Cognitive Assessment in Multiple Sclerosis: Towards the Development of a Cognitive Screening Pathway

Hannah Elwick, BSc (Hons), MSc

Thesis submitted to the University of Nottingham
for the degree of Doctor of Philosophy

September 2021

Abstract

Background

It is estimated that as many as 86% of people with Multiple Sclerosis (MS) can experience cognitive impairment, which can impact quality of life. Although the value of screening for cognitive problems is recognised, it is not routinely undertaken.

Methodology

The research comprises five studies to identify acceptable and appropriate methods of assessing cognitive impairment in people with MS. A systematic review (**Study I)** was conducted to determine the most frequently used cognitive assessments in MS research. **Study II** focused on one test, the Symbol Digit Modalities Test (SDMT), and explored the variability in scoring and interpretation in MS research. In **Study III** the domain structure of an MS-specific self-report measure of cognitive problems, and how it relates to an objective measure of information processing speed, was explored. Finally, two qualitative interview studies were conducted. **Study VI** examined stakeholders' experience and opinions of cognitive impairment and assessments, and explored perceived barriers and facilitators to implementation in clinical practice. In **Study V,** people with MS completed a digital test and were asked to 'talk through' the process.

Results

The systematic review (**Study I**) extracted data from 1,526 studies, using 5,665 cognitive assessments. The SDMT emerged as the most frequently used individual test. Focusing on this, there was variability in how results of this test are interpreted for people with MS noted in **Study II**. Whilst the SDMT is a promising objective measure of information processing speed, the results of **Study III** suggest that a reduced, multidimensional questionnaire has more promise than its unidimensional version and could be used in place of an objective test.

In **Study IV**, people with MS and clinicians reported that, despite being a prevalent symptom, cognitive impairment was not assessed routinely. Participants felt that symbol substitution tests and subjective self-reported questionnaires were acceptable and appropriate in assessment and could be completed digitally on personal mobile devices. This would also be feasible to implement in NHS clinical care. Based on these findings, a webapp was developed, and **Study V** identified usability problems, and barriers and facilitators to its use.

Conclusions

Future research is needed to assess the psychometric properties of the developed digital task. Normative data from non-MS control participants should also be collected to provide benchmarks for interpreting clinical data. Digitised, remote assessments can enable more people with MS to access screening, which can facilitate discussion and referral for cognitive rehabilitation. Overall, the results of this research indicate the potential of using digital, remote cognitive assessment with a symbol substitution task in MS clinical care.

List of publications and presentations arising from this thesis

Published Manuscripts

Elwick, H., Topcu, G., Allen, C. M., Drummond, A., Evangelou, N., & das Nair, R. das. (2021). Cognitive measures used in adults with multiple sclerosis: A systematic review. *Neuropsychological Rehabilitation*, 0(0), 1–18. https://doi.org/10.1080/09602011.2021.1936080

Elwick, H., Smith, L., Mhizha-Murira, J. R., Topcu, G., Leighton, P., Drummond, A., Evangelou, N., & Nair, R. (2021). Cognitive assessment in multiple sclerosis clinical care: A qualitative evaluation of stakeholder perceptions and preferences. *Neuropsychological Rehabilitation*, *0*(0), 1–19. https://doi.org/10.1080/09602011.2021.1899942

Smith, L., **Elwick, H**., Mhizha-Murira, J. R., Topcu, G., Bale, C., Evangelou, N., Timmons, S., Leighton, P., & das Nair, R. (2020). Developing a clinical pathway to identify and manage cognitive problems in Multiple Sclerosis: Qualitative findings from patients, family members, charity volunteers, clinicians and healthcare commissioners. *Multiple Sclerosis and Related Disorders*, *49*. https://doi.org/10.1016/j.msard.2020.102563

Manuscripts in Preparation

Elwick, H., Smith, L. Topcu, G. Mhiza-Murira, J. R. & das Nair, R. (in preparation). Interpretation of Symbol Digit Modality Test Results in Multiple Sclerosis. [Chapter 3]

Accepted Oral Presentations

Elwick, H., Smith, L., Mhiza-Murira, J. R., Topcu, G., Leighton, P., Drummond, A., Evangelou, N. & das Nair, R. (2020). Cognitive Assessment in Multiple Sclerosis: A Qualitative Study of Patient, Carer and Clinician Experience and Perceptions. Rehabilitation in Multiple Sclerosis (RIMS) conference, Leuven, Belgium.

Poster Presentations

Elwick, H., Topcu, G., Allen, C. M., Drummond, A., Evangelou, N. & das Nair, R. (2019). Cognitive Assessments Used in Multiple Sclerosis Research – A Systematic Review. Poster presented at MS Frontiers, Bath, UK. [Conference Poster Prize winner]

Elwick, H., Topcu, G., Allen, C. M., Drummond, A., Evangelou, N. & das Nair, R. (2019). The Use of Patient and Proxy Reported Outcome Measures to Assess Cognitive Impairment in Multiple Sclerosis. Poster presented at Patient Reported Outcome Measures Annual Research Conference, Leeds, UK.

Elwick, H., Topcu, G., Allen, C. M., Drummond, A., Evangelou, N. & das Nair, R. (2019). Cognitive Assessments Used in Multiple Sclerosis Research – A Systematic Review. Poster presented at Medicine and Health Sciences Faculty Postgraduate Research Forum, Nottingham, UK.

Statement of Contribution

This thesis was completed between April 2018 and September 2021. The project was funded by the MS Society as an independent PhD studentship, however, it was linked to two other trials and involved collaboration with other researchers. The research was co-designed and supervised by Professor Roshan das Nair, Professor Avril Drummond, Dr Nikos Evangelou and Dr Gogem Topcu.

I devised the design of the systematic review, ran the searches and conducted the initial screening and analysis of the systematic review in Chapter 2. Dr Chris Allen helped with the full-text data extraction from some of the included studies. I designed the scoping review in Chapter 3 and this was written as a paper with input from Dr Laura Smith and Dr Jacqueline Mhizha-Murira.

Cognitive Rehabilitation for Attention and Memory in People with Multiple Sclerosis (CRAMMS) was a randomised controlled trial funded by the National Institute for Health Research (NIHR; project number 12/190/05) that began in 2015. As the authors of this study collected screening and baseline data that was pertinent to the current research, I conducted a secondary analysis of this data as in Chapter 4. I designed and conducted the analysis and the write-up of this study. Professor Graham Law helped with the regression analyses.

Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS) is a six-year NIHR funded programme (project reference RP-PG-0218-20002) that began in May 2019. One of the study's aims is to develop a digitised cognitive assessment platform. The first work package of this programme collected qualitative data on experiences and opinions of cognitive assessment and rehabilitation. Four researchers collected qualitative data; I recruited and conducted interviews for all people with MS; family members; and charity volunteers and facilitated a focus group with people with MS. A research fellow, Dr Jacqueline Mhizha-Murira, recruited and conducted

interviews with clinicians, and two of the research leads for NEuRoMS, Professor Roshan das Nair and Dr Nikos Evangelou conducted interviews with National Health Service commissioners. All researchers used the same semi-structured interview schedules and constructed the framework for analysis and coded the data with input from Professor Paul Leighton. My focus was coding and analysing data related to cognitive assessment, and writing this for Chapter 5 and as a journal article (Elwick, et al., 2021).

Based on the findings of these studies, a cognitive assessment platform was developed by MindTech and the Digital Research Service of the University of Nottingham. Ethical approval was sought for usability testing of the technology in December 2019, and Dr Kyle Harrington conducted the Patient and Public Involvement pilot interviews at this time. Due to COVID-19, this project was delayed, and the design was amended due to restrictions on in-person data collection. I recruited participants and conducted the interviews in December 2020 and these interviews are described in Chapter 6. The technology development was advised by the results of these interviews, and I transcribed and analysed the data formally for the inclusion in this thesis.

The protocol of research for described in Chapter 7 was devised as part of this thesis and ethical approval was granted for this data collection in April 2020. It was anticipated this would begin in May 2020, however due to delays in the technology development, it was not feasible to complete this aspect of the PhD. This project will be conducted as part of NEuRoMS.

Acknowledgements

I would like to thank my supervisors, Professor Roshan das Nair, Professor Avril Drummond, Dr Nikos Evangelou and Dr Gogem Topcu for their supervision, guidance, knowledge, and encouragement. This research would also not have been possible without my colleagues, in particular Dr Jacqueline Mhizha-Murira and Dr Laura Smith. Thanks to my fellow PGR students at the Institute of Mental Health for the coffee breaks, lunches, and commiserations.

This PhD was funded by the MS Society, and I am grateful for their help and support. I would also like to thank the people with MS, and their family members involved in the research who kindly gave up their time to take part. The MS Patient and Public Involvement group at the University of Nottingham were always welcoming and encouraging.

All my love and gratitude to my family: Mum and Roger, Dad and Pam, and brothers Alex and Harry. Thank you for the support and encouragement. This PhD also immensely benefited from the ability to unwind with chats and beer with all my wonderful friends, especially Penny, Kelly, and Cat.

Without Adam willing to constantly provide encouragement in every way, this PhD would never have been possible. He was forever willing to provide emotional and practical support, listen to my grumbles, facilitate my coffee and vinyl habits, and unconditionally love me. I don’t know what I would do without you.

I will always remember writing my discussion chapter whilst Humphrey snoozed on my desk in the sunshine, the snuggles were the best. Moose and Scout kept me laughing, and ate a lot of books.

Table of Contents

Abstract ..... **i**

List of publications and presentations arising from this thesis ..... **iii**

Statement of Contribution ..... **i**

Acknowledgements ..... **iii**

Table of Contents ..... **iv**

Table of Figures ..... **i**

Table of Tables ..... **i**

Abbreviations ..... **i**

List of Appendices ..... **i**

Chapter 1 Introduction **2**

1.1 Chapter overview 2

1.2 Multiple sclerosis 2

1.3 Cognitive impairment in MS 4

1.3.1 Complex attention 6

1.3.2 Learning and memory 8

1.3.3 Executive function 10

1.3.4 Perceptual-motor function 11

1.3.5 Language 11

1.3.6 Social cognition 13

1.3.7 Neuropsychological profiles of cognitive deficits in MS 13

1.3.8 Factors related to cognitive impairment in MS 18

1.4 Measurement of cognitive impairment 21

1.4.1 Current cognitive assessment in MS 22

1.4.2 Validity and reliability 32

1.4.3 Digital neuropsychological assessment 34

1.5 Gaps and limitations of previous published work 38
1.6 Summary 41
1.7 Objectives 41
1.8 Research plan 42

Chapter 2 Cognitive Measures Used in Adults with Multiple Sclerosis: A Systematic Review 46

2.1 Chapter overview 46
2.2 Introduction 46
2.3 Methods 49
2.3.1 Search strategy and selection criteria 49
2.3.2 Data extraction and analysis 50
2.4 Results 51
2.4.1 Individual tests of cognition 53
2.5 Multidomain Tests and Batteries 55
2.5.1 Subjective measures of cognition 56
2.5.2 Uses of cognitive measures 57
2.5.3 Administration of cognitive measures 58
2.6 Discussion 58
2.7 Conclusions 63

Chapter 3 Variability in Administration, Scoring and Interpreting Symbol Digit Modalities Test Results in Multiple Sclerosis 65

3.1 Chapter overview 65
3.2 Introduction 65
3.3 Aims 66
3.4 Methods 66
3.5 Results 67
3.5.1 Administration and scoring 67

3.6 Discussion 69

3.7 Conclusions 71

Chapter 4 Exploring the factor structure of a self-reported cognitive problems questionnaire, and variables associated with processing speed deficits in people with MS 73

4.1 Chapter overview 73

4.2 Introduction 73

4.2.1 Objective and subjective measures of cognitive impairment 74

4.2.2 Processing speed and self-reported symptoms of multiple sclerosis 77

4.2.3 Cognitive Rehabilitation for Attention and Memory in People with Multiple Sclerosis 78

4.3 Aims 79

4.4 Methods 80

4.4.1 Data 80

4.4.2 Statistical analyses 81

4.5 Results 82

4.5.1 Factor structure of the MSNQ-P 82

4.5.2 Predictors of SDMT scores 87

4.5.3 Self-reported symptoms 88

4.5.4 Multivariate regression analysis 89

4.6 Discussion 90

4.7 Conclusions 95

Chapter 5 Cognitive Assessment in Multiple Sclerosis Clinical Care: A Qualitative Evaluation of Stakeholder Perceptions and Preferences 97

5.1 Overview of the Chapter 97

5.2 Introduction 97

5.3 Aim 99

5.4 Methods 100

5.4.1 Ethical Considerations 100

5.4.2 Design 101

5.4.3 Data collection 101

5.4.4 Participants 101

5.4.5 Materials 102

5.4.6 Procedure 103

5.4.7 Transcription and data analysis 104

5.5 Results 106

5.5.1 Sample characteristics 106

5.5.2 Context 108

5.5.3 Mechanisms 112

5.5.4 Outcomes 124

5.6 Discussion 127

5.7 Conclusions 131

Chapter 6 Usability Testing of a Cognitive Assessment Platform for Use with People with Multiple Sclerosis 133

6.1 Chapter overview 133

6.2 Introduction 133

6.2.1 Digital cognitive assessments for MS 134

6.2.2 Digitisation of cognitive assessments 135

6.2.3 Development of a cognitive assessment platform 138

6.3 Aims 139

6.4 Methods 139

6.4.1 Ethical considerations 139

6.4.2 Design 140

6.4.3 Participants 140

6.4.4 Materials 140

6.4.5 Procedure 143

6.4.6 Patient and public involvement 144

6.4.7 Cognitive interviewing procedure 145
6.4.8 Transcription and data analysis 145
6.5 Results 145
6.5.1 Pilot interviews 145
6.5.2 Cognitive interviews 146
6.5.3 Personal context 147
6.5.4 Quality of the digital assessment platform 148
6.5.5 Barriers and enablers to use 151
6.6 Discussion 154
6.7 Conclusions 159
Chapter 7 Protocol for the Validation of a Digital Cognitive Assessment for Use in MS 160
7.1 Chapter overview 160
7.2 Introduction 160
7.2.1 Inclusion of the Stroop task 162
7.2.2 Psychometric validation 163
7.3 Aims 164
7.4 Methods 164
7.4.1 Ethical considerations 164
7.4.2 Participants 165
7.4.3 Materials 165
7.4.4 Study I 167
7.4.5 Study II 169
7.4.6 Study III 170
7.5 Patient and public involvement 170
7.6 Recruitment 171
7.7 Analyses 171
7.7.1 Convergent validity 171
7.7.2 Cross modality differences 172
7.7.3 Reliability 173

7.7.4 Normative data analysis 173
7.8 Discussion 174
Chapter 8 Discussion 176
8.1 Summary of the research 176
8.2 General discussion 178
8.2.1 How the research relates and adds to the literature 180
8.3 Limitations of the research 182
8.4 Recommendations for future direction 184
8.5 Conclusions 189
References 191
Tables 245

Table of Figures

Figure 1: Six Key Neurocognitive Domains 6
Figure 2: Interrelation of cognitive domains in MS 16
Figure 3: Prisma flowchart for systematic review 53
Figure 4: Scree plot of components extracted from the MSNQ-P 84
Figure 5: Prototype of a symbol substitution task deployed on a PC. 141
Figure 6: DECSYS mobile platform showing the Symbol Substitution Task 143
Figure 7: Symbol Substitution Task instructions – Screen 1 309
Figure 8: Symbol Substitution Task instructions – Screen 2 310

Table of Tables

Table 1: Individual tests used in MS research 54
Table 2: Multidomain tests used in MS research 56
Table 3: Subjective measures used in MS research..... 57
Table 4: Frequency of uses of cognitive measures 58
Table 5: Demographics of people with MS who completed the MSNQ-P and SDMT at screening..... 83
Table 6: Initial eigenvalues and proportion of variance explained by the five-component solution..... 84
Table 7: Rotated component matrix of five extracted components of the MSNQ-P 85
Table 8: Conceptual components of the MSNQ-P..... 86
Table 9: Sample characteristics of people with MS who completed the FSS and MSIS-29 at baseline 88
Table 10: Self-reported questionnaire descriptive statistics..... 89
Table11: Correlation matrix of SDMT scores and self-reported variables 89
Table 12: Demographics of people with MS, family members and charity volunteers 106
Table 13: Sample characteristics information of people with MS interviewed 147
Table 14: Change log of amendments made due to minor usability issues……153

Abbreviations

ACE-R – Addenbrooke's Cognitive Examination

ADL – Activities of Daily Living

ATS - Automated Transcription Service

BICAMS – Brief International Cognitive Assessment for MS

BRB-N – Brief Repeatable Battery of Neuropsychological Tests

BVMT-R – Brief Visuospatial Memory Test-Revised

CAI – Cognitive Assessment Interview

CFA – Confirmatory Factor Analysis

CIS – Clinically Isolated Syndrome

COWAT – Controlled Oral Word Association Test

CRAMMS - Cognitive Rehabilitation for Attention and Memory in People with Multiple Sclerosis

CVLT – California Verbal Learning Test

DECSYS - Discrete and Ellipse-based Response Capture System

DKEFS – Delis-Kaplan Executive Function System

DSM - Diagnostic and Statistical Manual

EDSS – Expanded Disability Status Scale

EFA – Exploratory Factor Analysis

FSS - Fatigue Severity Scale

HCI – Human Computer Interaction

MACFIMS – Minimal Assessment of Cognitive Function in Multiple Sclerosis

MoCA – Montreal Cognitive Assessment

MRI - Magnetic Resonance Imaging

MS – Multiple Sclerosis

MSFC – Multiple Sclerosis Functional Composite

MSIS-29 - total Multiple Sclerosis Impact Scale

MSNQ – Multiple Sclerosis Neuropsychological Questionnaire

MSSE – Mini Mental State Examination

NEuRoMS - Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis

NHS – National Health Service

NICE – National Institute for Health and Care Excellence
NIHR – National Institute for Health Research
PASAT – Paced Auditory Serial Addition Test
PCA – Principal Components Analysis
PDQ – Perceived Deficits Questionnaire
PPI – Patient and Public Involvement
PPMS – Primary-Progressive MS
RCT – Randomised Controlled Trial
ROC – Receiver Operating Characteristic
RRMS – Relapsing-Remitting Multiple Sclerosis
SDMT – Symbol Digit Modalities Test
SPART - Spatial Recall Test
SPMS – Secondary-Progressive Multiple Sclerosis
TAM - Technology Acceptance Model
VFT – Verbal Fluency Test
WAIS – Wechsler Adult Intelligence Scale
WLG – Word List Generation
WLL – Word List Learning
WMS – Wechsler Adult Intelligence Scale

List of Appendices

1. Cognitive Screening Pathway and Rehabilitation Logic Model…...265
2. Qualitative Interviews – Adverts for Participants…….266
3. Qualitative Interviews – Participant Information Sheets…….268
4. Qualitative Interviews – Consent Forms……..283
5. Qualitative Interviews – Semi-Structured Interview Schedules…...286
6. Qualitative Interviews – Transcription Confidentiality Agreement ..289
7. Qualitative Interviews – Participant Facing Materials…….290
8. Qualitative Interviews – Ethical Approval……292
9. Usability Cognitive Interviews – Advert for Participants…….294
10. Usability Cognitive Interviews – Participant Information Sheet….295
11. Usability Cognitive Interviews – Verbal Consent Form……...299
12. Usability Cognitive Interviews – Interview Schedule ……...…..300
13.Usability Cognitive Interviews – Ethical Approval…….302
14.Usability Cognitive Interviews –Symbol Substitution Task ….……303
15. Validation Study – Ethical Approval……...305

Chapter 1 Introduction

1.1 Chapter overview

This chapter outlines the context in which the current research took place. It briefly discusses what multiple sclerosis (MS) is, how it is diagnosed, its progression, and its management from a UK healthcare perspective. I then explore the neuropsychological aspects of MS through the neurocognitive framework from the fifth edition of the Diagnostic and Statistical Manual (DSM-5, American Psychiatric Association, 2013). Using this framework of cognition, I describe some of the most frequently used cognitive assessments, the psychometric properties of cognitive assessments, and how digitisation can enable routine assessment within MS care. I then summarise the rationale and approaches used for the research presented in this thesis.

1.2 Multiple sclerosis

MS is a chronic neurological condition in which lesions are found in the brain and spinal cord, presenting with damage to the myelin and axons within the central nervous system (Compston & Coles, 2008). The extent of damage varies between individuals, and it has a variable course. In 2020, worldwide prevalence of MS was estimated at 2.8 million (Walton et al., 2020). Within the UK, MS affects over 130,000 people and it is the most common cause of serious disability in adults of working age (Mackenzie et al., 2014; MS Society, 2020).

The International Advisory Committee on Clinical Trials in MS has developed consensus definitions of four phenotypes that characterise the clinical course of MS (Lublin et al., 2014), as follows:

1. Clinically Isolated Syndrome (CIS): presents with symptoms that suggest demyelination, such as optic neuritis, but does not meet the criteria for a diagnosis of MS at that time, however 30-70% of CIS cases progress to a diagnosis of MS (Allen et al., 2020; Miller et al., 2005).

2. Relapsing-Remitting MS (RRMS): is the most common type of MS, with 85% of people with MS being initially diagnosed with this course (The Multiple Sclerosis International Federation, 2020). It is characterised by exacerbations of symptoms with periods of recovery between relapses.

3. Secondary-Progressive MS (SPMS): is characterised as a conversion from RRMS to a progressive course. SPMS can develop 10-15 years after the onset of RRMS, and 65% of RRMS cases convert to SPMS (Lublin et al., 2014).

