author advised outpatients via correspondence nut unable to confirm whether community dwelling.	N=115. (Test-retest N=14).	People with neurodegenerative disease or cognitive Impairment: AD (N=60); PD without dementia (N=12), MCI (N=24) Ctrl (N=19). Test-retest: AD only.	AD: 74.90 (7.11, nr) PD: 64.1 (11.9, nr) MCI: 71.67 (5.92, nr) Ctrl: 70.68 (8.21, nr) Total: 72.40 (7.52)*	AD: 45.00 PD: 58.33 MCI: 29.17 Ctrl: 42.11	AD: 22.55 (3.98, nr) PD: 27.2 (3.5, nr) MCI: 28.21 (1.06, nr) Ctrl: 29 (nr, nr)	AI-I AD: 9.20 (10.4, nr) PD: 8.00 (6.0, nr) MCI: 4.21 (8.6, nr) Ctrl: 1.05 (2.0, nr) AI-S AD: 3.74 (5.9, nr) PDexlD: 9.10 (8.3, nr) MCI: 2.47 (3.8, nr) Ctrl: 1.51 (2.9, nr)
[80]	AI-C	Portuguese	Internal consistency; Reliability (interrater); Hypothesis testing (convergent).	nr, but confirmed all community via correspondence	N=175.	Mixed sample: AD (N=55) MCI (N=35) Dep (N=32) PD (N=30) Ctrl (N=23)	AD: 78.4 (nr, 61 to 95) MCI: 69.1 (nr, 60 to 86) Dep: 69.7 (nr, 55 to 88) PD: 66.5 (nr, 42 to 84);	Total: 34.3%	AD: 16.8 (nr, 0 to 27) PD: 26.9 (nr, 18 to 20) Dep: 24.3 (nr, 16 to 30) MCI: 25.4 (nr, 22 to 27)	AI scores nr. Apathy 'diagnosis' according to Robert et al criteria: AD: 63.6% PD: 20% Dep: 68.8% MCI: 0% Ctrl: 0%

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
							Ctrl: 67.3 (nr, 52 to 88) Total: 71.45*		Ctrl: 29.1 (nr, 28 to 30) Total: 23.28*	
[72]	AI-C	French	Internal consistency; Hypothesis testing (convergent).	Outpatients – No confirmation from correspondence whether community-dwelling.	N=40	Cognitive Impairment AD (N=17); MCI (N=12); MD (N=8); VaD (N=2); DLB (N=1)	77.5 (8.01, nr)	45%*	20 (6.73, nr)	nr
[67]	AMI	English	Internal consistency; Hypothesis testing (convergent)	Outpatients – No confirmation from correspondence whether community-dwelling.	N=149	PD (N=102) Ctrl (N=147)	PD: 67.7 (8.1,nr) Ctrl: 66.1 (8.5, nr) All at least 18 to 80	PD: 77.5% Ctrl: 70.75%	ACE-III: PD: 89.4 (9.0, nr) All at least over 50 Ctrl: nr	PD: 35.29% apathetic in at least one AMI subscale
[90]	AS-I	Portuguese	Content validity	Outpatients – No confirmation from correspondence whether community-dwelling.	N=11	Dementia: AD (N=8); FtD (N=3);	AD: 78.3 (4.7) FtD: 55 (8.7) Total: 71.95 (5.59)*	AD: 50.00%* FtD: 33.33%* Total 45.45%*	nr for this sample. Total: 20.64 (3.85)*	22.8 (8.4, 12 to 39)
			Hypothesis testing (convergent & divergent)	Population random sample – No confirmation from correspondence whether community-dwelling.	N=20	Probable or Possible AD	84.1 (5.8)	30%	17.4 (SD=4.7)	23.6 (10.6; 9 to 40)
[59]	AS-S (14/13 item)	English^	Structural validity; Internal consistency.	nr, but confirmed all community via correspondence	N=226	Parkinson's Disease, without dementia.	65.02 (8.84, nr)	66.70%	(N=7) 29.14 (0.69, nr)	10.99 (6.26, nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[36]	AS-S	English^	Development	n/a, no participants	n/a	n/a	n/a	n/a	n/a	n/a
			Internal consistency; Reliability (interrater & test-retest); Hypothesis testing (known groups)	nr. Author unable to access the information.	N=50 (Reliability studies: N=11)	Parkinson's disease, grouped into sub-samples based on apathy and depression scores: PD, no apathy, no depression (PD; N=16) PD, with apathy, no depression, (PDa; N=6) PD, no apathy, with depression, (PDd; N=13) PD, with depression and apathy (PDa&d; N=15)	PD: 67 (9, nr) PDa: 69 (7, nr) PDd: 62 (12, nr) PDa&d: 69 (8, nr) Total: 66.54 (9.26)*	PD: 50% PDa: 66% PDd: 57% Pa&d: 73% Total: 62%*	PD: 28.7 (1.1, nr) PDa: 28.3 (1.2, nr) PDd: 26.3 (4.6, nr) PDa&d: 25.4 (4.5, nr) Total: 27.04 (3.06)*	PD: 7.3 (2.8, nr) PDa: 17.1 (4.0, nr) PDd: 10.0 (2.0, nr) PDa&d: 19.5 (3.3, nr) Total =12.84 (2.87)*
[58]	AS-S AS-HC	Japanese	Structural validity; Internal consistency; Hypothesis testing (divergent).	"Home-care" recipients. Assumed community-dwelling	N=122	Parkinson's Disease	70.9 (7.8, nr)	49.2%	nr	AS-S: 26.6 (8.12, nr) AS-S-11: 21.3 (6.88, nr)
[60]	AS-S	Norwegian	Structural validity; Internal consistency; Hypothesis testing (divergent).	nr. No confirmation from correspondence whether community-dwelling.	N=194	Parkinson's Disease	67.9 (9.0, nr)	59.3%	27.8 (2.3, nr)	15.5 (4.6, 4 to 29) (median =15.0).
[77]	AS - S	Spanish^	Internal consistency; Reliability (test-retest); Measurement error; Hypothesis testing (divergent & known-groups)	Outpatients – No confirmation from correspondence whether community-dwelling.	N=211 (test-retest: N=71)	Parkinson's Disease	67.5 (10.2, nr)	65.5%*	Short Portable Mental Status Questionnaire of Pfeiffer: 1.3 (1.6, nr).	12.7 (7.1, nr)
[61]	AS-S	English^	Structural validity; Internal consistency.	Outpatients. Confirmed community-dwelling via correspondence	N=233	Parkinson's Disease and healthy controls PD (N=157) Ctrl (N=76)	PD: 67.64 (8.27, nr) Ctrl: 66.95 (8.73, nr)	PD: 68.15%* Ctrl: 44.74%*	Mattis dementia rating scale: PD: 138.48 (3.88,nr) Ctrl: 140.46 (3.24,nr)	PD: 11.59 (5.36,nr) HC: 9.21 (4.67,nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[91]	AS-S	Spanish	Internal consistency; Hypothesis testing (convergent; divergent)	nr. Unknown to corresponding author as data not collected.	N=60	Advanced Parkinson's Disease	68.02 (7.43; 50 to 81)	60.70%	nr	11.55 (6.49, 1 to 24)
[42]	BMDS	English^	Development (item elicitation)	n/a, no participants	n/a	n/a	n/a	n/a	n/a	n/a
			Reliability (test-retest).	nr, but scale designed to assess people living in the community	N=38 (test-retest reliability N=18)	Dementia	76 (nr, 59 to 87)	23.68%	nr	24.95 (9.30, nr)
[43]	BSSD	English	Development (item elicitation and pilot)	Item elicitation: n/a no participants Pilot: nr	nr	nr	nr	nr	nr	nr
			Internal consistency; Reliability (interrater & test-retest); Hypothesis Testing (divergent & known groups)	Outpatients – No confirmation from correspondence whether community-dwelling.	N=106 (hypothesis testing: N=83 to 97; reliability: N=20 to 21)	Alzheimer's Disease	72.1 (9.8, 45 to 93)	35% male	Modified MMSE: 26.2 (13.8, 0 to 52)	Global apathy / indifference =31.1% absent; 50.0% minimal to mild; 18.8% moderate to severe. raw scores nr.
[37]	DAIR	English^	Development (item elicitation and pilot);	nr	nr	Mixed sample: People with AD, their carers and clinical researchers.	nr	nr	nr	nr
			Structural validity; Internal consistency; Hypothesis testing (convergent & divergent)	nr Designed to assess people living in environments whose daily activities are not structured, suggesting community-dwelling. No confirmation from correspondence whether	N=100	Alzheimer's Disease	75.00 (8.48; 52 to 92)	50%	18.55 (7.20; 3 to 29) (Unobtainable for 16%)	1.19 (0.69, 0 to 3)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
				community-dwelling.						
[38]	DAS	English (assumed)	Development study (item elicitation)	n/a no participants involved in item elicitation	n/a	n/a	n/a	n/a	n/a	n/a
[75]	DAS	English^	Internal consistency; Hypothesis testing (convergent; divergent)	Outpatients - all confirmed community via correspondence.	DAS-S N=68	Parkinson's Disease without dementia and healthy controls: PD (N=34) Ctrl (N=34)	PD: 68.2 (9.2, nr) Ctrl: 66.1 (9.2, nr)	44.12%	nr	PD: 25.8 (8.7, nr) Ctrl: 21.2 (7.0, nr)
					DAS-I N=60	(sub-sample of those above) PD (N=30) Ctrl (N=30)	nr for this sub-sample	nr for this sub-sample	nr	PD: 25.1 (12.8, nr) Ctrl: 19.7 (9.5, nr)
[74]	DAS	English^	Internal Consistency; Hypothesis testing (convergent & divergent)	Community-dwelling	N=157*	DAS-I Alzheimer's Disease and controls AD (N=102) Ctrl (N=55)	AD: 78.2 (8.5, nr) 82.4% aged 65 and over. Ctrl: 75.0 (6.1, nr)	AD: 51.0%* Ctrl: 50.9%*	AD (N=80): 22.0 (5.3, nr) Ctrl: nr	nr, but AES: AD: 51.7 (11.5, nr) Ctrl: 28.8 (5.2, nr)
						DAS-S AD (N=55, sub-sample of those above) Ctrl (same as above, n=55)	AD: 77.5 (7.9, nr) Ctrl: 75.0 (6.1, nr)	AD: 50.9%* Ctrl: 50.9%*	nr	nr, but AES: AD: 38.9 (9.0, nr)
[62]	DAS-S	Italian	Structural validity, Internal consistency, Hypothesis testing (convergent, divergent & known groups)	Outpatients - all confirmed community via correspondence.	N=207	Parkinson's Disease and controls PD (N=107) Ctrl (N=100)	PD: 66.02 (9.01,nr) Ctrl: 64.52 (8.79,nr)	PD: 60.75%*	PD: 27.63 (2.09,nr)	PD: 25.25 (12.76,nr) (Median (skewness)=23 (1.254)) Ctrl: 21.29 (8.35,nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[63]	bDAS	English	Structural validity	AD: Community-dwelling ALS: nr	N=204	Neurodegenerative Disease AD (N=102) ALS (N=102)	AD: 78.2 (8.5, nr) ALS: 63.8 (11.0, nr) Total: 71.0 (12.1, nr)	AD: 51%* ALS: 70%* Total: 60%*	AD: (N=80): 22.0 (5.3, nr) ALS: nr Total: nr	nr for bDAS AES: AD: 51.7 (11.5, nr) ALS: 33.2 (10.8, nr) Total: 42.4 (14.4, nr)
[76]	bDAS	English^	Internal consistency; Reliability (test-retest).	All confirmed community via correspondence.	N=53 (reliability N=43)	ALS	68.0 (7.5, nr)	83.01%*	ECAS cognitive score: 107.0 (14.1,nr)	nr for total score DAS-I subscales: Executive: 6.1 (4.8, nr) Emotional: 8.9 (4.2, nr) Initiation: 12.1 (5.5, nr) b-DAS Executive: 2.0 (2.0, nr) Emotional: 2.9 (1.9, nr) Initiation: 4.3 (2.6, nr)
[81]	DEX	Japanese	Reliability (test-retest); Hypothesis testing (convergent & divergent)	Outpatients.	N=122 (reliability N=44)	Alzheimer's Disease	72.0 (7.7, nr)	37.70%*	20.8 (2.0, nr)	nr
[68]	FrSBe-I	English^	Content validity (cognitive interview)	Outpatients - all confirmed community via correspondence.	N=10	People attending neuropsychological evaluation. 90% had memory complaints. Diagnoses nr.	nr	nr	nr	nr
			Structural validity; Internal consistency; hypothesis testing (groups & divergent);	Outpatients - all confirmed community via correspondence.	N=494	Mixed sample: Dementia: AD (19.3%*), VaD (4.9%); Dementia not otherwise specified (4.1%); MD (4.5%); FTD (4.1%); DLB (1.8%). PD (16.6%). MCI (12.5%). Cognitive disorder not otherwise specified (CDNOS, 8.8%). Frontal stroke (7.2%). Head injury (2.1%).	69.92 (13.96, 19 to 95)	47.04%*	nr	Original FrSBe-apathy: PD=33.29 (12.71); AD =37.24 (10.18); Frontal impairment =38.18 (10.35) Revised FrSBe-apathy: PD=27.24 (10.13); AD =29.71 (7.83); Frontal impairment =30.21 (8.08)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
						Other neurological disorder (<1%).				
[64]	FrSBe-I	English^	Structural validity; Internal consistency	Outpatients - all confirmed community via correspondence.	N=304	Older adults with memory complaints: Dementia (N=166) MCI (N=63) No definitive diagnosis (NDD; N=28) Ctrl (N=47)	79.12 (8.05; 52 to 99)	28.29%*	nr	86.12 (24.39)
[46,47]	GDS-30	English^	Development (Item elicitation and pilot study)	Item elicitation: n/a no participants Pilot: Community dwellers (N=20) and inpatients (N=51).	N=71	Healthy older adults (Ctrl: N=20) Depressed older pts (Dep: N=51)	nr. All over 55.	nr	nr	nr
[92]	GDS-3A	Dutch^	Hypothesis testing (convergent validity)	Community-dwelling	Study 1 N =427	Older adults with mild cognitive deficits	81.3 (4.6, nr) All at least 75 and over	39.8%*	*median (Q1 to Q3)* 26 (25 to 27)	GDS-3a score:0 =52.8%; 1=30.7%; 2=12.2%; 3=4.4% AS: 11.3 (4.7)
					Study 2 N=1118	Older adults with depressive symptoms	81.8 (4.9, nr) All at least 75 and over	38.9%*	*median (Q1 to Q3)* 28 (27 to 29)	GDS-3a: 0 =64.2%; 1 =25.6%; 2 =9.3%; 3 =0.89% AS: 7.5 (4.6, nr)
[69]	GDS-6A	English^	Internal consistency, Hypothesis testing (divergent & known groups)	Community-dwelling	N=140	Mixed sample: Dementia: AD (29.3%); VaD (29.3%); MD (13.6%) Cognitive disorder not specified or MCI (CNS-MCI, 17.1%) Other (6.4%); None (2.1%) (2.2% nr)	78.2 (7.23, nr) All at least 65 or over	35.0%*	24.86 (3.35, nr)	GDS-6a: 1.66 (1.39, nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[82]	GIP-a-s GIP-a-d	Dutch	Reliability (test-retest); Measurement error.	All confirmed community via correspondence.	N=109 Complete and analysed: N=56.	Mixed sample: Dementia: AD (82%); VaD (13%); Other dementia (3%); Other (affective disorder or other cognitive disorder, 2%)	*median (Q1 to Q3, range)* 80 (75.5 to 84, 53 to 96)	42.2%*	*median (Q1 to Q3, range)* Cognitive Screening test: 13.3 (10.4 to 16, 3.5 to 20) Amsterdam Dementia Screening test 3: 0 (-2 to 1, -5 to 4) Amsterdam Dementia Screening test 5: 1 (-1 to 3, -5 to 8).	N=56: GIP-a-s: 2.2 (2.3, 0 to 9) GIP-a-d: 2.8 (3.5, 0 to 15)
[49]	IMD	Italian^	Development (item elicitation)	n/a no participants	n/a	n/a	n/a	n/a	n/a	n/a
			Hypothesis testing (divergent)	Sample 1: Some Community-dwelling and some institutionalised. Author unable to confirm proportion.	N=236	nr, but at least some healthy older adults. Mild to moderate functional impairment (52.5%). Severe functional impairment (24.8%).	74.2 (6.8, nr)	40.6%*	19.4 (4.3, nr)	nr
				Sample 2: nr. Author unable to confirm.	N=203	Dementia	74.1 (5.56; 63 to 83)	33.99%*	19.7 (2.61, 15 to 23)	5.4 (3.15)
[50,104]	KBCI	English^	Development (item elicitation)	nr	nr	People with TBI, their carers, and TBI rehabilitation specialists.	nr	nr	nr	nr
			Development (item refinement)	panel1: nr. panel 2 & 3: n/a.	N=14	Panel 1: carers for people with TBI (N=4) Panel 2: clinical psychologists (N=3)	nr	nr	nr	nr

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
						Panel 3: clinical neuropsychologists (N=7)				
[93]	KBCI-a	English^	Hypothesis testing (divergent)	Outpatients. No reply from author.	N=97	Mixed sample: Ctrl (31%) MCI (18%) Probable AD (7%) Other (depression, CDNOS, PD, DLB, and possible AD)	72.34 (9.05, nr)	nr	26.89 (2.63, nr)	nr
[39]	LARS	French; English	Development	n/a – no participants involved.	n/a	n/a	n/a	n/a	n/a	n/a
[83]	LARS - C	Spanish	Reliability (interrater & test-retest); Hypothesis Testing (convergent)	Community-dwelling ("non-institutionalised")	N=151 (test-retest N=16, interrater N=21)	Dementia (Dem, N=101) and healthy controls AD (N=43) FtD (N=41) Primary Progressive Aphasia (N=17) Ctrl (N=50)	Dem: 74.3 (7.7, nr) Ctrl: 72.0 (9.7, nr)	Dem: 45.5%* Ctrl: 38%*	Dem: 21.59 (6.21, nr) Ctrl: 28.72 (1.42, nr)	Dem: -0.16 (18.50, nr) Ctrl: -29.54 (5.44, nr)
[70]	LARS-I	French^	Internal consistency; Reliability (interrater & test-retest); Hypothesis Testing (convergent)	Correspondence with author confirmed all community	N=60 (interrater N=34, test-retest N=29)	Parkinson's Disease: PD without dementia (PDexclD, N=43) PD with dementia (PDD, N=17)	PDexclD: 64.74 (9.29, nr) PDD: 69.53 (9.06, nr) Total: 66.10 (9.23)*	PDexclD: 67.44%* PDD: 35.29%*	nr	-16.18 (11.99, nr)
[65]	LARS - C	Spanish	Content validity; Structural validity; Internal consistency ; Reliability (interrater & test-retest); Hypothesis testing (convergent & divergent)	nr. No confirmation from correspondence whether community-dwelling.	N=200 (content validity and reliability N=30)	Parkinson's Disease and healthy controls PD (N=130) Ctrl (N=70)	PD: 71.6 (8.1, nr) Ctrl: 69.4 (8.7, nr)	PD: 60.0%* Ctrl: 55.7%*	MEC: PD: 30.7 (3.8, nr) Ctrl: 33.3 (1.7, nr)	PD: -14.5 (9.1, nr) Ctrl: -25.0 (5.5, nr)
[94]	NPI	Korean	Hypothesis Testing (known groups).	Assessment setting suggests outpatients. No	N=141 (test-retest N=29)	Dementia (N=92) and healthy controls: AD (N=43)	Dem: 67.5 (9.7, 38 to 85)	Dem: 47.8%* Ctrl: 34.7%*	Dem: 17.5 (6.8, 0 to 29)	NPI-apathy total nr. Dem: Prevalence: 77.2%.

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
				confirmation from correspondence whether community-dwelling.		VaD (N=32) FtD (N=11) Other dementia (N=6) Ctrl (N=49)	Ctrl: 66.9 (8.4, 51 to 82)		Ctrl: 26.3 (2.3,19 to 30)	Frequency: 2.52 (1.67; 0 to 4) Severity: 1.75 (1.18; 0 to 3) Ctrl: Prevalence =6.1%. Frequency =0.06 (0.24; 0 to 1) Severity =0.06 (0.24; 0 to 1)
[51]	NPI	English^	Development (item elicitation and Delphi study of comprehensiveness)	Item elicitation: n/a no participants Delphi study: n/a professionals	N=10	Geriatric psychiatrists, behavioural neurologists, and neuropsychologists	n/a	n/a	n/a	n/a
			Reliability (interrater & test-retest)	Community-dwelling	N=80 (interrater N=45, test-retest N=20)	Dementia (Dem) and healthy controls: AD (N=20) VaD (N=9) Other dementia (N=11) Ctrl (N=40)	75.7 (56 to 90)	Dem: 55.00%* Control: 50.00%*	Dem: 19.2 (0 to 29) Control: 28.4 (25 to 30)	NPI-apathy total nr. Frequency: 2.83 (1.55; 0 to 4) Severity: 1.35 (0.83; 0 to 3)
[85]	NPI	Icelandic	Reliability (test-retest); Hypothesis testing (known groups).	Community-dwelling	N=38 (test-retest N=6)	Dementia: AD (N=19) VaD (N=19)	78.84 (6.66; 59 to 89)	47%	19.26 (5.95; 1 to 29)	nr for total sample. Reported separately for two different severity groups (N in each group nr). Less severe dementia: 4.69 (3.72, nr) More severe dementia: 7.45 (4.45, nr)

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[73]	NPI	Farsi	Internal consistency; Reliability (interrater & test-retest); Hypothesis testing (convergent, divergent & known groups)	51% living with family, suggesting at least majority community dwellers. No confirmation from correspondence	N=100. (interrater N=50, test-retest reliability N=30, hypothesis testing N=50)	Dementia and healthy controls. Dem (N=100) Ctrl (N=49)	Dem: 74.5 (8.3, 60 to 90) Ctrl: 74.3yrs (8.5)	Dem: 47% Ctrl: 51%	nr for total sample. Hypothesis testing (N=50): Dem: 11.3 (7.5, nr) Ctrl: 29.4 (1.0, nr)	NPI-apathy total nr. Prevalence: 74% Frequency 2.5 (1.7, nr) Severity 1.6 (1.1, nr)
[79]	NPI	Spanish	Internal consistency; Reliability (interrater); Hypothesis testing (convergent)	Outpatients – No confirmation from correspondence whether community-dwelling.	Total N=63. (interrater N=39)	Mixed sample: Dem (N=44) Dep (N=6) Ctrl (N=13)	72.76 (9.67; 35 to 85)	49.21%*	nr	NPI-apathy total nr. Prevalence: 56%
[95]	NPI	Greek	Hypothesis testing (convergent)	Outpatients. Author correspondence confirmed all community.	N=29	Dementia	71.05 (5; 60 to 84)	60%	12.4 (6.0; 0 to 24)	5.8 (4.4, nr)
[86]	NPI	Chinese	Reliability	Community dwelling	N=91	Dementia and healthy controls. Dementia (Dem, N=62*): AD (N=41), VaD (N=16), Other (N=5) Ctrl (N=29)	Dem: 76.4 (7.0; 54 to 88). Ctrl: 74.9 (4.7; 68 to 86)	Dem: 22.58%* Ctrl: 72.41%*	Dem: 12.7 (5.9; 0 to 25.) Ctrl: 27.5 (2.2; 23 to 30.)	nr
[84]	NPI	Brazilian Portuguese	Reliability (interrater & test-retest)	Outpatients. Author correspondence confirmed all community	N=36	Alzheimer's Disease	78.78 (7.48)	22%*	7.06 (6.92)	NPI-apathy total nr. Severity: 5.31 (4.91) Frequency: 1 =33%, 2 =3%, 3 =64%.
[96]	NPI	Dutch	Hypothesis testing (divergent validity)	83.33% community-dwelling	N=24	Mixed sample: Dementia: AD (N=19), FtD (N=1), MD (N=1) Stroke (N=2) Amnestic disorder (N=1)	74.3 (10.4, nr)	33.33%*	21.5 (4.6; 12 to 29)."	nr

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[66]	NPI-A	English^	Structural Validity; Internal consistency.	Outpatients. Author was unable to confirm whether community-dwelling.	N=124	Dementia: AD (N=62) VaD (N=43) MD of AD+VaD (N=19)	79.8 (6.1; 61 to 91)	21.77%*	22.6 (3.5; 13 to 29)	8.89 (8.5, nr)
[87]	NPI-C	English^, French^, Greek^, Italian^, Hungarian^, Portuguese^, Spanish^	Content validity (further item elicitation and Delphi study)	Item elicitation: n/a no participants Delphi study: n/a professionals	Delphi study: N=8	Experts in dementia research	n/a	n/a	n/a	n/a
			Reliability (interrater); Hypothesis Testing (convergent)	79.5% community-dwelling	N=128	Alzheimer's Disease	75.7 (9.0; 54 to 94)	nr	17.6 (7.0; 0 to 28).	NPI-C-apathy total nr. AES (N=113): 33.1 (11.3; 0 to 51)
[88]	NPI-C	Portuguese	Reliability (interrater); Hypothesis Testing (convergent)	Author confirmed all community via correspondence	N=156	Dementia	76.7 (nr, nr)	26.28%*	17.2 (nr, nr)	NPI-C-apathy total nr. AI: 5.9 (nr, nr)
[52]	UPDRS	English	Development (item elicitation and review of comprehensibility)	n/a no participants involved	n/a	n/a	n/a	n/a	n/a	n/a
[100]	UPDRS	Spanish^	Hypothesis Testing (convergent)	Outpatients – No confirmation from correspondence whether community-dwelling.	N=168 (convergent validity N=164)	Parkinson's Disease	65.9 (9.8, nr)	57%	24.4 (5.4, nr)	nr
[99]	UPDRS	Norwegian^	Hypothesis Testing (convergent)	nr. Participants were assessed in outpatient clinics, at home and in nursing homes. No confirmation from correspondence regarding proportion of community-dwellers.	N=89 (convergent N=58)	Parkinson' Disease (41.4% with cognitive impairment)	74.2 (8.8, nr)	44.8%	23.0 (7.2, nr)	UPDRS-apathy item nr. 17% had apathy according to diagnostic criteria.

Reference	Measure	Language of measure	Measurement properties investigated	Residential status	N	Population (N of each subgroup, or % where N not possible to calculate)	Mean age (SD, range)	Gender (% Male)	Cognitive status Mean MMSE (SD, range) unless otherwise stated	Apathy score Mean (SD, range)
[98]	UPDRS	English^	Hypothesis Testing (convergent)	Outpatients. Confirmed all community via correspondence with authors	N=301	Parkinson's Disease	67.8 (10.6; 30 to 90)	63%	nr	1.14 (1.1; 0 to 4) AS =13.7 (6.9) range =0 to 31. AS≥14: 50%
[105,106]	mds-UPDRS	English	Development (Item elicitation [including adaptation of items from UPDRS to create mds-UPDRS], Pilot study)	nr	nr	Item elicitation: nr. Pilot study: Part 1: Patients (PD, N=80), carers (N=nr) and professionals (N=nr) Part 2: Patients (N=32) and professionals (N=14)	nr	nr	nr	nr
[97]	mds-UPDRS	Hungarian	Hypothesis testing (convergent)	nr. Correspondence with author confirmed majority community.	N=584	Parkinson's Disease PD with neurocognitive disorder (N=310) PD with depression (N=217) Apathy status: No apathy (N=477), Apathy (N=107)	*median (Q1 to Q3)* No apathy: 67 (61 to 73. Apathy: 68 (61 to 75)	No apathy: 60.2% Apathy: 52.3%	*median (Q1 to Q3)* No apathy: 28, (27 to 29) Apathy: 27 (24 to 28)	*median (Q1 to Q3)* LARS: No apathy: -26 (-30 to -21) Apathy: -15 (-22 to 5)

Note: Where the study had used secondary data, the primary data sources were sought to gain the necessary information where it was not available in the article in question.

^ Assumed based on location of study and/ or nationality of participants.

*Calculated by authors

Abbreviations: AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; ACE, Addenbrooke's Cognitive Examination; AD, Alzheimer's Disease; AES-12PD, Apathy Evaluation Scale 12-item Parkinson's Disease; AES-C, Apathy Evaluation Scale Clinician; AES-I, Apathy Evaluation Scale Informant; AES-S, Apathy Evaluation Scale Self; AI, Apathy Inventory; AI-C, Apathy Inventory Clinician; AI-I, Apathy Inventory Informant; ALS, Amyotrophic Lateral Sclerosis; AMI, Apathy Motivation Index; AS-S, Apathy Scale Self; AS-I, Apathy Scale Informant; bDAS, brief Dementia Apathy Scale; BMDS, Behavioural and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; CD, Current Depression; CDNOS, Cognitive Disorder Not Otherwise Specified; Ctrl, Healthy Controls; DAIR, Dementia Apathy Interview Rating; DAS, Dementia Apathy Scale; DAS-I, Dementia Apathy Scale Informant; DAS-S, Dementia Apathy Scale Self; Dem, Dementia; Dep, Depression; DEX, Dysexecutive Questionnaire; DLB, Dementia with Lewy Bodies; FrSBe-I, Frontal Systems Behavior Scale Informant; FtD, Frontotemporal Dementia; GDS, Geriatric Depression Scale; GIP, Behavioral Rating Scale for Psychogeriatric Inpatients ; IMD, Index of Mental Decline; KBCI, Key Behaviors Change Inventory; LARS, Lille Apathy Rating Scale; LARS-C, Lille Apathy Rating Scale Clinician; LARS-I, Lille Apathy Rating Scale Informant; LHS, Left Hemisphere Stroke; MCI, Mild Cognitive Impairment; MCS, Mild Cognitive Symptoms; MD, Mixed Dementia; mds-UPDRS, Movement disorder Society Unified Parkinson's Disease Rating Scale; NPI, Neuropsychiatric Inventory; NPI-A, Neuropsychiatric Inventory Alternative; NPI-C, Neuropsychiatric Inventory Clinician; nr, not reported; PD, Parkinson's Disease; PDa&d, Parkinson's Disease with apathy and depression; PDa, Parkinson's Disease with apathy; PDD, Parkinson's Disease Dementia; PDd, Parkinson's Disease with depression; PDDd, Parkinson's Disease with dementia and depression; PDexclD, Parkinson's Disease without dementia;

PDexclDd, Parkinson's Disease without dementia or depression; PS, Parkinsonian Symptoms; RD, Remitted Depression; RHS, Right Hemisphere Stroke; SCD, Subjective Cognitive Decline; UPDRS, Unified Parkinson's Disease Rating Scale; VaD, Vascular Dementia.

Risk of Bias and results of development and content validity studies

Systematic Review - Risk of bias and results of development and content validity studies

Reference	Measure	Met criteria? (Y/N)	Description	Relevance		Comprehensiveness		Comprehensibility	
				Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[40,41]	AD-RD	Y	Development study: qualitative interviews for concept elicitation and expert review to refine the measure.	Apathy subscale: Inadequate	Construct of apathy is not clear. Items were all based on their mention by at least two carers (informal or formal) in qualitative interviews about how people with dementia express their mood. No justification was provided for the response options or recall period. (1?)	Doubtful	Patients or carers were not asked specifically about the measure. Expert review lead to reduction of items to avoid repetition. However, it was unclear what professionals were asked. (1?).	Doubtful	Patients or carers were not asked specifically about the measure. Expert review lead to modified instructions. However, it was unclear what professionals were asked (1?).
[33]	AES	Y	Development study and pilot study.	Inadequate	Construct of apathy is clear. Items were developed from the literature, professionals, and authors' observations and opinions of people with apathy, but participants not involved in eliciting items and observations not reported on. (1?).			Doubtful	Unclear what participants were asked. 14 items were removed from the preliminary item pool due to poor comprehensibility. (1?).
[34]	AI	Y	Development study.	Inadequate	Construct of apathy is clear. Items were developed from the literature and diagnostic criteria, but participants not involved in eliciting items. (1?).				
[35]	AMI	N	Development study.	Inadequate	Construct of apathy is clear. Items were developed from the relevant items of the LARS and by professionals. Participants were not involved in eliciting items. (1?).				
[36]	AS	Y	Development study (Adaptation of AES to make AS.)	Inadequate	Construct of apathy is clear. Participants not involved in eliciting items. Most relevant items of AES were selected by 2 professionals (S. Starkstein personal, communication, October 01, 2018). (1?).	Doubtful	Pilot study conducted with participants with neurological disorders, but not published, so unable to rate. New items were included by 2 professionals (S. Starkstein personal, communication, October 01, 2018). (1?).	Inadequate	Pilot study conducted with participants with neurological disorders, but not published, so unable to rate. Some items were modified by 2 professionals (S. Starkstein personal, communication, October 01, 2018). (1?).
[90]	AS-I	Y	Content validity study.					Doubtful	Unclear what participants were asked. Participants showed good understanding and no modifications were required (1?).
[42]	BMDS	Y	Development study.	Inadequate	Constructs of behaviour and mood, and apathy were not clear. Items were developed from the literature and author				

Reference	Measure	Met criteria? (Y/N)	Description	Relevance		Comprehensiveness		Comprehensibility	
				Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
					opinion, but participants not involved in eliciting items. (1?).				
[43]	BSSD	Y	Development study and pilot study.	Inadequate	Items were developed from professionals and previous measures, but participants not involved in eliciting items. (1?).	Doubtful	Multiple pilot studies conducted to refine scale, but methods and results not reported. (1?).	Doubtful	Multiple pilot studies conducted to refine scale, but methods and results not reported. (1?).
[37]	DAIR	Y	Development study and pilot study.	Doubtful	Construct of apathy is clear. Items refer to apathy, and were developed with participation from people with dementia and carers. No justification was provided for the response options or recall period. (1+/-).	Doubtful	Unclear what participants were asked. (1?).	Doubtful	Unclear what participants were asked. (1?).
[38]	DAS	Y	Development study.	Inadequate	Items were developed from existing scales and experts, but participants not involved in eliciting items. (1?).				
[68]	FrSBe-11a	Y	Content validity: cognitive interviewing study					Doubtful	27% items had no discrepancies, with 82% of items having acceptable discrepancy*. However, participants do not appear to have been asked about the comprehensibility of instructions or response options. (1?)
[68]	FrSBe-14a	Y	Content validity: cognitive interviewing study					Doubtful	21% items had no discrepancies, with 86% of items having acceptable discrepancy*. However, participants do not appear to have been asked about the comprehensibility of instructions or response options. (1?)
[46,47]	GDS	N	Development and pilot study (as a measure of depression)	Inadequate	Items were developed from professionals, but participants not involved in eliciting items. (1?).			Doubtful	Reported that patients accepted the measure, but methods by which this was ascertained were unclear. (1?)
[49]	IMD	Y	Development	Inadequate	Items were developed from existing measures and professionals, but participants not involved in eliciting items. (1?).				

Reference	Measure	Met criteria? *(Y/N)*	Description	Relevance		Comprehensiveness		Comprehensibility	
				Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[50,104]	KBCI	N	Development and pilot	Doubtful	Construct of apathy clear. Items were developed from the literature and interviews with patients, carers and professionals. Methods not clear. No justification for response options and recall period not clear. Patients and carers were later asked to rate the importance of items, and the majority were rated very or extremely important, but exact ratings not reported. (1+/-).	Doubtful	Patients and carers did not suggest any additional items. However, items were later removed after another phase in the development, so comprehensiveness may have changed. Method not clear. (1?)	Doubtful.	Patients and carers were asked about comprehensibility and no changes were suggested. Professionals were asked about comprehensibility and 15 items were re-worded. Methods and focus not clear (e.g. whether they were asked about each item, response options and recall period) (2?)
[39]	LARS	N	Development	Inadequate	Items were developed from Marin's concept of apathy and authors' clinical experience, but no systematic process and participants not involved in eliciting items. (1?).				
[65]	LARS	Y	Pilot study	Doubtful	Participants asked about relevance, but results not reported. Methods and focus not clear (e.g. whether they were asked about each item, response options and recall period) (1?)			Doubtful	Participants asked about comprehensibility and format. Methods and focus not clear (e.g. whether they were asked about comprehensibility of instructions and response options as well as items) (1?)
[51]	NPI	N	Development and Delphi study	Inadequate	Items developed from the literature, but participants not involved in eliciting items. (1?).	Doubtful	Delphi panel of 10 professionals. Assessed "whether the essential elements of the behavior were captured" in each domain by rating screening and sub questions from 1 (well assessed) to 4 (poorly assessed). Apathy: screening questions mean score = 1.3; sub-questions mean score = 1.4. No assessment of comprehensiveness by participants. (1?)		
[87]	NPI-C	Y	Content validity (adaptation)	Doubtful	New items added from symptoms listed by alternative measures. Items were selected that were consistent with diagnostic	Doubtful	Delphi panel of 8 professionals. Unclear what was asked. (1?)	Doubtful	Delphi panel of 8 professionals. Unclear what was asked. (1?)

Reference	Measure	Met criteria? (Y/N)	Description	Relevance		Comprehensiveness		Comprehensibility	
				Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
					criteria 2009. Participants not involved in eliciting new items. (1?)				
[52]	UPDRS	N	Development study	Inadequate	Expert group elicited items from existing measures, but participants not involved in eliciting items. (1?).			Inadequate	Authors reviewed comprehensiveness of preliminary items. Changes were made and final version does not appear to have been reviewed. (1?)
[105,106]	mds-UPDRS	N	Development (Adaptation of UPDRS but involved new item elicitation and pilot study)	Apathy subscale: Inadequate	Expert group elicited items from literature, existing measures, clinical experience and participant survey, though methods not described in sufficient detail. Justification provided for response options but not recall period. (1?).			Doubtful	Comprehensiveness of preliminary items was reviewed by participants and professionals in a qualitative, then quantitative study. Items, instructions and response options were assessed. Unsure if recall period discussed. Changes were made in the first round and then again in the second round. (1?)

Note: Studies only listed if they assessed content validity in some way or were a study describing the development of a measure. Some studies have multiple citations as multiple articles or similar (e.g. PhD thesis) were published on the same study. Blank cells indicate this measurement property was not investigated by the study.

Quality of measurement property: Number of studies in parenthesis followed by rating: +, Sufficient; +/-, Inconsistent; -, Insufficient; ?, Indeterminate.

* Acceptable discrepancy was defined by the authors of the study as less than 30% of participants interpreting the items meaning in the way it was intended [68].

Abbreviations: AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; AES Apathy Evaluation Scale; AMI, Apathy Motivation Index; AI, Apathy Inventory; AS, Apathy Scale; AS-I, Apathy Scale Informant; BMDS, Behavioural and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; DAIR, Dementia Apathy Interview Rating; DAS, Dimensional Apathy Scale; FrSBe-11a, Frontal Systems Behavior Scale 11 item apathy subscale; FrSBe-14a, Frontal Systems Behavior Scale 14 item apathy subscale; GDS, Geriatric Depression Scale; IMD, Index of Mental Decline; KBCI, Key Behaviors Change Inventory; LARS, Lille Apathy Rating Scale; mds-UPDRS, Movement Disorder Society Unified Parkinson's Disease Rating Scale; NPI, Neuropsychiatric Inventory; NPI-C, Neuropsychiatric Inventory Clinician; UPDRS, Unified Parkinson's Disease Rating Scale.

Unable to obtain development articles for: Dysexecutive Questionnaire (DEX), FrSBe and Behavioral Rating Scale for Psychogeriatric Inpatients (GIP).

Reviewer Rated Content Validity

Systematic Review - Reviewer rating of content validity

Measure	Relevance		Comprehensiveness (quality rating)	Comprehensibility (quality rating)	Overall validity	
	Older adults (quality rating)	Dementia & MCI (quality rating)			Older adults	Dementia & MCI
AD-RD	Unable to obtain the full list of items and instructions.					
AES	94% relevant to apathy. 100% relevant to older adults. 94% relevant to research context. Response options appropriate. Suggested recall period too long, but personalised recall period also possible. (1+).	AES-I & AES-S: 94% relevant to apathy. 100% relevant to people with dementia. 94% relevant to research context. Response options appropriate. Suggested recall period too long, but personalised recall period also possible. (1+). AES-C: 94% relevant to apathy. 78% relevant to people with dementia, as some items based on where some items are rated based on patient free-recall. 94% relevant to research context. Response options appropriate. Suggested recall period too long, but personalised recall period also possible. (1+/-).	3 domains of apathy included. (1+).	AES-I & AES-S: 94% appropriately worded. 72% match response options. (1+/-). AES-C: has additional guidance around this so AES-C response options deemed appropriate. (1+).	Sufficient (AES-I & AES-S: 2+, 1+/-; AES-C: 3+)	Sufficient (2+, 1+/-)
AI	100% items relevant to apathy, older adults and the research context. Response options appropriate for AI-C and AI-I, but not for AI-S. Recall period referencing onset of disease not appropriate for older adults, but personalised recall period possible. (Using the given recall period: 1+/-. Using the personalised recall period: 1+.)	100% items relevant to apathy, people with dementia and the research context. Response options appropriate for AI-C and AI-I, but not for AI-S. Recall period of since onset of disease too long for people with dementia, but personalised recall period possible. (Using the given recall period: 1+/-. Using the personalised recall period: 1+.)	3 domains of apathy included. (1+).	0% of items appropriately worded. (1-)	Inconsistent (Given recall period: 1+, 1-, 1+/-; Personalised recall period: 2+, 1-)	Inconsistent (Given recall period: 1+, 1-, 1+/-; Personalised recall period: 2+, 1-)
AMI	78% relevant to apathy. 100% relevant to older adults. 100% relevant to research context. Response options and recall period appropriate. (1+/-).	78% relevant to apathy and to older adults. 100% relevant to research context. Response options and recall period appropriate. (1+/-).	3 domains of apathy included. (1+).	100% of items appropriately worded. 100% match response options. (1+).	Sufficient (2+, 1+/-)	Sufficient (2+, 1+/-)
AS	93% relevant to apathy. 93% relevant to older adults.100% relevant to research	93% relevant to apathy. 100% relevant to people with dementia and the research	3 domains of apathy included. (1+).	93% of items appropriately worded. 57% match response options (1+/-)	Sufficient (2+, 1+/-)	Sufficient (2+, 1+/-)

Measure	Relevance		Comprehensiveness (quality rating)	Comprehensibility (quality rating)	Overall validity	
	Older adults (quality rating)	Dementia & MCI (quality rating)			Older adults	Dementia & MCI
	context. Response options appropriate. Recall period too long. (1+).	context. Response options appropriate. Recall period too long. (1+)				
BMDS	55% relevant to apathy. 100% relevant to older adults and the research context. Response options appropriate. Recall period uncertain. (1+/-).	55% relevant to apathy. 100% relevant to people with dementia and the research context. Response options appropriate. Recall period uncertain. (1+/-).	Emotional dimension missing. (1-).	100% of items appropriately worded, but combination with response options produces double negatives. (1+/-).	Inconsistent (1-, 2+/-)	Inconsistent (1-, 2+/-)
BSSD	71% relevant to apathy. 100% relevant to older adults and research context. 14% response options appropriate. Recall period appropriate. (1+/-)	71% relevant to apathy. 100% relevant to people with dementia and research context. 14% response options appropriate. Recall period appropriate. (1+/-)	3 domains of apathy included. (1+).	86% of items (questions directed at informants) appropriately worded. 100% match response options. (1+).	Sufficient (2+, 1+/-)	Sufficient (2+, 1+/-)
DAIR	94% items relevant to apathy. 0% relevant for healthy older adults due to mandatory follow-up question relating to "illness". Response options appropriate. Recall period too long. (1+/-).	94% items relevant to apathy. 100% relevant for people with dementia. Response options appropriate. Recall period too long. (1+).	3 domains of apathy included. (1+).	100% items appropriately worded. 81% match the response options. (1+/-).	Inconsistent (1+, 2+/-).	Sufficient (2+, 1+/-).
DAS	DAS: 79% items relevant to apathy. bDAS: 67% items relevant to apathy Both versions: 100% relevant to older adults. Response options appropriate. Recall period too long. (1+/-).	DAS: 79% items relevant to apathy. bDAS: 67% items relevant to apathy Both versions: 100% relevant to people with dementia. Response options appropriate. Recall period too long. (1+/-).	3 domains of apathy included. (1+).	100% of items appropriately worded. 100% match response options. (1+).	Sufficient (2+, 1+/-).	Sufficient (2+, 1+/-).
DEX	63% items relevant to apathy. 100% relevant to older adults and research context. Complete response options not available. Recall period appropriate. (1+/-).*	63% items relevant to apathy. 100% relevant to people with dementia and research context. Complete response options not available. Recall period appropriate. (1+/-).*	3 domains of apathy included. (1+).*	Full wording not available, but 75% of items appear appropriately worded. Complete response options not known. (1?).	Inconsistent (1+, 1+/-, 1?)*	Inconsistent (1+/-, 1+/-, 1?)*

Measure	Relevance		Comprehensiveness (quality rating)	Comprehensibility (quality rating)	Overall validity	
	Older adults (quality rating)	Dementia & MCI (quality rating)			Older adults	Dementia & MCI
FrSBe	FrSBe-6a: 83% relevant to apathy. 100% relevant to older adults. FrSBe-11a: 82% relevant to apathy. 91% relevant to older adults. FrSBe-14a: 86% relevant to apathy. 93% relevant to older adults. And all versions: 100% relevant to research context. Response options not available. Recall period not appropriate for older adults. (1+/-).*	FrSBe-6a: 83% relevant to apathy. 100% relevant to older adults FrSBe-11a: 82% relevant to apathy. 91% relevant to people with mild dementia. FrSBe-14a: 86% relevant to apathy. 93% relevant to mild dementia. And all versions: 100% relevant to research context. Response options not available. Recall period not appropriate for people with dementia. (1+/-).*	All versions: 3 domains of apathy included. (1+).*	6a: Full wording not available, but items suggests that 67% appropriately worded. Response options not available. (1?). 11a: Full wording not available, but items suggests that 91% appropriately worded. Response options not available. (1?). 14a: Full wording not available, but items suggests that 86% appropriately worded. Response options not available. (1?).	Inconsistent (1+, 1+/-, 1?)*	Inconsistent (1+, 1+/-, 1?)*
GDS-3a	67% of items are relevant to apathy. All items relevant to older adults and the research context. Dichotomous response options not appropriate. Recall period appropriate. (1+/-).	67% of items are relevant to apathy. All items relevant to people with dementia and the research context. Dichotomous response options not appropriate. Recall period appropriate. (1+/-).	Emotional dimension of apathy is missing (1-).	100% appropriately worded and match response options. (1+).	Inconsistent (1+, 1-, 1+/-)	Inconsistent (1+, 1-, 1+/-)
GDS-6a	50% of items are relevant to apathy. All items relevant to older adults and the research context. Dichotomous response options not appropriate. Recall period appropriate. (1+/-).	50% of items are relevant to apathy. All items relevant to older adults and the research context. Dichotomous response options not appropriate. Recall period appropriate. (1+/-).	3 domains of apathy included. (1+).	100% appropriately worded and match response options. (1+).	Sufficient (2+, 1+/-)	Sufficient (2+, 1+/-)
GIP-9a	44% of items relevant to apathy. 89% relevant to older adults in the community. 100% relevant to research context. Recall period appropriate. Response options not available. (1+/-).*	44% of items relevant to apathy. 89% of items relevant to people with dementia in the community. 100% relevant to research context. . Recall period appropriate. Response options not available. (1+/-).*	Emotional dimension of apathy is missing. (1-).*	Full wording and official English translation of items not available, but authors translation suggest 89% appropriately worded. Response options not available. (1?).*	Inconsistent (1-, 1+/- 1?)	Inconsistent (1-, 1+/- 1?)
IMD	100% of items relevant to apathy, older adults and the research context. Response options and recall period not available. (1?).	100% of items relevant to apathy, people with dementia and the research context. Response options and recall period not available. (1?).	3 domains of apathy included. (1+).	Full wording not available, but items suggest 33% appropriately worded. Response options not available. (1?).	Indeterminate (1+, 2?)	Indeterminate (1+, 2?)
KBCI-10a	90% of items relevant to apathy. 80% of items relevant to older adults. All items relevant to research context. Response options appropriate. Recall period not available. (1+/-).	90% of items relevant to apathy. 80% of items relevant to people with dementia. All items relevant to research context. Response options appropriate. Recall period not available. (1+/-).	3 domains of apathy included. (1+).	80% of items appropriately worded. 100% match response options. (1+/-).	Inconsistent (1+, 2+/-)	Inconsistent (1+, 2+/-)