4. Primary-Progressive MS (PPMS): is characterised by a gradual worsening clinical course. 10-15% of people with MS are diagnosed with PPMS at onset (The Multiple Sclerosis International Federation, 2020).

MS is diagnosed using the revised McDonald criteria (Thompson et al., 2018). This is an assessment of the neurological episodes and the clinical presentation of lesions consistent with MS. The Expanded Disability Status Scale (EDSS; Kurtzke, 1983) is the predominant clinical measurement of disability and symptom progression in MS, and describes eight functional systems: visual, brainstem, pyramidal, cerebellar, sensory, bowel and bladder, cerebral, and ambulation. Symptoms of MS can be both physical and psychological, and both vary across individuals and the progression of MS. Common symptoms include pain, fatigue, depression, changes in sensation, cognitive impairment, stiffness, muscle spasm, vestibular disturbance, bladder and bowel problems, impaired speech and swallowing, and impaired vision (Compston & Coles, 2008).

There is no cure, but there are several treatments available. Disease-modifying treatments can reduce relapses or progressions of MS, and symptomatic treatments of medical and rehabilitative therapies can alleviate symptoms (Gholamzad et al., 2019). The use of both drug treatment and neurorehabilitation has shown promise in improving independence and activities of daily living, leading to an improvement in

quality of life (Kesselring & Beer, 2005). In the UK, the National Institute for Health and Care Excellence (NICE; 2019) guidelines recommend a multidisciplinary approach to MS care that can facilitate this combined approach to treatment. The involvement of not only neurologists and MS nurses, but also other specialists including occupational therapists, speech and language therapists, psychologists, and physiotherapists can enable an inclusive approach to the management of symptoms.

Despite advances in treatment, MS is still a disabling disease. Using the EQ-5D-5L (Herdman et al., 2011) to measure health state, a recent European study reported that people with MS had a low health utility score, and disability and progressive courses of MS were related to poorer quality of life (Visser et al., 2021).

1.3 Cognitive impairment in MS

Cognitive impairment is a common symptom of MS, with a recent review estimating 34% to 65% of people with MS show deficits in one or more domains of cognition (Benedict et al., 2020). Additionally, in a large scale European survey (*n*=13,186), 71% of people with MS reported experience of cognitive difficulties as a symptom of MS (Kobelt et al., 2006). Basso et al. (1996) reported that prevalence of impairment is higher than previously thought, likely due to the growth in research into cognition, along with greater availability of comprehensive and sensitive assessment methods. Prevalence estimates can also vary given the difference in assessments used, along with how impairment is classified (Fischer et al., 2014). Cognition is often described in terms of the domains that are compromised.

A systematic review that aimed to profile cognitive dysfunction in MS reported that the most frequently impaired domains are memory, executive function, and attention (Ferreira, 2010). This is echoed in other literature, with impairment mainly documented within the domains of memory, information processing speed and executive function (e.g., Chiaravalloti & DeLuca, 2008; Sumowski et al., 2018). Focusing on

these domains can influence the assessments chosen to describe cognition in MS (this is explored further in section 1.4).

Previous categorisations of impairment have used a range of methodologies with no clear consensus of how to operationalise ‘cognitive impairment’ (as explored in Chapter 3), and used a bottom-up approach of describing impairment using results on previously developed batteries. However, taking approaches from other conditions that present with cognitive impairment (for example, epilepsy as described by Norman et al., 2021), some recent research (Hancock et al., 2022) has used the fifth edition of the DSM to provide a framework for classifying cognitive impairment. The DSM-5 describes mild and major neurocognitive disorders (American Psychiatric Association, 2013) and attempts to standardise how cognitive function can be described through six main domains, and multiple subdomains, of cognitive functioning (as shown in Figure 1).This includes domains that are not frequently assessed in batteries developed for MS, such as social cognition, and therefore has potential for presenting a ‘top-down’ approach to cognitive impairment in MS.

The mild and major cognitive impairment described through this framework represent a decline of cognitive functioning in a clinical condition (Sachdev et al., 2014) and therefore it is appropriate to use these descriptors for conditions such as MS. Previous research has found that using this framework categorises a greater number of people with MS as impaired (Hancock et al., 2022) than previous methods used in MS, however this is unsurprising given the greater range of domains that are now included. This is the framework I have used to describe and explore the profile of cognitive functioning in MS, and how it can be assessed. Where appropriate, domains have been examined using their subdomains. These will now be presented and discussed.

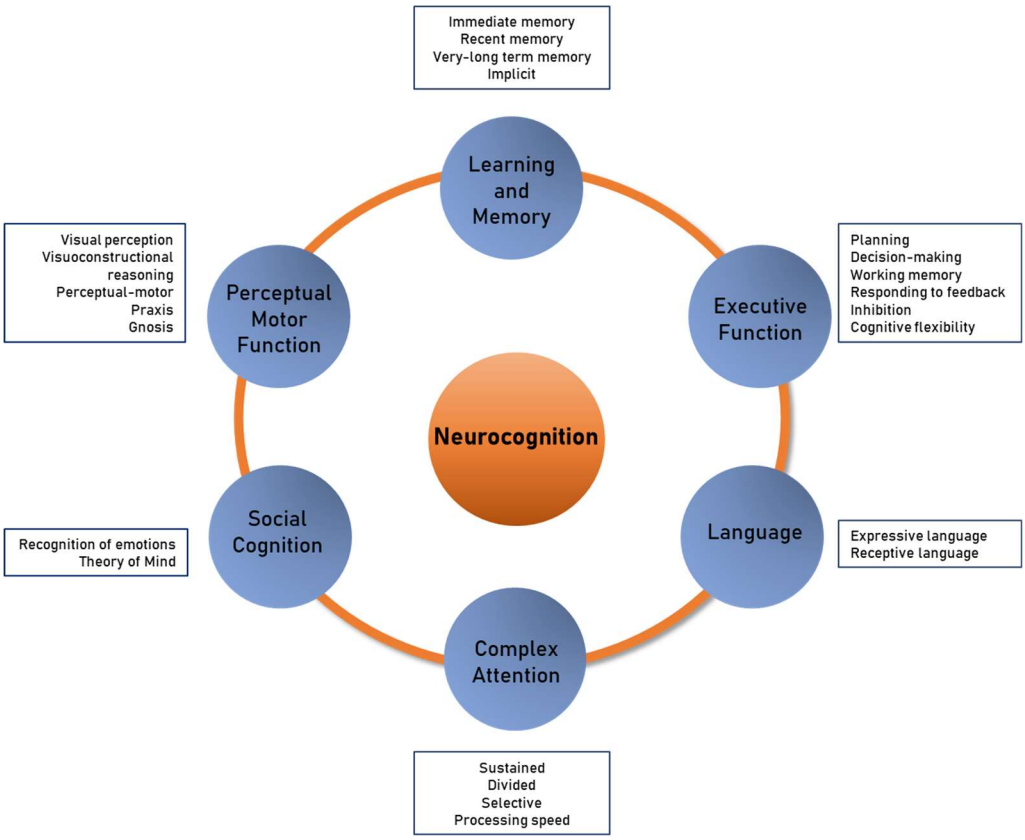


Figure 1: Six Key Neurocognitive Domains (adapted from the DSM-5, American Psychiatric Association, 2013)

1.3.1 **Complex attention**

The DSM-5 defines complex attention through four principal subdomains: sustained attention, divided attention, selective attention, and processing speed. Whilst all of these are thought to be impaired (Oreja-Guevara et al., 2019), information processing speed is frequently cited as the subdomain most frequently impaired in MS, with estimates of 50-70% of people with MS experiencing slowed processing speed (Chiaravalloti & Deluca, 2009; Hansen & Lautenbacher, 2017). It is thought to be the subdomain that may prompt impairment in other domains (Deluca, 2008).

1.3.1.1 *Sustained, selective and divided attention*

Sustained attention is the preservation of attention over time, selective attention is the preservation of attention through competing stimuli, and divided attention refers to attention to two stimuli at the same time (American Psychiatric Association, 2013). Estimates suggest

20-50% of people with MS have difficulties with attention, however this could be because of dysfunction in other subdomains, in particular working memory, and information processing speed (Oreja-Guevara et al., 2019). Attention may also be moderated by fatigue, with longer tasks of sustained attention provoking greater fatigue, and thus poor performance on the tasks (Chiaravalloti & Deluca, 2009).

1.3.1.2 *Processing speed*

A rudimentary description of processing speed is the time taken to complete a cognitive task (Deluca, 2008). Information processing is perceived to be one of the most commonly impaired subdomains of cognition in MS (Costa et al., 2017; DeLuca et al., 2004), and as such is often the focus of cognitive assessment in MS research. A recent largescale study in Germany, of 1094 people with MS, found that information processing speed, as measured by the Symbol Digit Modalities Test (SDMT), was the most impaired of three tests within a frequently recommended battery, the Brief International Cognitive Assessment for MS (BICAMS; Renner et al., 2020). Furthermore, studies examining magnetic resonance imaging (MRI) have found associations between SDMT scores and cortical lesions (Benedict et al., 2017; Rao et al., 2014).

A review of deficits in information processing speed attempted to synthesise how processing speed has been defined and conceptualised in studies (Costa et al., 2017). This review underlined that other domain of cognition, including working memory and executive functions, and poorer outcomes in quality of life, activities of daily living and employment are correlated with information processing. Furthermore, in this review, Costa et al. (2017) found that the disparities in definitions reported by studies may suggest a tri-factor model of information processing speed deficits in people with MS. These relate to the speed at which people with MS can execute a task based on deficits in three distinct factors: sensorial, cognitive, or motor.

1.3.2 **Learning and memory**

Learning and memory is considered another frequently impaired domain, with prevalence estimates of 40-60% of people with MS affected (Calabrese, 2006). The DSM-5 describes four subdomains of learning and memory: immediate, recent, very long-term memory, and implicit memory. The purpose of this model is to describe functioning across memory and learning, with the recognition that in most cases of neurocognitive impairment, dysfunction is more likely to occur across immediate, recent memory, however dysfunction in very long-term memory may be variable across conditions (American Psychiatric Association, 2013).

Memory is typically assessed using recall (either free or cued) or recognition. Impairment in recalling information is more frequent than recognition (Calabrese, 2006), and it has been suggested that encoding is largely unaffected, but people with MS have difficulties in retrieving information (Zakzanis, 2000). However, this is an area of controversary, with some research finding greater impairment in encoding than retrieval (DeLuca et al., 1998). An early meta-analysis of 36 studies investigating the difference between people with MS and healthy controls on measures of memory concluded that people with MS were significantly more likely to be impaired across all memory domains than healthy control participants, but long-term memory showed greater deficits than immediate or recent memory (Thornton & Raz, 1997).

1.3.2.1 *Immediate and recent memory*

In the DSM-5, immediate memory (also referred to as short term memory in other literature) refers to an individual's ability to hold a small amount of information for a very short period of time, often only seconds. This domain is frequently also described as working memory which falls under the DSM-5's domain of executive function and is described in section 1.3.3. Recent memory is associated with immediate memory but encapsulates the encoding of information. This may be assessed through asking individuals to recall as many words as possible they can remember from a list they heard or saw very recently,

either with free recall or cued recall (for example, asked to recall all words that belong to a certain category). Some research has found that people with MS show little discrepancy between immediate and recent recall of information, and suggests that with more time to encode information, people with MS can improve their performance on tests of recall (Oreja-Guevara et al., 2019).

1.3.2.2 *Very long-term memory*

Very long-term memory is the ability to store information for retrieval at a longer time than in recent memory, for example after hours, or years after encoding the information. This subdomain of memory is concerned with explicit memory and the conscious recall of information, which could be autobiographical (such personal experiences), or semantic (such as facts and skills). Long-term memory can also be conceptualised temporally into two subdomains: retrospective or prospective (Graf & Uttl, 2001).

Retrospective memory can be assessed using some of the methods detailed above, whereby people with MS encode and recall information, such as a list of words, after a longer delay or recall aspects of their semantic or autobiographical memory. Autobiographical memory is not regularly assessed in MS, however there is some evidence that it is impaired (Ernst et al., 2013; Ernst, 2020). Ernst (2020) also demonstrated that individuals with RRMS were aware of difficulties with their autobiographical memory, and the adverse impact it had on their everyday functioning. Similarly, prospective memory, related to difficulties in executing a planned activity, such as remembering to attend an appointment, is also largely precluded from neuropsychological assessment in MS. Research has nevertheless documented deficits in this aspect of very long-term memory, along with people's awareness of prospective memory dysfunction (Rouleau et al., 2018).

Another component of very long-term memory is that of implicit memory, the unconscious recall of information in the long-term. For example, remembering how to ride a bike. This has been assessed in

people with MS using word stem completion, where participants are given the first few letters of a word and asked to complete the word (Scarrabelotti & Carroll, 1999). Implicit memory in MS has been found largely intact (Calabrese, 2006).

1.3.3 **Executive function**

Executive functioning concerns the high-order functioning that facilitates other cognitive functions and behaviour (Diamond, 2013). In the DSM-5, executive function is defined through six subdomains: planning, decision-making, working memory, response to feedback, inhibition, and mental flexibility. Inhibition, mental flexibility and working memory in particular may be compromised (Drew et al., 2008), and some research has found planning, decision-making and responding to feedback can be impaired (Calabrese, 2006).

Due to these subdomains being considered higher-order cognition, it is likely that there are strong associations between other subdomains. Drew et al. (2008) recommended a broad assessment of this domain, given the likelihood of executive dysfunction affecting real-life situations, such as how impaired planning and decision-making may impact vocational activities for people with MS. Working memory and its relationship to information processing speed, is of particular interest in MS.

1.3.3.1 *Working memory*

Baddeley (2012) modelled working memory as a four-component system, including those that store the information: the phonological loop, the visuospatial sketchpad, the episode buffer, and the central executive, which processes and manipulates the information. As discussed in section 1.3.2.1, immediate memory can be assessed through asking individuals to repeat a list of words immediately. In this working memory model, the auditory stimuli of the words would be stored in the phonological loop; the central executive would concern the manipulation of this information, such as ordering those words semantically, which would use semantic knowledge stored in very long-

term memory. The commonality between memory and information processing speed is clear, and working memory is frequently grouped with these subdomains as predominantly impaired in MS (Chiaravalloti & Deluca, 2009; Sumowski et al., 2018).

1.3.4 **Perceptual-motor function**

Perceptual-motor function includes visual perception, visuoconstructural ability, perceptual-motor function, praxis, and gnosis. Deficits in performance on visual perception tasks have been found in 25% of people with MS (Vleugels et al., 2000). However, these subdomains can be affected by common deficits in visual acuity, dexterity, and motor ability unrelated to cognition in MS.

MS research into this domain has often used tests of visual perception and visuoconstruction, assessed by figure drawing or copying designs of blocks. Whilst these are predominantly used as cognitive tests, they include a motor and dexterity element that could be compromised in people with MS (Lau et al., 1998). It can, therefore, be difficult to disentangle elements of cognitively impairment and impairment in other functions. Likewise, perceptual-motor ability can be impaired in MS due to motor function and often assessed as part of arm and hand dexterity assessments such as the Nine-Hole Peg Test (Polman & Rudick, 2010). Although predominantly used as a test of dexterity, it is possible that it includes elements of executive functioning and information processing speed associated with comprehension of the instructions and performing the task.

Gnosis refers to the recognition of previously learned information, such as recognising faces, whereas praxis refers to learned movements. These are closely linked to the recall of implicit memory, and are rarely impaired in MS (McNicholas et al., 2018).

1.3.5 **Language**

Deficits in language is conceptualised into two subdomains: expressive and receptive. Expressive language includes aspects of naming, word finding, fluency, grammar, and syntax, whereas receptive

language is concerned with comprehension. Receptive and most aspects of expressive (naming, grammar and syntax) remain largely unimpaired, however verbal fluency can be compromised (Bagert et al., 2002).

Many reviews focusing on neuropsychological impairment state that language is relatively unaffected cognitively (e.g., Calabrese, 2006; Chiaravalloti & DeLuca, 2008). Furthermore, research into language and communication difficulties in MS has often focused on the motor speech impairment of dysarthria (Renauld et al., 2016), which can occur in 25-50% of people with MS (Hartelius et al., 2000). Like perceptual-motor function deficits, determining how cognitive impairment could impact on language deficits, in the presence of physical deficits, is complex (Feenaughty et al., 2018, 2021). Literature suggests that some aspects of language are cognitively compromised, but it is unclear as to whether these are due to dysfunctions in related domains. Verbal fluency is commonly assessed and found to be compromised in MS, and although less frequently examined, comprehension may also impaired (Renauld et al., 2016). However, both may more accurately represent impairment in executive function. Similarly, deficits in comprehension may identify difficulties with long-term memory retrieval (Renauld et al., 2016).

A meta-analysis examining phonemic and semantic fluency deficits found that people with MS were significantly more impaired on verbal fluency than neurologically healthy individuals (Henry & Beatty, 2006). However, studies reporting verbal fluency must be considered with caution. Prioritising assessments of verbal fluency over other language assessments in a study limits our understanding about deficits in language abilities unrelated to cognition. Additionally, verbal fluency relies on several neurocognitive domains other than speech and language, such as executive functioning and information processing speed (Henry & Beatty, 2006).

1.3.6 **Social cognition**

Using the DSM-5's framework, social cognition includes both Theory of Mind (the ability to infer someone's intentions and beliefs) and facial, vocal, and bodily affect recognition (the ability to determine emotional states based on their face, voice, or body). Deficits in social cognition could have implications on interpersonal skills, which could negatively impact social life and quality of life (Cotter et al., 2016).

This domain is not a widely reported symptom of cognitive impairment in MS. However, two meta-analyses on social cognition in MS examined Theory of Mind and Facial Emotion Recognition in particular (Bora et al., 2016; Cotter et al., 2016). These studies found that people with MS showed greater deficits in both domains of social cognitive in comparison to healthy control participants. The authors also reported an association between social cognition and other domains, including attention, executive function, and visual perception, and poorer social cognition in progressive courses. There is mixed evidence on the association between social cognition and depression, fatigue, and quality of life (Genova et al., 2020; Grothe et al., 2021). There are fewer studies examining this domain than other cognitive domains affected in MS, and it is rarely included in neuropsychological batteries, so these results are not conclusive. Social cognitive impairment is an expanding area of interest in MS, and future studies may add further insight into dysfunction in this domain.

1.3.7 **Neuropsychological profiles of cognitive deficits in MS**

Attempts to succinctly and comprehensively profile cognitive impairment in MS has proved difficult, with one meta-analysis emphasising that "research often yields an ambiguous mix of results – decidedly significant, suggestive, convincingly null, and hopelessly inconclusive, perhaps reflecting the random distribution of white matter demyelinating lesions and their influence on cognitive processes" (Zakzanis, 2000, p. 131).

A meta-analysis quantifying cognitive impairment in individuals with RRMS examined effect sizes reported in research (Prakash et al., 2008). The 57 studies (*n*=3891) analysed suggested that overall cognition was negatively affected in RRMS, with significant differences in impairment between neurologically healthy participants and those with MS and moderate cognitive impairment in RRMS (*g*=-0.54). Performance in memory and learning (*g*=-0.61), complex attention and executive functioning (*g*=-0.56), showed predominant deficits in people with MS. The exploration of the neuropsychology of MS points to key subdomains that are impaired, and some of these, particularly information processing speed, may be the basis for impairment in other domains, such as working memory.

However, literature on cognition in MS has used a range of assessments to assess cognitive dysfunction, some of which may not be appropriate for the domain under investigation. The variance in how domains are operationalised, the assessments used, and how test results are interpreted could also contribute towards the inconsistency in how cognitive impairment is described in MS. This was remarked upon in the review by Chiaravalloti and DeLuca (2008) that, in particular, studies purporting to measure complex attention may use tests that would more accurately be described as tests of executive function. A recent secondary analysis also demonstrated this through the authors' attempts to identify which domains are impaired in MS (Zurawski et al., 2020). In their sample of 60 people with MS, their six functional domains were determined using 21 different tests. Some tests included in their overall attention/executive functioning domain may have been more appropriate for their memory domain, such as Digit Span tests.

Assessments used to assess cognitive impairment may also evaluate multiple subdomains from different domains; for example, the Paced Auditory Serial Addition Test (PASAT) is considered to assess both working memory and information processing speed. As a result, it

can be a challenge to comprehensively describe domain-specific impairment.

Cognitive domains are also not independent of each other. DeLuca et al. (2004) described two models relating to the association between cognitive domains in people with MS. The authors conceived these models to explore the interaction between information processing speed and other cognitive domains. The Independent Consequent Model, in which information processing speed deficits do not solely account for deficits in other domains (such as working memory), and the Relative Consequence Model, in which impaired processing speed does account for deficits in other domains. Using the Wechsler Adult Intelligence Scale (WAIS) and the Wechsler Memory Scale (WMS) to assess cognition, the authors found support for the Relative Consequence Model. This suggests that impairment in working memory is a consequence of information processing deficits.

Working memory involves simultaneously processing and storing information, and therefore bridges the overarching domains of attention and learning and memory. An example of the interrelatedness of some of the domains of learning and attention taken from a literature review of cognition in MS (Chiaravalloti & DeLuca, 2008) is shown in Figure 2. Clearly, this is not exhaustive, but demonstrates how complex attention can impact on other subdomains, including those of executive functioning, and memory and learning. In this example, slowed information processing speeds can impact sustained attention, which can impact the encoding of information (learning), and in turn affect decision-making, a component of executive function.