Measure	Relevance		Comprehensiveness (quality rating)	Comprehensibility (quality rating)	Overall validity	
	Older adults (quality rating)	Dementia & MCI (quality rating)			Older adults	Dementia & MCI
LARS	94% of items relevant to apathy. 100% relevant to older adults. Response options appropriate. Recall period too long. (1+).	94% of items relevant to apathy. 94% relevant to people with dementia. Response options appropriate. Recall period too long. (1+).	3 domains of apathy included. (1+).	87% appropriately worded. 100% match response options. (1+).	Sufficient (3+)	Sufficient (3+)
NPI (original)	100% of items relevant to apathy, older adults and the research context. Response options appropriate. Suggested recall period too long, but personalised recall period also possible. (1+).	100% of items relevant to apathy, people with dementia and the research context. Response options appropriate. Suggested recall period too long, but personalised recall period also possible. (1+).	Emotional dimension of apathy is missing from the screening questions. No dimensions are rated separately. (1-).	Assessments of frequency and severity are based on multiple symptoms, so could be considered a double barrelled question and therefore not appropriately worded. However carers are advised to rate the worst one. 100% match the response options. (1+).	Inconsistent (2+, 1-)	Inconsistent (2+, 1-)
NPI-A	Unable to obtain full instructions and guidance.					
NPI-C	100% of items relevant to apathy, older adults and the research context. Response options appropriate. Recall period too long. (1+).	100% of items relevant to apathy, people with dementia and the research context. Response options appropriate. Recall period too long. (1+).	3 domains of apathy included. (1+).	Assessments of frequency and severity are based on multiple symptoms, so could be considered a double barrelled question and therefore not appropriately worded. However carers are advised to rate the worst one. 100% match the response options. (1+).	Sufficient (3+)	Sufficient (3+)
UPDRS	100% relevant to apathy, older adults and research context. (Note: only 1 item). Response options appropriate. Recall period not clear. (1+).	100% relevant to apathy, people with dementia and research context. (Note: only 1 item). Response options appropriate. Recall period not clear. (1+).	Emotional domain of apathy missing. Cognitive and Behavioural elements included but not rated separately. (1-).	Item wording is not given, or could not be obtained; only the heading is provided, so it is unclear if it matches the response options. (1?).	Inconsistent (1+, 1-, 1?)	Inconsistent (1+, 1-, 1?)
mds-UPDRS	100% relevant to apathy, older adults and the research context. (Note: only 1 item). Response options and recall period appropriate. (1+).	100% relevant to apathy, people with dementia and the research context. (Note: only 1 item). Response options and recall period appropriate. (1+).	Emotional domain of apathy missing. Cognitive and Behavioural elements included but not rated separately. (1-).	100% appropriate worded and match response options. (1+).	Inconsistent (2+, 1-)	Inconsistent (2+, 1-)

*based on list of apathy items presented by another publication (DEX [81]; FrsBE [64,68]; GIP [107])

Quality of measurement property: Number of studies in parenthesis followed by rating: +, Sufficient; +/-, Inconsistent; -, Insufficient; ?, Indeterminate.

Abbreviations: AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; AES Apathy Evaluation Scale; AES-C, Apathy Evaluation Scale Clinician; AES-I, Apathy Evaluation Scale Informant; AES-S, Apathy Evaluation Scale Self; AI, Apathy Inventory; AI-C, Apathy Inventory Clinician; AI-I, Apathy Inventory Informant; AI-S, Apathy Inventory Self; AMI, Apathy Motivation Index; AS, Apathy Scale; b-DAS, brief-Dimensional Apathy Scale; BMDS, Behavioural and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; DAIR, Dementia Apathy Interview Rating; DAS, Dimensional Apathy Scale; DEX, Dysexecutive Questionnaire; FrSBe, Frontal Systems Behavior Scale; FrSBe-6a, Frontal Systems Behavior Scale 6-item apathy subscale; FrSBe-11a, Frontal Systems Behavior Scale 11-item apathy subscale; FrSBe-14a, Frontal Systems Behavior Scale 14-item apathy subscale; GDS-3a, Geriatric Depression Scale 3 item apathy

subscale; GDS-6a, Geriatric Depression Scale 6 item apathy subscale; GIP-9a, Behavioral Rating Scale for Psychogeriatric Inpatients 9 item apathy subscale; IMD, Index of Mental Decline; KBCI-10a, Key Behaviors Change Inventory 10 item apathy subscale; LARS, Lille Apathy Rating Scale; MDS-UPDRS, Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale; NPI, Neuropsychiatric Inventory; NPI-A, Neuropsychiatric Inventory Alternative; NPI-C, Neuropsychiatric Inventory Clinician; UPDRS, Unified Parkinson's Disease Rating Scale

Risk of Bias and Results for remaining measurement properties

Systematic Review - Risk of bias and results of studies of remaining measurement properties

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) [quality rating]	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[40]	AD-RD					1 Doubtful.	r=.72 (1+).				
[78]	AES-12PD			3 Very good.	α=.90 to .92 (3+)					3 Adequate. 1 Very Good.	3 met hypothesis (3+). 1 did not meet hypothesis (1-).
[53]	AES-C	1 Doubtful.	3 factors (57.06%): Apathy (40.02%); Novelty Seeking (9.35%); Insight & social (7.68%). [1+]	1 Very good.	α=.90. (1+).	2 Doubtful.	r=.88 to .86 (2+).			2 Inadequate. 2 Very Good.	4 met hypothesis (4+).
[54]	AES-C	1 Doubtful.	2 factors (51.1%): Apathy (42.4%); Interest (8.7%). [1+]							1 Inadequate. 1 Doubtful. 2 Adequate.	2 met hypothesis (2+). 2 did not meet hypothesis (2-).
[55]	AES-C	1 Inadequate.	3 factors (84.17^): Interest & Motivation (39.72%^); Task Completion (29.67%^); Insight (14.78%^). [1-]	1 Very good.	α=.93. (1+).					1 Inadequate. 1 Doubtful. 1 Very Good.	3 met hypothesis (3+).
[33]	AES-C	1 Inadequate.	3 factors: Apathy (32-53%); Novelty Seeking (5-10%); Insight & dependency (7-8%). [1?]	1 Very good.	α=.90. (1+).	1 Doubtful. 1 Adequate.	r=.88 (1+). ICC= .94 (+).			3 Inadequate. 1 Doubtful. 1 Adequate. 4 Very Good.	5 met hypothesis (5+). 1 did not meet hypothesis (1-). 3 insufficient information (3?).
[54]	AES-I	1 Doubtful.	2 factors (54.4%): Interest (45.1%); Apathy (9.3%). [1+]							1 Inadequate. 1 Doubtful. 2 Adequate.	2 met hypothesis (2+). 2 did not meet hypothesis (2-).
[55]	AES-I			1 Very good.	α=.89. (1+).					1 Inadequate. 1 Doubtful. 1 Very Good.	3 met hypothesis (3+).
[33]	AES-I	1 Inadequate.	3 factors: Apathy (32-53%); Novelty Seeking (5-10%); Insight & dependency (7-8%). [1?]	1 Very good.	α=.94. (1+).	1 Doubtful.	r=.94 (1+).			3 Inadequate. 1 Doubtful. 1 Adequate. 4 Very Good.	4 met hypothesis (4+). 2 did not meet hypothesis (2-). 3 insufficient information (3?).

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) [quality rating]	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[56]	AES-I	1 Doubtful. 1 Adequate.	2 factors (62.56%^): Factor 1 (56.2%); Factor 2 (6.36%). [1+]. 1 factor (62.8%). [1+].	1 Very good.	α=.95. (1+).			n/a	SEM=2.9. (1?).	2 Very Good.	2 met hypothesis (2+).
[89]	AES-I									1 Adequate. 2 Very good.	1 met hypothesis (1+). 2 did not meet hypothesis (2-).
[71]	AES-I			1 Very good.	α=.88. (1+).						
[71]	AES-I-16			1 Very good.	α=.90. (1+).					1 Adequate. 1 Very Good	1 met hypothesis (1+). 1 did not meet hypothesis (1-).
[55]	AES-S			1 Very good.	α=.90. (1+).					1 Inadequate. 1 Doubtful. 1 Very Good.	1 met hypothesis (1+). 1 did not meet hypothesis (1-). 1 insufficient information (1?)
[54]	AES-S	1 Doubtful.	2 factors (43.3%^): Apathy (36.4%); Other (6.9%) [1+]							1 Inadequate. 1 Doubtful. 2 Adequate.	2 met hypothesis (2+). 2 did not meet hypothesis (2-).
[33]	AES-S	1 Inadequate.	3 factors: Apathy (32-53%); Novelty Seeking (5-10%); Insight & dependency (7-8%). [1?]	1 Very good.	α=.86. (1+).	1 Doubtful.	r=.76 (1+).			3 Inadequate. 1 Doubtful. 1 Adequate. 4 Very Good.	5 met hypothesis (5+). 1 did not meet hypothesis (1-). 3 insufficient information (3?)
[56]	AES-S	1 Doubtful. 1 Adequate.	2 factors (61.69%^): Factor 1 (55.37%); Factor 2 (6.32%). [1+] 1 factor (61.2%). [1+].	1 Very good.	α=.95. (1+).			n/a	SEM=2.7. (1?).	2 Very Good.	2 met hypothesis (2+).
[89]	AES-S									1 Adequate. 2 Very good.	1 met hypothesis (1+). 2 did not meet hypothesis (2-).
[57]	AES-S	2 Doubtful.	3 factors (58%): Apathy (38.27%); Friendship (10.86%); Other (8.88%) [1+]. 3 factors (59.54%; variance explained per	2 Very good.	α=.90 to .92. (2+).					2 Doubtful. 4 Adequate. 4 Very Good.	5 met hypothesis (5+). 3 did not meet hypothesis (3-).

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) [quality rating]	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
			factor not reported.) [1?]								
[80]	AI-C					1 Doubtful.	ICC=.97 (1+).			1 Inadequate.	1 met hypothesis (1+).
[72]	AI-C			1 Doubtful.	α=.83. (1?).						
[34]	AI-I			1 Doubtful.	α=.84. (1?).	1 Doubtful. 1 Inadequate.	Kappa= .96 to .99 (2+).			1 Adequate. 3 Very Good.	3 met hypothesis (3+). 1 did not meet hypothesis (1-).
[72]	AI-I			1 Doubtful.	α=.83. (1?).						
[34]	AI-S									3 Very Good.	1 met hypothesis. (1+). 2 did not (2-)
[72]	AI-S			1 Doubtful.	α=.61. (1?).						
[67]	AMI			*	α=.86					2 Adequate.	2 did not meet hypothesis (2-).
[58]	AS-HC	1 Very Good.	1 factor CFI=1.00, RMSEA=0.00. [1+]	1 Very Good	α=.94. (1+).					1 Very Good.	1 did not meet hypothesis (1-).
[90]	AS-I									1 Inadequate. 1 Doubtful. 1 Very Good.	1 met hypothesis (1+). 1 did not meet hypothesis (1-). 1 insufficient information available (1?).
[59]	AS-S	1 Doubtful.	13-item: 3 factors (55.61%). Variance explained per factor not reported. [1?]	2 Doubtful.	14 item version: α=.82. 13 item version: α=.85. (2?).						
[36]	AS-S			1 Doubtful.	α=.76. (1?).	2 Doubtful.	r=.81 to .90. (2+).			1 Doubtful.	1 met hypothesis (1+).
[58]	AS-S	1 Very Good.	1 factor. CFI=1.00, RMSEA=0.00. [1+].								
[60]	AS-S	2 Adequate.	14-item: 2 factors (57.7%): Cognitive-Behavioural (24.2%);	2 Doubtful.	14 item: α=.69. 13 item: α=.74. (2?).					1 Adequate. 2 Very Good.	3 met hypothesis (3+).

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) [quality rating]	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
			Apathy and insight (15.05%). [1-]. 13-item: 2 factors (41.7%) Variance explained per factor not reported. [1?]								
[77]	AS-S			1 Inadequate.	Guttman's λ = .89. (1?).	1 Inadequate.	ICC=.78 (1+).	n/a	SEM = 2.34. (1?).	1 Doubtful. 2 Very Good.	1 met hypothesis (1+). 1 did not meet hypothesis (1-). 1 insufficient information (1?).
[61]	AS-S	1 Very Good. 1 Adequate.	AS-S: 3 factors (nr). [1+/-]. 11 item: 2 factors: 54.1% of variance explained. [1-].	11-item: 1 Inadequate.	11 item: α=.77 (1?)						
[91]	AS-S			1 Doubtful.	α=.78. (1?).					1 Inadequate. 2 Doubtful. 1 Very good.	3 met hypothesis (3+). 1 did not meet hypothesis (1-).
[42]	BMDS					1 Inadequate.	r=.90. (1+)				
[43]	BSSD			1 Doubtful.	α=.82 to .83 (1?)	1 Inadequate. 3 Doubtful.	ICC=.65 to .85. (2+, 2-).			2 Inadequate. 1 Doubtful. 1 Very Good.	1 met hypothesis (1+). 1 did not meet hypothesis (1-). 2 insufficient information (2?).
[37]	DAIR	1 Adequate.	1 factor (38%) [1+]	1 Very Good.	α=.89. (1+).	1 Inadequate.	r=.85 (1+)	1 Doubtful.	100% agreement (1+).	2 Inadequate. 2 Very Good.	3 met hypothesis (3+). 1 did not meet hypothesis (1+).
[75]	DAS-I			*	α=.92					2 Adequate.	1 met hypothesis (1+). 1 did not meet hypothesis (1-).
[74]	DAS-I			*	α=.93					2 Adequate.	2 met hypothesis (2+).
[75]	DAS-S			*	α=.84					2 Adequate.	1 met hypothesis (1+). 1 did not meet hypothesis (1-).
[74]	DAS-S			*	α=.85					2 Adequate.	2 met hypothesis (2+).

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) [quality rating]	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[62]	DAS-S	*	3 factors (45.87^) Organisation & planning (28.21%); Initiation (9.76%); Emotional (7.90%).	*	α=.87					4 Adequate. 2 Very Good.	4 met hypothesis (4+). 2 did not meet hypothesis (2-).
[63]	bDAS	*	Item Hi=.40 to .76. No other fit measures reported.								
[76]	bDAS			*	α=.81.	1 Inadequate.	ICC=.84 (1+).				
[81]	DEX					1 Doubtful.	ICC=.93 (1+).			1 Inadequate. 1 Adequate. 2 Very Good.	2 met hypothesis (2+). 2 did not meet hypothesis (2+)
[64]	FrSBe-6a			1 Doubtful.	α=.88. (1?).						
[68]	FrSBe-11a			1 Doubtful.	α=.83. (1?).						
[68]	FrSBe-14a			1 Doubtful.	α=.88. (1?).					6 Doubtful.	5 met hypothesis (5+). 1 did not meet hypothesis (1-).
[64]	FrSBe-14a	1 Inadequate.	1 Factor specified: 12 out of 14 items had loadings >.40. (nr). [1?].	1 Doubtful.	α=.80. (1?).						
[92]	GDS-3a									2 Adequate.	2 did not meet hypothesis. (2+).
[69]	GDS-6a			1 Doubtful.	α=.51 (1?).					1 Doubtful. 2 Adequate.	3 met hypothesis (3+).
[82]	GIP-apathy subscale					1 Doubtful.	ICC=.72 (1+).	n/a	SEM=1.22. (1?)		
[82]	GIP-apathy domain					1 Doubtful.	ICC=.83 (1+.)	n/a	SEM=1.38. (1?)		
[49]	IMD									1 Inadequate. 3 Doubtful.	3 met hypothesis (3+). 1 insufficient information (1?).

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[93]	KBCI									1 Inadequate. 1 Doubtful. 5 Adequate.	6 met hypothesis (6+). 1 did not meet hypothesis (1-).
[83]	LARS-C					2 Doubtful.	ICC=.94 to .99 (2+).			2 Inadequate. 2 Doubtful. 5 Very Good.	7 met hypothesis (7+). 1 did not meet hypothesis (1-). 1 insufficient information (1?).
[65]	LARS-C	*	4 factors (67.5%): intellectual curiosity (nr); emotion (nr); action-initiation (nr); self awareness (nr).	*	α=.81. (*).	2 Doubtful.	ICC= .97. (1+). Kappa = .93 (1+).			1 Inadequate. 2 Adequate.	2 met hypothesis (2+). 1 did not meet hypothesis (1-).
[70]	LARS-I			*	α=.87. (*).	2 Doubtful.	ICC =.99. (1+). . (1+)ICC =.99. (1+).			2 Adequate.	2 met hypothesis (2+).
[84]	NPI					1 Doubtful. 1 Inadequate.	ICC = .67 (1-). r_s= .53 (1-).				
[94]	NPI									2 Doubtful.	2 insufficient information (2?).
[51]	NPI									1 Doubtful.	1 insufficient information (1?).
[85]	NPI					1 Inadequate	r=.96 (1+).			1 Doubtful.	1 insufficient information (1?).
[73]	NPI			1 Doubtful.	α=.82 (1?)	1 Doubtful. 1 Inadequate.	ICC=.87. (1+). r=.76 (1+).			1 Inadequate. 1 Doubtful. 1 Very Good.	1 met hypothesis (1+). 1 did not meet hypothesis (1-). 1 insufficient information (1?).
[95]	NPI									1 Inadequate.	1 did not meet hypothesis (1-).
[79]	NPI			1 Doubtful.	α=.83 (1?)	1 Doubtful.	Kendell CC= 1.00 (1+).			1 Inadequate.	1 did not meet hypothesis (1-).
[96]	NPI									1 Very Good.	1 did not meet hypothesis (1-).
[86]	NPI					1 Doubtful.	ICC=.99 (1+).				
[66]	NPI-A	1 Adequate.	1 factor (66%). [1+].	1 Very Good.	α=.91 (1+)						

Reference	Measure	Structural validity		Internal consistency		Reliability		Measurement error		Hypothesis testing	
		Methodological quality	Result (% variance explained) (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)	Methodological quality	Result (quality rating)
[87]	NPI-C					1 Doubtful.	Item ICC= .74 to .89 (1+).			1 Adequate.	1 did not meet hypothesis (1-).
[88]	NPI-C					1 Doubtful.	ICC=.87 (1+).			1 Adequate.	1 met hypothesis (1+).
[97]	mds-UPDRS									1 Very Good.	1 met hypothesis (1+).
[98]	UPDRS									1 Adequate.	1 met hypothesis (1+).
[99]	UPDRS									1 Inadequate. 2 Very Good.	3 did not meet hypothesis (3-).
[100]	UPDRS									1 Very Good.	1 met hypothesis (1+).

Blank cells indicate this measurement property was not investigated.

*Was assessed by the study, but methodological quality rating nor quality rating of result conducted, as the measure is based on a formative model.

^ Value calculated by review team based on information provided in the article.

Quality of measurement property: Number of studies in parenthesis followed by rating: +, Sufficient; +/-, Inconsistent; -, Insufficient; ?, Indeterminate.

Abbreviations: +, Sufficient; -, Insufficient; ?, Indeterminate; AD-RD, Alzheimer's Disease and Related Dementias Mood Scale; AES-12PD, Apathy Evaluation Scale for Parkinson Disease; AES-C, Apathy Evaluation Scale Clinician; AES-I, Apathy Evaluation Scale Informant; AES-I-16, Apathy Evaluation Scale Informant 16 item version; AES-S, Apathy Evaluation Scale Self; AI-C, Apathy Inventory Clinician; AI-I, Apathy Inventory Informant; AI-S, Apathy Inventory Self; AMI, Apathy Motivation Index; AS-HC, Apathy Scale Home Care; AS-I, Apathy Scale Informant; AS-S, Apathy Scale Self; b-DAS, brief-Dimensional Apathy Scale; BMDS, Behavioural and Mood Disturbance Scale; BSSD, Behavioral Syndromes Scale for Dementia; DAIR, Dementia Apathy Interview Rating; DAS-I, Dimensional Apathy Scale Informant; DAS-S, Dimensional Apathy Scale Self; DEX, Dysexecutive Questionnaire; FrSBe-6a, Frontal Systems Behavior Scale 6-item apathy subscale; FrSBe-11a, Frontal Systems Behavior Scale 11-item apathy subscale; FrSBe-14a, Frontal Systems Behavior Scale 14-item apathy subscale; GDS-3a, Geriatric Depression Scale 3 item apathy subscale; GDS-6a, Geriatric Depression Scale 6 item apathy subscale GIP, Behavioral Rating Scale for Psychogeriatric Inpatients; IMD, Index of Mental Decline; KBCI, Key Behaviors Change Inventory; LARS-C, Lille Apathy Rating Scale Clinician; LARS-I, Lille Apathy Rating Scale Informant; mds-UPDRS, Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale; NPI, Neuropsychiatric Inventory; NPI-A, Neuropsychiatric Inventory Alternative; NPI-C, Neuropsychiatric Inventory Clinician; nr, not reported; UPDRS, Unified Parkinson's Disease Rating Scale

Where there is no rating available for the researcher, this means it was not possible to obtain sufficient information regarding the measure to assess its content validity. Ratings of content validity are for both people with dementia or MCI and older adults unless otherwise specified.

References for systematic review tables in appendix

References

1 Robert PH, Lanctôt KL, Agüera-Ortiz L, Aalten P, Bremond F, Defrancesco M, et al. Is it time to revise the diagnostic criteria for apathy in brain disorders? The 2018 international consensus group. European Psychiatry. 2018 Oct;54:71–6.

2 Bergh S, Selbæk G. The prevalence and the course of neuropsychiatric symptoms in patients with dementia. 1. 2012 Nov;22(2). DOI: 10.5324/nje.v22i2.1570

3 Cipriani G, Lucetti C, Danti S, Nuti A. Apathy and dementia. Nosology, assessment and management. J Nerv Ment Dis. 2014 Oct;202(10):718–24.

4 Monastero R, Mangialasche F, Camarda C, Ercolani S, Camarda R. A Systematic Review of Neuropsychiatric Symptoms in Mild Cognitive Impairment. Journal of Alzheimer's Disease. 2009 Jan;18(1):11–30.

5 Martin E, Velayudhan L. Neuropsychiatric Symptoms in Mild Cognitive Impairment: A Literature Review. DEM. 2020;49(2):146–55.

6 Lam LCW, Tam CWC, Chiu HFK, Lui VWC. Depression and apathy affect functioning in community active subjects with questionable dementia and mild Alzheimer's disease. Int J Geriatr Psychiatry. 2007 May;22(5):431–7.

7 Benoit M, Andrieu S, Lechowski L, Gillette-Guyonnet S, Robert PH, Vellas B, et al. Apathy and depression in Alzheimer's disease are associated with functional deficit and psychotropic prescription. Int J Geriatr Psychiatry. 2008 Apr;23(4):409–14.

8 Mukherjee A, Biswas A, Roy A, Biswas S, Gangopadhyay G, Das SK. Behavioural and Psychological Symptoms of Dementia: Correlates and Impact on Caregiver Distress. DEE. 2017;7(3):354–65.

9 Sultzer DL. Why Apathy in Alzheimer's Matters. Am J Psychiatry. 2018 01;175(2):99–100.

10 Armstrong N, Schupf N, Grafman J, Huey ED. Caregiver Burden in Frontotemporal Degeneration and Corticobasal Syndrome. DEM. 2013;36(5–6):310–8.

11 Ready RE, Ott BR, Grace J, Cahn-Weiner DA. Apathy and executive dysfunction in mild cognitive impairment and Alzheimer disease. Am J Geriatr Psychiatry. 2003 Apr;11(2):222–8.

12 Stuss DT, Van Reekum R, Murphy KJ. Differentiation of states and causes of apathy. The neuropsychology of emotion. New York, NY, US: Oxford University Press; 2000; pp 340–63.

13 Yeager CA, Hyer L. Apathy in dementia: relations with depression, functional competence, and quality of life. Psychol Rep. 2008 Jun;102(3):718–22.

14 Pickett J, Bird C, Ballard C, Banerjee S, Brayne C, Cowan K, et al. A roadmap to advance dementia research in prevention, diagnosis, intervention, and care by 2025. International journal of geriatric psychiatry. 2018 Jul;33(7):900–6.

15 Onyike CU, Sheppard J-ME, Tschanz JT, Norton MC, Green RC, Steinberg M, et al. Epidemiology of apathy in older adults: the Cache County Study. Am J Geriatr Psychiatry. 2007 May;15(5):365–75.

16 Chilovi BV, Conti M, Zanetti M, Mazzù I, Rozzini L, Padovani A. Differential Impact of Apathy and Depression in the Development of Dementia in Mild Cognitive Impairment Patients. DEM. 2009;27(4):390–8.

17 Starkstein SE, Jorge R, Mizrahi R, Robinson RG. A prospective longitudinal study of apathy in Alzheimer's disease. J Neurol Neurosurg Psychiatr. 2006 Jan;77(1):8–11.

18 Wadsworth LP, Lorius N, Donovan NJ, Locascio JJ, Rentz DM, Johnson KA, et al. Neuropsychiatric Symptoms and Global Functional Impairment along the Alzheimer's Continuum. DEM. 2012;34(2):96–111.

19 Gilmore-Bykovskyi A. Commentary on Apathy as a Model for Investigating Behavioral and Psychological Symptoms in Dementia. Journal of the American Geriatrics Society. 2018;66(S1):S13–6.

20 Clarke DE, Ko JY, Kuhl EA, van Reekum R, Salvador R, Marin RS. Are the available apathy measures reliable and valid? A review of the psychometric evidence. Journal of psychosomatic research. 2011;70(1):73–97.

21 Radakovic R, Harley C, Abrahams S, Starr JM. A systematic review of the validity and reliability of apathy scales in neurodegenerative conditions. Int Psychogeriatr. 2015;27(6):903–23.

22 Mohammad D, Ellis C, Rau A, Rosenberg PB, Mintzer J, Ruthirakuhan M, et al. Psychometric Properties of Apathy Scales in Dementia: A Systematic Review. J Alzheimers Dis. 2018;66(3):1065–82.

23 Whiting P, Rutjes AW, Reitsma JB, Bossuyt PM, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Medical Research Methodology. 2003 Nov;3(1):25.

24 Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, de Vet HCW, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Quality of Life Research. 2018 Feb DOI: 10.1007/s11136-018-1798-3

25 Terwee CB, Prinsen CAC, Chiarotto A, Westerman MJ, Patrick DL, Alonso J, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2018 Mar DOI: 10.1007/s11136-018-1829-0

26 Burgon C, Goldberg S, van der Wardt V, Harwood RH. A systematic review of measures of apathy for older adults: validity, reliability and conceptualisation of apathy. PrAISED Discussion Paper Series. 2019 Jul;(5).

27 Mokkink LB, de Vet HCW, Prinsen CAC, Patrick DL, Alonso J, Bouter LM, et al. COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2017 Dec DOI: 10.1007/s11136-017-1765-4

28 de Vet HCW, Terwee CB, Mokkink LB, Knol DL. Measurement in Medicine: A Practical Guide. Cambridge: Cambridge University Press; 2011. DOI: 10.1017/CBO9780511996214

29 Park MS, Kang KJ, Jang SJ, Lee JY, Chang SJ. Evaluating test-retest reliability in patient-reported outcome measures for older people: A systematic review. International Journal of Nursing Studies. 2018 Mar;79:58–69.

30 Terwee CB, Jansma EP, Riphagen II, de Vet HCW. Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. Quality of Life Research. 2009 27 01/23/received 08/06/accepted;18(8):1115–23.

31 Mokkink LB, Terwee CB, Patrick DL, Alonso J, Stratford PW, Knol DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010 Jul;63(7):737–45.

32 Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. Trials. 2016 Sep;17(1):449.

33 Marin RS, Biedrzycki RC, Firinciogullari S. Reliability and validity of the Apathy Evaluation Scale. Psychiatry Res. 1991 Aug;38(2):143–62.

34 Robert PH, Clairet S, Benoit M, Koutaich J, Bertogliati C, Tible O, et al. The apathy inventory: assessment of apathy and awareness in Alzheimer's disease, Parkinson's disease and mild cognitive impairment. Int J Geriatr Psychiatry. 2002 Dec;17(12):1099–105.

35 Ang Y-S, Lockwood P, Apps MAJ, Muhammed K, Husain M. Distinct Subtypes of Apathy Revealed by the Apathy Motivation Index. PLOS ONE. 2017;12(1):1–15.

36 Starkstein SE, Mayberg HS, Preziosi TJ, Andrezejewski P, Leiguarda R, Robinson RG. Reliability, validity, and clinical correlates of apathy in Parkinson's disease. J Neuropsychiatry Clin Neurosci. 1992;4(2):134–9.

37 Strauss ME, Sperry SD. An informant-based assessment of apathy in Alzheimer disease. Neuropsychiatry Neuropsychol Behav Neurol. 2002 Sep;15(3):176–83.

38 Radakovic R, Abrahams S. Developing a new apathy measurement scale: Dimensional Apathy Scale. Psychiatry Research. 2014 Nov;219(3):658–63.

39 Sockeel P, Dujardin K, Devos D, Denève C, Destée A, Defebvre L. The Lille apathy rating scale (LARS), a new instrument for detecting and quantifying apathy: validation in Parkinson's disease. J Neurol Neurosurg Psychiatry. 2006 May;77(5):579–84.

40 Tappen RM, Williams CL. Development and Testing of the Alzheimer's Disease and Related Dementias Mood Scale. Nurs Res. 2008;57(6):426–35.

41 Tappen RM, Williams C. Attribution of emotion in advanced Alzheimer's disease: Family and caregiver perspectives. American Journal of Alzheimer's Disease. 1998 Sep;13(5):257–64.

42 Greene JG, Smith R, Gardiner M, Timbury GC. Measuring behavioural disturbance of elderly demented patients in the community and its effects on relatives: a factor analytic study. Age Ageing. 1982 May;11(2):121–6.

43 Devanand DP, Brockington CD, Moody BJ, Brown RP, Mayeux R, Endicott J, et al. Behavioral syndromes in Alzheimer's disease. Int Psychogeriatr. 1992;4 Suppl 2:161–84.

44 Burgess PW, Alderman N, Wilson BA, Evans JJ, Emslie H. The Dysexecutive Questionnaire. Behavioural assessment of the dysexecutive syndrome. Bury St Edmunds, UK: Thames Valley Test Company; 1996.

45 Grace J, Malloy PF. Frontal Systems Behavior Scale (FrSBe): Professional manual. Lutz, FL: Psychological Assessment Resources; 2001.

46 Brink TL, Yesavage JA, Lum O, Heersema PH, Adey M, Rose TL. Screening tests for geriatric depression. Clinical Gerontologist: The Journal of Aging and Mental Health. 1982;1(1):37–43.

47 Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, et al. Development and validation of a geriatric depression screening scale: A preliminary report. Journal of Psychiatric Research. 1982 Jan;17(1):37–49.

48 Verstraten P, Van Eekelen C. Handleiding voor de GIP: Gedragsobservatieschaal voor de Intramurale Psychogeriatrie. Deventer: Van Loghum Slaterus; 1987.

49 Cucinotta D, Ambrosoli L, Poli A, Martorelli M, Savorani GC, Anzivino F. Clinical assessment of mental decline in elderly people: A proposal for a new quantitative index. Aging Clin Exp Res. 1995 Feb;7(1):29–34.

50 Kolitz BP, Vanderploeg RD, Curtiss G. Development of the Key Behaviors Change Inventory: a traumatic brain injury behavioral outcome assessment instrument. Arch Phys Med Rehabil. 2003 Feb;84(2):277–84.

51 Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The Neuropsychiatric Inventory: comprehensive assessment of psychopathology in dementia. Neurology. 1994 Dec;44(12):2308–14.

52 Fahn S, Elton RL, Members of the UPDRS Development Committee. Unified Parkinson's Disease Rating Scale. Recent Developments in Parkinson's Disease Volume II. Macmillan Healthcare Information; 1987; pp 153–63.

53 Lee S-H, Wen M-C, Chao C-C, Chen YJ, Yen C-F. Apathy in late-life depression among Taiwanese patients. Int Psychogeriatr. 2008 Apr;20(2):328–37.

54 Clarke DE, Reekum R van, Simard M, Streiner DL, Freedman M, Conn D. Apathy in dementia: an examination of the psychometric properties of the apathy evaluation scale. J Neuropsychiatry Clin Neurosci. 2007;19(1):57–64.

55 Guercio BJ, Donovan NJ, Munro CE, Aghjayan SL, Wigman SE, Locascio JJ, et al. The Apathy Evaluation Scale: A Comparison of Subject, Informant, and Clinician Report in Cognitively Normal Elderly and Mild Cognitive Impairment. J Alzheimers Dis. 2015;47(2):421–32.

56 Johansson M, Johansson P, Stomrud E, Hagell P, Hansson O. Psychometric testing of a Swedish version of the Apathy Evaluation Scale. Nord J Psychiatry. 2017 Aug;71(6):477–84.

57 Lueken U, Evens R, Balzer-Geldsetzer M, Baudrexel S, Dodel R, Gräber-Sultan S, et al. Psychometric properties of the apathy evaluation scale in patients with Parkinson's disease. Int J Methods Psychiatr Res. 2017;26(4). DOI: 10.1002/mpr.1564

58 Morita H, Kannari K. Reliability and validity assessment of an apathy scale for home-care patients with Parkinson's disease: a structural equation modeling analysis. J Phys Ther Sci. 2016 Jun;28(6):1724–7.

59 Kay DB, Kirsch-Darrow L, Zahodne LB, Okun MS, Bowers D. Dimensions of apathy in Parkinson's disease: exploratory factor analysis of the apathy scale. J Parkinsons Dis. 2012;2(2):161–6.

60 Pedersen KF, Alves G, Larsen JP, Tysnes O-B, Møller SG, Brønnick K. Psychometric properties of the Starkstein Apathy Scale in patients with early untreated Parkinson disease. Am J Geriatr Psychiatry. 2012 Feb;20(2):142–8.

61 Lopez FV, Eglit GML, Schiehser DM, Pirogovsky-Turk E, Litvan I, Lessig S, et al. Factor Analysis of the Apathy Scale in Parkinson's Disease. Movement Disorders Clinical Practice. 2019;6(5):379–86.

62 Santangelo G, D'Iorio A, Piscopo F, Cuoco S, Longo K, Amboni M, et al. Assessment of apathy minimising the effect of motor dysfunctions in Parkinson's disease: a validation study of the dimensional apathy scale. Qual Life Res. 2017;26(9):2533–40.

63 Radakovic R, McGrory S, Chandran S, Swingler R, Pal S, Stephenson L, et al. The brief Dimensional Apathy Scale: A short clinical assessment of apathy. Clin Neuropsychol. 2020 Feb;34(2):423–35.

64 Midden AJ, Mast BT. IRT Analysis of the Frontal Systems Behavior Scale: Identifying the Best Items for Use with Geriatric Patients in Primary Care. Clinical Gerontologist. 2017 Aug;40(4):295–306.

65 García-Ramos R, Villanueva Iza C, Catalán MJ, Reig-Ferrer A, Matías-Guíu J. Validation of a Spanish Version of the Lille Apathy Rating Scale for Parkinson's Disease [Internet]. The Scientific World Journal. 2014 Oct;2014:e849834.

66 Gallo JL, Schmidt KS, Libon DJ. An itemized approach to assessing behavioral and psychological symptoms in dementia. Am J Alzheimers Dis Other Demen. 2009 May;24(2):163–8.

67 Ang Y-S, Lockwood PL, Kienast A, Plant O, Drew D, Slavkova E, et al. Differential impact of behavioral, social, and emotional apathy on Parkinson's disease. Annals of Clinical and Translational Neurology. 2018;5(10):1286–91.

68 Carvalho JO. Revision of the frontal systems behavior scale (FrSBe). Doctoral Dissertations Available from Proquest. 2011 Jan;1–107.

69 Davis TE. The Effectiveness of the Geriatric Depression Scale to Distinguish Apathy From Depression in Alzheimer's Disease and Related Dementias. 2008

70 Dujardin K, Sockeel P, Delliaux M, Destée A, Defebvre L. The Lille Apathy Rating Scale: validation of a caregiver-based version. Mov Disord. 2008 Apr;23(6):845–9.

71 Köller L, Knebel M, Haberstroh J, Krause K, Sahlender S, Jakob M, et al. Apathy in dementia care: An investigation in community-dwelling persons using the German informant version of the Apathy Evaluation Scale. GeroPsych: The Journal of Gerontopsychology and Geriatric Psychiatry. 2016;29(3):129–39.

72 Leone E, Lafont V, Filleau C, Baudu C, Benoit M, Deudon A, et al. Validation de la version soignant de l'Inventaire Apathie. Revue Geriatr. 2008;33(9):777–85.

73 Malakouti SK, Panaghi L, Foroughan M, Salehi M, Zandi T. Farsi version of the Neuropsychiatric Inventory: validity and reliability study among Iranian elderly with dementia. Int Psychogeriatr. 2012 Feb;24(2):223–30.

74 Radakovic R, Starr JM, Abrahams S. A Novel Assessment and Profiling of Multidimensional Apathy in Alzheimer's Disease. J Alzheimers Dis. 2017;60(1):57–67.

75 Radakovic R, Davenport R, Starr JM, Abrahams S. Apathy dimensions in Parkinson's disease. Int J Geriatr Psychiatry. 2018 Jan;33(1):151–8.

76 Radakovic R, Gray D, Dudley K, Mioshi E, Dick D, Melchiorre G, et al. Reliability and validity of the brie dimensional apathy scale. Arch Clin Neuropsychol. 2020 Jul;35(5):539–44.

77 Serrano-Dueñas M, Martínez-Martín P, Merchán T, Bravo R, Serrano M. Properties of the Apathy Scale (AS) for use on Parkinson's patients. Advances in Parkinson's Disease. 2013 May;2(2):53–7.

78 Stankevich Y, Lueken U, Balzer-Geldsetzer M, Dodel R, Gräber-Sultan S, Berg D, et al. Psychometric Properties of an Abbreviated Version of the Apathy Evaluation Scale for Parkinson Disease (AES-12PD). The American Journal of Geriatric Psychiatry. 2018 Oct;26(10):1079–90.

79 Vilalta-Franch J, Lozano-Gallego M, Hernández-Ferrándiz M, Llinàs-Reglà J, López-Pousa S, López OL. [The Neuropsychiatric Inventory. Psychometric properties of its adaptation into Spanish]. Rev Neurol. 1999 Jul;29(1):15–9.

80 Stella F, Andrade LP de, Garuffi M, Vital TM, Hernández SSS, Ruocco M, et al. Validation of the Brazilian version of the Apathy Inventory. International Journal of Geriatric Psychiatry. 2013;28(9):979–86.

81 Shinagawa Y, Nakaaki S, Hongo J, Murata Y, Sato J, Matsui T, et al. Reliability and validity of the Japanese version of the Dysexecutive Questionnaire (DEX) in Alzheimer's disease: validation of a behavioral rating scale to assess dysexecutive symptoms in Japanese patients with Alzheimer's disease. Int J Geriatr Psychiatry. 2007 Oct;22(10):951–6.

82 Diesfeldt HFA. [Interpreting change scores of the Behavioural Rating Scale for Geriatric Inpatients (GIP)]. Tijdschr Gerontol Geriatr. 2013 Sep;44(4):166–74.

83 Fernández-Matarrubia M, Matías-Guiu JA, Moreno-Ramos T, Valles-Salgado M, Marcos-Dolado A, García-Ramos R, et al. Validation of the Lille's Apathy Rating Scale in Very Mild to Moderate Dementia. The American Journal of Geriatric Psychiatry. 2016 Jul;24(7):517–27.

84 Camozzato AL, Kochhann R, Simeoni C, Konrath CA, Franz AP, Carvalho A, et al. Reliability of the Brazilian Portuguese version of the Neuropsychiatric Inventory (NPI) for patients with Alzheimer's disease and their caregivers. International Psychogeriatrics. 2008 Apr;20(2):383–93.

85 Davidsdottir SR, Snaedal J, Karlsdottir G, Atladottir I, Hannesdottir K. Validation of the Icelandic version of the Neuropsychiatric Inventory with Caregiver Distress (NPI-D). Nord J Psychiatry. 2012 Feb;66(1):26–32.

86 Leung VP, Lam LC, Chiu HF, Cummings JL, Chen QL. Validation study of the Chinese version of the neuropsychiatric inventory (CNPI). Int J Geriatr Psychiatry. 2001 Aug;16(8):789–93.

87 de Medeiros K, Robert P, Gauthier S, Stella F, Politis A, Leoutsakos J, et al. The Neuropsychiatric Inventory-Clinician rating scale (NPI-C): reliability and validity of a revised assessment of neuropsychiatric symptoms in dementia. Int Psychogeriatr. 2010 Sep;22(6):984–94.

88 Stella F, Forlenza OV, Laks J, de Andrade LP, Ljubetic Avendaño MA, Gasparetto Sé EV, et al. The Brazilian version of the Neuropsychiatric Inventory-Clinician rating (NPI-C): Reliability and validity in dementia. Int Psychogeriatr. 2013 Sep;25(9):1503–11.

89 Valentino V, Iavarone A, Amboni M, Moschiano F, Picillo M, Petretta V, et al. Apathy in Parkinson's disease: differences between caregiver's report and self-evaluation. Funct Neurol. 2018 Mar;33(1):31–5.

90 Guimarães HC, Fialho PPA, Carvalho VA, Santos EL dos, Caramelli P, Guimarães HC, et al. Brazilian caregiver version of the Apathy Scale. Dementia & Neuropsychologia. 2009 Dec;3(4):321–6.

91 Wetmore JB, Arbelo JM, Catalán MJ, Valldeoriola F, Rodriguez-Blazquez C, Martinez-Martin P. Psychometric Properties of the Apathy Scale in Advanced Parkinson's Disease [Internet]. Parkinson’s Disease. 2019 Mar;2019:e1965394.

92 Bertens AS, Moonen JEF, de Waal MWM, Foster-Dingley JC, de Ruijter W, Gussekloo J, et al. Validity of the three apathy items of the Geriatric Depression Scale (GDS-3A) in measuring apathy in older persons. Int J Geriatr Psychiatry. 2017;32(4):421–8.

93 Belanger HG, Brown LM, Crowell TA, Vanderploeg RD, Curtiss G. The Key Behaviors Change Inventory and executive functioning in an elderly clinic sample. Clin Neuropsychol. 2002 Aug;16(3):251–7.

94 Choi SH, Na DL, Kwon HM, Yoon SJ, Jeong JH, Ha CK. The Korean version of the neuropsychiatric inventory: a scoring tool for neuropsychiatric disturbance in dementia patients. J Korean Med Sci. 2000 Dec;15(6):609–15.

95 Politis AM, Mayer LS, Passa M, Maillis A, Lyketsos CG. Validity and reliability of the newly translated Hellenic Neuropsychiatric Inventory (H-NPI) applied to Greek outpatients with Alzheimer's disease: a study of disturbing behaviors among referrals to a memory clinic. Int J Geriatr Psychiatry. 2004 Mar;19(3):203–8.

96 Kat MG, de Jonghe JFM, Aalten P, Kalisvaart CJ, Dröes RM, Verhey FRJ. [Neuropsychiatric symptoms of dementia: psychometric aspects of the Dutch Neuropsychiatric Inventory (NPI)]. Tijdschr Gerontol Geriatr. 2002 Sep;33(4):150–5.

97 Weintraut R, Karádi K, Lucza T, Kovács M, Makkos A, Janszky J, et al. Lille Apathy Rating Scale and MDS-UPDRS for Screening Apathy in Parkinson's Disease. J Parkinsons Dis. 2016;6(1):257–65.

98 Kirsch-Darrow L, Zahodne LB, Hass C, Mikos A, Okun MS, Fernandez HH, et al. How Cautious Should We Be When Assessing Apathy with the Unified Parkinson's Disease Rating Scale? Mov Disord. 2009 Apr;24(5):684–8.

99 Pedersen KF, Larsen JP, Aarsland D. Validation of the Unified Parkinson's Disease Rating Scale (UPDRS) section I as a screening and diagnostic instrument for apathy in patients with Parkinson's disease. Parkinsonism & Related Disorders. 2008 Apr;14(3):183–6.

100 Starkstein SE, Merello M. The Unified Parkinson's Disease Rating Scale: Validation study of the mentation, behavior, and mood section. Movement Disorders. 2007;22(15):2156–61.

101 Cummings JL, Friedman JH, Garibaldi G, Jones M, Macfadden W, Marsh L, et al. Apathy in Neurodegenerative Diseases: Recommendations on the Design of Clinical Trials. J Geriatr Psychiatry Neurol. 2015 Sep;28(3):159–73.

102 Marin RS. Apathy: a neuropsychiatric syndrome. JNP. 1991 Aug;3(3):243–54.

103 König A, Aalten P, Verhey F, Bensadoun G, Petit P-D, Robert PH, et al. A review of current information and communication technologies: can they be used to assess apathy? Int J Geriatr Psychiatry. 2014 Apr;29(4):345–58.

104 Kolitz BP. Development of the Key Behaviors Change Inventory (KBCI), a TBI behavioral outcome assessment instrument. 2000 May

105 Goetz CG, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stebbins GT, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Process, format, and clinimetric testing plan. Mov Disord. 2007 Jan;22(1):41–7.

106 Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. Mov Disord. 2008 Nov;23(15):2129–70.

107 GIP-28 Instruction Manual. [GIP-28 Hanleiding]. [cited 2018 Oct 19]. Available from: https://www.werkenindeouderengeneeskunde.nl/wp-content/uploads/2011/04/GIP-handleiding.pdf

Appendix 7 Node labels and descriptions prior to item reduction

Recoding and labelling variables:

Some data was arbitrarily reversed/ reflected for purposes of network analysis (but not descriptive data), to avoid confusing negative associations being present in the graph or inappropriate merging of negative and positive relationships within centrality measures.