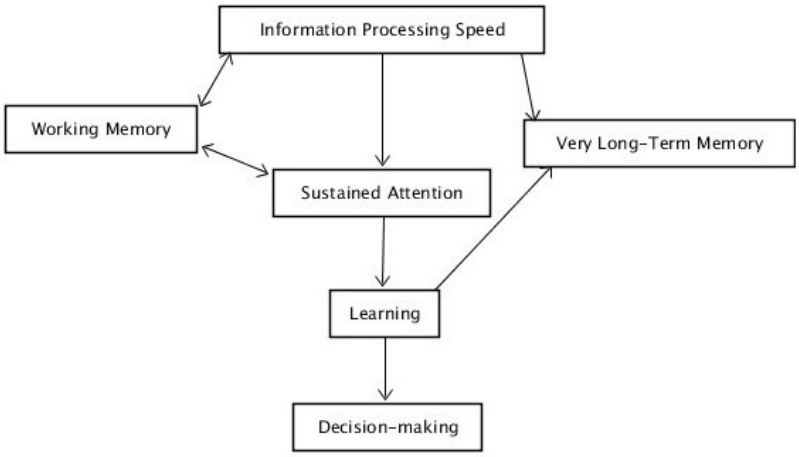


Figure 2: Interrelation of cognitive domains in MS (adapted from results in Chiaravalloti & DeLuca, 2008)

The use of different assessments and different interpretations of scores give a wide range of domain-specific prevalence statistics. Studies have attempted to profile cognitive functioning using novel methods of categorisation. Leavitt et al. (2018) described different classifications depending on which domains were impaired; processing speed impaired, memory impaired, or processing speed and memory impaired. A large Italian study has recently furthered this approach (De Meo et al., 2021). Using latent profile analysis on scores of the Brief Repeatable Battery of Neuropsychological Tests (BRB-N) and a Stroop task from 1212 people with MS and control participants, the authors used five different classifications of impairment:

1. No cognitive impairment
2. Mild verbal memory/semantic fluency impairment
3. Mild multidomain impairment
4. Severe executive function/attention impairment
5. Severe multidomain impairment

Both of these studies used descriptions of impairment that were taken from what the tests purport to measure and assess limited cognitive domains. However, approaches such as these allow for a deeper insight into the differences in impairment in MS.

Research and recommendations for exploring cognitive dysfunction in MS have favoured objective assessment (such as the

recommendations in Benedict et al., 2002 and Rao, 1990) above how people with MS describe their own experiences of impairment. Some subjective measures of cognition are described in section 1.4.1.3, and studies often find that self-reported cognition is not significantly associated with objective tests (Rosti-Otajärvi et al., 2014). However, some domains may be more suited for self-reports than others. A systematic review suggested that self-reported executive functioning may show stronger associations with objective measures than other domains (Mazancieux et al., 2019). This is worth further investigation, especially with domains neglected from standard neuropsychological assessment, such as autobiographical and prospective memory.

Metacognition may provide valuable insights into how people with MS perceive and describe how cognition impacts their everyday functioning and quality of life. One method of self-reported cognition, the Cognitive Assessment Interview (CAI; Ventura et al., 2010), uses a semi-structured interview procedure. A study using the CAI found that cognitively impaired people with MS reported difficulties with information processing speed (85% of participants), working memory (80%), verbal learning and memory (73%), and attention and concentration (59%). Social cognition (37%), and reasoning and problem solving (28%) were less likely to be reported as impaired (Eilam-Stock, et al., 2021). The domains identified are similar to those reported by studies using objective assessments, and show promise in different methods to explore the cognitive profile in MS.

Identifying impairment using the framework of neurocognition suggested in the DSM-5 could enable common assessment of these domains, including a more holistic representation of dysfunction and the methods to describe them. Of course, comprehensive appraisals using a range of techniques may be out of the scope in clinical practice, however the interrelatedness of the neurocognitive domains suggests that assessment could focus on the higher-level cognitive domains that underpin deficits in other domains. One recommendation suggested administering the SDMT to assess information processing speed,

alongside self-report assessments (Benedict et al., 2008). Approaches such as these could enable briefer and specific screening.

1.3.8 **Factors related to cognitive impairment in MS**

Cognition is related to other factors, including individual differences and other aspects of MS. Co-morbidities, including physical and psychological symptoms, pain, and fatigue can be a risk factor for cognitive impairment in MS (Amato et al., 2019), and it is unclear how factors relate to each other, and relate to cognitive impairment.

1.3.8.1 *Individual differences*

The association between demographics and cognitive impairment have been explored in MS. As we would expect cognitive decline in older age, it can be difficult to determine whether impairment is due to normal ageing, or the duration or progression of MS. Age has been found to be a significant predictor of cognitive decline in people with MS (Eijlers et al., 2018). Older individuals (aged 60 or over) show worse performance on tests of information processing speed and memory than young or middle aged people with MS (Baird et al., 2019). However, whilst people with MS show poorer information processing speed than control participants, the progression of decline across age is similar in both groups (Bodling et al., 2009).

There is disagreement as to gender differences in cognitive functioning. Some studies have reported that men with MS perform significantly worse on cognitive tests than women with MS, across the domains of attention (including information processing speed), memory, and executive functions (Lin et al., 2017; Schoonheim et al., 2012). However the progression of severity of impairment over time in men with MS is contentious (Amato et al., 2019).

Low education attainment in people with MS has been found to be associated with poor performance on cognitive tasks, particularly in individuals who have been diagnosed with MS for longer than five years (Rimkus et al., 2018). Higher educational attainment may be associated with cognitive reserve (Santangelo et al., 2019), whereby genetic or

environmental factors may protect against cognitive decline (Sumowski & Leavitt, 2013). Expanding past educational attainment, cognitive reserve further incorporates other factors that may protect people with MS from cognitive impairment, including lifestyle and occupation (Stern, 2002). Cognitive reserve is frequently assessed using questionnaires that ask respondents to report their education, leisure and vocational activities (such as the Cognitive Reserve Index; Nucci et al., 2011).

1.3.8.2 *MS-related differences*

Poor cognitive function is often found to be weakly related to general MS-related disability, predominantly measured by EDSS (Rao et al., 1991), however cognitive impairment can be detected in people with MS without physical disability (Marrie & Horwitz, 2010). The profile of cognitive impairment can vary across disability, duration of MS and phenotype. In a five-year study of 234 people with MS, 28% of participants showed decline in their cognitive functioning over the course of the study, and progressive forms were more likely to show decline (Eijlers et al., 2018). In this study, duration of MS was not shown to be predictor of further impairment but did have a strong correlation with cognition.

Cognitive dysfunction is apparent in all phenotypes of MS, but it can vary across relapsing and progressive forms. Individuals with progressive forms of MS have been shown to have poorer cognitive functioning than those with relapsing remitting forms (De Meo et al., 2021; Renner et al., 2020), and some research has suggested that secondary progressive courses may be associated with the lowest cognitive function (Amato et al., 2006). This longitudinal study also found that cognitive problems are unlikely to remit significantly across the course of MS, and the progression of MS is likely to be related to poorer cognitive outcomes. Additionally, there may be domain-based differences between MS phenotypes. Verbal fluency deficits have been found to be greater in people with progressive forms of MS, than relapsing-remitting (Henry & Beatty, 2006). However, the authors noted that progressive forms are more likely to have greater neurological

disability, longer duration of MS, and this could explain worsening cognitive impairment.

1.3.8.3 *Mood and fatigue*

Approximately 24 to 50% of people with MS show signs of depression, and which is associated with poorer quality of life and greater fatigue (Marrie et al., 2015), and dysfunction in the domains of working memory, information processing speed and overall executive functioning (Feinstein et al., 2014). Fatigue is also a prevalent symptom of MS, described as debilitating, and affects approximately 80% of people with MS (Krupp, 2006). Similarly, some studies have found associations between fatigue and aspects of cognition, in particular with complex attention. The relationship between these common symptoms and cognition is uncertain; for example, both depression and fatigue could limit in activities that help cognitive reserve in MS, affecting cognitive performance.

Acknowledging the potential confounding impact on cognition in MS, research has attempted to tease out how performance on assessments can be affected. Self-reported depression and fatigue are able to predict poorer performance on objective cognitive assessments and self-reported cognition (Hughes et al., 2019). These factors also predict that people with MS with severe cognitive impairment will underestimate their own cognitive function through discrepancies between objective and subjective cognitive impairment (Kinsinger et al., 2010). Although improvements in fatigue and depression did not influence objective cognitive performance, this particular study had implications on our understanding of self-reported cognition (as explored in section 1.4.1.3).

1.3.8.4 *Quality of life and activities of daily living*

People with MS with cognitive dysfunction have significantly lower quality of life than people with MS who are cognitively preserved (Campbell et al., 2017). In particular, deficits in information processing speed have been associated with poorer outcomes in quality of life and activities of daily living (Costa et al., 2017), and lower scores on

processing speed measures (such as the SDMT) can predict employment status with better accuracy than measures of other domains (Clemens & Langdon, 2018). This systematic review found that poor scores on objective cognitive assessments were related to unemployment and reduced working hours in people with MS. Cognitive impairment is also associated with adherence to MS treatments (Bruce et al., 2010), social support, and physical wellbeing even in early MS (Glanz et al., 2010).

Objective tests of everyday functioning have been used in research to explore how cognitive deficits contribute to day-to-day life for people with MS. Poor cognitive functioning can predict performance on tasks of 'actual reality' and executive function (Goverover & DeLuca, 2015; Kalmar et al., 2008). Using 'actual reality' tasks involving naturalistic activities, such as buying tickets online, people with MS performed worse than control participants, and weak performance on memory and information processing predicted more errors and necessitated more cues to successfully complete activities.

1.4 Measurement of cognitive impairment

The first diagnostic criteria of mild and severe neurocognitive disorders from the DSM-5 are impairment on a standardised neuropsychological assessment, or reports of cognitive deficits from the patient, clinician, or an informant. In literature and clinical practice, cognitive assessment can be defined as a comprehensive examination to evaluate cognition, which can include, but is not limited to, formal testing procedures. Assessments can include clinical interviews, file reviews and patient history. A recent critical review found inconsistent results of the incremental value that full neuropsychological assessment adds to clinical practice (Donders, 2020). In this thesis, the term 'cognitive assessment' will be used to describe neuropsychological testing specifically, and this term encompasses objective assessments of individual tests, multidomain batteries and cognitive screening tests; and questionnaires completed by people with MS or proxies such as clinicians or family members.

Some tests may have been developed as condition- or disease-specific tests but have been validated with other populations. Most assessments have been developed to be administered as in-person tests, completed with pen-and-paper, or verbally, and which are administered and scored by technicians. Technicians can refer to a wide range of clinical staff that administer cognitive assessments. Although test publishers may specify what qualifications are needed to administer and interpret tests, the level of expertise and experience can vary between assessments and administrations. Therefore, in this thesis, 'technician' is used as a common term for any person who administers cognitive assessments.

In choosing assessments to use, clinicians and researchers must consider factors such as available resources (including time, staff, and their expertise), appropriateness for the population, assessment modalities (such as written, verbal, or digital), and availability of assessments. Many assessments may be limited by the language they were developed in, and the availability of norming data for the country in which the measure is to be used. Increasingly, there have been developments in digitised assessments. Digital assessments can be newly developed tests which lend themselves to digital administration, or adaptations of validated tests. These can mitigate some of the factors that hinder assessment, however, require evidence of their equivalence to their originally developed version, and appropriateness for the change in administration.

1.4.1 **Current cognitive assessment in MS**

Cerebral functions involving cognition are one of the eight functional systems used in the EDSS to assess MS disability (Kurtzke, 1983). In examinations of neurological functions, cognitive function is based on clinical judgement rather than an objective measurement. A chart review of people with MS sought to identify the accuracy of neurologists' predictions of cognitive impairment in their patients through standard neurological examination and clinical interview (Romero et al., 2015). Standard clinical interviews were not found to

accurately identify cognitive deficits, as defined by low corresponding scores on an MS-specific cognitive battery (X^2=1.25), and there is clear value in routine neuropsychological assessment with valid tools, to control for this inaccuracy in clinician-reported deficit.

Both generalised and MS-specific cognitive assessments are used in research and clinical practice. Recommendations have stressed the importance of routine cognitive assessment for all people with MS (Bakirtzis et al., 2018; Sumowski et al., 2018), however this is not corroborated in clinical practice. A survey of US and Canadian neuropsychologists reported that 47.5% never or rarely worked with people with MS (Rabin et al., 2005). Therefore, it is possible that assessment may be conducted by other clinicians. However in a further US survey of healthcare professionals working in MS, 49% had no formal procedure for cognitive assessment (Foley et al., 2012). In a more recent survey of MS clinicians in the UK, only a third of respondents routinely assessed cognition (Klein et al., 2019). The low uptake of assessment could be due to the lack of time and resources available to conduct such tests. Administering a full neuropsychological battery may take up to two hours, and relies on clinics having the staff, expertise, capacity, space, and access to such assessments.

There is inconclusive evidence as to how much value a full neuropsychological examination adds to clinical practice (Donders, 2020). This suggests that easily accessible, brief screening may be more beneficial for clinical practice than a full neuropsychological examination in the first instance, which has considerable resources and costs associated with it. Recommendations for MS have endorsed an approach of routine screening with brief measures, followed by a comprehensive examination if the results indicate impairment. Batteries and screening tests are constructed using individual tests to measure one or more domains of cognition. Further details of individual tests and batteries and screens used regularly in MS are explored in Chapter 2.

1.4.1.1 *Individual tests of cognition in MS*

Individual tests can be used for very brief screening of cognitive impairment, or as part of multidomain assessments including batteries. This is explored further in the systematic review in Chapter 2, but the following paragraphs will provide key background on some of the most frequently recommended and used individual tests. These tend to focus on the commonly identified domains of information processing speed, memory, and executive function. Stimuli of tests are predominantly visual or auditory, and responses given are verbal or motor (mostly written). For example, a respondent may receive a spoken instruction, and then write their response. Results of cognitive tests can include the amount of time taken to complete a task, correct and incorrect responses given, and the time taken to respond to each individual stimulus.

The PASAT was developed in 1958 as a measure of attention and immediate memory (Sampson, 1958), however more recently, information processing speed has been considered the primary domain underlying performance (Forn et al., 2008). The PASAT forms part of a number of neuropsychological batteries, including the BRB-N and MACFIMS, and the cognitive test of the Multiple Sclerosis Functional Composite (MSFC, Rudick et al., 2002), a clinical outcome measure of the severity of MS. During administration of the PASAT, a series of single digits are presented at intervals and respondents are asked to sum the last two digits together and say the answer. Whilst the PASAT has good psychometric properties and sensitivity (Tombaugh, 2006), its use is contentious. Premorbid mathematical ability and age can influence performance (Tombaugh, 2006) and people with MS perceive completing it as unpleasant (Walker et al., 2012), with multiple studies finding that people with MS may decline to complete the testing procedure (Aupperle, et al., 2002; Cortés-Martínez et al., 2019).

Whilst use of the PASAT has been ubiquitous, the SDMT is often heralded as the gold standard of brief cognitive screening in MS, and recently there have been numerous efforts to ascertain its validity and

promote its use (Benedict et al., 2017; Kalb et al., 2018; Strober et al., 2018). It has also been recommended as a replacement for the PASAT in the MSFC (Drake et al., 2010). The SDMT consists of a key of symbols matched with a number from 1 to 9, presented with a grid underneath showing 120 symbols. Participants have 90 seconds to complete the grid using the key to match the symbol with the correct number, either by writing or saying the correct number. In the original administration, both written and oral tests are completed by the respondent (Smith, 1982). However, in MS research, the oral administration is predominantly used to reduce the confounding effect of poor manual dexterity for people with MS (Benedict et al., 2017). The validity and reliability of the SDMT is adequate (as demonstrated in section 1.4.2), and it is perceived as the least demanding test of the BICAMS, with people with MS finding it acceptable to complete (Mortensen et al., 2020). A longitudinal study, across ten years, found that people with MS who had low SDMT scores showed greater deterioration in performance in comparison with people with MS with better SDMT performance at baseline (Healy et al., 2021). This suggests that those with lower scores may benefit from early intervention to prevent worsening of their information processing speed.

Both the SDMT and PASAT are understood to assess information processing speed. It can be difficult to extricate which of three processing speeds is impaired: sensorial, cognitive or motor (Costa et al., 2017). For example, in a written SDMT, all three factors may lead to poor results as the person with MS has to process the visual information of the test (sensorial speed), the matching of the symbols (cognitive speed) and use a pencil to write the answer (motor speed). The oral version may also have similar associated difficulties with oral-motor slowing common in MS (Arnett et al., 2008). There may be other aspects of cognition involved in performance, with some studies using these tests as measures of executive function, or immediate or working memory. A recent study concluded that three cognitive domains contribute towards performance of the oral SDMT: memory, information

processing speed and a ‘lexical access speed’, distinct from information processing (Sandry et al., 2021). This research suggests that the SDMT is appropriate as a more global cognitive assessment than previously believed.

Other frequently used tests are also susceptible to oral-motor slowing when assessing the speed at which stimuli are produced by the respondent. Verbal fluency Tests (VFT), such as the Controlled Oral Word Association Test (COWAT, Benton et al., 1994), involve respondents saying aloud as many words they can think of that belong to a phonemic category (such as words beginning with ‘F’) or a semantic category (such as ‘animals’) within a specified time period. VFTs provide insight not only in accessing semantic memory, but executive functioning in the way people with MS plan their responses, and as such they are sensitive to impairment (Ruff et al., 1996). Other verbal tasks involve word list learning, such as the California Verbal Learning Test II (CVLT-II, Delis et al., 2000), in which participants are read or hear a list of words and are asked to recall or recognise the words. There are a number of recalls that can be assessed, including over an immediate, short or long time period; cued or free recall; or recognition of the words heard. People with MS perform significantly worse on this test than control participants, and it also distinguishes between employed and unemployed people with MS, suggesting ecological validity (Stegen et al., 2009).

1.4.1.2 *Multidomain assessments of cognition in MS*

Multidomain assessments are defined sets of individual tests, which assess multiple domains of cognitive function. These can be brief screening assessments, such as the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), or comprehensive batteries, such as the WAIS (Matarazzo, 1978). There have been several attempts to develop standardised batteries for use in MS. There are descriptions of additional batteries used in MS as part of Chapter 2, but here I will discuss the development and administration of three frequently used MS-specific batteries: the BRB-N, the MACFIMS, and the BICAMS.

The BRB-N began with 23 tests included (Rao et al., 1991), and was eventually reduced to shorter battery of the five most sensitive tests, including the 10-36 Spatial Recall Test (SPART), the Selective Reminding Test (SRT), and PASAT, with the addition of the SDMT and a VFT task (Bever et al., 1995). Whilst overall it is regarded as a sensitive battery, there are challenges with the inclusion of the WLG task, which is not able to distinguish between control participants and people with MS (Sepulcre et al., 2006) and the PASAT, due to respondents' dislike of this test (Tombaugh, 2006). It also has a relatively long administration time of 30 minutes.

In 2001, a committee of neuropsychologists and psychologists proposed the Minimal Assessment of Cognitive Function in MS (MACFIMS; Benedict et al., 2002). The aim of this development was to construct a comprehensive battery to improve routine assessment of cognitive impairment in MS. The MACFIMS covers multiple domains, is psychometrically robust, and consists of seven tests: the Brief Visuospatial Memory Test-Revised (BVMT-R), COWAT, CVLT-II, PASAT, SDMT, a sorting test from the Delis-Kaplan Executive Function System (DKEFS; Delis et al., 2001) and a visual perception task of Judgement of Line Orientation (JLO; Benedict et al., 2006). However, the 90 minutes taken to administer the battery is a barrier for use in routine clinical care. An abbreviated five test battery, the aMACFIMS (consisting of the BVMT-R, COWAT, CVLT-II, SDMT, and card sorting), was further developed, which included a measure of executive functioning but its specificity may be compromised (Gromisch et al., 2018).

The BICAMS was constructed in 2011 by another consensus panel of seven neurologists and five neuropsychologists with expertise in MS (Langdon et al., 2012). The panel proposed a three-test battery (consisting of SDMT, CVLT-II and BVMT-R) that can be used in clinical practice with ease. A recent systematic review and meta-analysis found 16 validation studies as of 2018 and found that people with MS scored significantly lower on all three tests compared to control participants,

with the SDMT having the largest effect size (*g*=0.934; Corfield & Langdon, 2018).

All three batteries described show promise but have not had the uptake in clinical MS care originally anticipated. This could be due to a number of practical difficulties. The named tests in all three are published by different copyright holders, and require clinicians who wish to use the batteries to approach each publisher separately, causing financial and administrative burdens (as experienced by Beier et al., 2019). Scoring the tests also requires expertise to calculate and interpret results, which may be out of the scope of clinicians administering the batteries.

Studies using Principal Components Analysis (PCA) on the batteries developed for MS have found two common domains for both the BRB-N and MACFIMS: attention, and memory. The MACFIMS indicates additional factors, including executive function and refinement of memory into visuospatial and verbal (Golan et al., 2019; Ruano et al., 2020), whereas the BRB-N shows a discrepancy whether memory is impaired in a single overall domain, or multiple domains (Benedict et al., 2006; Damasceno et al., 2020). Whilst this provides some insight, these neuropsychological batteries were constructed for use in MS, and as such the components extracted are unsurprising. Future approaches could aim to move away from these preconstructed batteries, and instead use a more focused approach. Hansen and Lautenbacher (2017) endorse the use of a range of psychometric measures, including tests that are not routinely included in MS-specific batteries, such as social cognition. Batteries could potentially be adapted to different diagnostic requirements, for example to assess domains associated with everyday functioning, and include other measures that may confound performance on cognitive tests, such as mood and fatigue.