AES

Item	Item description	Node label	Originally, low score=	Now, low score=
AES_1	S/he is interested in things.	Interest	Less apathy	More apathy
AES_2	S/he gets things done during the day	GetDone	Less apathy	More apathy
AES_3	Getting things started on his/her own is important to him/her	Started	Less apathy	More apathy
AES_4	S/he is interested in having new experiences	NewExp	Less apathy	More apathy
AES_5	S/he is interested in learning new things	LearnNew	Less apathy	More apathy
AES_6	S/he puts little effort into anything	Effort	Less apathy*	More apathy
AES_7	S/he approaches life with intensity	Intensity	Less apathy	More apathy
AES_8	Seeing a job through to the end is important to her/him	Completion	Less apathy	More apathy
AES_9	S/he spends time doing things that interest her/him	TimeInterest	Less apathy	More apathy
AES_10	Someone has to tell her/him what to do each day	WhatToDo	Less apathy*	More apathy
AES_11	S/he is less concerned about her/his problems than s/he should be	SelfConcern	Less apathy*	More apathy
AES_12	S/he has friends	HasFriends	Less apathy	More apathy
AES_13	Getting together with friends is important to her/him	GetTogether	Less apathy	More apathy
AES_14	When something good happens, s/he gets excited	Excited	Less apathy	More apathy
AES_15	S/he has an accurate understanding of her/his problems	Insight	Less apathy	More apathy
AES_16	Getting things done during the day is important to her/him	ImportDone	Less apathy	More apathy
AES_17	S/he has initiative.	Initiative	Less apathy	More apathy
AES_18	S/he has motivation.	Motivation	Less apathy	More apathy

*= reverse scored as part of data input.

Depression and other

Item	Item description	Node label	Originally, low score=	Now, low score=
HADS_2	I feel as if I am slowed down	NotSlowed	Less depression*	More depression
HADS_6	I have lost interest in my appearance	InterestinAppear	Less depression*	More depression
HADS_3	I still enjoy the things I used to enjoy	StillEnjoy	Less depression	More depression†
HADS_7	I can laugh and see the funny side of things	Laugh	Less depression	More depression†
HADS_10	I look forward with enjoyment to things	LookForward	Less depression	More depression†
HADS_14	I can enjoy a good book or radio or TV programme	EnjoyMedia	Less depression	More depression†
DEMQOL_P_1	Cheerful?	Cheerful	Less cheerful*	Less cheerful
DEMQOL_P_2	Worried or Anxious?	NotWorried	More worry	More worry
DEMQOL_P_4	Full of Energy?	Energy	Less energy*	Less energy
DEMQOL_P_5	Sad?	NotSad	More sadness	More sadness
DEMQOL_P_9	Irritable?	NotIrritable	More irritability	More irritability
DEMQOL_8	Lonely?	NotLonely	More loneliness	More loneliness
DAD_Initiation_Score	Calculated from initiation subQs, accounting for	ADLInit	Worse ADL	Worse ADL
DAD_Planning_Score	Calculated from Planning & organization subQs	ADLPlan	Worse ADL	Worse ADL
DAD_Performance_Score	Calculated from performance sub Qs	ADLPerf	Worse ADL	Worse ADL
MoCA_Score	MoCA scored as per manual	MoCA	Worse EF	Worse EF
Age	Participant Age in years	Age	Lower age	Lower age
VF_correct	Number of animals correctly named	FluencyEF	Worse EF	Worse EF
BBS_Score	Berg balance score (sum of item scores)	Balance	Worse balance	Worse balance
TUG_Trial	TUG score (time to walk in seconds)	WalkSpeed	Better mobility	Worse mobility
SHARE_FI	SHARE frailty index scored as per manual	Frailty	Less frail	More frail

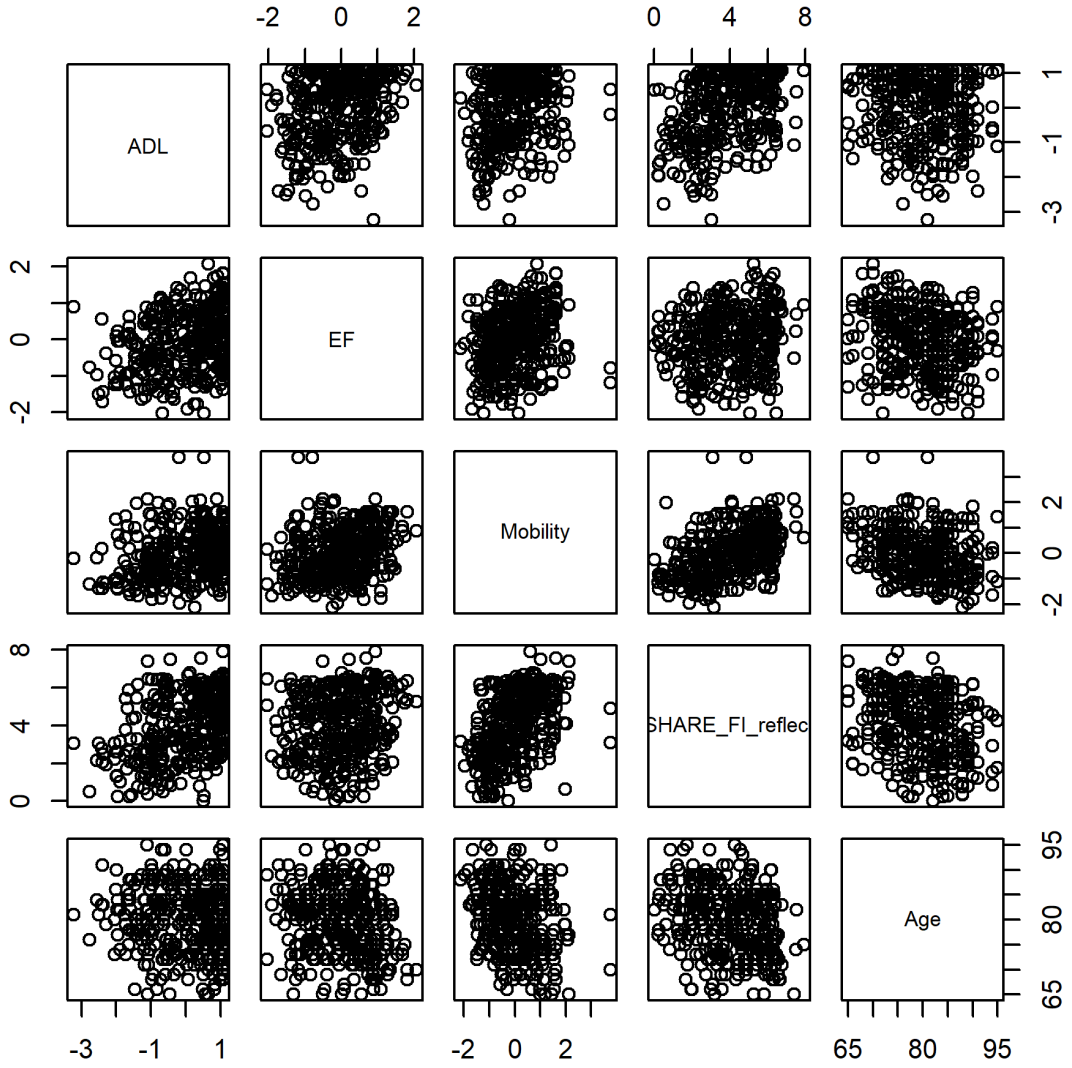
*= reverse scored as part of data input.

Appendix 8 Transcription Notations

Notation	Meaning
[…]	Removal of part of the transcript for purposes of clarity or conciseness
…	Pause in speech
-	A break in speech. Speaker was cut off or cut themselves off mid-sentence.
<>	Action
[ ]	Text that has been added or altered for clarity or anonymity

Appendix 9 Scatterplot matrix

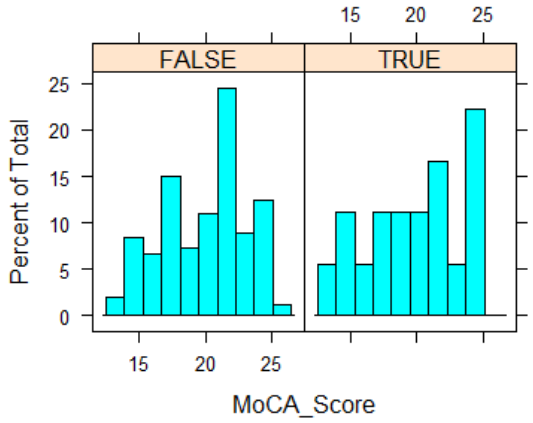
Scatterplot matrix shows continuous data after transformation and merging variables with topological overlap.



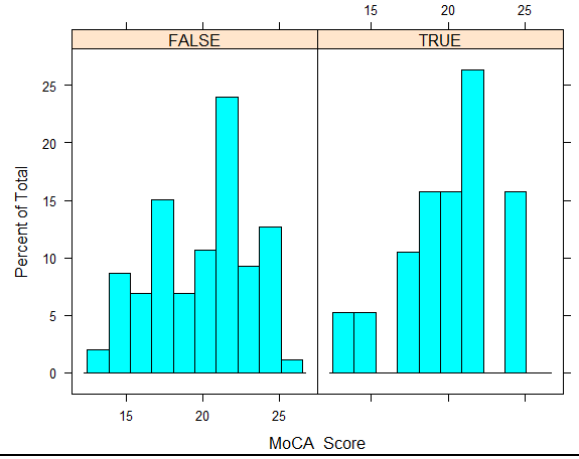
Appendix 10Missing Data histograms

MoCA score distribution, split by whether key variables were missing. ‘TRUE’ indicates the variable was missing for at least one item, where ‘FALSE indicates that all items were present.

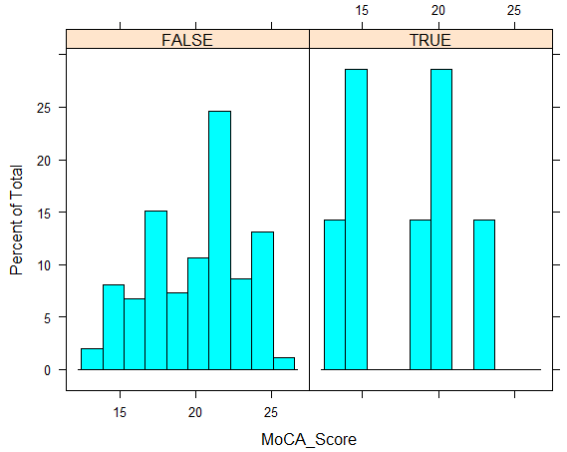
AES



BBS

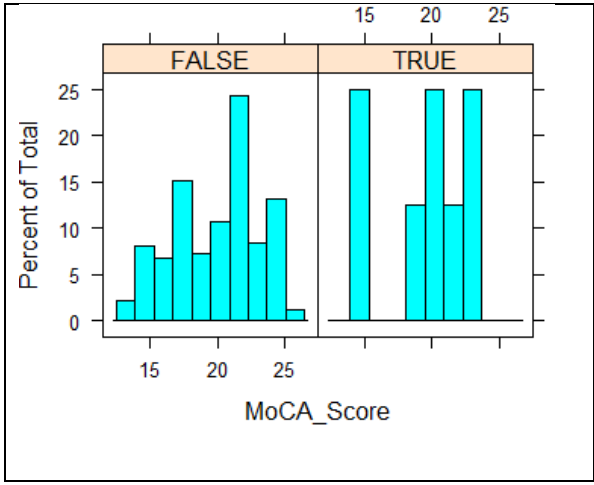


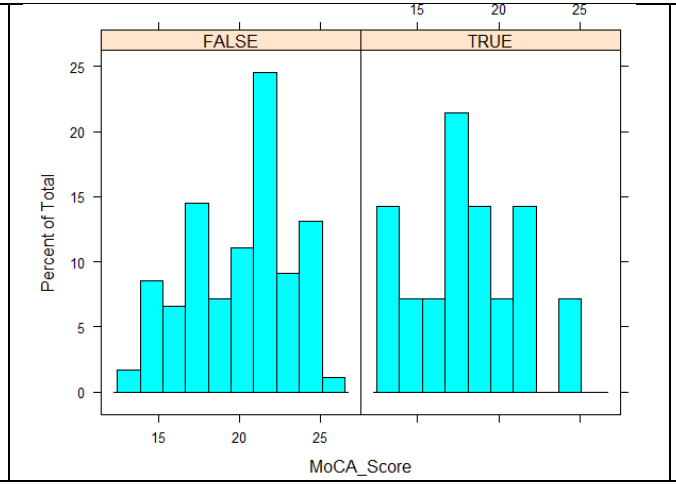
HADS-D



SHARE-FI

TUG

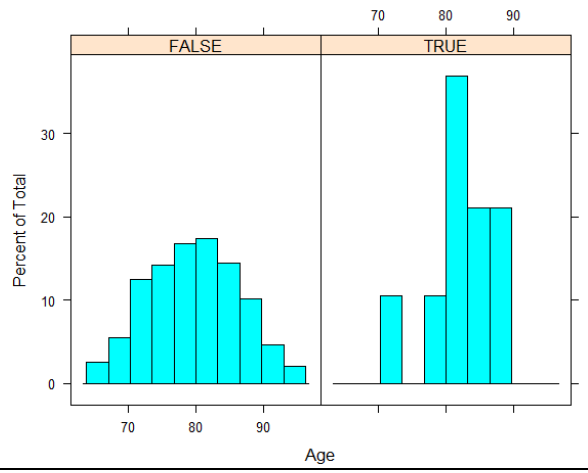




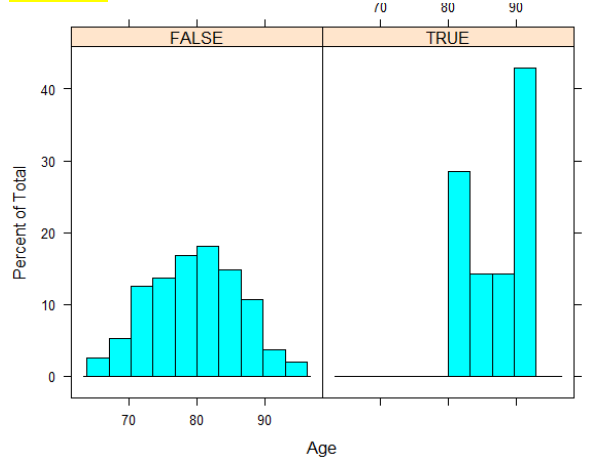
Age distribution, split by whether key variables were missing. ‘TRUE’ indicates the variable was missing for at least one item, where ‘FALSE indicates that all items were present.

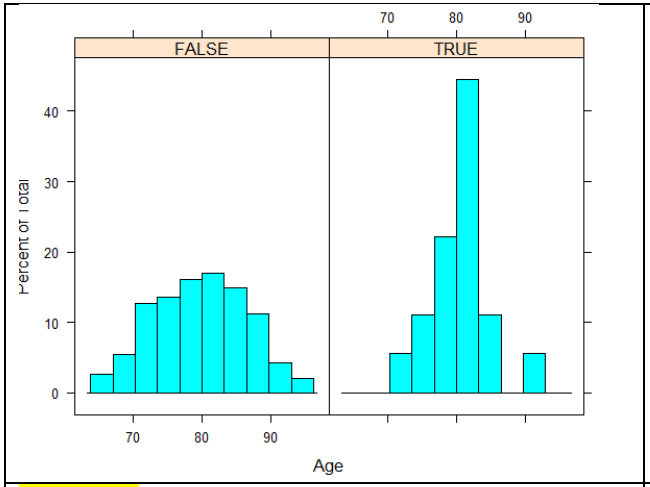
AES

BBS

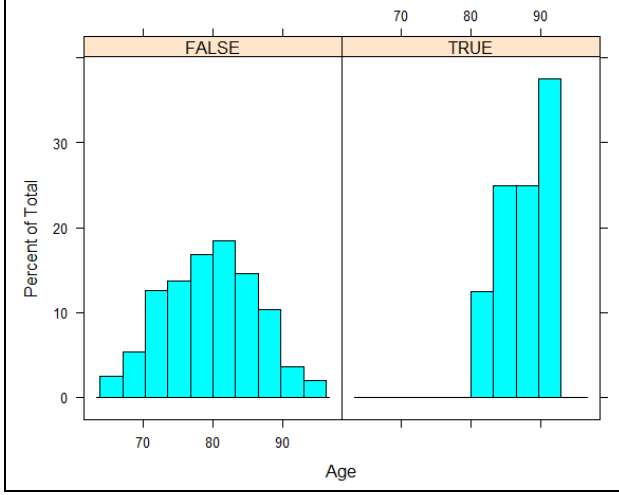


HADS-D

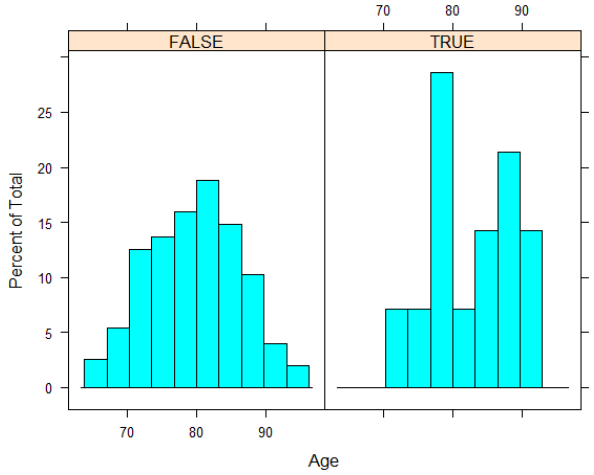



FALSE
TRUE
70
80
90
0
10
20
30
40
Age

SHARE-FI

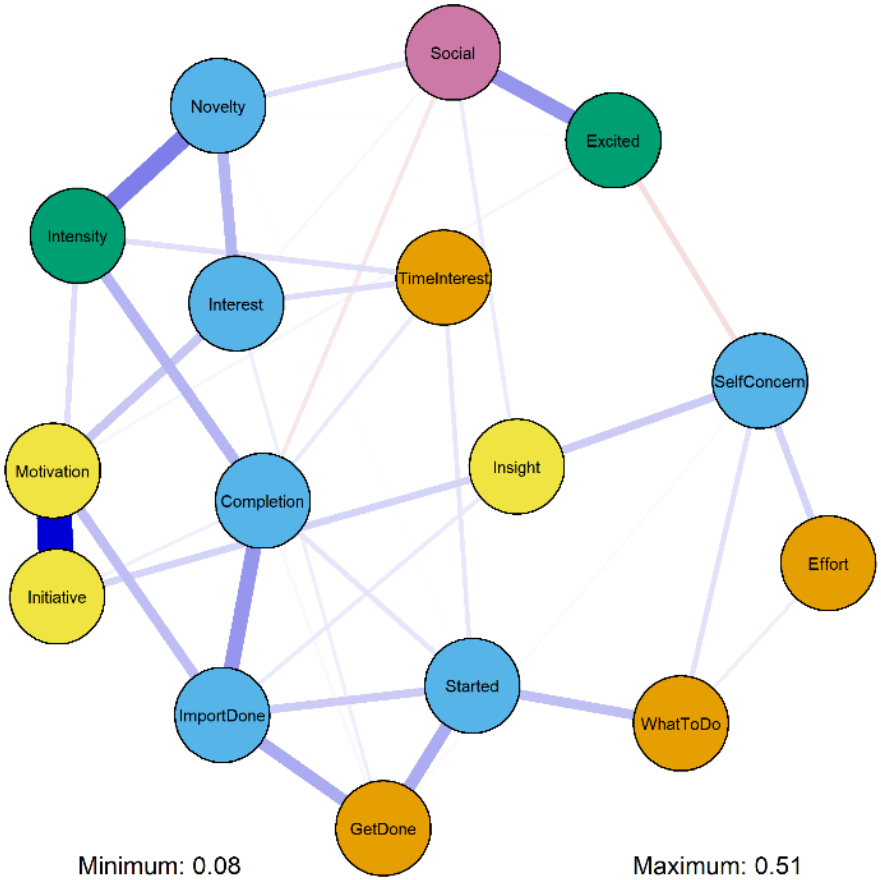

FALSE
TRUE
70
80
90
0
10
20
30
Percent of Total
Age

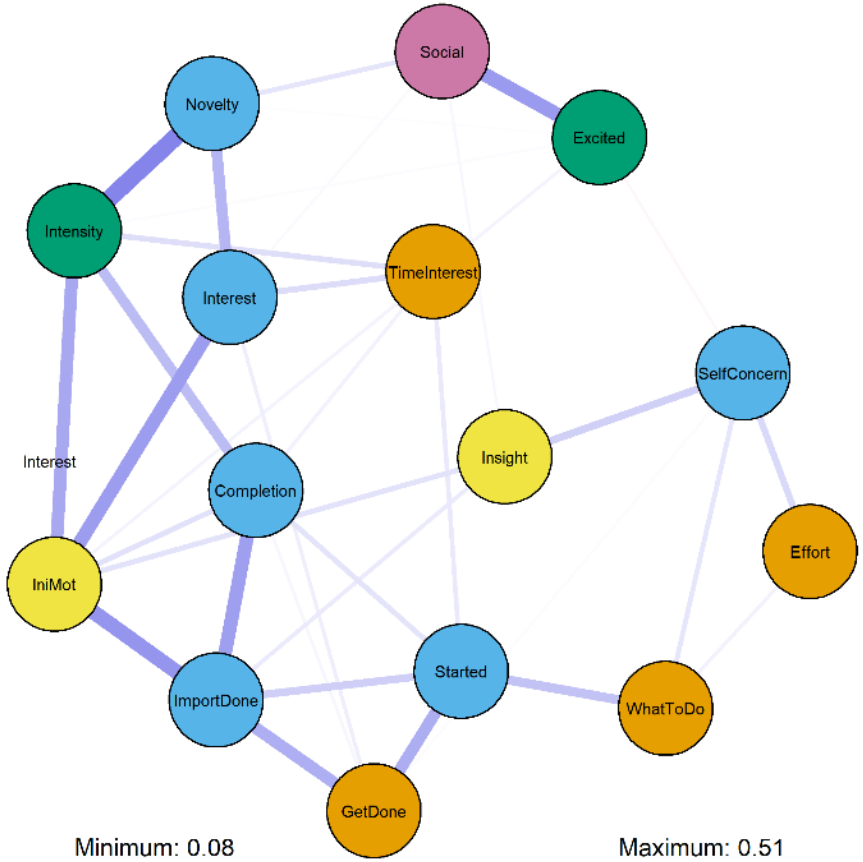
TUG


FALSE
TRUE
70
80
90
0
5
10
15
20
25
Percent of Total
Age

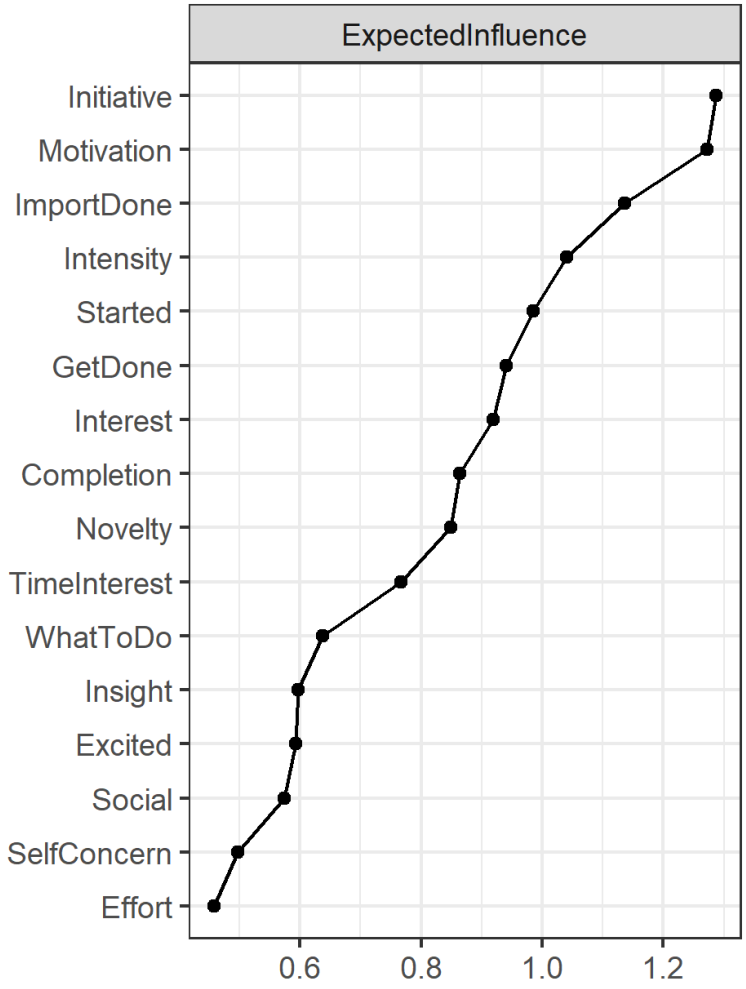
Appendix 11Effect of combining of motivation and initiative in Phase 1b