1.4.1.3 *Subjective measures of cognition in MS*

There are a number of subjective measures of cognition used in MS, including the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ; Benedict et al., 2003), the Perceived Deficits Questionnaire

(PDQ; Sullivan et al., 1990), the Multiple Sclerosis Inventory of Cognition (MUSIC; Kalbe et al., 2013) and the CAI (Ventura et al., 2010). Subjective measures are most often completed by the person with MS; however, some measures have proxy, or informant measures, whereby a family member or clinician reports on the individual's cognitive functioning. The CAI uses a semi-structured interview conducted by a clinician to quantify cognitive function.

Subjective measures of cognition have some benefits over objective measures. They generally require fewer resources in terms of administration, for example, a technician is not needed; administration and scoring are simpler; responses can be collected over the telephone, and they are easier to adapt to digital versions. They are likely to be shorter and cheaper, and more easily accessible for both clinicians and patients. Respondents may also prefer completing subjective measures over objective measures due to the lowered burden, and lessened anxiety that some assessments can provoke (Aupperle et al., 2002). However, their use is contentious with many studies finding weak associations with performance on objective tests (Benedict et al., 2003; 2004). Subjective measures may also be perceived as less accurate owing to confounding factors (such as mood) on scores (Kinsinger et al., 2010; Rosti-Otajärvi et al., 2014).

One of the most frequently used questionnaires in MS, the MSNQ, has been shown to have good psychometric validity and reliability (Benedict et al., 2003) and consists of an informant-reported version (the MSNQ-I) and a patient-reported measure (the MSNQ-P). Its developers have recommended its use, together with the SDMT, for regular screening in MS (Benedict et al., 2008). However, some research has shown a stronger correlation of MS-reported cognition with depression (Hughes et al., 2019; Rosti-Otajärvi et al., 2014), than with performance on objective cognitive assessments. As such, there is concern regarding the sensitivity and specificity of the MSNQ-P (Nauta et al., 2018). Some of these studies have shown that the MSNQ-I has a stronger relationship with objective measures, but people with MS may

not have an appropriate informant to complete the questionnaire or may feel uncomfortable with this. A semi-structured interview approach, the CAI, has shown some promise in its relationship with objective measures of cognition (Eilam-Stock et al., 2021). This interview-based assessment was shown to correlate with objective measures of executive functioning, and the overall scores of the MSNQ-P.

Subjective measures used in MS can often be brief and focus on capturing an overall assessment of cognition (such as the MSNQ-P and PDQ). It is possible that particular cognitive domains may be more sensitive to subjective assessment than others, and this could indicate the value of developing and using multidomain questionnaires over unidimensional ones.

1.4.1.4 *Classification and interpretation of cognitive assessment*

Some studies rely on a binary interpretation of cognitively impaired or unimpaired to categorise cognitive function. It is also possible to classify cognitive impairment in terms of which cognitive domains are affected, either through domain-based; or a composition of domains; for example, measures of information processing speed and selective attention could be combined into a 'complex attention' classification (Leavitt et al., 2018). A more detailed profile can complement understanding of cognitive functioning in MS and examine associations between different domains.

An additional consideration for neuropsychological assessments for people with MS is that of the threshold, or cut-off scores, available for discerning those who are cognitively impaired from non-impaired individuals. A common method of evaluation of scores uses raw score conversion of clinical scores into z scores, and a comparison against normative data of control participants. A cut-off can then be derived using standard deviations below the normative means. The DSM-5 (American Psychiatric Association, 2013) defines the difference between mild and major neurocognitive disorder using standard deviations (SDs) from the normative means. Mild impairment is defined as 1-2 *SDs* below the normative mean, and major is 2 *SD* below the

normative mean. However, these depend on the tests being used (and their predictive values, explored in section 1.4.1.5) and factors such as premorbid level of cognition, and the availability of relevant norming data that can be applied.

Recent research has used a cognitive functioning-based z-index, in which cognition can be presented as a continuous variable, as opposed to the classic method of a cut-off score derived from norming data, to discern cognitive impairment (Pitteri et al., 2019). By using this approach, the authors found a greater number of newly diagnosed RRMS patients would be considered cognitively impaired, and advocate for early, routine neuropsychological assessment. However, this approach requires greater statistical analysis to be applied to the data of people with MS and control data, which may not be possible within routine assessment in healthcare settings.

Using multidomain tests, impairment can be defined through several tests considered 'failed' using the cut-offs for each test. Alternatively, a single or composite score below the cut-off can be used as the differentiation between impaired and non-impaired. In using multidomain batteries of assessments, it may be possible to ascertain a more detailed profile of cognitive functioning than the binary of impaired or unimpaired for each test. Factor analysis of neuropsychological batteries can further our understanding of what tests measure and the relationships between domains. For example, Ruano et al. (2017) found that in the four components extracted from a PCA on the BRB-N, the SDMT bridged two components – those of information processing speed and visuospatial memory. Leavitt et al. (2018) proposed that cognitive impairment in MS could be described through domain-based categories, or 'cognitive phenotype'-based (using a combination of domains). A similar approach was also used by de Meo et al. (2021), which reported five categories based on the combinations of tests from the BRB-N and the Stroop task. However, these approaches are scarce and require further validation with a greater range of assessments.

1.4.1.5 *Sensitivity and Specificity*

Cognitive assessment can be assessed for their clinical utility, using sensitivity and specificity (Lalkhen & McCluskey, 2008). Sensitive cognitive assessments will correctly identify people with MS with cognitive impairment. Correspondingly, specific cognitive assessments can identify those who do not have any cognitive impairment. For example, if a cognitive assessment had 90% sensitivity and 20% specificity, many people with MS would be correctly identified as being impaired (a true positive), but many who are not impaired would also be categorised as impaired (a false positive).

In a review of the sensitivity and specificity of the SDMT, Parmenter et al. (2007) found that a cut-off score of 55 or lower provided an adequate sensitivity of 82% and a specificity of 60% in classifying cognitive impairment in people with MS. However, alternative cut-offs could be used to find different levels of sensitivity and specificity if required.

1.4.2 **Validity and reliability**

1.4.2.1 *Validity*

Validity refers to the extent that the test measures what it proposes to, and there are a number of facets to this (Cook & Beckman, 2006). Face validity is often one of the first validations undertaken with an assessment. Interviewing the proposed end-users of tests can explore whether the test is subjectively perceived as suitable and appropriate. This form of validity is frequently neglected in evaluating neuropsychological tests, and as some authors argue that it is not an important component of validity given discrepancies between objective and subjective cognitive impairment (Walker et al., 2012). However, there is still value in assessing and reporting face validity to ensure acceptability and relevance to those administering and completing a test.

Ecological validity evaluates how a test relates to everyday activities and functioning. The ability of performance on a test to predict

performance or behaviour in real-life situations implies ecological validity, and increases the clinical utility of tests (Weber et al., 2019). Poor performance on the SDMT has been found in a few studies to predict vocational status in people with MS, suggesting its ecological validity for this real-world outcome (Benedict et al., 2005).

One of the most frequently assessed forms of validity in neuropsychological assessment is convergent validity. Finding a strong relationship between the test under evaluation and another measure expected to be strongly related helps to confirm that the test is valid. Administering the SDMT and other tests of cognitive functioning, significant correlations were found between other tests in the BRB-N, indicating convergent validity through measuring the same and similar constructs of cognition, such as information processing speed and working memory (Sonder et al., 2014). Known-groups validity can be determined through comparisons of groups we would expect to perform differently. The SDMT has shown to be valid through its ability to discriminate between people with MS and control participants, with people with MS performing significantly worse (Benedict et al., 2017).

Multidomain batteries of cognitive assessment may be judged by their content validity. As demonstrated, many batteries used in MS do not include tests of cognitive domains that may be impaired, so batteries could be evaluated as to whether they represent all aspects of cognition. As remarked, batteries such as the BRB-N, are not comprehensive, and may lack tests of commonly impaired but not commonly assessed domains, such as social cognition or prospective memory (Cotter et al., 2016; Rouleau et al., 2018). Therefore, these batteries may lack content validity.

1.4.2.2 *Reliability*

Reliability refers to the consistency of scores of an assessment (Cook & Beckman, 2006), and test-retest reliability is routinely examined. It would be expected that respondents would score the same result on a test across a brief period of time, providing there has been no neurological change. The SDMT demonstrates strong

correlations between repeat administrations over two weeks (r=0.97; Benedict et al., 2017), supporting its test-retest reliability. However, many tests can be subject to a practice effect in which respondents improve their scores. This can be mitigated using alternate forms.

Inter-rater reliability is also a key index of reliability with technician-administered assessments. This examines the extent of agreement between different raters on rating the performance on a test. Whilst this is not particularly contentious for the results of the SDMT, another frequently used assessment, the BVMT-R, demonstrates the importance of evaluating this for some tests. The BVMT-R is scored through assessments of drawings of geometric figures from memory, and thus raters may have a subjective judgement of the responses (Gaines et al., 2008). A strong correlation between all raters' scores provides inter-rater reliability evidence.

Internal consistency can also assess the reliability of test batteries. This measures the scores of all tests within a battery to determine the degree of relationships between them, implying that they are measuring related constructs, demonstrated through a high Cronbach's alpha (Cronbach, 1951). Validation studies of BICAMS report the internal consistency, such as the translation and validation study by Filser et al. (2018), which found good level of agreement between the three tests in the battery.

1.4.3 **Digital neuropsychological assessment**

The Topol Review (Topol, 2020) outlined how digital health initiatives implemented within the National Health Service (NHS) could have patient benefits. One of the recommendations was for patient-facing applications to be provided remotely on a user's personal device. In 2020, 96% of all adults in the UK had access to the Internet, with 46.6% using the Internet daily, and 15% using the Internet to access online health services (Office of National Statistics, 2020), and these figures are projected to increase. There are several benefits of digital assessments, including improving accessibility, reducing the resources

needed in traditional administrations such as time, staff and cost, greater precision in recording responses, automated scoring and interpretation, and integration with other systems, such as electronic health records. There may also be greater acceptability by providing assessments that can be completed in the naturalistic and comfortable setting of a respondent's own home, which may reduce anxiety associated with attending a clinical appointment.

The Topol Review was released before the COVID-19 pandemic; however, it became especially pertinent with the necessity of telemedicine during this time. Rapid guidance was released on how teleneuropsychology can be continued despite physical distancing necessitated by the pandemic (Bilder et al., 2020). However, neuropsychological technology developments have generally been slow (Miller & Barr, 2017), and there are a number of considerations to account for this, including:

1. What device the assessment is deployed on: a personal computer (PC), or a mobile device (such as a smartphone, tablet, or Personal Digital Assistant [PDA]), and how the participant responds to stimuli (peripheral inputs such as a keyboard and mouse; a touchscreen using a finger or stylus; or a verbal response). If a smartphone or tablet is used, the assessment may be a native application, or a web application accessed through an internet browser.
2. Who administers the assessment: if it is self-administered or administered or guided by a technician.
3. Where the assessment takes place: in-clinic or remotely.
4. How the assessment is scored: by a technician or automatically.

Sumowski et al. (2018) identified key priorities for the assessment of cognitive impairment, one of which being routine integration into existing clinical MS care. Indeed, the US National MS Society's recommendations suggested that a self-administered cognitive screening test can be used in routine MS management to make access to screening easier and improve uptake (Kalb et al., 2018). However, while calls for digitisation of assessments have been prevalent for a

long time (for example, Achiron & Barak, 2006), application in MS clinical care has been protracted.

There have been several digital developments of cognitive assessments for use in MS, but many of these are limited in their use, and evidence as to how they could be used in clinical practice is lacking. A recent systematic review of digital cognitive assessments for people with MS found that most digital assessments use a PC with a peripheral mouse and keyboard, administered in clinic, and require the presence of a technician (Wojcik et al., 2019). Self-administered assessments included in this review nevertheless showed equivalence to technician-administrations.

Test developers and copyright holders may have little incentive to develop digital translations of their tests, and commercial endeavours undertaken by private software companies, often sacrifice psychometric standards (Miller & Barr, 2017). As experienced by the development of the digitised BICAMS, the iCAMS, copyright holders of tests may also deny permissions for digital developments of their tests (Beier et al., 2019).

Many of the assessments commonly used in MS are challenging to digitise. The BVMT-R, included in both BICAMS and MACFIMS, requires participants to draw geometric line figures from memory (Benedict, 1997). Technicians then use a rubric to score these drawings, which has clear difficulties for digitisation. Attempts to develop digital versions of the BICAMS have used iPad administrations, and necessitate a technician to score the drawings of the BVMT-R Birchmeier and Studer (2019) approached this obstacle using machine learning to try to automate scoring. Although accuracy showed promise, this was with a small sample of neurologist-scored drawings, and certainly required further refinement and testing to ensure equivalence with technician ratings. Using a drawing-based task also further restricts the devices that assessments can be deployed on, and the differences in device size may restrict the ability to draw accurately. Another digital development of the BICAMS did not include the BVMT-R test due to the

difficulties of translating such a test to a digital format (Barcellos et al., 2020).

To mitigate such problems, it would be judicious to approach future digital development using valid tests that lend themselves to the modality. Of course, this is a trade-off between the advantages of digitised tests, and the disadvantage of not being able to use many of the sensitive tests that are recommended for use in MS care. However, the opportunities associated with providing routine assessment for people with MS in a cost-effective and efficient way may be too valuable to disregard.

All forms of digital assessment can have benefits over pen-and-paper versions, but there is particular potential for self-administered assessments that can be completed remotely as this can help to protect clinic time and resources, and being accepted by people with MS (Patel et al., 2019; Settle et al., 2015). Recommendations and guidance has begun to emerge for neuropsychological assessment with vigour in light of COVID-19's restraint on in-person clinical practice (Bilder et al., 2020). Given the difficulties of attending clinics during the pandemic, and the impracticality of providing standardised technology to all patients to complete at home, a solution may be to deploy assessments on personal mobile devices. Assessments can be developed for all types of devices, and patients can use their own device to access digitised cognitive assessments.

In Great Britain, 79% of adults use a smartphone to access the internet “on the go”, and many people show a preference for using their smartphone instead of a computer to complete day-to-day activities (Office of National Statistics, 2020). This growing use and familiarity with mobile devices have the potential to increase the uptake of digital health initiatives. This approach can enable people with MS to complete assessments routinely in their own home; reducing time spent in clinics and increasing accessibility (Casaletto & Heaton, 2017). There are difficulties to overcome, and assessments require future-proofing to ensure that they do not become obsolete, or cease functioning on

newer devices. Casaletto and Heaton (2017) urge developers of digital assessments to gain input from neuropsychology, rather than being driven by software development. Combining qualitative Human Computer Interaction (HCI) methods and psychometric approaches can support the development of digitised cognitive assessments that are usable, acceptable, and appropriate, and psychometrically robust. Digital assessments are new assessments and they therefore require the same evidence as required for new assessments, and new normative data (Bauer et al., 2012).

1.5 Gaps and limitations of previous published work

There have been several reviews and recommendations for cognitive assessment in MS (Hansen & Lautenbacher, 2017; Kalb et al., 2018; Langdon et al., 2012; Sumowski et al., 2018), but there is still uncertainty about which tests and multidomain batteries are being used to examine cognitive dysfunction in MS. Surveys of neuropsychological clinical practice have shown that assessment can be sporadic and not specific to MS (Foley et al., 2012; Klein et al., 2019). It is unclear whether this is also the case in MS research, and whether recommendations have had an impact on the uptake of specific tests and batteries. There is also uncertainty as to the prevalence of digital assessments. Whilst a systematic review identified validations of digital developments (Wojcik et al., 2019), the extent to which these are being used, and in which contexts, is unknown. The aims of Chapter 2 builds upon the work of the described surveys and previous reviews of digital administrations. This systematically establishes which tests are used in MS research, in which contexts, and using which modes of administration.

The SDMT has emerged as one of the most promising objective cognitive assessments for brief screening in MS recently (Benedict et al., 2017). Ubiquitous use of a single test can allow comparisons across studies; however, research often use different methods of interpreting scores. There have been recent interpretation approaches that move beyond a simplistic impaired-unimpaired categorisation (for example in

De Meo et al., 2021; and Leavitt et al., 2018). However, there is little guidance of on how the scoring and interpretation of the SDMT can extend beyond the traditional scoring to provide greater insight. Therefore, in Chapter 3, a scoping review aimed to identify and describe these different interpretative methodologies.

Along with the SDMT, the MSNQ is recommended as a self-report questionnaire of cognitive problems (Benedict et al., 2008). Whilst there is contention with regards to the relationship between subjective and objective scores (Hughes et al., 2019), there is some promising research that indicates that some cognitive domains, including executive functioning, are more associated with subjective ratings (Eilam-Stock et al., 2021; Mazancieux et al., 2019). The research described in Chapter 4 explored multidimensional questionnaires and their relationship to overall cognition, or specific domains, and also explored how results of the SDMT can predict other MS symptoms.

Research determining how to measure cognitive impairment in MS has predominantly used quantitative approaches. There are some promising studies exploring opinions and experiences of assessment using a qualitative research methodology, however there is only one study examining this, with a small sample of people with MS in Denmark (*n*=12; Mortensen et al., 2020). This study found that people with MS valued routine assessment with brief assessments such as the SDMT. Despite recommendations, cognitive assessment is not routinely conducted with people with MS, and the standardised batteries and tests proposed for use in MS care are not being used in clinical practice (Klein et al., 2019), and certainly, digitised assessment has not had the uptake anticipated by such studies as Achiron and Barak (2006). Understanding how cognitive assessment can be used in clinical practice, and how to promote engagement through appropriate tests and methods of administrations could be crucial to implementing routine assessment in clinical MS care. Thus, this was the aim of a large qualitative study described in Chapter 5.

Technological advances have a potential to increase uptake, by making assessment less resource intensive and more acceptable. Miller and Barr (2017) advocated for technological advances in teleneuropsychology through digital cognitive assessments but acknowledged that this has been slow, and the COVID-19 pandemic revealed the gaps in how to provide cognitive assessment remotely (Bilder et al., 2020). Digital adaptations of brief assessments have been restricted to specific devices, such as iPads (Beier et al., 2019; Middleton et al., 2020; Rao et al., 2017; van Dongen et al., 2020). Whilst these may be practical for use in research where devices can be provided to participants, accessibility may be an issue for healthcare delivery, with clinics and patients not necessarily having access to the specified devices.

A systematic review of mobile health initiatives in MS found a number of barriers, including those related to MS symptoms (Gromisch et al., 2020). For example, small text/font size could be an issue for people with MS with poor visual acuity, and small buttons to press may be affected by fine motor dysfunction. This review emphasised the importance of including users in the design of mobile technologies throughout the process. Especially in the development of device-agnostic assessments, a range of devices can be used in this process to identify any device-related differences. However, this is rare in research. Two promising developments of smartphone versions of the SDMT did not integrate usability testing into their development (Pham et al., 2021; Van Oirschot et al., 2020). Whilst their quantitative data is promising in showing discriminant and convergent validity, their development has not yet been researched with regard to aspects of face validity and user acceptance. The design and presentation of neuropsychological tests can change the behaviour of the end user, and ensuring clear and simple designs and instructions can have a pronounced impact (Germine et al., 2019). Therefore, the research described in Chapter 6 applied methodologies predominantly used in Human Computer Interaction research to establish the usability of a

developed cognitive screening platform, and user acceptance. The next stages of development of this technology require further evidence of its validity and reliability and generating data to interpret the results.

This introductory chapter has highlighted that cognitive deficits, across a range of domains, are a frequent symptom of MS, and there is a need for effective ways of assessing these. Having valid assessments that can be easily implemented within clinical practice has numerous benefits. Initial routine assessment can reveal cognitive deficits at an early stage, enable engagement with rehabilitation, monitor progression of impairment, and be used as an outcome measurement to appraise the efficacy of interventions.

1.6 Summary

Previous research has examined how individual tests and generic multidomain batteries can measure deficits (such as the batteries described in section 1.4.1.2). Recommendations have predominantly used expert consensus panels, and results from batteries, to advocate domains to be assessed, and with which tests. The lack of a common framework has often led to under-researched domains such as social cognition, autobiographical and prospective memory, which are increasingly being acknowledged as impaired in MS (Cotter et al., 2016; Rouleau et al., 2018). The availability of an acceptable, easily administered platform could have implications for the future of cognitive assessments in MS. Including methodological approaches, such as precise scoring data, can increase capability of assessment, and provide batteries that are relevant, acceptable, and accessible to people with MS. This could support a more comprehensive profile of impairment that furthers our understanding of the profile of cognitive dysfunction in MS.

1.7 Objectives

The objective of the research described in this thesis is to examine how cognitive assessment can be conducted in the setting of UK clinical care. The research conducted as part of this PhD are the

first stages in establishing the value and feasibility in providing cognitive screening to all people with MS as part of their routine care. The objectives were to explore current cognitive assessment practice, determine how cognitive assessment could work in routine healthcare, and based on these, evaluate a digital cognitive assessment platform for people with MS. A single task, based on the SDMT, was developed in its first iteration, but this opened the potential for further tests and analysis, including self-report questionnaires, and personalised batteries. Moving past expert consensus approaches, I have examined cognitive assessment from a patient-centred perspective and investigated how it can be implemented successfully using this approach. Where research has focused on recommendations and psychometric validation, one of the driving forces in this thesis is ensuring acceptability of a cognitive screening programme for both people with MS and clinicians working in MS.

1.8 Research plan

The first phase of this research was a systematic review (described in Chapter 2), to determine the breadth and purposes of cognitive assessments that have been used in peer-reviewed research with people with MS, systematically identifying which cognitive assessments have been used in research. I also aimed to explore whether recommendations for cognitive assessments have influenced the measures being used. This reports how frequently individual tests, multidomain tests and subjective measures have been used, in what research setting, and for what purpose. This is the first step in examining how cognition is currently assessed in MS. These findings, along with consensus recommendations were used to inform the next stages of the research.

The frequent use of the SDMT within research and the recommendations of its use as a screening tool, and part of a battery, was examined further. Whilst the scoring of the SDMT is simple, recommendations have not specified how to interpret the meaning of these scores. There are multiple ways of interpreting raw scores, and

categorising people with MS into potential categories to describe their impairment. I undertook a scoping review in Chapter 3 to explore some of these interpretation methods. I highlight these diverse approaches through summaries of the most common methods of categorising cognitive impairment, demonstrated by a sample of studies using these methods. This also highlights some of the opportunities of digital administration and automated scoring and interpretation of the SDMT.

The data analysis in Chapter 4 proceeds to determine the dimensionality of a self-report questionnaire, the MSNQ-P, and also examine how scores on the SDMT can be predicted by self-reported physical and psychological symptoms. Some people with MS may not wish to complete objective cognitive assessments (Cortés-Martínez et al., 2019), and the availability of a self-report questionnaire of cognitive problems may be more acceptable to some people with MS. This analysis uses a Principal Components Analysis to determine the dimensionality of the MSNQ-P and how extracted domains of subjective cognitive impairment, and other self-reported aspects of MS, relate to objective assessment. The aim of these analyses was to establish the utility of the SDMT, and self-reported cognition, for routine cognitive assessment.

The US National MS Society's recommendations promoted the practical nature of routine screening and the role that digitised assessment can have in accessible and brief screening (Kalb et al., 2018). Blandford (2019) has approached digital health development using five stages of iterative design. Neuropsychological assessment has often not involved people with MS and clinicians working in MS, and this is particularly apparent in digital developments (Miller & Barr, 2017). Implementation of a routine screening programme can only be successful if it is accessible and acceptable to those completing or administering the assessment. People with MS perceiving the value of assessment and being motivated to take part in routine programme could be critical to its success. Conversely, if assessments provoke anxiety, and are inconvenient to access, implementation may not occur,

as shown by patients' reluctance to complete the PASAT (Cortés-Martínez et al., 2019).

In the next study, described in Chapter 5, I conducted a series of interviews and a focus group with people with MS, carers/family members, and MS charity volunteers also living with MS, and three other researchers interviewed clinicians working in MS and healthcare commissioners. We developed a theoretical framework which was iteratively refined. The aims of this study were to explore previous experience with cognitive impairment and current assessment in MS, and the context of how to deliver a cognitive assessment programme. We asked participants about the types of assessment, the resources required, the modality of assessments, and the benefits of such a programme. This identified barriers and facilitators to implementing cognitive assessment in clinical practice.

Where a complementary approach of health research and HCI methods have been used for digital health intervention development, its application has potential for digital assessments. The evidence gathered from the systematic review and consensus recommendations alongside a person-centred approach can develop groundwork for a successful digital assessment programme.

Based on the conclusions from previous stages, prototypes of assessments deemed to be acceptable and feasible were developed that could be deployed on computers, tablets and smartphones. My next study in Chapter 6 used HCI methods to collect further qualitative data to evaluate the usability, accessibility, and acceptance of the technology. Usability testing can provide insight into the problems that people with MS may encounter with the technology, and the self-administration nature of digital, remote cognitive assessment.

Pilot interviews were first conducted with Patient and Public Involvement (PPI) members to refine a cognitive interviewing procedure and followed by usability interviews with people with MS through video conferencing, with participants using their own mobile devices. Due to

this methodology, the study used a modified think-aloud approach to cognitive interviewing. Data were analysed using a change log to capture iterative changes to the technology, and a framework analysis to capture mechanisms such as facilitators and barriers to engagement with the digital screening platform.

A final study was planned that aimed to include a further test, validate the battery, and collect normative data. Unfortunately, due to the COVID-19 pandemic, the technology development was delayed significantly and therefore it was not feasible to conduct this study. A protocol is included in this thesis as Chapter 7.

Chapter 2 Cognitive Measures Used in Adults with Multiple Sclerosis: A Systematic Review

2.1 Chapter overview

This chapter presents a systematic review that was conducted to scope the breadth of cognitive assessments used with people with MS and appraise the most commonly used individual tests and batteries. Aspects of this chapter has been published in Neuropsychological Rehabilitation (Elwick et al., 2021).

2.2 Introduction

Cognitive impairment can have an adverse effect on many aspects of an individual's life, including the ability to work and socialise (Rao et al., 1991), treatment adherence (Bruce et al., 2010), and overall quality of life (Glanz et al., 2010). As discussed in section 1.3 in the introduction predominantly impaired cognitive domains are memory and learning, (including immediate and delayed recall and recognition), complex attention including information processing speed, and executive functions (Chiaravalloti & Deluca, 2009; Ferreira, 2010; Prakash et al., 2008). A recent recommendation has supported domain-specific management of cognitive problems as opposed to more generalised approaches (DeLuca et al., 2020), and novel and improved methods of rehabilitation require appropriate outcome measures to determine their efficacy. Valid and reliable cognitive tests can additionally be used to identify cognitive issues at an early stage of MS, to assess progression of cognitive impairment.

Reviews by neurologists may not identify specific details of cognitive functioning and their significance (Romero et al., 2015), and appropriate, standardised cognitive measurement tools for use with people with MS are therefore required to detect changes to cognitive function accurately. A recent narrative review identified a need for

incorporating neuropsychological assessment into routine care for people with MS (Sumowski et al., 2018). However, due to the high prevalence of cognitive impairment in MS and limited resources in clinics, it may not always be feasible to conduct a full neuropsychological assessment for all patients at routine clinic visits. Therefore, there is a need for both comprehensive test batteries and shorter tests that can reduce the resources needed, including the level of expertise required to use and interpret results. Tests also need to be acceptable to clinicians and patients.

A survey of North American clinical neuropsychologists (Rabin et al., 2005) reported that the most frequently used measures across different clinical populations were Wechsler Adult Intelligence Scale (WAIS), the Wechsler Memory Scale (WMS), the Trail Making Test (TMT), and the California Verbal Learning Test (CVLT). However, as this survey did not specify tests used with different neuropsychological conditions, it is not possible to determine whether this reflects what psychologists are using with people with MS in particular. A more recent UK-wide survey (Klein et al., 2019), which specifically focused on cognitive assessments for people with MS, found that the most commonly used measures by healthcare professionals working clinically with people with MS (including neuropsychologists, neurologists, and occupational therapists) are the Montreal Cognitive Assessment (MoCA) and the Addenbrooke's Cognitive Examination (ACE-R). Generic measures may measure cognitive domains not commonly impaired in MS, such as orientation. Conversely, they might not include measurements of commonly affected domains, such as the lack of a measure of information processing speed in the MoCA. These studies show that generic measures are prevalently used in clinical practice, and it is unclear why some measures are chosen over others.

In addition to generic measures that are not specific to MS (such as the MoCA), a number of MS-specific batteries of cognitive measures have been developed, and both comprehensive and brief screening batteries exist, as explored in Section 1.4.1. Recommendations have

promoted the use of the proprietary Symbol Digit Modalities Test (SDMT) as a brief, sensitive measure processing speed impairment in people with MS, which is one of the most predominantly impaired domains (Kalb et al., 2018). The recommendations for cognitive measures, screens and batteries discussed above have used consensus approaches. These used expert panels comprising neuropsychologists, neurologists, clinical psychologists, researchers and Patient and Public Involvement to make recommendations, including that of full batteries (MACFIMS, Benedict et al., 2003), and more brief assessment (Kalb et al., 2018; Langdon et al., 2012). All approaches stated the need for full neuropsychological batteries to be administered for further insights if needed but highlighted the necessity for screening tools at a minimum, to promote routine assessment.

Following on from consensus recommendations, it would be useful to know which cognitive assessments are being used in research and clinical practice with people with MS. This might influence future consensus recommendations. There are ongoing concerns whether MS-specific batteries and tests (i) are able to evaluate the range of cognitive domains impaired in people with MS, particularly executive function, processing speed and memory (Hansen & Lautenbacher, 2017), and (ii) may lack ecological validity (Korakas & Tsolaki, 2016). Furthermore, in some healthcare and research settings, the use of proprietary measures may not be practical due to the financial implications of using such measures.

The aim of this review was to systematically compile and describe the neuropsychological measures that have been used in research with people with MS, and to examine the modality of administration and contexts in which they have been used. Details of the most common tests and multidomain batteries are provided. This review is the first step in determining an evidence-based approach to selecting measures for use in cognitive assessment with people with MS and complements existing consensus-based approaches.

2.3 Methods

The study protocol was prospectively registered with PROSPERO International Prospective Register of Systematic Reviews (Registration number CRD42018103384).

2.3.1 **Search strategy and selection criteria**

The systematic review followed the guidelines for the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA, Moher, 2009). We searched the following databases: MEDLINE, EMBASE, CINAHL, PsycINFO and Web of Science, for studies published in English from inception to February 2020 and using the relevant key words including, but not limited to: Multiple Sclerosis, Cogniti*, Neuropsych*, Screen, Assessment, Test, Outcome. I developed the search strategy for all databases, which was then independently reviewed by four members of the supervisory team[1]. The database searches were limited to adult, human studies, and texts in the English language.

Studies were included if they used psychometric measures to assess cognitive functioning in MS or screen participants into a study. We included measures that were performed in-person, and measures which were administered electronically or via telephone. Studies were excluded using the following hierarchy:

1. Not in English
2. Non-Human studies
3. Reviews, protocols, qualitative studies, case studies, translations, and validation studies.
4. MS patients do not constitute 50% or higher of the population included.
5. Population is under 18

[1] I use the term ‘we’ in this thesis where other researchers were involved in the data collection, analysis, or interpretation processes. In line with the ‘Team Science’ approach to research, use of this term (we) acknowledges the team’s input into this research.

6. Does not use a measure of cognition

The target population for studies was people (aged 18 and over) who had been diagnosed with any form of MS. We included individual measures, screening tests, and batteries of cognition, measures of single and multiple domains of cognition, and both subjective and objective measures of cognition. There were no constraints for the setting in which the cognitive assessment took place.

I completed the initial title and abstract screening, and full texts were divided between myself and another reviewer for data extraction - given the volume of the literature. Any disagreements were resolved by discussion and arbitration with a third reviewer. Studies were excluded based on the hierarchy shown in Figure 3.

Whilst we initially planned to conduct an assessment for quality and risk of bias using Critical Appraisal Skills Programme checklists (as detailed in PROSPERO), a formal quality assessment was not conducted. This was due to the absence of a single appropriate assessment for all studies, and the impracticality of undertaking this given the considerable number of studies included upon extraction. As we aimed to capture the breadth of measures being used and their frequency of use in a variety of contexts, including a range of research methodologies, the quality of studies would have been difficult to compare, and this would therefore have been of limited use. However, we did capture pertinent issues, including whether studies specified the names of measures used, or how measures were administered.

Despite restricting the searches to adult and English language papers, some papers that did not meet these criteria were still present in the citations extracted and were therefore removed in cases where two or more publications reported results of the same dataset, only the most recent study was included for data extraction.

2.3.2 **Data extraction and analysis**

The main interest was to identify the measures of cognition being used with people with MS, including how the measure was used (e.g.,

as a screening measure, or as an independent or dependent variable), and the modality the measure was administered in (e.g., in-person, electronically). We also extracted information about the participants' type of MS, the country in which the study took place, if it was a single centre or multicentre study, how many respondents completed the measure, and if it was a patient- or proxy-reported measure. For the purpose of analysis, the different measures were characterised into a number of classes for tests so that we could determine what measures were used most often, despite the names and ascriptions given. Examples of the classes used are serial addition tests in which respondents add up numbers as they are presented (e.g., the Paced Auditory Serial Addition Test (PASAT)), substitutional tests, where respondents are presented with an array of stimuli (usually digits, symbols, or letters) and asked to match them correctly with similar stimuli using a key (e.g., SDMT). The classes, named examples of these, and a description of the class are described in Table 1. These classes were determined through an assessment of how the assessments were described conceptually in relevant literature, or as described within the study. Once all individual assessments used in studies were classified, total numbers and frequency of use was calculated for each class, or multidomain test including batteries.

2.4 Results

The searches produced a total of 11,854 papers to be screened from which data from 1,526 studies were extracted for analysis. The PRISMA flowchart is shown in Figure 3. An excerpt from the data extraction format is shown in Table A in the appendix. Full-text review was conducted by two reviewers on a subsample of randomly chosen (n=50) manuscripts, to ensure uniformity of data extraction. We found a high level of agreement qualitatively.

The studies were conducted in 52 countries (50% conducted in Europe, and 39% from North America), and included 29 multicentre studies. Included studies collected data from 316,053 people with MS. All phenotypes of MS were reflected in the studies; most of the studies

(66%) used a mixed sample of people with MS, and 29% of studies enrolled people with Relapsing Remitting MS (RRMS). Only 4% of the studies specifically enrolled people with progressive forms of MS.

There was a total of 5,665 cognitive measures used. Of these measures, 93% (n=5,260) were ‘objective’ measures (for example, memory or processing speed tests administered by a researcher or clinician), and 7% (n=370) were ‘subjective’ (for example, self-report or proxy questionnaires). Thirty-five studies did not specify the names or provide sufficient description of the measures or procedures they used to classify them. However, some of these unnamed measures did report whether they were objective or subjective and were included in the overall counts. Approximately 18% (*n*=912) were named batteries or screening tests of cognitive functions consisting of two or more individual assessments (for example, the MoCA is a brief screening test but includes a trail making task, a verbal fluency task, word list learning among others).

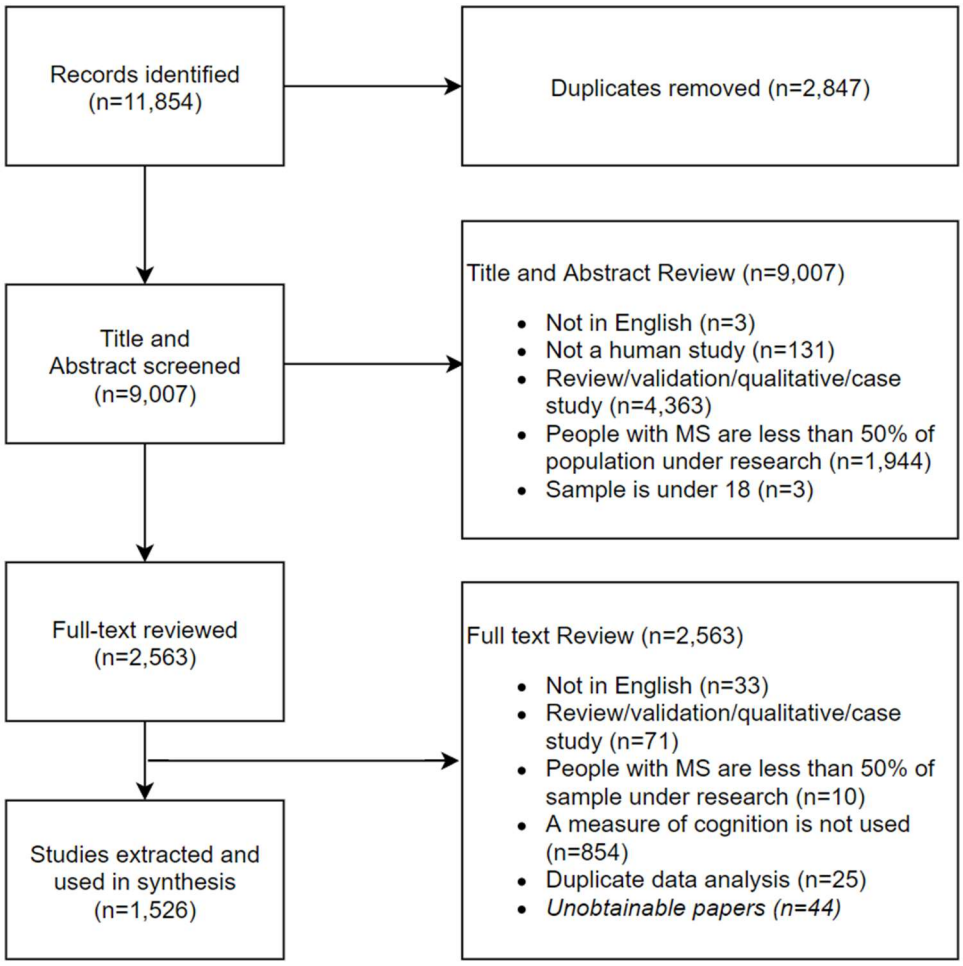


Figure 3: Prisma flowchart for systematic review

2.4.1 **Individual tests of cognition**

There were 504 different individual measures used, however many of these overlapped considerably. For example, verbal fluency tests that used a variety of stimuli, such as different letters of the alphabet, were given different names, including the Controlled Oral Word Association Test (COWAT), F-A-S test, or letter fluency tests.

Table 1: Individual tests used in MS research

Class	Examples of Tests	Number of Uses	% of Uses
ST- Substitution tests	Symbol Digit Modalities Test (SDMT (Smith, 1982), Digit Symbol (DS), Letter Digit Substitution Test	511	11.80%
SAT – Serial Addition Test	Paced Auditory Serial Addition Test (PASAT), Paced Visual Serial Addition Test (PVSAT)	476	11.00%
WLL - Word List Learning	California Verbal Learning Test (CVLT), Hopkins Verbal Learning Test (HVLT), Rey's Auditory Verbal Learning Test (RAVLT)	438	10.20%
VFT - Verbal Fluency Test	Controlled Oral Word Association Test (COWAT), or Word List Generation (WLG), FAS test	325	7.50%
ST - Stroop Test	Color and Word Test	242	5.60%
TMT - Trail Making Test	Colour Trails, Oral Trails, Making. Trails Test, Letter Number Sequencing	232	5.30%
DS - Digit Span	Digit Span, Forward Span, Backwards Span,	207	4.80%
ST - Sorting Test	Wisconsin Card Sorting Task (WCST) Modified Card Sorting, Weigl Test	164	3.80%
VR – Visual Reproduction Test	ROCFT - Rey-Osterrieth Complex Figure Test (ROCFT), Brief Visuospatial Memory Test (BVMT), Figural Learning Tests, Taylor Complex Figure (TCF)	120	2.80%

The most frequently used measures were the substitution-based tests (12% of all individual tests) such as the SDMT, and serial addition tests (11%) such as the PASAT, followed by tests which require participants to learn a list of words and recall them (immediately and/or after a delay; 7.8%) and verbal fluency tests (7.5%), such as the COWAT, in which participants are asked to name as many words that belong to a certain category (such as animals), or words beginning with a specific letter of the alphabet (see Table 1). The domains of cognitive impairment measured by the most frequently used measures relate to those experienced in MS, including information processing speed, attention, executive function, and aspects of memory including working

and short-term memory. Description of the tests, the domains they assess, and their advantages and disadvantages are provided in Table B in the Appendix.

2.5 Multidomain Tests and Batteries

Measures were operationally defined as a named collection of tests if they were a named screening test (e.g., MoCA) or a battery (e.g., BRB-N) that included one or more measure that is considered an individual test. Sixty-one different multidomain neuropsychological tests or batteries were reported in the included studies, and these were used 913 times in included studies. The most frequently used multidomain test used was the BRB-N, followed by the Mini–Mental State Examination (MMSE) and the WAIS (See Table 2). Description of the multidomain tests and batteries, the domains they assess, and their advantages and disadvantages are provided in Table C in the appendix. The Multiple Sclerosis Functional Composite (MSFC) only includes one neuropsychological measure (the PASAT) and was therefore included as an individual measure rather than a battery. The MoCA was frequently used to screen participants into studies (37%), or to report on cognitive functioning or between group differences (38%). However, 25% of the studies using the MMSE used it as a dependent or independent variable.

Consensus recommendations have advised the use of the MACFIMS and BICAMS as cognitive batteries for use with people with MS. The MACFIMS was recommended by consensus in 2003 and the BICAMS in 2011. The BRB-N was developed in 1990 and was used 10 times before the recommendation of the MACFIMS in 2003. Since then, the BRB-N has continued to be used frequently, with 217 studies using it, compared to 72 studies using MACFIMS. Even since the publication of BICAMS, the BRB-N has been used regularly, being used in 151 studies, whereas the BICAMS was used in 62 studies.

Table 2: Multidomain tests used in MS research

Abbreviation	Name	Type	Number of Uses	% of Uses
BRB-N	Brief Repeatable Battery of Neuropsychological Tests	Battery	210	23.00%
MMSE	Mini-Mental State Examination	Screen	148	16.20%
WAIS	Wechsler Adult Intelligence Scale	Battery	86	9.40%
MACFIMS	Minimal Assessment of Cognitive Function in MS	Battery	72	7.90%
WMS, WMS-III, WMS-IV	Wechsler Memory Scale, Wechsler Memory Scale Fourth Edition	Battery	52	5.90%
BICAMS	Brief International Cognitive Assessment for Multiple Sclerosis	Screen	62	6.80%
MOCA	Montreal Cognitive Assessment	Screen	30	3.30%

2.5.1 **Subjective measures of cognition**

A breakdown of the subjective, patient (or proxy) reported measures and their frequency of use is shown in Table 3. There was a total of 370 uses of named subjective measures in the studies, 87% were completed by the people with MS, and 11% were proxy reported (8% were carer/family reported, and 3% were clinician reported). Eighteen studies did not specify who completed the subjective measure of cognition. The most frequently used subjective measures were the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ), the Perceived Deficits Questionnaire (PDQ), and Fatigue Scale for Motor and Cognitive Functions (FSMC). There were 29 studies that used a self-report measure that was not reported by name, or not a formally named measure (for example, a Likert scale or visual analogue scale of cognitive functioning not attributed to a specific measure).