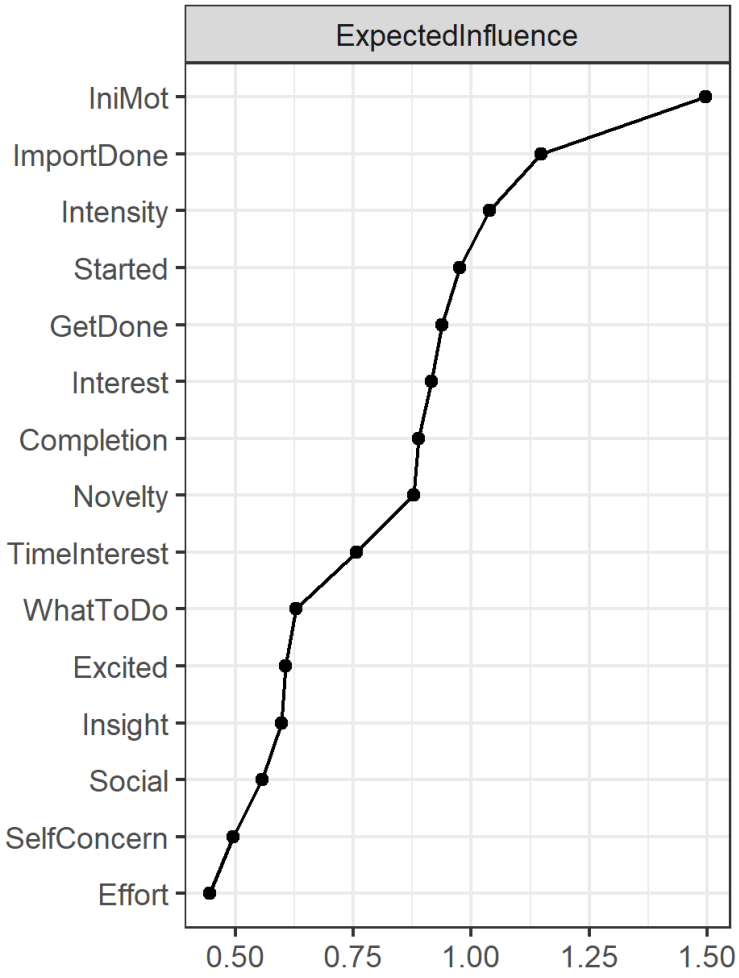
Network plots with motivation and initiative as separate nodes, and then combined





Centrality with motivation and initiative as separate nodes, and then combined



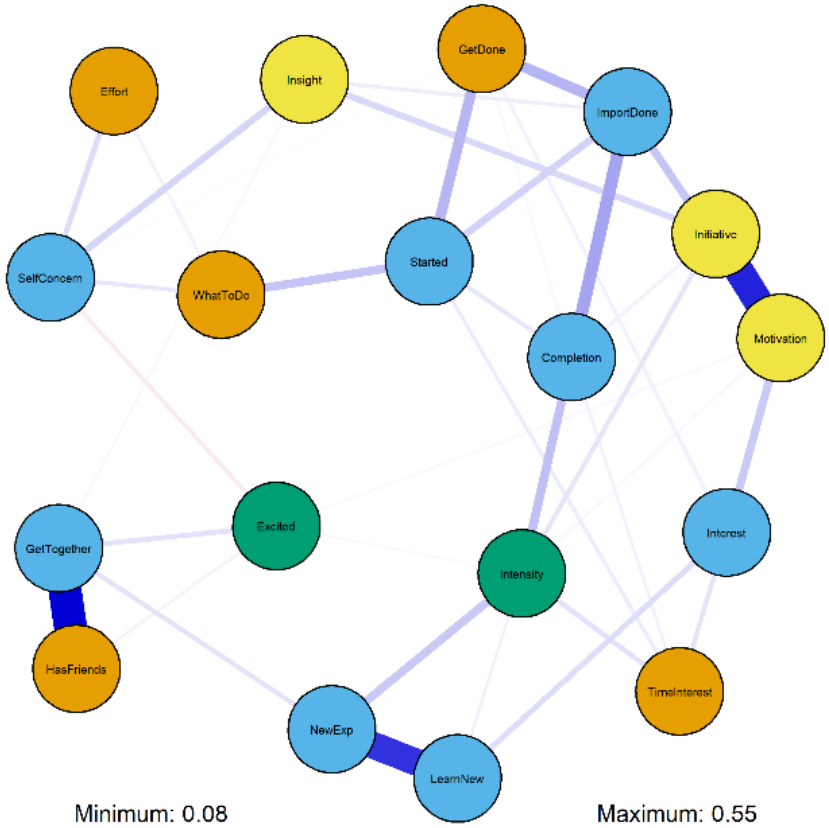


Appendix 12Pre and Post UVA comparison

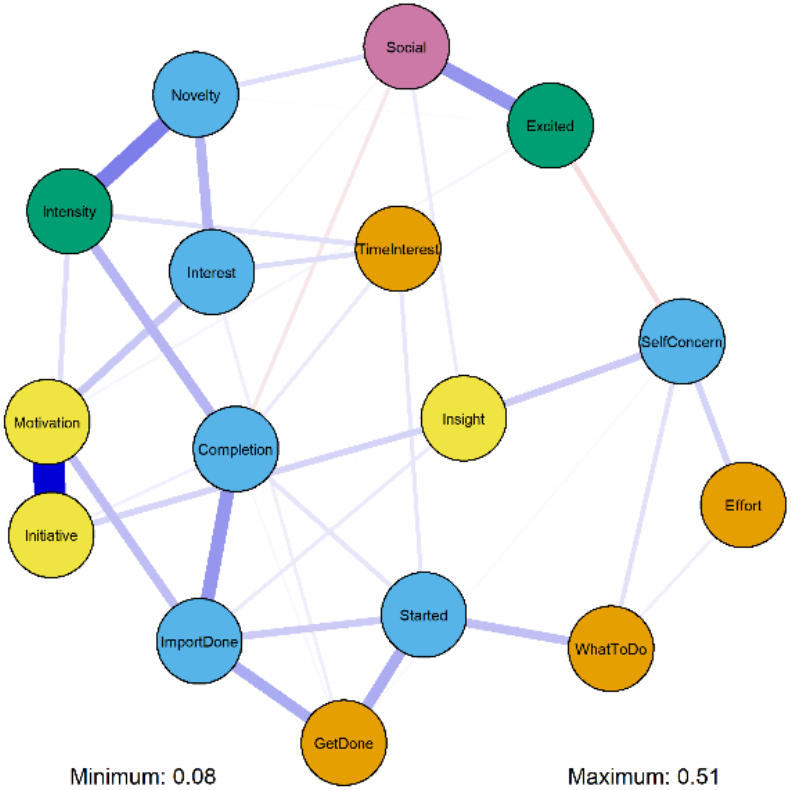
Each phase was separated into before UVA (Part A) and after UVA (Part B). It is worth noting that part B was performed before moving on to the next phase, so that the next phase contained any merged variables from the previous phase. There was no part B for Phase 2, as no additional nodes were reduced in this phase.

Comparison of Network plots between Phase 1A (left) and Phase 1B (right)

Note that interpretations should not be made regarding node placement, as slight differences in the network results in large differences in locations of nodes. An average layout could not be applied due to differing number of nodes.

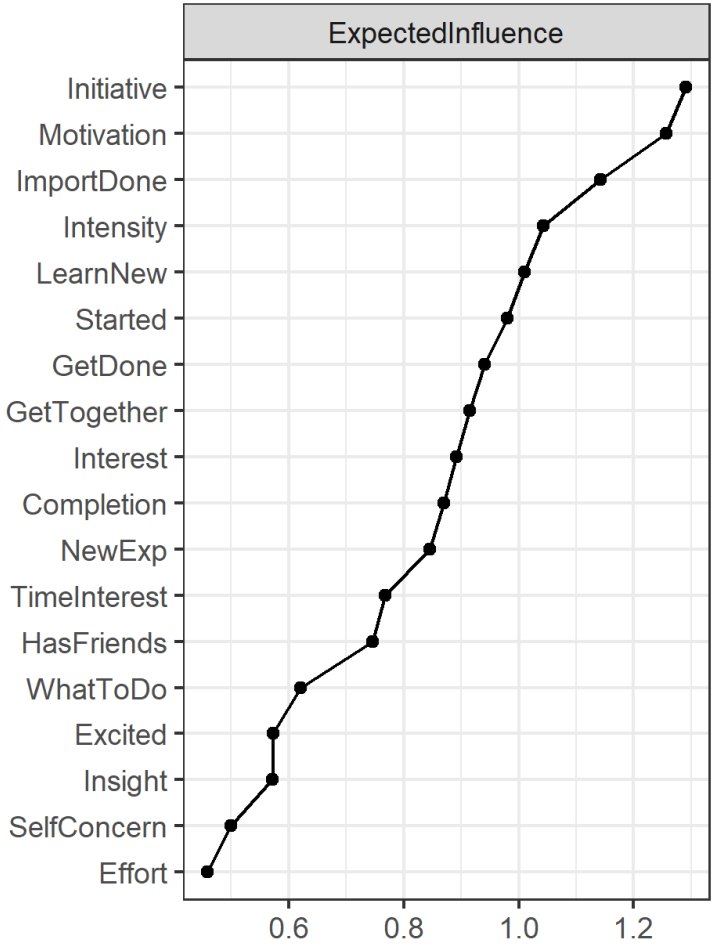


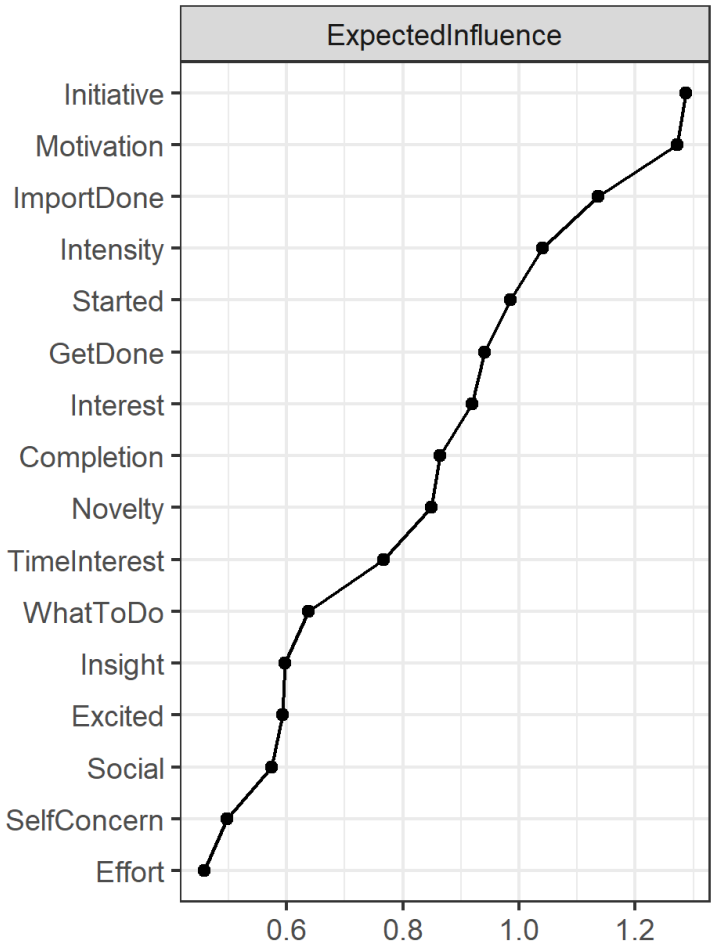
Minimum: 0.08 Maximum: 0.55



Minimum: 0.08 Maximum: 0.51

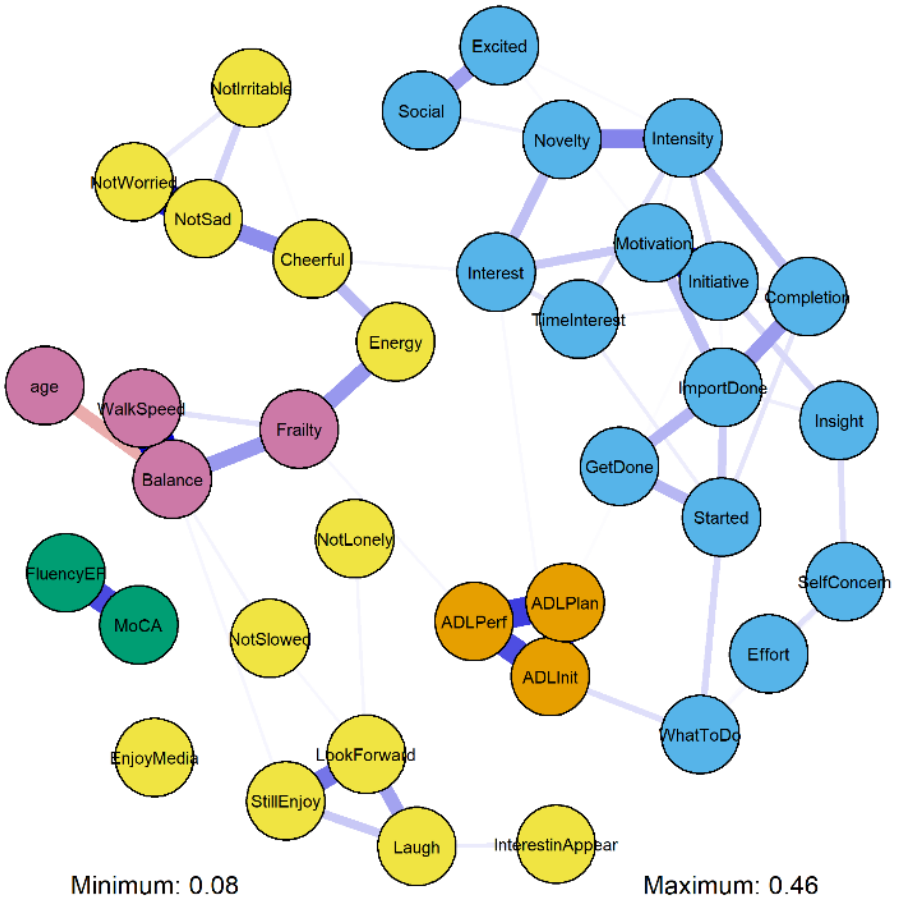
Comparison of centrality estimates between phase 1a (left) and phase 1b (right)

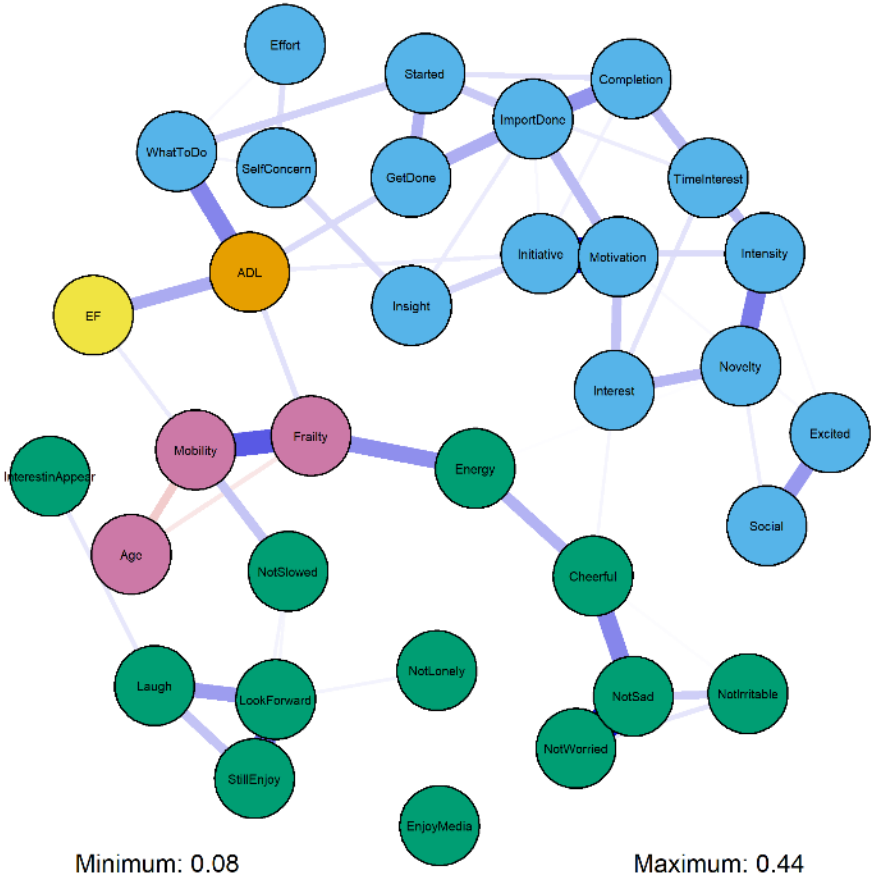




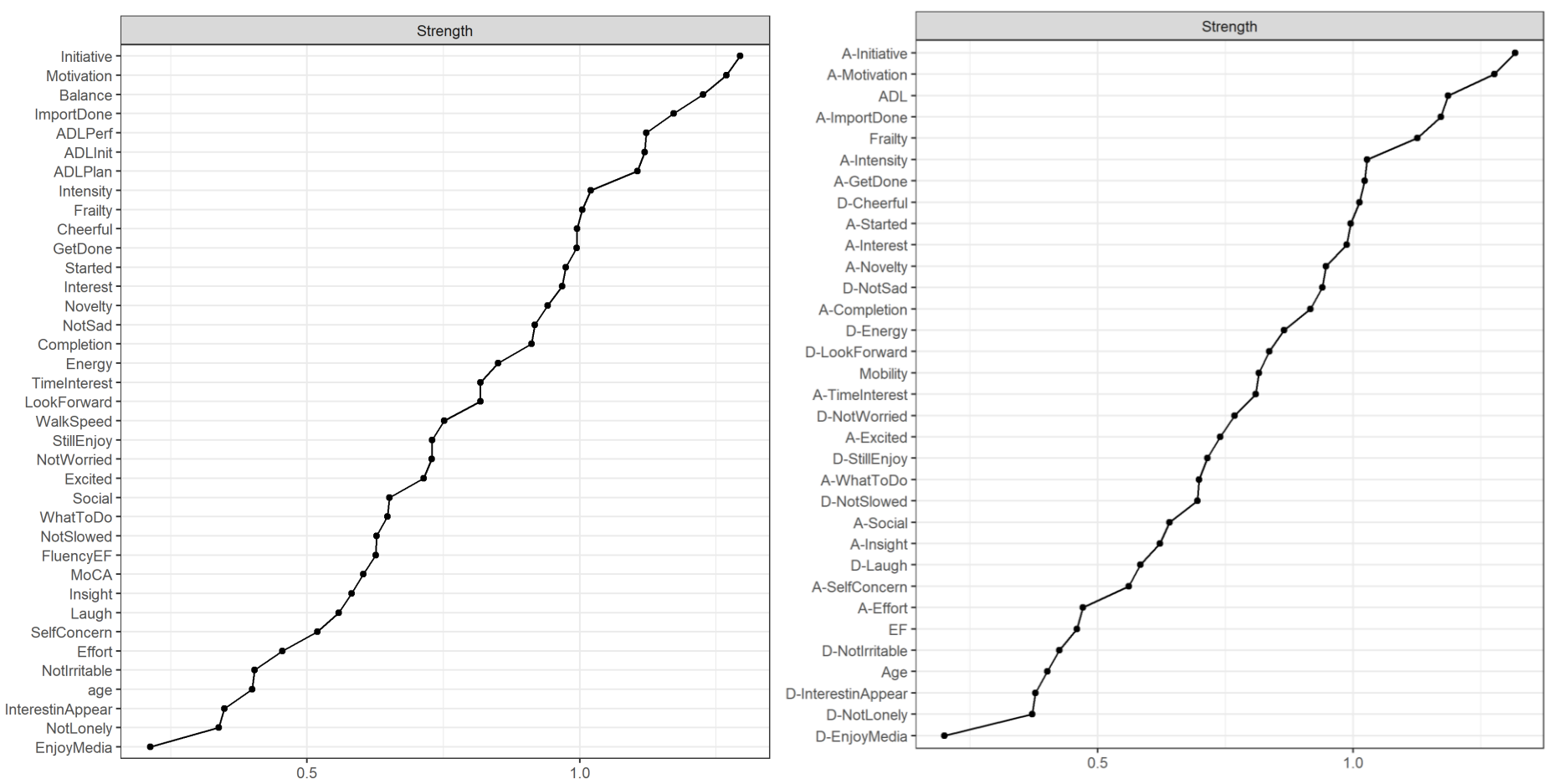
Comparison of Network plots between Phase 3A (left) and Phase 3B (right)

Note that interpretations should not be made regarding node placement, as slight differences in the network results in large differences in locations of nodes. Note that nodes are also colour coded differently here.