Table 3: Subjective measures used in MS research

Subjective Measure Name	Frequency of Uses	% of Uses
MSNQ - Multiple Sclerosis Neuropsychological Questionnaire	77	20.50%
PDQ - Perceived Deficits Questionnaire	34	9.00%
Self-report questions – not specified	*29*	*7.70%*
FSMC - Fatigue Scale for Motor and Cognitive Functions	21	5.60%
NeuroQoL	16	4.30%
FAMS - Functional Assessment of MS	14	3.70%
Global Impression Scale - Cognition	12	302%
EMQ - Everyday Memory Questionnaire	10	2.70%
MUSIC - Multiple Sclerosis Inventory of Cognition	9	2.40%
DEX - Dysexecutive Questionnaire	8	2.10%

2.5.2 **Uses of cognitive measures**

Cognitive measures were used for myriad reasons, and these were grouped into four categories (see Table 4). Many of the measures were used for two or more purposes in the study (for example to screen participants into the study, and as a dependent variable). Cognitive measures were predominantly used to report on cognitive functioning or as a between-groups measure in people with MS, and as a dependent variable.

There were 27 studies that used measures in ways that did not fall into these predefined categories, and these were recorded separately. Twelve were recorded as using the cognitive measures in a training programme for cognitive rehabilitation for people with MS, and other uses included matching healthy controls with people with MS, power calculations, and to derive the minimal clinically important difference. Some studies did not report on the cognitive measures they administered.

Table 4: Frequency of uses of cognitive measures

Use of cognitive measure	Frequency of Uses*	% of Uses
As a screening tool for inclusion/exclusion into a study	415	5.20%
To report on cognitive functioning in people with MS, for example as a between groups analysis	3757	47.03%
As a dependent variable	2305	28.86%
As a predictor, independent or correlational variable.	1484	18.58%
Other	27	0.34%

*Some measures were used for two or more purposes in a study

2.5.3 **Administration of cognitive measures**

Where available, we extracted data on how the cognitive measures were administered. Approximately half of the 5,670 measures used in the studies were not described in enough detail to extrapolate how they were administered. Of the remaining 2,680 measures, the majority were administered in-person (87%) by a healthcare professional, technician, or researcher, and 11% used a computerised administration. Remaining modalities consisted of postal and telephone administrations. The 53 measures completed by post were all subjective, self or proxy-reported measures, however, the 12 telephone administered measures were a mix of subjective and objective measures, including verbal fluency tests, word list learning and digit span tests.

2.6 Discussion

We identified a wide range of different individual cognitive tasks, screens and batteries that cover several cognitive domains, and despite recommendations for the use of specific measures in research with people with MS, their use is still far from being ubiquitous. Recommendations have focused on minimum cognitive screening for all people with MS, given the high prevalence of impairment (Kalb et al., 2018; Korakas & Tsolaki, 2016). Assessment should be able to identify impairment at an early stage, and becoming a part of routine monitoring. It is therefore important for assessments to be relevant,

available and accessible, and acceptable to both people with MS and their clinicians.

Only 34% of the studies that used a neuropsychological multidomain test or battery used one specifically developed for people with MS. Two of the most frequently used of these are screening assessments which are not specific to MS, the MMSE and the MoCA, which have been criticised for floor and ceiling effects, and therefore not being sensitive or specific enough to fully identify and reflect changes in cognitive impairment in people with MS (Beatty & Goodkin, 1990). The MMSE does not include measures of processing speed or executive functioning and is not considered an adequate assessment for use in MS clinical practice (Tobin, 2019), particularly due to inability to detect mild cognitive impairment.

Consensus-based recommendations for use of cognitive tests in MS have stated the need for both brief screening tests with the availability of comprehensive neuropsychological assessment in MS where needed. Quick screening tests of cognition (MMSE, BICAMS and MoCA, which take less than 30 minutes to complete), and longer assessments were reflected in the results, with eight batteries taking between 30 and 100 minutes to administer. While longer assessment may not be practical due to the high prevalence of fatigue in MS (Hadjimichael et al., 2008) and the limited time available in clinics, the MS-specific batteries are important to capture a full neuropsychological evaluation of cognitive impairment.

The BICAMS has been translated into several languages since being recommended, however the extent of the validations is limited due to its recency, and a lack of development of translations and linguistic validation could mean that uptake of this battery has been limited. Conversely, the availability of the BRB-N in European countries, shown by the range of published translations and validations, is likely to make it more appealing to researchers and perhaps explain its high use in Europe (76% of uses of the BRB-N).

An expert consensus recently recommended the use of the SDMT as a minimum in people with MS (Kalb et al., 2018), and the high volume of use of this measure suggests that it is both acceptable and an adequate measure of processing speed for people with MS. The SDMT has been shown to have good validity and reliability (Benedict et al., 2017), and measuring one of the most frequently affected cognitive domains, that of information processing speed. There is some evidence that the results of the SDMT are related to factors affected by cognitive impairment, including employment status and activities of daily living (ADLs) in people with MS (Benedict et al., 2005). However, the extent of this association has yet to be systematically reviewed. In the studies included in our review, there were several substitution style tests used with different names (for example, the Symbol Search Test; the Digit Symbol Coding Test, Coding Test), which may have been chosen over the SDMT due to cost or licencing requirements, over the psychometric properties. The omission of information about the administration of substitution style tasks in many of the included studies also demonstrates the importance of including pertinent issues about the outcome measures chosen. The use of variants and different administrations of these measures could mean that the generalisability of individual study findings is brought into question unless cross-validation has been considered. Norming data for the SDMT would be of limited value for similar tasks or different administrations, and so researchers and clinicians may use invalid comparisons to a normative sample.

The high usage of serial addition measures in research is unsurprising, given the inclusion of the PASAT in the MSFC. However more recent endorsements have moved away from using this measure, owing to its potential for inducing anxiety (Aupperle et al., 2002), stress and frustration (Locke et al., 2011). Results of this task can be affected by IQ, and mathematical ability which cannot be corrected without further and extensive data collection (Wills & Leathem, 2004). It is also likely that using the PASAT would lead to a high proportion of missing

data due to participants refusing to undergo or complete the task. Participants who do not wish to complete this task have been shown to be associated with higher levels of cognitive impairment (Cortés-Martínez et al., 2019), so using the PASAT could lead to significant missing data from some of the most cognitively impaired. These studies affirm the importance of choosing measures that are acceptable and accessible to people with MS.

Subjective measures of cognitive impairment in the review made up approximately 7% of the measures. Research has found that self-report measures of cognitive impairment in people with MS can show stronger associations with measures of depression than performance on objective cognitive measures (Benedict et al., 2003), however proxy-reported measures tends to correlate higher with cognitive measures and may be valuable in assessing cognitive functioning (Benedict et al., 2004). Despite this, only 11% of subjective measures were proxy-reported, and it is possible that researchers and clinicians do not trust subjective reports of cognitive impairment. However, subjective measures can overcome many of the disadvantages of objective measures. Self- and proxy-reported measures can be less onerous on the participant and administrator, have fewer variables to control, may be more acceptable to the participant, and could reflect a global aspect of cognition that is distinct from that assessed in other ways.

Just under half of the studies had no administration details of the assessments used (for example, verbal, written, or electronic), and thirty-five of the cognitive measures were not described well enough to ascertain the name, or what class of measures was used. Replication of these studies would be challenging. There was also considerable difficulty in classifying many of the measures due to the lack of information given in the papers, and this may have had an impact on the frequency of the classes and batteries. Both visual acuity and motor dexterity can be impaired in MS (Lau et al., 1998), and cognitive assessments often involve the visual processing of stimuli, with participants giving motor (for example, written) or verbal responses.

The written administrations of assessments can be affected by manual dexterity issues in MS, and while the verbal administrations are often used to overcome this, such as the oral administration of the SDMT, studies often did not specify this. Whilst some studies excluded participants with problems with vision and dexterity, many did not report how such issues would be overcome for the assessments. The missing information would make replication of these studies difficult.

Our finding associated with the frequency of use of tests should be treated with caution. This is based on simple counts, and the selection criteria for this review, which excluded translation and validation papers, may have meant that some papers were not counted here. Another limitation of this review is the missing information on the measures that were extracted for each study. There were some aspects of the use of cognitive measures in research that were outside of the scope for this review, such as the domains purported to be measured in the studies. Future search strategies could include specific cognitive domains and the names of measures, but this may inflate the number of articles that will need to be searched.

A formal quality assessment was not conducted as we wanted to capture the breadth and context of cognitive measures. An unexpected finding was that, for many studies, we were unable to ascertain what measures were used, or how the measures were administered, for example, data regarding the types of electronic devices used, or where the respondents completed the measure was missing in many primary studies. Devising a quality assessment checklist for future similar reviews may add further insight into this replicability issues. A further limitation of this review was that we only included papers that were published in English. Research conducted and published in other languages may have provided pertinent information regarding measures that are used globally. Additionally, whilst we specified in the PROSPERO registration that grey literature would be searched, initial searches were not found to be relevant and therefore these were excluded from the systematic review.

It was clear from the results of the systematic review that there are many cognitive measures being used in research with people with MS, however some, for example substitution and serial addition tasks, are prominent. The measures used ranged from short, individual screening tests to lengthy batteries that require trained staff to administer, and subjective measures that could be completed by the people with MS, or a proxy. Cognitive measures were administered in a variety of modalities. One of the impacts of COVID-19 has been a shift towards telemedicine, and for neuropsychology to respond to this shift, research is needed on how cognitive assessments perform in remote administrations. This shift will likely be long lasting, beyond the COVID-19 pandemic. Evidence of validity and reliability for computerised cognitive tests and batteries for people with MS is emerging (Wojcik et al., 2019), but these do not guarantee equivalence between all potential modalities (for example, self-administered vs. technician administered; keyboard and mouse vs touchscreen), and new normative data may be required (Bauer et al., 2012). It is recommended that researchers and clinicians ensure that the measures used in all settings are appropriate to measuring cognitive impairment in MS, and being valid and reliable.

More research is needed to determine the acceptability of MS-specific tests from clinicians' and patients' perspectives, which may increase the use of and response to specific tests. Computerised adaptations especially require further evidence of their psychometric properties, and acceptability to patients. Systematic reviews are also needed to examine the association between the cognitive tests and functional outcomes, to improve clinicians’ and patients' engagement with these tools.

2.7 Conclusions

The SDMT is the most frequently used individual assessment, and is recommended as a minimal test for cognitive impairment in MS (Kalb et al., 2018), due to its brief administration, good psychometric properties (Benedict et al., 2017; Strober et al., 2018), and acceptability to people with MS (Mortensen et al., 2020). It is therefore unsurprising

that it is frequently used and forms part of many of the batteries used in MS research. Based on the evidence of the SDMT's utility, I have examined how the SDMT, and similar symbol digit tests can be used in MS clinical care. This includes a quantitative analysis of the results of the SDMT, and the development of a digital version which forms the latter part of the thesis. Prior to this, it was important to explore how one of the practical aspects of using the SDMT that is, how to score it to enable meaningful interpretation within clinical care. As found in the current chapter, studies often do not state key details of the tests chosen, and how the results of tests have been interpreted. This presents difficulties for the generalisability of results, and may make using such tests in clinical practice difficult. Whilst the manual for the SDMT (Smith, 1982) details normative data, this may be outdated or not applicable.

The range of interpretation approaches, and the normative data available, may impact uptake of a test, and therefore the next study examined how scoring and interpretation have been used, and how these methods have been used to categorise impairment.

Chapter 3 Variability in Administration, Scoring and Interpreting Symbol Digit Modalities Test Results in Multiple Sclerosis

3.1 Chapter overview

This chapter describes a scoping review, which aimed to explore and outline the different methods of how the Symbol Digit Modalities Test (SDMT) has been interpreted in multiple sclerosis (MS) research. I present some of the challenges related to interpreting the SDMT and raise questions around generalisability of findings and future implementation.

3.2 Introduction

The Symbol Digit Modalities Test (SDMT) is a brief measure used to assess information processing speed, memory, and lexical access speed (Sandry et al., 2021). As explored in Chapter 2, it is the most frequently used individual test in MS research, and the US National MS Society advocates its use for cognitive screening and monitoring within MS clinics (Kalb et al., 2018). There are multitudinous advantages to using this individual test, including its acceptability to people with MS (Mortensen et al., 2020) ease of administration, and it is sensitive to identifying cognitive problems amongst people with MS (Pitteri et al., 2019). It is also included in many MS-specific batteries, including BICAMS, MACFIMS and BRB-N (as shown in Table C). Reviews of the use of the SDMT have found evidence as to its validity, reliability, and clinical relevance for MS (Benedict, et al., 2017). On the basis of the recommendations and potential for its use in MS, I have chosen to examine the SDMT in greater detail in this thesis, in order to ascertain how it can be used in clinical practice.

The user manual details standardised instructions for delivery and scoring (Smith, 1982). However, whilst considerable variability exists in terms of how the SDMT is used, scored, and interpreted, discussion of these methodologies is lacking in the literature. This variability creates

barriers for clinical decision making and limits the replicability and generalisability of research findings, in turn impacting the quality of care provided to people with MS for their cognitive problems.

3.3 Aims

The aim of this scoping review was to explore and outline the different approaches of how the SDMT has been interpreted in MS research. Methods in administration, scoring and interpretation of the SDMT apparent in the literature are described.

3.4 Methods

This scoping review was to rapidly synthesise evidence on the administration, scoring and interpretation of the SDMT in MS. I developed the search strategy which was reviewed by the supervisory team. Searches were limited to published and indexed articles. PubMed and Google Scholar were searched in July 2020 for the Medical Subject Headings terms “multiple sclerosis, AND symbol digit modalities test AND cut-offs OR thresholds”. Studies were considered if they administered the SDMT with people with MS, used either verbal or pen-and-paper administration, and used and described a method of interpretation of the results of the SDMT. Papers were not limited if they used the test as part of a battery or as an individual test, or by how they used the SDMT. Digital versions of the SDMT were excluded from consideration as there is no formally available version of this test. Relevant papers were read, and common interpretation methods were extracted from a range of relevant studies, and the team discussed and formulated the methodologies used to categorise scores. In this scoping review, we present the delivery format of the SDMT, if scores were adjusted using normative data, and the source of normative data. This was not a systematic review of the various aspects of the SDMT, so we have selected studies that illustrate each one of the methodologies described in this chapter (for example, z-scores, percentiles, statistical comparison). Each highlights the heterogeneity

we describe, to show single examples of each methodology we found that interpreted neuropsychological results of the SDMT."

3.5 Results

Table D in the appendix provides examples through the inclusion of eleven studies that have used the SDMT in the context of MS. Extracted data include study design, SDMT format, scoring, source of normative data (if used), any reported cut-offs, and the categories of classification (if given).

3.5.1 **Administration and scoring**

One source of variability is related to how the SDMT is administered. The SDMT can be completed as part of a battery (in-table references 1, 2, 3, 4, 6 and 7) such as the Brief Repeatable Battery of Neuropsychological Tests (BRB-N), as an individual test (8, 9, 10 and 11), and combined with another test (5). The most common scoring convention for the SDMT is the number of correct responses in 90 seconds (Smith, 1982). However, verbal administrations differ from written responses, and scoring cannot be compared across formats: the time taken to respond to stimuli will differ between writing and saying a number aloud. Written responses may be further complicated by dexterity problems faced by people with MS. Reliance on raw SDMT scores, such as impairment defined by scores below 40 (9), is convenient, but does not consider demographic differences (e.g., age and education attainment), which may need to be adjusted for.

Another challenge is the lack of consistency and reporting of the scoring and norming of the SDMT. Scoring traditionally focuses on the number of correct responses within 90 seconds, but longer administrations are sometimes used, and scoring can also include how many errors were made. However, such information is not presented in some studies. Indeed, nine of our example studies did not report how the SDMT was scored.

Scores can be compared to a normative mean to determine whether they are suggestive of an impairment. The manual provides

normative data from 1307 adults collected in the USA in 1975. Scores that fall 1 to 1.5 SDs below the normative mean are suggestive of impairment (Smith, 1982). However, the generalisability of these normative scores is likely to be limited because of the Flynn effect (Dickinson & Hiscock, 2011). Higher neuropsychological test scores are often observed over time, potentially resulting in norm obsolesce (Trahan et al., 2014); and by the size and diversity of the sample (Sheridan et al., 2006).

The example studies in Table D in the appendix highlight three key approaches identified in the studies to scoring and interpreting results of the SDMT:

1. Raw scores: descriptive statistics, such as mean, SD and range of scores, are reported and significant differences and/or correlations analysed (7).

2. Standardised scores: Z-scores depend on normally distributed data and are calculated by evaluating the difference between an individual's raw test score and the distribution mean, divided by the distribution standard deviation (SD). Z-scores can be adjusted for demographic differences where normative samples are large enough. Participants were considered cognitively impaired if their score was 0.5 to 2SDs below the mean (1, 3, 4, 5, 6, 8, and 11). Using demographically adjusted normative data, alternatively, T-scores have been used for smaller samples. One advantage to standardised scores is the conversion to cut-offs, enabling straight-forward, though simplistic, interpretation (e.g., 'impaired' or 'unimpaired').

3. Quantiles: Z-scores and raw scores can be transformed into quantile ranks of the normal distribution, such as percentiles, to express the percentage of raw scores that fall at or below the score under investigation, or tertiles, which divide the dataset into thirds. This interpretation can be useful for non-normally distributed or limited normative data to express how frequently a clinical score occurs. Cut-off scores have been suggested for percentiles, however their

application needs to be assessed depending on the administration and/or version of test used (Mitrushina et al., 2005). Compared to normative data, if participants score less than a defined centile, this suggests impairment. Studies can apply the threshold of below the fifth centile of normative data to define impairment (2, 9).

3.6 Discussion

Guidance from the SDMT manual (dichotomous categorisation of impaired or non-impaired) is frequently used to interpret scores. A raw score cut-off can be determined depending on the Receiving Operating Characteristics (ROC) deemed acceptable by the study; for example, a score of 35 (7) or 55 (8). Although most studies applied dichotomous categorisation (e.g., impaired or unimpaired), some studies categorised impairment into cognitive phenotypes through combinations of impaired tests; for example, three categories of memory and processing speed impaired (5), or further conceptual categorisation such as ‘mild’ and ‘severe’ cognitive phenotypes by using a latent profile analysis (3). Studies also stratified people with MS into one of multiple levels depending on z-scores but may not label these (4).

If normative distributions are non-normal, percentile ranks are more appropriate for interpretation. Where data on age, gender and education level are available, analyses should adjust for these factors, and regression techniques can be used for continuous variables (e.g., age, education level). Parmenter et al. (2010) first used a regression-based model of SDMT norms to examine this, and further studies have confirmed the reliability of this approach (Burggraaff et al., 2017).

Alternative approaches to interpreting SDMT scores, such as ROC and regression models have been applied to extend the classification of impaired/non-impaired by generating continuous norms (Van der Elst et al., 2013). These can be implemented where prevalence rates are known (ROC) and with sufficient samples for continuous demographic variables (regression). Adjustments made on the basis of these regression models can provide valuable insight and

nuances for SDMT performance, and improve sensitivity and specificity, but are not routinely applied. Applying regression-based norms to SDMT scores requires manual calculations that require expertise that may not be feasible in clinics. A digitised version of the Brief International Cognitive Assessment for MS, the iCAMS, used an a technician-administered iPad application to automatically score this battery (Beier, et al., 2019). Using applications such as these would allow for more precise, regression-based interpretation to account for demographically adjusted norms.

Frequently, authors do not state how they scored or interpreted the SDMT and omit key details that would enable replication, similar to the findings in Chapter 2. Reporting which modality of the SDMT was used, how it was scored, whether normative data were used, and the dataset used, and how researchers interpreted the scores, such as using standardised or quantile scores may enable researchers to identify available normative datasets and interpret the data with greater ease. This can also enable more precise comparisons between studies.

Future research could examine how the SDMT is currently being implemented within clinical settings, for example for brief screening/monitoring in MS clinics, or as a part of a detailed neuropsychological evaluation, and how interpretation could be facilitated amongst clinicians and people with MS. Recent research has begun to use regression-based models to adjust for continuous variables such as age and education level. However, calculating standardised scores with demographic variables such as age and education can be challenging in a time-pressured clinical setting. Digital administration affords the opportunity to build these calculations into the technology and to provide an easily understandable output.

COVID-19-related restrictions have underscored the importance of having valid and reliable neuropsychological assessments that can be completed remotely. Telephone administrations are valid (Eilam-Stock, Shaw, et al., 2021), but a fully digitised version of the SDMT would be valuable. Whilst there are several conceptually similar tests, an official,

licenced digitised SDMT has not yet been developed. Digital administrations are considered to be ‘new’ tests, and therefore need evidence of their validity, reliability, and new normative data. Different administration practices, sources of normative data, inclusion of demographically adjusted scores, and the chosen ROC might explain the heterogeneity in the range of cut-off scores and different methods of interpretation in the literature. Conceptually similar substitution style tests, including digital administrations exist, however scoring and interpretation should not be assumed equivalent between these. Therefore, using the same model of interpretation across modalities and versions is not possible. Normative data used should be specific to the administration procedures applied and the population under investigation. When comparing clinical scores to normative samples, normality, effects of demographic variables, and acceptable ROC need to be considered.

We advocate for further research that will enable the classification of dysfunction into further categories, such as mild, moderate, and severe; or phenotypes as exemplified by some studies (e.g., 3 and 5). These could highlight the severity of impairment, and the domains in which dysfunction occurs. Approaches such as these may help in triaging people with MS to receive appropriate further testing and/or cognitive rehabilitation earlier in their care pathway. This is particularly relevant in health services that do not have the resources to see all those deemed cognitively ‘impaired’

3.7 Conclusions

In this chapter, the methodologies for the scoring and interpretation of the results of this test have been explored, to document how research categorises cognitive impairment using this test. Through the systematic review in Chapter 2 and the current chapter, I found that studies often do not offer administration details of how assessments were used and scored, making replication of these studies challenging. This also presents difficulties in choosing and administering the SDMT in clinical care, despite it being the most widely

used in research, and a frequently recommended minimal assessment of cognitive impairment.

The aims of the other studies in this thesis are to evaluate how cognitive screening can be conducted in a clinical setting. Having established how the SDMT can be scored and interpreted in people with MS, the next chapter broadens the scope of cognitive screening, by examining how the SDMT correlates with self-reported cognition and symptoms in MS (including physical and psychological impacts). The aim of this analysis is to determine if self-reported cognition can be used to supplement or replace the SDMT, and to also examine if poor SDMT scores can predict other MS symptoms. In future chapters, I detail how one digitised symbol substitution based on the SDMT was developed and its validity evidenced, through stakeholder input, usability testing and quantitative support.

Chapter 4 Exploring the factor structure of a self-reported cognitive problems questionnaire, and variables associated with processing speed deficits in people with MS

4.1 Chapter overview

This chapter details a secondary data analysis of data obtained from a clinical trial evaluating the effectiveness of cognitive rehabilitation in MS. The Symbol Digit Modalities Test (SDMT) is recommended as a minimum assessment for people with MS, and as found in the systematic review, is the most commonly used cognitive assessment in MS research. This chapter examines how the SDMT relates to a subjective measure of cognitive impairment and MS symptoms.

The dataset used included the screening-in data from a wide range of people with MS who completed the SDMT and the patient version of the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ-P). The first analysis examined the underlying factor structure of the MSNQ-P to explore its potential multidimensionality, and the relationship between an objective and a subjective measure of cognitive impairment in MS.

A second analysis included data from a large sample of people with MS with cognitive impairment, who completed the SDMT and self-report questionnaires of fatigue, physical symptoms, and psychological symptoms at baseline for a clinical trial. Linear regression analyses were conducted to examine whether self-reported symptoms predict scores on the SDMT.

4.2 Introduction

As found in the systematic review in Chapter 2, the SDMT is the most frequently used measurement of cognitive impairment in MS. It is

included in the Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS; Langdon et al., 2012). A meta-analysis of the BICAMS found that people with MS had significantly lower scores on the SDMT than control participants, and the effect size of this was large (*g*=0.934), and greater than the other two tests in this battery (Corfield & Langdon, 2018). Qualitative research supports the use of the SDMT as a stand-alone test, through suggesting that people with MS find it relevant to their symptoms and the least demanding and fatiguing of the tests in the BICAMS (Mortensen et al., 2020).

Information processing speed is the most commonly impaired domain of cognition, and is often the first to be detected (Oreja-Guevara et al., 2019). Being clinically relevant, acceptable to people with MS, quick and easily administered, administering the SDMT would be valuable as a minimal, routine assessment. However, it is possible that some individuals may be unwilling to complete any objective tests and may prefer self-report questionnaires to report their cognitive impairment. Subjective reports may also provide insight into how people with MS experience and perceive their cognitive functioning. Clinical practice may routinely collect both objective and subjective measures to evaluate both, and provide an alternative measure for those who may reject the SDMT.

4.2.1 **Objective and subjective measures of cognitive impairment**

Research shows that people with MS who do not wish to complete an objective measure of cognitive impairment, such as the Paced Auditory Serial Addition Test, were more likely to be those with more severe cognitive impairment (Cortés-Martínez et al., 2019). The interview study conducted as part of this thesis (described in Chapter 5) revealed that some people with MS may be reluctant to complete objective cognitive screening measures, even those that were seen to be less anxiety inducing, such as the SDMT. A cognitive screening programme would need to consider that some people with MS may not be willing to complete any objective cognitive screening, therefore an alternative acceptable method of capturing cognitive impairment could

be offered. This can ensure that people with MS are not excluded from potential treatments for cognitive impairment, especially if they are more likely to show greater impairment. If some people with MS prefer to complete a self-report questionnaire about their cognitive impairment, there are several questionnaires that aim to evaluate this.

With a shift to teleneuropsychological services, pushed forward by the COVID-19 pandemic, some people with MS may lack the technology to complete objective cognitive assessments, and clinics may not have the necessary resources to administer their assessments remotely (Bilder et al., 2020). As explored by the development of a digital cognitive assessment platform in this thesis (described in Chapters 5-7), digital administrations require evidence of their validity, reliability, equivalence to their in-person modalities and the use of normative data for comparisons. While objective cognitive measures are perceived as more accurate in assessing cognition (Middleton et al., 2006), self-report questionnaires can potentially overcome the practical challenges of administering cognitive assessment remotely. There is more flexibility to administer these over the telephone, or they can be posted/emailed and returned by patients. Consequently, it would be of benefit to ascertain how strongly a frequently used subjective measure correlates with a frequently used objective measure of cognition, the SDMT, and whether it can be used as an alternative to the SDMT. Subjective measures may also provide additional insight into cognitive impairment in MS by examining the impact of impairment on Activities of Daily Life (ADLs) and quality of life; and an individual's awareness of their cognitive impairment, which is not captured by objective cognitive assessment.

The Multiple Sclerosis Neuropsychological Questionnaire self-report form (MSNQ-P [P for patient-reported, as the MSNQ also has an informant-reported version]) is a 15-item questionnaire which has been developed as a brief and reliable self-report of cognitive impairment, consisting of a single overall scale (Benedict et al., 2003). There has been mixed evidence of how well the MSNQ-P correlates with objective

measures. For example, in the original development, the MSNQ-P was found to correlate with self-report measures of depression, but there were no statistically significant relationships with the objective measures of cognition, which however, did not include the SDMT (Benedict et al., 2003). However, a follow up study with a larger sample (n=85) found a significant correlation between the SDMT and MSNQ-P (r=-0.45, p<0.001), although again MSNQ-P scores were correlated more strongly with depression scores (Benedict et al., 2004). This study also demonstrated a low specificity of the MSNQ-P, indicating a high likelihood of false positives caused by symptoms of depression. Further studies (Migliore et al., 2021; O'Brien et al., 2007) have also shown an inconsistent relationship between scores on the MSNQ-P and SDMT (e.g., r=0.03), although one study concluded that the MSNQ-P could still be a useful indicator of cognitive impairment or emotional distress (Morrow et al., 2010).

Assessing the results of the MSNQ-P and the SDMT with a large UK dataset could provide additional evidence of the relationship between these two measures and explore the suitability of offering a self-report questionnaire to those who do not wish to complete such screening. A large study (n=660) examining cognitive impairment in people with MS treated with natalizumab administered both measures monthly across 48 weeks, allowing for a retrospective analysis on the scores of the SDMT and MSNQ-P across time (Morrow et al., 2010). The authors found that scores on both measures showed stability and good test-retest reliability over time and concluded that both tests are acceptable measures to effectively assess cognition at different time points, however this study did not report the association between these two measures.

There have also been mixed results as to whether the MSNQ-P comprises a single, overall domain, or whether it has an aspect of multidimensionality to it, reported in some studies (Konstantinopoulou et al., 2018; Sonder et al., 2012). The original development (Benedict et al., 2003) used an *a priori* four domain structure (attention and

processing speed; memory; other cognitive ability; and personality and behaviour) to compile items, and did not conduct an Exploratory (EFA) or Confirmatory Factor Analysis (CFA) to ascertain whether these domains emerged as subscales from the data. A small number of studies have investigated the factor structure, with mixed results. Sonder et al. (2012), found two potential latent factors with eigenvalues over one, however resolved to use the MSNQ-P as a unidimensional scale as they were unable to replicate this structure using a parallel analysis. A more recent translation and validation study (Konstantinopoulou et al., 2018) of 103 people with MS in Greece found that a three-factor solution (prospective memory, memory and attention, and executive functioning and problem solving) explained greater variance than the suggested single overall factor (65% compared to 49%). It would therefore benefit from further exploration with a larger sample size to explore this potential multidimensionality, what conceptual domains the MSNQ-P may comprise, and how these relate to the original development's conceptual factors. Due to the inconsistent findings in small samples in translation studies (Konstantinopoulou et al., 2018; Migliore et al., 2021; Sonder et al., 2012), running a further Principal Components Analysis (PCA) on a larger dataset can explore whether the scale is multidimensional and examine the potential subscales within the MSNQ-P.

4.2.2 **Processing speed and self-reported symptoms of multiple sclerosis**

Fatigue is an ubiquitous symptom of MS, affecting 75-90% of people with MS (Kos et al., 2008), and in a study of health status, 69% of people with MS reported it as the worst or one of their worst symptoms. Fatigue can further impact mental health and general health (Fisk et al., 1994). People with MS who experience fatigue have been found to score significantly worse on the SDMT than those reporting less fatigue (Yigit et al., 2020). Negative correlations have been found between scores on the SDMT and the Fatigue Severity Scale (FSS; Krupp et al., 1988), indicating that greater levels of fatigue are

associated with slower processing speed). However the strength of these relationships tends to be weak; (*r*=-0.21, - 0.22; Drake et al., 2010; Hughes et al., 2019). Fatigue in MS has been reported to be positively related with age and duration of MS, and with progressive forms type of MS (Newland et al., 2016).

Similarly, higher levels of self-reported depression is associated with slower information processing speeds in people with MS (Landrø et al., 2004). One study (Schäffler et al., 2013) found a significant negative correlation between SDMT scores and total Multiple Sclerosis Impact Scale (MSIS-29; Hobart et al., 2001) score and the physical subscale (*r*=-0.32) but not the psychological subscale (*r*=0.19). The psychological scale includes items regarding mood, fatigue, and cognitive problems. Both subscales have been found to be significantly and positively associated with age, and duration of MS, and progressive phenotypes (Jones et al., 2013). Additionally, a study found that men report greater psychological and physical symptoms of their MS as measured by the MSIS-29 (Anens et al., 2014).

4.2.3 **Cognitive Rehabilitation for Attention and Memory in People with Multiple Sclerosis**

CRAMMS began in 2015 and aimed to evaluate the impact of group cognitive rehabilitation programme had on people with MS (Lincoln et al., 2020). The aim of the trial was to assess the impact of a cognitive rehabilitation to improve attention and memory in people with MS. Participants were recruited from five NHS trusts across England. Participants were eligible if they were aged between 18 and 69 years and had a diagnosis of relapsing-remitting or progressive forms of MS at least three months before the screening. An initial screening assessment included the MSNQ-P and the BRB-N.

Eligibility was determined by a score of >27 on the MSNQ-P and a score of more than one standard deviation (SD) below the mean of age and education adjusted control data on one or more test of the BRB-N.

The BRB-N includes the SDMT, which is used in the analyses described in the current chapter.

Two of the self-reported baseline assessments in the trial were the MSIS-29 and the FSS and these were collected again at 6 months and 12 months post-randomisation. The MSIS-29 is a 29-item self-report measure comprising two subtests: 9 items measure psychological impact, and 20 items measure physical impact of MS, and has been found to be reliable and valid (Hobart et al., 2004). The FSS was developed as a nine-item self-report questionnaire, however Rasch analyses have recommended a reduced version of 5 items (Mills et al., 2009), which was used in CRAMMS.

Of the people with MS who were screened into the study through completion of these screening assessments (*n*=579), 449 participants were individually randomised to receive their usual care and participate in group rehabilitation (*n*=245) or only receive their usual care (*n*=204). The group cognitive rehabilitation included restitution strategies and compensation strategies.

The primary outcome measurement was the MSIS-29 psychological subscale, which includes items relating to mood and cognitive impairment. Whilst there was a small difference on scores for the cognitive rehabilitation group at 6 months, this difference was not apparent at 12 months.

4.3 Aims

The aims were to:

1. Explore the factor structure of the MSNQ-P using an EFA, and the correlation between scores on the MSNQ-P and SDMT in a large sample of people with MS.
2. If multidimensionality is found, examine how these relate to SDMT scores.
3. Examine how self-reported cognitive impairment, physical functioning, psychological impacts of MS, and fatigue predict

objective information processing speed as measured by the SDMT, in people with MS considered to be cognitively impaired.

4.4 Methods

4.4.1 **Data**

Two datasets from the CRAMMS trial were used in this analysis. This analysis was covered by the ethical approval for the CRAMMS study, obtained from the National Research Ethics Service (West Midlands, South Birmingham Committee; reference number 14/WM/1083). The first dataset is from all participants who were screened (*n*=579). The second uses the data from participants screened into the trial who completed baseline assessments (*n*=479).

There was some missing data from both datasets, as some participants did not complete some or all the measures.

4.4.1.1 *Screening data*

Screening data included the following variables: MS type, gender, age, years diagnosed with MS, SDMT score, MSNQ-P item scores, and overall MSNQ-P score.

The MSNQ-P consists of 15 individual items. Each item has 5 response options, from ‘Never’ to ‘Very often’, which are given values from 0-4. The sum of these 15 items provides an overall score, from 0-60. Higher scores indicate greater reported cognitive impairment. The SDMT is scored by the number of correct responses given in 90 seconds.

Twenty participants did not complete the SDMT, and one participant did not complete the MSNQ-P in its entirety, therefore 21 participants were excluded from the analysis (*n*=558).

4.4.1.2 *Baseline data*

Baseline data were taken from participants who consented to take part in the trial, and met the inclusion criteria, which included participants considered cognitively impaired.

This dataset included the following variables: MS phenotype, gender, age, years diagnosed with MS, years of education, SDMT score, individual FSS scores (5 items), overall FSS score, individual MSIS-29 scores (29 items), and overall scores on the physical and psychological subscales of the MSIS-29.

Participants with missing data were removed from the analysis (*n*=48), leaving an overall sample of 431.

4.4.2 **Statistical analyses**

All analyses were conducted in IBM's Statistical Package for the Social Sciences (SPSS) 26th edition.

4.4.2.1 *Factor structure of the MSNQ-P*

A varimax PCA was used to determine the factor structure of a scale. The suitability of the data was assessed using the Kaiser-Meyer-Olkin (KMO) value, which showed the proportion of common variance in the variables (expecting a result of 0.5 or greater). The Eigenvalues of the data upon running the PCA would suggest whether there is multidimensionality and the number of factors, and a scree plot was drawn to confirm this. A parallel analysis was also run, which was to confirm that any factors were not being generated from random noise. A factor loading of 0.3 or greater in two or more factors implies multidimensionality. If the items load on one factor only, this implies that the scale is unidimensional. Given the *a priori* domains used to develop the questionnaire items and small studies finding two or three factors, I expected to find more than the single factor structure, which could explain greater variance within the model.

4.4.2.2 *Association between the SDMT and MSNQ-P*

A Pearson product-moment correlation analysis was then conducted to examine the relationship between the MSNQ-P and the SDMT. If the questionnaire did exhibit multidimensionality, I expected that correlations between each of the discrete factors and processing speed may indicate that one subscale was better at measuring cognitive impairment than others. I anticipated that any latent factors

found may suggest whether any one factor or subscale is more appropriate for measuring self-reported cognitive impairment. As one of the *a priori* factors used in the original development was “personality and behaviour” (Benedict et al., 2003), items loading onto a similar latent factor could explain the strong relationship between reported depression and MSNQ-P scores. If this is the case, I predicted that these items could be removed to investigate the use of a modified questionnaire that could be used as a cognitive assessment with people with MS.

4.4.2.3 *Association between SDMT scores and self-reported symptoms in cognitively impaired people with MS*

Directed acyclic graphs were constructed to identify which covariates to adjust for in each model. Categorical covariates (gender and MS phenotype) were transformed to dummy variables to include in the model. For each component, age, gender, duration of MS, MS phenotype, and years of education were included as covariates in each model. Using the enter method, SDMT was entered as the dependent variable, with overall component scores as the exposure variable. The differences in SDMT score by MS phenotype, duration of MS, gender, age, and years of education were assessed. Linear regression models adjusted for covariates were then run to examine whether any extracted components of the MSNQ-P, or fatigue, psychological, and physical symptoms predicted processing speed in people with MS. The models provide the percentage of variance explained in processing speed scores by self-reported variables and their significance at a 0.05 level.

4.5 Results

4.5.1 **Factor structure of the MSNQ-P**

4.5.1.1 *Descriptive statistics*

Of the 579 participants who completed the MSNQ-P and SDMT at screening, 21 participants with missing data were removed from the analysis The demographics for the remaining 558 participants are presented in Table 5. Data collected for age and length of MS diagnosis

was in years and collected when participants were screened into the trial between 2015 and 2017, and therefore the years of descriptive statistics are given at the time the participant was screened.

Participants who were screened were predominantly women (72.8%), and had a mean age of 49.53, ranging from 18 to 69. Most participants had a relapsing-remitting phenotype of MS (65.2%), and the remaining participants had progressive forms (34.6%). One participant reported benign MS. The average length of time since MS diagnosis was 11.48 years.

Table 5: Demographics of people with MS who completed the MSNQ-P and SDMT at screening

Age at screening (years)	*Mean (SD)*		49.53 (9.82)
	Min		18
	Max		69
Gender	*Woman*	*n (%)*	406 (72.8)
	Man	*n (%)*	152 (27.2)
MS phenotype	*Relapsing-remitting*	*n (%)*	364 (65.2)
	Secondary Progressive	*n (%)*	137 (24.6)
	Primary Progressive	*n (%)*	56 (10)
	Benign	*n (%)*	1 (0.2)
Length of MS diagnosis at screening (years)	*Mean (SD)*		11.48 (8.40)
	Min		0
	Max		40

The mean SDMT score for people with MS who were screened was 37.43 (SD 12.58), and scores ranged from 3 to 78.

4.5.1.2 *Exploratory factor analysis*

The data were suitable for conducting a PCA; the KMO Measure of Sampling Adequacy was 0.77, and Bartlett's Test of Sphericity was significant ($p<0.01$), indicating that the items are related.

Components were determined by items loading onto a factor with an eigenvalue of 1 or more. Five components were extracted, which explained 55.77% of variance within the data. A forced one-factor

extraction, as based on the original 15-item questionnaire structure, explained 21.45% of variance. Eigenvalues and the variance explained by the five-component structure is given in Table 6.

Table 6: Initial eigenvalues and proportion of variance explained by the five-component solution

Component	Initial Eigenvalues		
	Total	% of Variance	Cumulative %
1	3.22	21.45	21.45
2	1.54	10.26	31.71
3	1.36	9.04	40.75
4	1.2	8.02	48.77
5	1.05	7	55.77

The scree plot in Figure 4 confirms the extraction of these five components, or factors. A parallel analysis also supported the five-component solution.

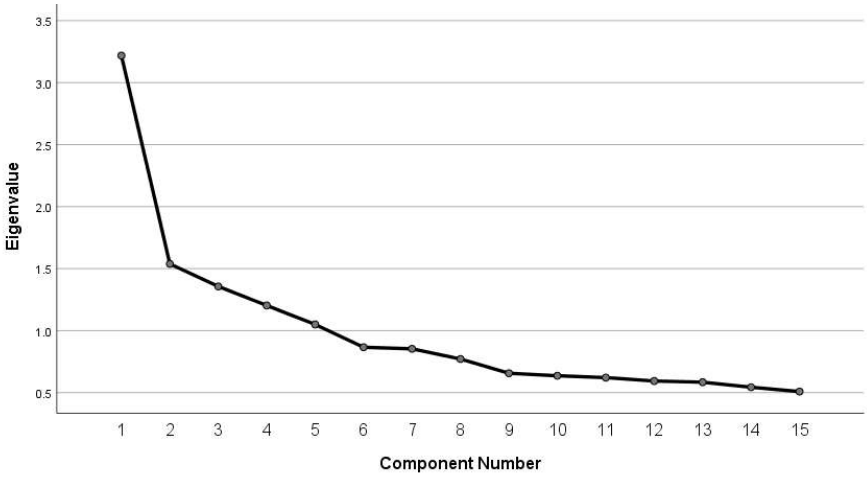


Figure 4: Scree plot of components extracted from the MSNQ-P

The items and their proposed five factor structure solution is presented in rotated component matrix of Table 7. No items showed a loading of less than 0.5 on their component.

Table 7: Rotated component matrix of five extracted components of the MSNQ-P

Item	Component				
	1	2	3	4	5
MSNQ11 - Do you have difficulty keeping track of two things at once?	0.69				
MSNQ10 - Do you have difficulty answering questions?	0.65				
MSNQ07 - Do you need to have instructions repeated?	0.6				
MSNQ03 - Are you slow when trying to solve problems?	0.59				
MSNQ12 - Do you miss the point of what someone is trying to say?	0.57				
MSNQ04 - Do you forget appointments?		0.75			
MSNQ08 - Do you have to be reminded to do tasks?		0.7			
MSNQ09 - Do you forget errands that were planned?		0.69			
MSNQ13 - Do you have difficulty controlling impulses?			0.76		
MSNQ15 - Do you talk excessively or focus too much on your own interests?			0.66		
MSNQ14 - Do you laugh or cry with little cause?			0.57		
MSNQ05 - Do you forget what you read?				0.77	
MSNQ06 - Do you have trouble describing shows or programs recently watched?				0.74	
MSNQ01 - Are you easily distracted?					0.83
MSNQ02 - Do you lose your thoughts while listening to somebody speak?					0.66

Conceptually, items load onto similar domains shown in Table 8. The components cover the domains of complex attention, memory, and executive function.