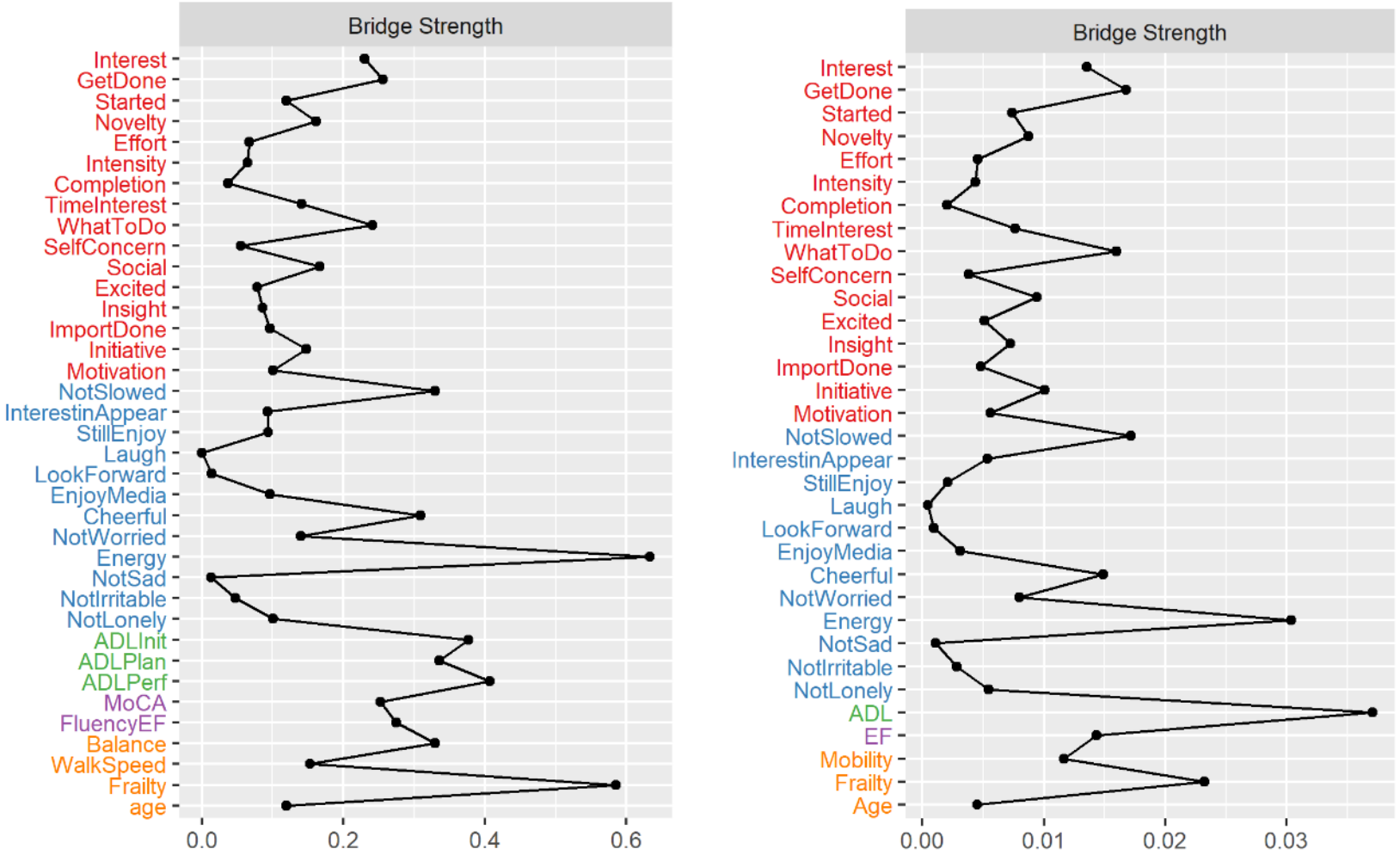


Comparison of centrality estimates between Phase 3a (left) and Phase 3b (right)



Phase 3a and Phase 3b centrality plots.

Comparison of bridge centrality estimates between Phase 3a (left) and Phase 3b (right)



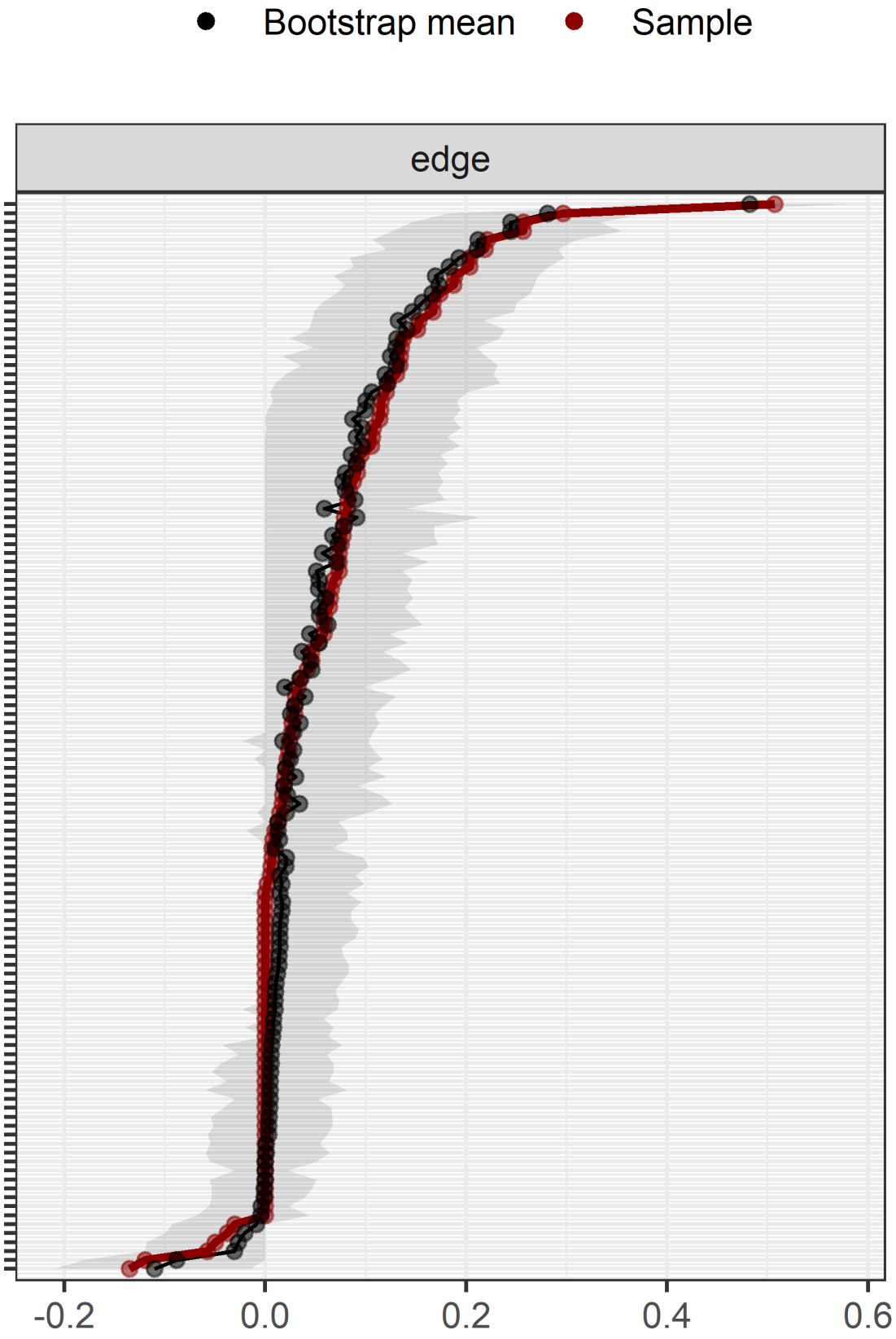
Appendix 13 Phase 1 network analysis

This appendix contains: a table of edge weights and edge weight stability plot , for phase 1 GGM.

P1b Edge weights

Node	AES_1r	AES_2r	AES_3r	Novelty	AES_6r	AES_7r	AES_8r	AES_9r	AES_10r	AES_11r	Social	AES_14r	AES_15r	AES_16r	AES_17r	AES_18r
AES_1r	0	0.107	0	0.203	0.0004	0	0	0.137	0.017	0.068	0.091	0.042	0	0.006	0.073	0.174
AES_2r	0.107	0	0.219	0	0.046	0	0.091	0	0.075	0.087	0.01	0	0.019	0.221	0.064	0
AES_3r	0	0.219	0	0.082	0	0	0.122	0.12	0.187	0	0	0	0.035	0.167	0.053	0
Novelty	0.203	0	0.082	0	0	0.297	0.002	0.059	-0.05	0.033	0.134	0.084	-0.058	0	0	0.064
AES_6r	0.0004	0.046	0	0	0	0.008	0.046	0	0.108	0.151	0	0	0	0.026	0	0.073
AES_7r	0	0	0	0.297	0.008	0	0.203	0.134	0.019	-0.038	0.066	0.078	0.029	0.026	0.135	0.082
AES_8r	0	0.091	0.122	0.002	0.046	0.203	0	0.115	0	0.02	-0.119	0.005	0	0.257	0.105	0.017
AES_9r	0.137	0	0.12	0.059	0	0.134	0.115	0	0.03	0	0.014	0.077	0.021	-0.031	0.03	0.059
AES_10r	0.017	0.075	0.187	-0.05	0.108	0.019	0	0.03	0	0.13	0	0.024	0	0.023	0.074	0
AES_11r	0.068	0.087	0	0.033	0.151	-0.038	0.02	0	0.13	0	0	-0.135	0.167	0	0	0.013
Social	0.091	0.01	0	0.134	0	0.066	-0.119	0.014	0	0	0	0.257	0.114	0	0.007	0
AES_14r	0.042	0	0	0.084	0	0.078	0.005	0.077	0.024	-0.135	0.257	0	0	0.058	0.006	0.095
AES_15r	0	0.019	0.035	-0.058	0	0.029	0	0.021	0	0.167	0.114	0	0	0.115	0.153	0
AES_16r	0.006	0.221	0.167	0	0.026	0.026	0.257	-0.031	0.023	0	0	0.058	0.115	0	0.078	0.189
AES_17r	0.073	0.064	0.053	0	0	0.135	0.105	0.03	0.074	0	0.007	0.006	0.153	0.078	0	0.507
AES_18r	0.174	0	0	0.064	0.073	0.082	0.017	0.059	0	0.013	0	0.095	0	0.189	0.507	0

P1b Edge weight stability Plot



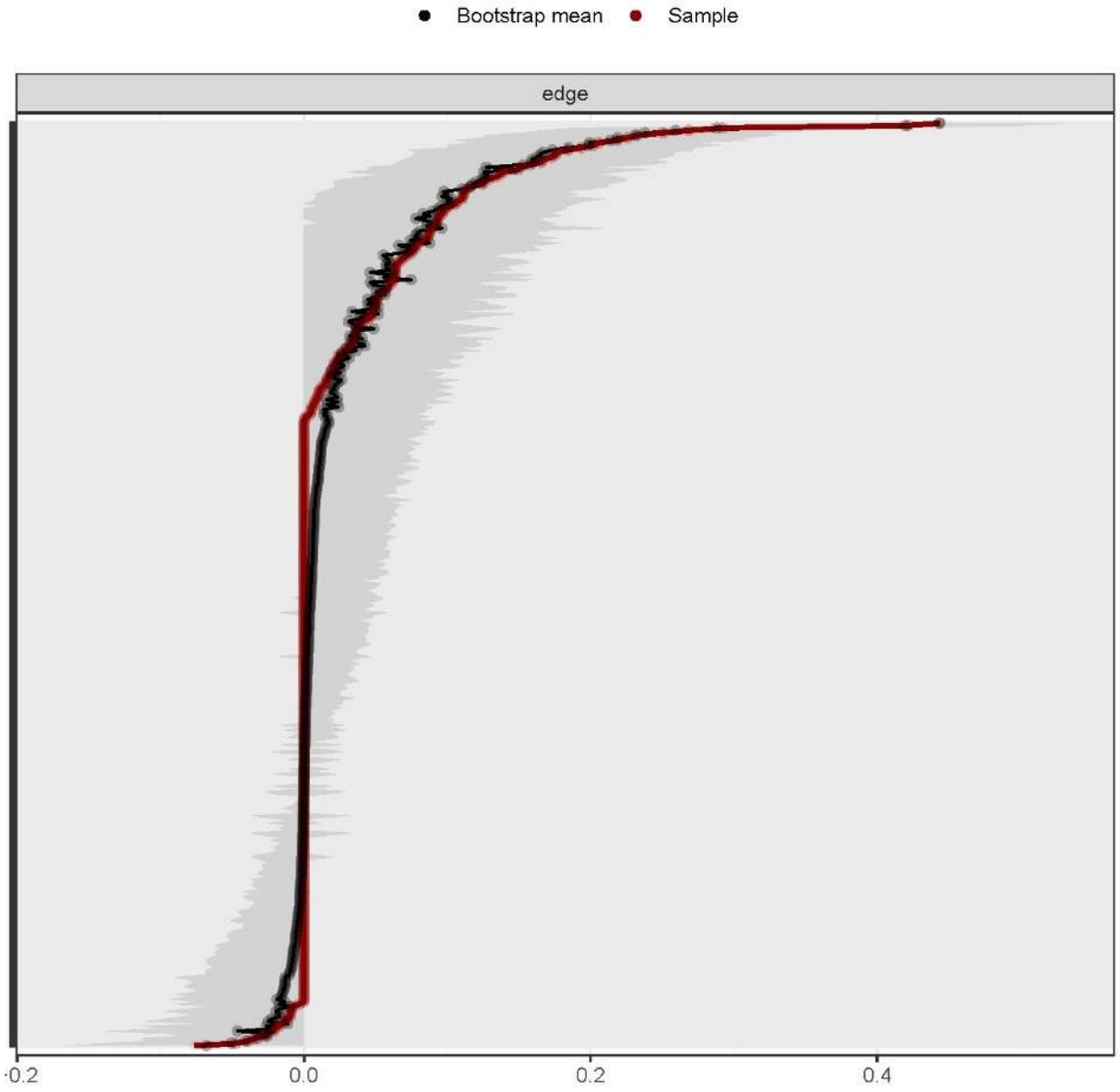
Appendix 14 Phase 2 Network analysis

This appendix contains: a table of edge weights and edge weight stability plot, for phase 2 GGM

P2a Edge Weights

Column1	AES_1r	AES_2r	AES_3r	Novelty	AES_6r	AES_7r	AES_8r	AES_9r	AES_10r	AES_11r	Social	AES_14r	AES_15r	AES_16r	AES_17r	AES_18r	HADS_2r	HADS_6r	HADS_3r	HADS_7r	HADS_10r	HADS_14r	DEMQOL_P_1	DEMQOL_P_2	DEMQOL_P_4	DEMQOL_P_5	DEMQOL_P_9	DEMQOL_8
AES_1r	0	0.094	0	0.184	0	0	0	0.125	0.011	0.045	0.064	0.035	0	0.015	0.082	0.166	0	0	0	0	0.01	-0.035	0.093	0	0.021	0	0	0
AES_2r	0.094	0	0.206	0	0.037	0	0.091	0	0.076	0.078	0	0	0.024	0.202	0.064	0.001	0	0	0	0	0	0	0.005	0	0.112	0	0	0
AES_3r	0	0.206	0	0.062	0	0	0.125	0.111	0.172	0	0	0	0.034	0.158	0.058	0.02	0	0.06	0	0	0	0	0	0	0	0	0	-0.008
Novelty	0.184	0	0.062	0	0	0.269	0	0.053	0	0	0.106	0.078	0	0	0	0.051	-0.027	0	0	0	0	0	0.036	0	0.086	0	0	0
AES_6r	0	0.037	0	0	0	0	0.041	0	0.094	0.129	0	0	0	0.026	0.007	0.064	0.036	0	0	0	0	0	0	0	0.047	0	0	0
AES_7r	0	0	0	0.269	0	0	0.177	0.134	0.004	0	0.024	0.088	0.016	0.036	0.137	0.092	0	0	0	0	0	0	0	0	0.049	0	0	0
AES_8r	0	0.091	0.125	0	0.041	0.177	0	0.097	0	0.005	0	0	0	0.233	0.101	0.029	0	0	0	0	0	0	0	-0.0001	0.021	0	0	0
AES_9r	0.125	0	0.111	0.053	0	0.134	0.097	0	0.021	0	0	0.066	0.012	0	0.034	0.057	0	0.018	0	0.006	0	0	0.064	-0.012	0	0	0	0
AES_10r	0.011	0.076	0.172	0	0.094	0.004	0	0.021	0	0.11	0	0	0	0.03	0.072	0.008	-0.009	0	0	0	0	0	0	0	0	0	0	0
AES_11r	0.045	0.078	0	0	0.129	0	0.005	0	0.11	0	0	-0.021	0.139	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.048
Social	0.064	0	0	0.106	0	0.024	0	0	0	0	0	0.226	0.067	0	0	0	0	0	0	0	0	0	0.019	0	0.063	0	0.036	0.031
AES_14r	0.035	0	0	0.078	0	0.088	0	0.066	0	-0.021	0.226	0	0	0.04	0.011	0.088	0	0	0	0	0	0	0.05	-0.039	0	0	0	0
AES_15r	0	0.024	0.034	0	0	0.016	0	0.012	0	0.139	0.067	0	0	0.111	0.151	0	0	0	0	-0.003	0	0	0	0	0	0	0.015	0
AES_16r	0.015	0.202	0.158	0	0.026	0.036	0.233	0	0.03	0	0	0.04	0.111	0	0.089	0.175	0	0	0	0	0	-0.01	0	-0.028	0.023	0	0	0
AES_17r	0.082	0.064	0.058	0	0.007	0.137	0.101	0.034	0.072	0	0	0.011	0.151	0.089	0	0.444	-0.008	0	0	0	0	0	0.048	0	0	0	0	0
AES_18r	0.166	0.001	0.02	0.051	0.064	0.092	0.029	0.057	0.008	0	0	0.088	0	0.175	0.444	0	-0.016	0	0	0	0	0	0	0	0.07	0	0	0
HADS_2r	0	0	0	-0.027	0.036	0	0	0	-0.009	0	0	0	0	0	-0.008	-0.016	0	0.047	0.107	0	0.099	0	0	0	0.08	0	0.016	0.051
HADS_6r	0	0	0.06	0	0	0	0	0.018	0	0	0	0	0	0	0	0	0.047	0	0.049	0.113	0.063	0	0	0	0	0	0	0.008
HADS_3r	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.107	0.049	0	0.166	0.292	-0.076	0	0	0	0	0	0
HADS_7r	0	0	0	0	0	0	0	0.006	0	0	0	0	-0.003	0	0	0	0	0.113	0.166	0	0.219	-0.018	0.051	0	0	0	0	0.01
HADS_10r	0.01	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.099	0.063	0.292	0.219	0	0	0.056	0	0.001	0	0	0.097
HADS_14r	-0.035	0	0	0	0	0	0	0	0	0	0	0	0	-0.01	0	0	0	0	-0.076	-0.018	0	0	-0.01	0	0	-0.021	0	-0.013
DEMQOL_P_1	0.093	0.005	0	0.036	0	0	0	0.064	0	0	0.019	0.05	0	0	0.048	0	0	0	0	0.051	0.056	-0.01	0	0.062	0.194	0.25	0.086	0
DEMQOL_P_2	0	0	0	0	0	0	-0.0001	-0.012	0	0	0	-0.039	0	-0.028	0	0	0	0	0	0	0	0	0.062	0	0	0.422	0.116	0
DEMQOL_P_4	0.021	0.112	0	0.086	0.047	0.049	0.021	0	0	0	0.063	0	0	0.023	0	0.07	0.08	0	0	0	0.001	0	0.194	0	0	0	0	0
DEMQOL_P_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	-0.021	0.25	0.422	0	0	0.149	0.075
DEMQOL_P_9	0	0	0	0	0	0	0	0	0	0	0.036	0	0.015	0	0	0	0.016	0	0	0	0	0	0.086	0.116	0	0.149	0	0
DEMQOL_8	0	0	-0.008	0	0	0	0	0	0	-0.048	0.031	0	0	0	0	0	0.051	0.008	0	0.01	0.097	-0.013	0	0	0	0.075	0	0

P2a Edge weights stability Plot



Appendix 15 Phase 3 network analysis

This appendix contains: a table of edge weights and edge weight stability plot, for phase 3 GGM

P3b Edge Weights

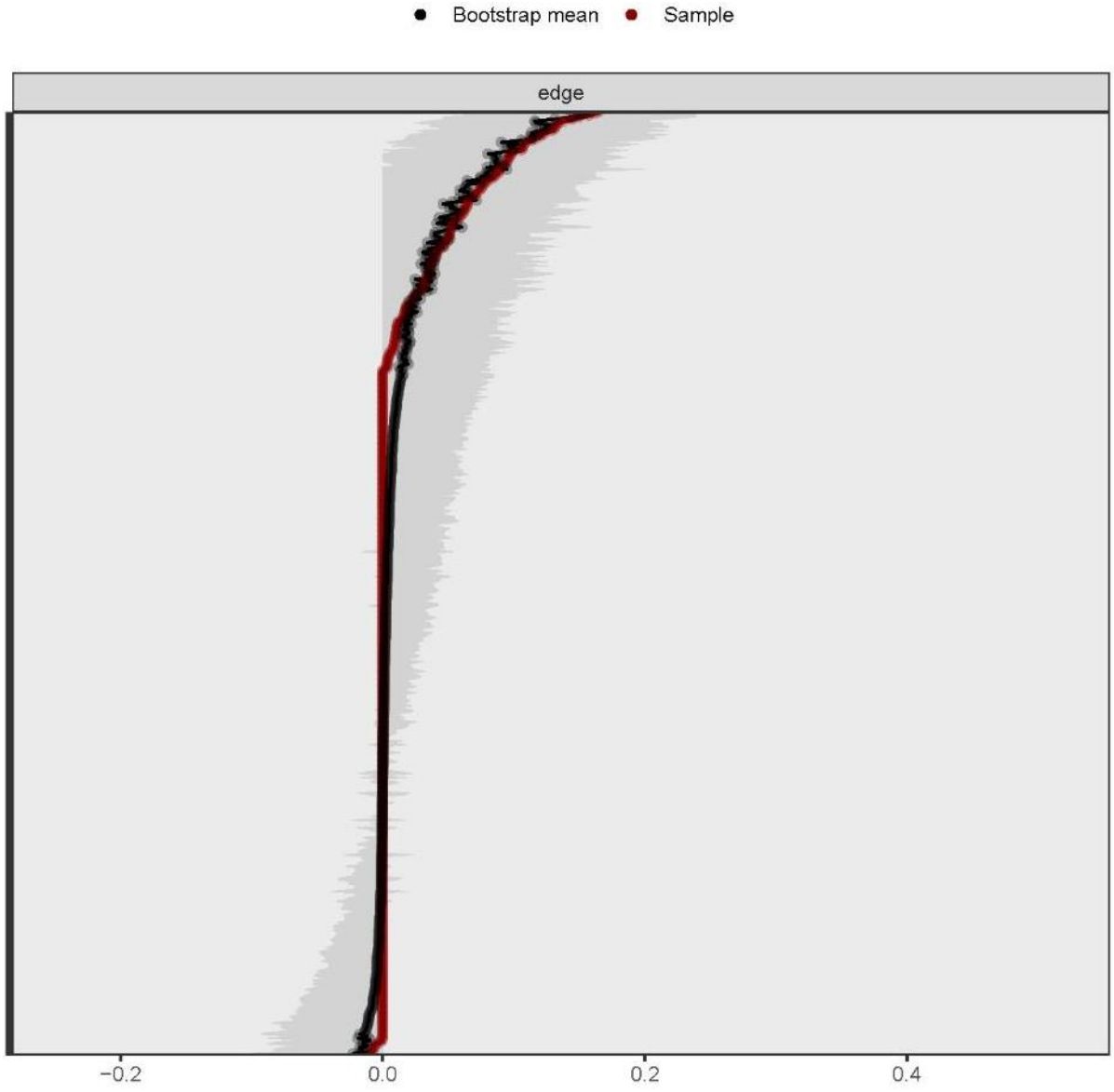
Part 1

Column1	AES_1	AES_2	AES_3	Novel	AES_6	AES_7	AES_8r	AES_9r	AES_1(	AES_1	Social	AES_1	AES_15	AES_16	AES_17	AES_18r
AES_1r	0	0.075	0	0.183	0	0	0	0.119	0	0.039	0.062	0.035	0	0.011	0.069	0.165
AES_2r	0.075	0	0.189	0	0.036	0	0.085	0	0.03	0.073	0	0	0.009	0.194	0.046	0
AES_3r	0	0.189	0	0.058	0	0	0.122	0.107	0.145	0	0	0	0.024	0.157	0.05	0.018
Novelty	0.183	0	0.058	0	0	0.268	0	0.053	0	0	0.106	0.078	0	0	0	0.052
AES_6r	0	0.036	0	0	0	0	0.041	0	0.09	0.129	0	0	0	0.027	0.007	0.064
AES_7r	0	0	0	0.268	0	0	0.175	0.131	0	0	0.023	0.088	0.012	0.034	0.131	0.092
AES_8r	0	0.085	0.122	0	0.041	0.175	0	0.095	0	0.003	0	0	0	0.233	0.099	0.03
AES_9r	0.119	0	0.107	0.053	0	0.131	0.095	0	0.011	0	0	0.066	0.009	0	0.032	0.056
AES_10r	0	0.03	0.145	0	0.09	0	0	0.011	0	0.096	0	0	0	0.022	0.034	0
AES_11r	0.039	0.073	0	0	0.129	0	0.003	0	0.096	0	0	-0.022	0.134	0	0	0
Social	0.062	0	0	0.106	0	0.023	0	0	0	0	0	0.225	0.065	0	0	0
AES_14r	0.035	0	0	0.078	0	0.088	0	0.066	0	-0.022	0.225	0	0	0.04	0.011	0.088
AES_15r	0	0.009	0.024	0	0	0.012	0	0.009	0	0.134	0.065	0	0	0.108	0.138	0
AES_16r	0.011	0.194	0.157	0	0.027	0.034	0.233	0	0.022	0	0	0.04	0.108	0	0.089	0.176
AES_17r	0.069	0.046	0.05	0	0.007	0.131	0.099	0.032	0.034	0	0	0.011	0.138	0.089	0	0.441
AES_18r	0.165	0	0.018	0.052	0.064	0.092	0.03	0.056	0	0	0	0.088	0	0.176	0.441	0
HADS_2r	0	0	0	-0.03	0.034	0	0	0	-0.019	0	0	0	-0.002	0	-0.017	-0.021
HADS_6r	0	0	0.051	0	0	0	0	0.011	0	0	0	0	0	0	0	0
HADS_3r	0	0	0	0	0	0	0	0	0	0	0	0	-0.0004	0	0	0
HADS_7r	0	0	0	0	0	0	0	0.003	0	0	0	0	-0.005	0	0	0
HADS_10r	0.004	0	0	0	0	0	0	0	0	0	0	0	-0.0001	0	0	0
HADS_14r	-0.033	0	0	0	0	0	0	0	0	0	0	0	0	-0.006	0	0
DEMQOL_P_1	0.094	0.005	0	0.036	0	0	0	0.061	0	0	0.019	0.05	0	0	0.047	0
DEMQOL_P_2	0	0	0	0	0	0	0	-0.007	0	0	0	-0.037	0	-0.02	0	0
DEMQOL_P_4	0.012	0.075	0	0.083	0.044	0.042	0.016	0	0	0	0.059	0	0	0.004	0	0.064
DEMQOL_P_5	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
DEMQOL_P_9	0	0	0	0	0	0	0	0	0	0	0.036	0	0.014	0	0	0
DEMQOL_8	0	0	0	0	0	0	0	0	0	-0.041	0.037	0	0	0	0	0
ADL	0.079	0.133	0.071	0	0	0.032	0.017	0.01	0.252	0.024	0.01	0	0.064	0.0001	0.107	0.01
EF	0	0	0	0	0	0	0	0	0	0	0	0	0.032	0	0	0
Mobility	0	0	0	0	0	0	0	0	0	0	0	0	0.004	0	0	0
SHARE_FI_refl	0.008	0.072	0.003	0	0	0	0	0.037	0	0	0	0	0	0.051	0	0
Age	0	0	0	0	0	0	0	0	0	0	0	-0.0003	0	0	0	0

Part 2:

Column1	HADS_2r	HADS_6	HADS_3	HADS_7	HADS_1	HADS_14r	DEMQOL_P_1	DEMQOL_P_2	DEMQOL_P_4	DEMQOL_P_5	DEMQOL_P_9	DEMQOL_8	ADL	EF	Mobility	SHARE_FI_reflect	Age
AES_1r	0	0	0	0	0.004	-0.033	0.094	0	0.012	0	0	0	0.079	0	0	0.008	0
AES_2r	0	0	0	0	0	0	0.005	0	0.075	0	0	0	0.133	0	0	0.072	0
AES_3r	0	0.051	0	0	0	0	0	0	0	0	0	0	0.071	0	0	0.003	0
Novelty	-0.03	0	0	0	0	0	0.036	0	0.083	0	0	0	0	0	0	0	0
AES_6r	0.034	0	0	0	0	0	0	0	0.044	0	0	0	0	0	0	0	0
AES_7r	0	0	0	0	0	0	0	0	0.042	0	0	0	0.032	0	0	0	0
AES_8r	0	0	0	0	0	0	0	0	0.016	0	0	0	0.017	0	0	0	0
AES_9r	0	0.011	0	0.003	0	0	0.061	-0.007	0	0	0	0	0.01	0	0	0.037	0
AES_10r	-0.019	0	0	0	0	0	0	0	0	0	0	0	0.252	0	0	0	0
AES_11r	0	0	0	0	0	0	0	0	0	0	0	-0.041	0.024	0	0	0	0
Social	0	0	0	0	0	0	0.019	0	0.059	0	0.036	0.037	0.01	0	0	0	0
AES_14r	0	0	0	0	0	0	0.05	-0.037	0	0	0	0	0	0	0	0	-0.0003
AES_15r	-0.002	0	-0.0004	-0.005	-0.0001	0	0	0	0	0	0.014	0	0.064	0.032	0.004	0	0
AES_16r	0	0	0	0	0	-0.006	0	-0.02	0.004	0	0	0	0.0001	0	0	0.051	0
AES_17r	-0.017	0	0	0	0	0	0.047	0	0	0	0	0	0.107	0	0	0	0
AES_18r	-0.021	0	0	0	0	0	0	0	0.064	0	0	0	0.01	0	0	0	0
HADS_2r	0	0.037	0.097	0	0.094	0	0	0	0.04	0	0.016	0.051	0	-0.023	0.162	0.053	0
HADS_6r	0.037	0	0.046	0.113	0.062	0	0	0	0	0	0	0.009	0	0	0.019	0.031	0
HADS_3r	0.097	0.046	0	0.165	0.29	-0.074	0	0	0	0	0	0.0002	0	0	0.003	0.039	0
HADS_7r	0	0.113	0.165	0	0.219	-0.017	0.05	0	0	0	0	0.01	0	0	0	0	0
HADS_10r	0.094	0.062	0.29	0.219	0	0	0.054	0	0	0	0	0.098	0	0	0	0.016	0
HADS_14r	0	0	-0.074	-0.017	0	0	-0.008	0	0	-0.021	0	-0.013	0	0	-0.026	0	0
DEMQOL_P_1	0	0	0	0.05	0.054	-0.008	0	0.065	0.187	0.251	0.085	0	0	0	0	0	0
DEMQOL_P_2	0	0	0	0	0	0	0.065	0	0	0.419	0.116	0	0	-0.027	-0.007	-0.016	0.053
DEMQOL_P_4	0.04	0	0	0	0	0	0.187	0	0	0	0	0	0	0	0	0.239	0
DEMQOL_P_5	0	0	0	0	0	-0.021	0.251	0.419	0	0	0.149	0.077	0	0.005	0	0	0.017
DEMQOL_P_9	0.016	0	0	0	0	0	0.085	0.116	0	0.149	0	0	0	0	0	0.009	0
DEMQOL_8	0.051	0.009	0.0002	0.01	0.098	-0.013	0	0	0	0.077	0	0	-0.035	0	0	0	0
ADL	0	0	0	0	0	0	0	0	0	0	0	-0.035	0	0.199	0.021	0.123	0
EF	-0.023	0	0	0	0	0	0	-0.027	0	0.005	0	0	0.199	0	0.108	0	-0.065
Mobility	0.162	0.019	0.003	0	0	-0.026	0	-0.007	0	0	0	0	0.021	0.108	0	0.315	-0.152
SHARE_FI_refl	0.053	0.031	0.039	0	0.016	0	0	-0.016	0.239	0	0.009	0	0.123	0	0.315	0	-0.115
Age	0	0	0	0	0	0	0	0.053	0	0.017	0	0	0	-0.065	-0.152	-0.115	0

P3b Edge Weight Stability Plot



Appendix 16 MGM comparison

Phase 1 GGM and MGM networks

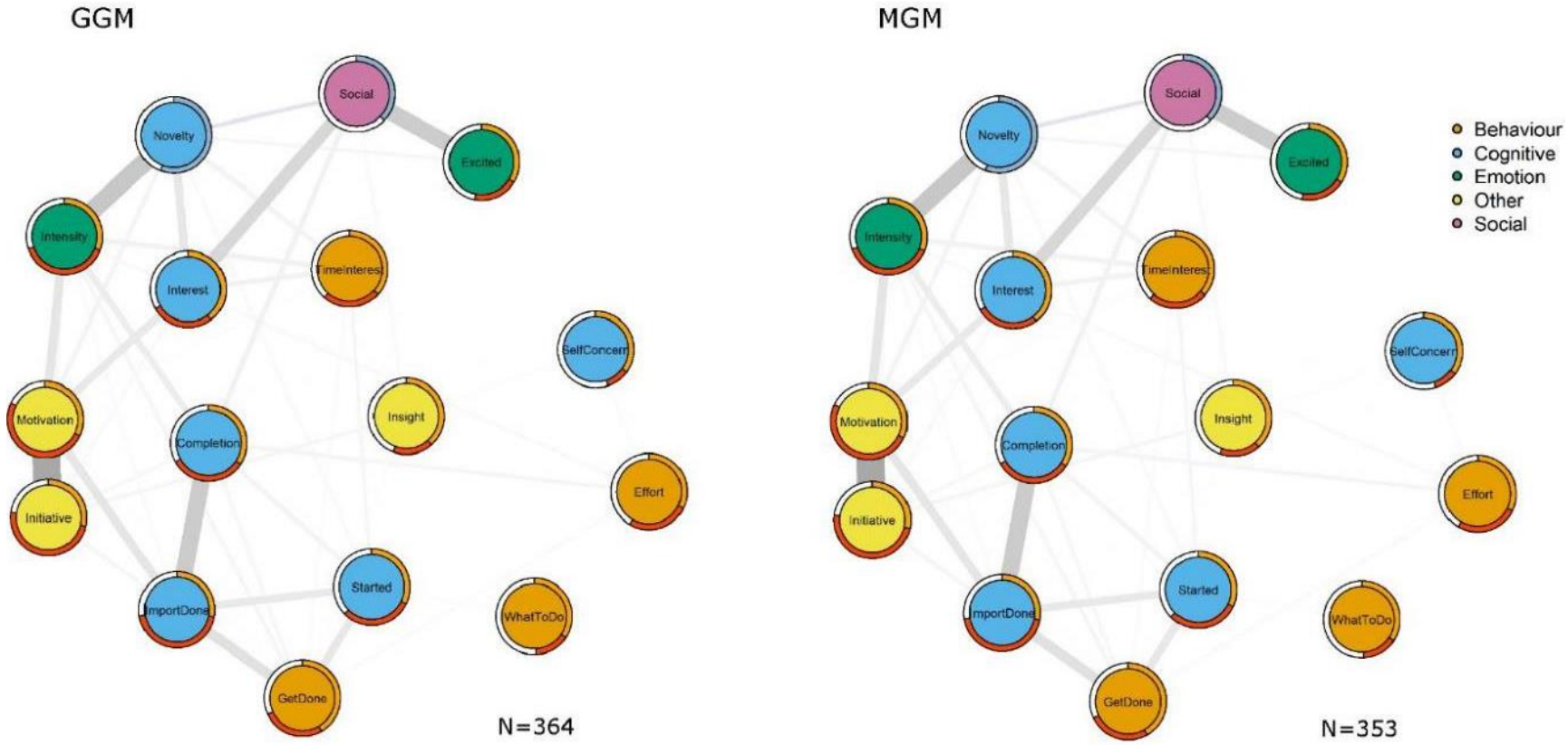


Figure shows GGM and MGM after phase 1. Maximum set at .79 (which was the maximum of all graphs), minimum set at .08, and MGM layout constrained to GGM layout, to aid visual interpretation and comparison

Phase 2 GGM and MGM networks

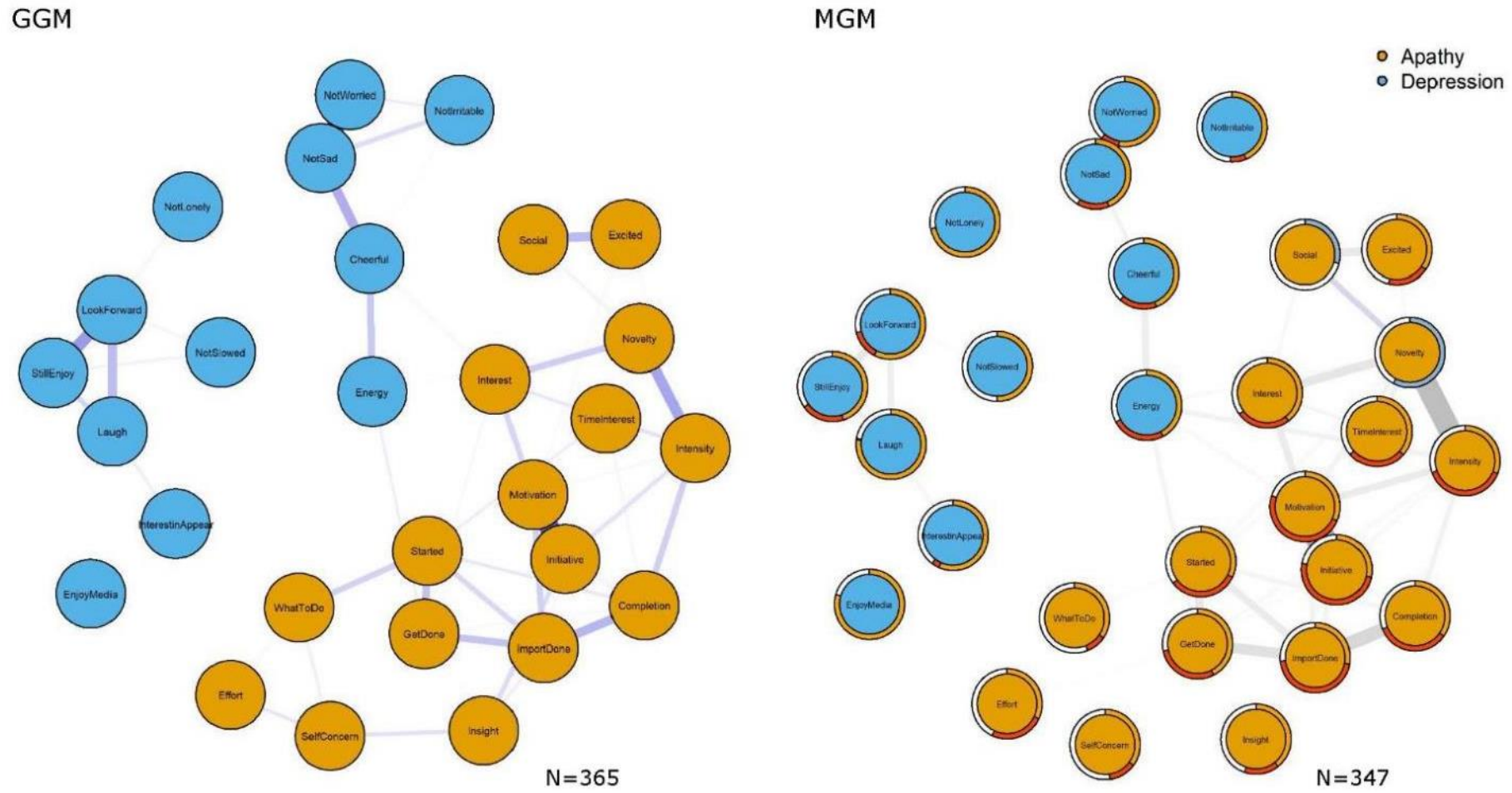


Figure shows GGM and MGM after phase 2. Maximum set at .63 (which was the maximum of all graphs), minimum set at .08, and MGM layout constrained to GGM layout, to aid visual interpretation and comparison.

Phase 3 GGM and MGM networks

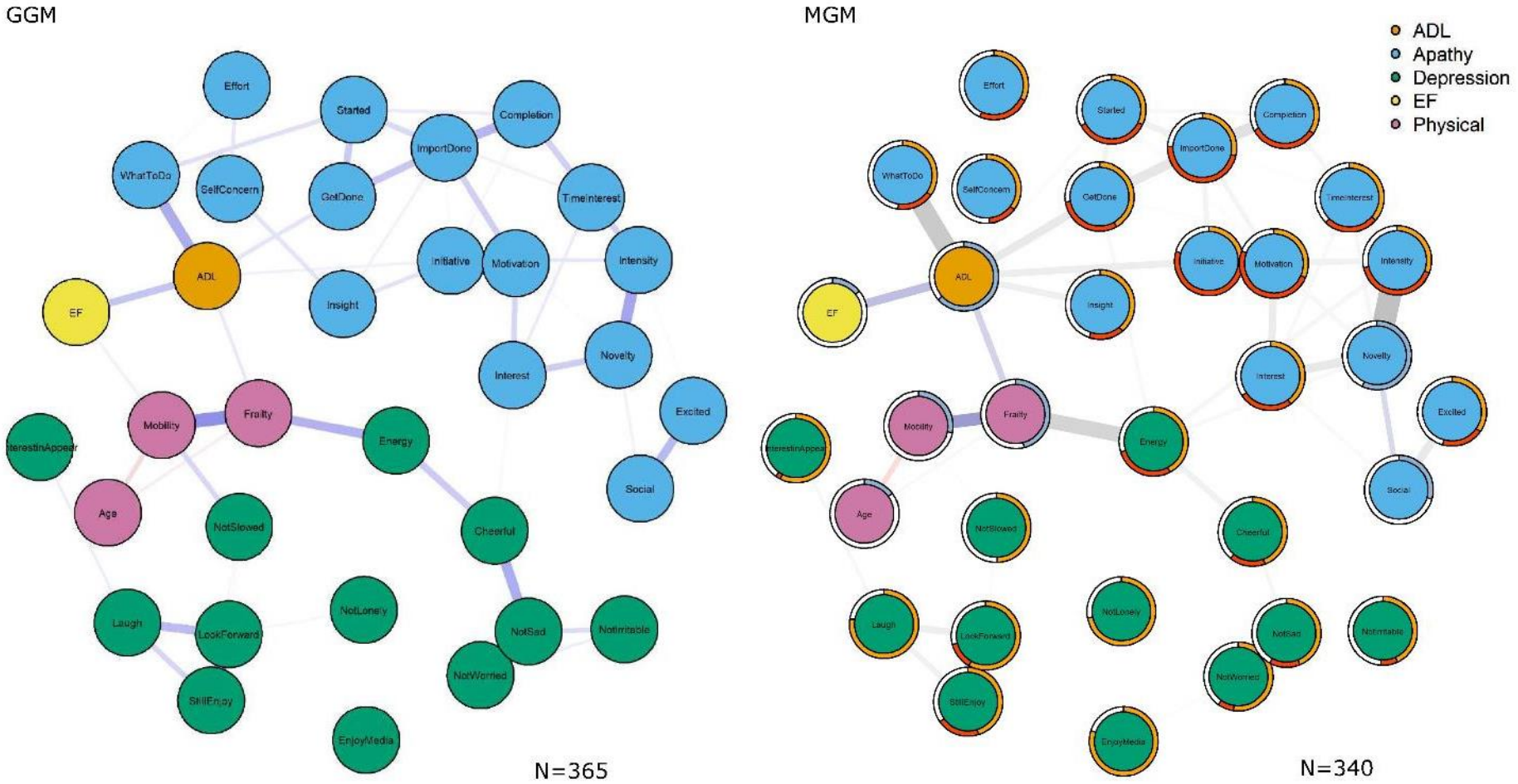


Figure shows GGM and MGM after phase 3. Maximum set at .61 (which was the maximum of all graphs), minimum set at .08, and MGM layout constrained to GGM layout, to aid visual interpretation and comparison.

Appendix 17 Predictability

Phase 1 Predictability

Node	R2	nCC	CC	CCmarg	Beyond marginal (CC-CCmarg)
Interest		0.446	0.666	0.397	0.269
GetDone		0.451	0.68	0.417	0.263
Started		0.448	0.626	0.322	0.304
Novelty	0.559				
Effort		0.389	0.586	0.322	0.264
Intensity		0.561	0.697	0.31	0.387
Completion		0.496	0.669	0.343	0.326
TimeInterest		0.393	0.615	0.366	0.249
WhatToDo		0.228	0.493	0.343	0.15
SelfConcern		0.153	0.45	0.351	0.099
Social	0.374				
Excited		0.288	0.53	0.34	0.19
Insight		0.281	0.558	0.385	0.173
ImportDone		0.618	0.725	0.28	0.445
Initiative		0.679	0.771	0.287	0.484
Motivation		0.738	0.822	0.321	0.501

Phase 2 Predictability

Node	R2	nCC	CC	CCmarg	Beyond marginal (CC-CCmarg)
Interest		0.416	0.648	0.397	0.251
GetDone		0.51	0.718	0.424	0.294
Started		0.477	0.646	0.323	0.323
Novelty	0.579				
Effort		0.369	0.571	0.32	0.251
Intensity		0.558	0.695	0.31	0.385
Completion		0.524	0.689	0.347	0.342
TimeInterest		0.427	0.637	0.366	0.271
WhatToDo		0.141	0.438	0.346	0.092
SelfConcern		0.196	0.478	0.351	0.127
Social	0.291				
Excited		0.295	0.539	0.346	0.193
Insight		0.265	0.553	0.392	0.161
ImportDone		0.632	0.735	0.28	0.455
Initiative		0.694	0.781	0.284	0.497
Motivation		0.709	0.801	0.316	0.485
NotSlowed		0	0.501	0.501	0
InterestinAppear		0.08	0.602	0.567	0.035
StillEnjoy		0.366	0.651	0.45	0.201
Laugh		0.025	0.778	0.772	0.006
LookForward		0.32	0.712	0.576	0.136
EnjoyMedia		0	0.795	0.795	0
Cheerful		0.314	0.617	0.442	0.175
NotWorried		0.183	0.614	0.528	0.086
Energy		0.43	0.671	0.423	0.248
NotSad		0.267	0.588	0.438	0.15
NotIrritable		0.133	0.51	0.435	0.075
NotLonely		0	0.72	0.72	0

Phase 3 Predictability

Node	R2	nCC	CC	CCmarg	Beyond marginal (CC-CCmarg)
Interest		0.422	0.653	0.4	0.253
GetDone		0.525	0.724	0.419	0.305
Started		0.519	0.674	0.322	0.352
Novelty	0.571				
Effort		0.361	0.568	0.324	0.244
Intensity		0.587	0.715	0.31	0.405
Completion		0.498	0.671	0.345	0.326
TimeInterest		0.403	0.621	0.365	0.256
WhatToDo		0.27	0.524	0.348	0.176
SelfConcern		0.209	0.488	0.353	0.135
Social	0.295				
Excited		0.302	0.544	0.347	0.197
Insight		0.269	0.553	0.389	0.164
ImportDone		0.658	0.756	0.287	0.469
Initiative		0.713	0.794	0.282	0.512
Motivation		0.72	0.809	0.318	0.491
NotSlowed		0	0.497	0.497	0
InterestinAppear		0.056	0.6	0.576	0.024
StillEnjoy		0.369	0.653	0.45	0.203
Laugh		0.013	0.776	0.773	0.003
LookForward		0.303	0.709	0.582	0.127
EnjoyMedia		0	0.8	0.8	0
Cheerful		0.305	0.612	0.442	0.17
NotWorried		0.16	0.6	0.524	0.076
Energy		0.482	0.7	0.421	0.279
NotSad		0.257	0.582	0.437	0.145
NotIrritable		0.144	0.512	0.43	0.082
NotLonely		0	0.718	0.718	0
ADL	0.635				
EF	0.141				
Mobility	0.282				
Frailty	0.46				
Age	0.151				