Table 8: Conceptual components of the MSNQ-P

Component number	Conceptual label	Original item number	Item text
1	Complex attention (divided, and information processing speed)	3	Are you slow when trying to solve problems?
		7	Do you need to have instructions repeated?
		10	Do you have difficulty answering questions?
		11	Do you have difficulty keeping track of two things at once?
		12	Do you miss the point of what someone is trying to say?
2	Memory (long-term - prospective)	4	Do you forget appointments?
		8	Do you have to be reminded to do tasks?
		9	Do you forget errands that were planned?
3	Impulse control	13	Do you have difficulty controlling impulses?
		14	Do you laugh or cry with little cause?
		15	Do you talk excessively or focus too much on your own interests?
4	Memory (long-term - retrospective)	5	Do you forget what you read?
		6	Do you have trouble describing shows or programs recently watched?
5	Complex attention (selective, and sustained)	1	Are you easily distracted?
		2	Do you lose your thoughts while listening to somebody speak?

4.5.1.3 *Associations between objective and subjective measures of cognitive impairment*

The five individual components were calculated as subscale scores by adding the scores of each item together for each participant.

In testing for normality using Kolmogorov-Smirnov statistic, the overall score of the SDMT was normally distributed $D(558)=0.40$, $p>0.05$, however the overall score on the MSNQ-P and the individual components, were not normally distributed, therefore a Spearman's rank order correlation was used to determine the relationships between the measures.

There was no significant correlation between scores on the SDMT and overall score on the MSNQ-P using a Spearman rho correlation (r=-0.05, p<0.05). However, two of the newly extracted components showed weak but significant, negative correlations between their overall score and SDMT score; Component 1, (complex attention including information processing speed, r=-0.13, p<0.05); and Component 2 (prospective memory; r=-0.10, p<0.05).

Combining the items of both of these extracted components into an overall score of eight items also suggested a significant but weak negative correlation between a brief version of the MSNQ-P and the SDMT: r=-0.13, p<0.05.

4.5.2 **Predictors of SDMT scores**

Multivariate linear regression analyses were used to identify whether any of the extracted components predicted scores on the SDMT when covariates were adjusted for. Visual inspections of scatterplots of each variable showed linearity and homoscedasticity. Multicollinearity was not found between any two variables as indicated by a high correlation of r=0.9 or higher.

The only two components of the MSNQ-P that significantly predicted SDMT score when adjusted were those relating to information processing speed (adjusted R^2=0.10); F(8, 549)=8.14, p<0.05, β=-0.13, t=-3.27, p<0.05; and prospective memory (adjusted R^2=0.09); F(8, 549)=7.83, p<0.05, β=-0.12, t=-2.90, p<0.05.

4.5.2.1 *Sample characteristics*

Of the 449 participants who completed the FSS and MSIS-29 baseline assessments, 49 were removed from the analysis due to missing data. The demographics for the remaining 430 participants are given in Table 9..

The demographic data for participants at baseline was similar to the screening data, with a sample that had women predominantly (71.7%) and had a mean age of 49.55 (SD 9.91). Most participants had a relapsing-remitting phenotype of MS (67.3%), and the remaining

participants had progressive forms (32.7%). The average length of time since MS diagnosis was 11.06 (SD 8.04) years.

Table 9: Sample characteristics of people with MS who completed the FSS and MSIS-29 at baseline

Age (at time of consent)	*Mean (SD)*		49.73 (9.87)
	Min		18
	Max		70
Gender	*Woman*	*n (%)*	308 (71.6)
	Man	*n (%)*	122 (28.4)
MS phenotype	*Relapsing-remitting*	*n (%)*	289 (67.2)
	Secondary Progressive	*n (%)*	101 (23.5)
	Primary Progressive	*n (%)*	40 (9.3)
MS diagnosis (years)	*Mean (SD)*		11.05 (8.05)
	Min		0
	Max		40

The average score of correct responses within 90 seconds on the SDMT was 37.16 (SD 11.73). A one-way ANOVA examining the effect of gender and score on the SDMT showed a significant difference ($F(1, 428) = 5.01$, $p<0.05$), with men scoring significantly lower than women. A one-way ANOVA with a post-hoc Bonferroni test showed there was a significant difference in SDMT scores between MS phenotypes, $F(2,430) =13.23$, $p<0.05$; respondents with progressive MS scored significantly lower than those with RRMS.

Using Pearson's correlation coefficients, SDMT score was found to be significantly, negatively correlated with participants' age (r=-0.11, $p<0.05$), however no significant correlations were found between years of education (r=0.08, $p>0.05$), or years lived with MS (r=-0.04, $p>0.05$).

4.5.3 **Self-reported symptoms**

Descriptive statistics of scores on the FSS, two subscales of the MSIS (physical and psychological) are shown in Table 10.

Table 10: Self-reported questionnaire descriptive statistics

Self-reported Questionnaire	Mean	Std. Deviation
Multiple Sclerosis Impact Scale 29 - Physical subscale	52.16	13.5
Multiple Sclerosis Impact Scale 29 - Psychological subscale	23.76	5.98
Fatigue Severity Scale	27.71	6.85

4.5.4 **Multivariate regression analysis**

A histogram showed that the SDMT score was normally distributed for the analysis. The SDMT was significantly and negatively correlated with the MSIS physical subscale, (r=-0.30, p<0.05); the MSIS psychological subscale, (r=-0.17, p<0.05); and the MSNQ-P subscale of prospective memory (r=-0.14, p<0.05). The MSNQ-P subscale of information processing speed (r=-0.08, p>0.05) was not significantly correlated in this analysis of cognitively impaired participants.

The correlation matrix of all self-reported symptom variables and the SDMT is shown in Table 11.

Table11: Correlation matrix of SDMT scores and self-reported variables

	M	SD	*1*	*2*	*3*	*4*	*5*	*6*
1. SDMT	37.16	11.73	-	-0.08	-0.14*	-0.30**	-0.17**	-0.09
2. MSNQ - IPS	13.37	3.30	-0.08	-	0.25**	0.23**	0.22**	0.17**
3. MSNQ - Memory	7.83	2.63	-0.14*	0.25**	-	0.14**	0.17**	0.17**
4. MSIS - Physical	52.16	13.50	-0.30**	0.23**	0.14**	-	0.60**	0.47**
5.MSIS - Psychological	23.76	5.98	-0.17**	0.22**	0.74**	0.60**	-	0.43**
6. FSS	27.71	6.85	-0.09	0.17*	0.17**	0.47**	0.43**	-

**Correlation is significant at the 0.01 level

*Correlation is significant at the 0.05 level

Multivariate linear regression analyses were used to identify variables that predicted scores on the SDMT. Visual inspections of scatterplots of each variable showed linearity and homoscedasticity. Multicollinearity was not found between any two variables as indicated by a high correlation of r=0.9 or higher. Directed acyclic graphs were used to identify which covariates to adjust for in each model.

Categorical covariates (gender and MS phenotype) were transformed to dummy variables to include in the model. For each scale, age, gender, duration of MS, and MS phenotype were included as covariates in each model (as described in section 4.2.2). For the MSNQ-P subscales, education was also entered as a covariate. Using the enter method, SDMT was entered as the dependent variable, with overall questionnaire scores as the exposure variable.

Whilst the overall model was significant (adjusted R^2=0.06); F(6, 423)=5.69, p<0.05; fatigue did not significantly predict SDMT scores (β=0.89, t=-1.90, p>0.05). The model examining the psychological symptoms was significant (adjusted R^2=0.11); F(6, 423)=7.84, p<0.05, and this MSIS-29 psychological subscale significantly predicted SDMT scores, (β=-0.18, t=-3.96, p>0.05). Similarly, the physical symptom model was significant (adjusted R^2=0.11); F(6, 423)=9.57, p<0.05, and the MSIS-29 physical subscale significantly predicted SDMT scores, (β=-0.24, t=-5.04, p>0.05). This indicates that higher reported psychological and physical symptoms predict worse performance on the SDMT.

For the subscales of the MSNQ-P, years of education was additionally entered as a covariate. For the overall model of information processing speed, the overall model was significant (adjusted R^2=0.06); F(7, 422)=4.90, p<0.05; however information processing speed did not significantly predict SDMT scores (β=0.07, t=-1.51, p>0.05). The model examining the prospective memory subscale was significant (adjusted R^2=0.08); F(7, 422)=4.43, p<0.05; and prospective memory was found to significantly predict SDMT scores (β=0.11, t=-2.40 p<0.05).

4.6 Discussion

The MSNQ-P was developed and has been used as a unidimensional, single factor questionnaire, consisting of 15 items. In exploring the factor structure through a PCA, a multidimensional five-component structure was found to explain a greater percentage of variance than a single scale (55.77% against 21.45%). This suggests

that it may be more appropriate to evaluate MSNQ-P responses as individual subscales or exclude items that do not perform as well as others. In comparison to a EFA as part of a Greek validation study (Konstantinopoulou et al., 2018) the factor structure of prospective memory (Component 2; items 4, 8 and 9) also emerged, with similarly high factor loadings However, the two other factors extracted herein this study did not correspond as well to the ones reported here. Konstantinopoulou et al. (2018) also found a high amount of variance explained by a component-based version of the MSNQ-P (65% of variance) as found in the current study (56%).

Research has found inconsistent results regarding the relationship between the SDMT and MSNQ-P (Benedict et al., 2003; Benedict et al., 2004), suggesting limited use of the MSNQ-P as an alternative self-reported measure of information processing speed. However, the results of this current research are promising. The administration of a brief version of the MSNQ-P containing two subscales of attention and prospective memory may be a better indicator of cognitive impairment than the full 15-item questionnaire. The original development of the MSNQ included items related to a 'personality and behaviour' which, although may be associated with cognitive impairment, are not considered as facets of cognition in themselves (American Psychiatric Association, 2013). The three items in the component 'impulse control' relate to inhibition, pseudobulbar affect, and excessive egocentric speech (Benedict et al., 2003). Research has found that these three items exhibit the lowest factor loadings on the overall scale (Sonder et al., 2012). People with MS may be less likely to exhibit these symptoms; for example pseudobulbar affect is thought to only impact 10% of people with MS, and occurs more frequently in advanced stages of MS (Pinkston et al., 2007). Therefore, it is unsurprising that these items are less likely to be endorsed and may suggest redundancy in retaining these items in the questionnaire, in particular for a screening programme.

There were significant but weak correlations between the overall score on the SDMT and two components that were extracted from the MSNQ-P through a PCA – those related to complex attention (including information processing speed) and prospective memory. This suggests that lower scores on the SDMT are associated with poorer self-reported impairments in attention and prospective memory. The significant relationship between lower SDMT scores and higher prospective memory subscale scores is unsurprising. Individuals are possibly more likely to recall and self-report occasions when they have forgotten to complete an action, rather than self-reporting forgetting something that they have learned (Ernst, 2020). Comparisons with other objective tests may provide further insight into how self-reported cognitive impairment correlates with other cognitive domains.

A further analysis examined how self-reported cognitive problems may predict processing speed scores as measured by the SDMT. The analysis of predictor variables only used a sample of people with MS who were considered cognitively impaired, and who participated in baseline assessment in the CRAMMS trial. Using linear regression models to analyse how self-reported variables predict SDMT, only items related to prospective memory significantly predicted scores on the SDMT for cognitively impaired people with MS when adjusted for covariates; items relating to information processing speed were not found to significantly predict scores on the SDMT. This may suggest that the MSNQ-P is more valid as a self-reported information processing speed for people with MS with lower levels of cognitive impairment, however further research is needed.

The SDMT is seen as a gold standard in minimal cognitive assessment, and these results do not provide sufficient evidence that a brief version of the MSNQ-P could be used instead of the SDMT. However, for those people with MS who do not wish to or cannot complete objective assessments, it has promise of capturing some cognitive impairment. This could be important in capturing the cognitive functioning of those individuals and ensure that they will not be

excluded by a cognitive screening programme that only includes objective tests. It has been suggested that people with MS with lower cognitive functioning may be more likely to be unwilling to complete objective measures (Cortés-Martínez et al., 2019). Thus, as the information processing speed subscale of the MSNQ-P is a non-significant predictor of SDMT for people with MS exhibiting cognitive dysfunction, this subscale may have limited utility. Future research could analyse the relationship of the MSNQ-P's subscales to other objective assessments assessing different domains to explore this further.

As the original MSNQ-P was designed to be a brief, unidimensional questionnaire (Benedict et al., 2003), concentrating on items relating to these components of cognition; including complex attention and prospective memory, and including items related to domains such as social cognition (typically neglected in cognitive assessment in MS), a new self-report questionnaire may show more promise than previously developed self-report measures. This may create a more sensitive self-report measure of cognitive functioning which encapsulates greater relevant domains.

In the regression analysis of Aim III, both physical and psychological subscales of the MSIS-29 predicted SDMT score, with higher scores of physical and psychological symptoms associated with decreased processing speed. However, fatigue, measured by the self-reported questionnaire the FSS, did not significantly predict SDMT scores. A cross-sectional and longitudinal study found similar results when examining the association between fatigue and performance on neuropsychological tests (Morrow et al., 2009). This study was conducted with a large sample of people with MS (*n*=465) both with and without cognitive impairment. The current analysis suggests that fatigue as measured by the FSS does not predict processing speed for cognitively impaired people with MS. It has been suggested that poor complex attention may be associated with fatigue due to an attention-

deficit (Amato et al., 2019), and this could be explored further using other tasks of attention.

Whilst mood disorders are prevalent in MS, its relationship to cognitive impairment in MS is complex. Feinstein (2006) proposed a model that depression may diminish cognitive capacity in people with MS, especially in relation to executive functioning. Depression and anxiety have been found to be factors associated with deficits in executive function and information processing speed, potentially though lowering cognitive reserve (Amato et al., 2019), which may be shown by the significant relationship between psychological symptoms and SDMT scores in the current analysis.

Similarly, physical symptoms were found to predict slowed information processing speed in cognitively impaired participants with MS. This is potentially due to the progressive nature of MS, with increasing physical symptoms being associated with greater cognitive impairment (Amato et al., 2006), however this could also be associated with slowed motor function.

For this analysis, data were taken from a completed RCT. There are some differences between when the data used were collected. Participants could complete the MSNQ-P during a screening telephone call or at the screening visit at which the BRB-N would be administered. Cognitive problems including information processing speed can fluctuate daily (Kratz et al., 2017), and therefore an individual's cognitive function may have differed across the two time periods they may have completed the assessments.

If the participant was determined eligible, the researchers gave participants the baseline assessments of the MSIS-29 and FSS to complete at any time in the two weeks before their baseline assessment visit. If they had not completed these before the visit, they would complete it at the baseline visit. Therefore, in this analysis participants had completed different self-report data at various time points. Fatigue can fluctuate for people with MS and can be influenced

by the time of day, physical exertion, and daily stressors (Powell et al., 2017). Whilst the MSIS-29 has good test-retest reliability (McGuigan & Hutchinson, 2004), this reliability assumes stability in other variables, which cannot be adequately captured.

The dataset for the third aim consisted of people with MS who had already been identified as cognitively impaired. Cognitive impairment was defined through failure on two or more of the tasks within the BRB-N, one of which could have been the SDMT. Although the sample was still normally distributed, scores on the SDMT were higher than those as baseline, as expected. The BRB-N assesses other domains of cognition, including memory. Some participants may show impairment in any one of the tests, and may only have impairment in one domain and not of information processing speed.

The findings of Aim III used self-report questionnaires that were chosen for use in CRAMMS. The MSIS-29 psychological subscale has been criticised for not distinguishing between anxiety and depression (Jones et al., 2013), and as such it is difficult to conclude what particular psychological symptoms may be more likely to predict processing speed deficits. Further analyses could examine the relationship between other self-reported questionnaires to evaluate how information processing speed may be predicted by other variables, such as general quality of life.

4.7 Conclusions

This chapter has shown that a component-based, reduced-item MSNQ-P may have more utility as a self-reported cognitive questionnaire than the original unidimensional version due to its significant relationship with the objective measure of information processing speed (the SDMT). However, the correlations were weak, and the proposed reduced questionnaire would be eight items, or two subscales of five and three-items. This still does not assess some cognitive domains missing from the original scale, such as social cognition which may lend themselves to a self-report assessment.

Additionally, these subscales may not be able to predict slowed information processing speed in people with MS who are cognitively impaired, and this warrants additional research.

Previous chapters have determined that the SDMT is frequently used, and is valid, and reliable for use in people with MS, and it is recommended as a brief screening assessment. However, previous research using qualitative methodologies to determine whether SDMT or other brief cognitive assessments are appropriate and acceptable is insufficient. Therefore, the next step in developing a method of screening for cognitive problems is to evaluate how it can be implemented in practice in the context of UK clinical care. The research described in Chapter 5 explores how cognitive screening with brief tests, or self-report measures such as the MSNQ can be used in this situation, and what can facilitate the uptake of cognitive screening. The following chapters will detail how a cognitive screening programme, which may use objective and subjective measures, could be implemented to ensure acceptability by examining attitudes to cognitive assessment.

Chapter 5 Cognitive Assessment in Multiple Sclerosis Clinical Care: A Qualitative Evaluation of Stakeholder Perceptions and Preferences

5.1 Overview of the Chapter

This chapter presents a qualitative study which aimed to develop and refine a cognitive screening programme through semi-structured interviews and focus groups with people with MS, and other stakeholders including family members, charity volunteers, clinicians, and healthcare commissioners. This chapter presents the views of the stakeholders regarding current practice of the identification of cognitive impairment within UK healthcare, and how a cognitive screening could feasibly and effectively operate within this setting.

This chapter is an extended version of a journal article published in Neuropsychological Rehabilitation (Elwick, et al., 2021). Methodology and results have been extended in this chapter to fully describe the framework that was used to analyse the data.

5.2 Introduction

Qualitative interviews with people with MS have found that cognition is a 'neglected symptom' in their MS care, and although they placed importance on this symptom, this is rarely addressed by their MS clinicians, unless it is severe (McAuliffe & Hynes, 2019). Although a full neuropsychological evaluation within clinical care would provide detailed insight into cognitive impairment, research has shown that screening for cognitive problems is not routine in the UK neurology services (Klein et al., 2019). The feasibility of delivering such a programme may be limited due to the lack of time, staffing, and other resources within clinics. Emphasis is therefore placed on briefer assessments that are still effective at measuring cognitive impairment (Cullen et al., 2007) to ensure that cognitive assessment is practicable and efficient (Amato & Zipoli, 2003)

Approaches to standardised cognitive assessments for people with MS have used expert consensus recommendations (e.g., Benedict et al., 2002; Kalb et al., 2018; Lensch et al., 2006) which do not appear to take into account the preferences of people with MS, their family members, or clinicians. Indeed, despite these recommendations, a recent survey found that commonly used assessments for people with MS in UK healthcare included generic measures or measures developed for dementia, such as (MoCA) and the Addenbrooke's Cognitive Examination (ACE-R), and were not standardised across clinical services (Klein et al., 2018). Furthermore, as discussed in Chapter 4, people with MS often do not wish to complete one of the most commonly used and recommended measures in MS, the Paced Auditory Serial Addition Test (PASAT; Cortés-Martínez et al., 2019). People with MS with greater cognitive impairment are more likely to decline completing the PASAT, causing challenges in being able to effectively identify and monitor cognitive problems and provide appropriate care.

The Topol Report emphasises the importance of involving patients in developing digital medicine for the UK National Health Service (NHS) to improve their own experience and ensure that it is appropriate for their needs (Topol, 2020). A proof of concept study suggested that people with MS prefer a self-administered, computerised tests to the traditional technician-administered Symbol Digit Modalities Test (SDMT, Patel et al., 2019). Some assessments can be developed as an electronic test more readily than others. However, adapting tests to an electronic format may require substantial procedural changes, such as how stimuli are responded to, and will need to ensure that people are not put off or disadvantaged by the testing modality.

These issues can be addressed by understanding stakeholders' motivations and concerns about cognitive assessments being used in usual clinical practice. However, few studies have evaluated 'how' specific tests or batteries are used in practice. An investigation into patient preference for routine cognitive screening found the people with

MS were supportive of routine cognitive screening and for their clinicians to manage both physical and cognitive symptoms (Mortensen et al., 2020). Participants perceived the SDMT to be less demanding and fatiguing than the other brief assessments included in the BICAMS (Brief International Cognitive Assessment for MS). Although valuable, this research used a small select sample of Danish patients, some of whom were already volunteering in a study of cognitive screening, which only included participants who scored between 20-60 on the SDMT. Additionally, this study did not include patients who declined to complete the SDMT or elicit the opinions of those who scored outside of the score boundaries mentioned, which could have been valuable in determining how to engage all patients in cognitive screening. Furthermore, this study only focussed on people with MS, and did not include other key stakeholders such as clinicians.

5.3 Aim

The aim of the current research was to explore which cognitive assessments may be acceptable to administer, in which modality (for example electronically), and in which settings, amongst key stakeholders based in the UK.

Key stakeholders are defined as those who are involved in the process of cognitive screening to ensure its success within the UK healthcare system. These include: (i) people with MS who need to be capable, motivated, and have the opportunity to complete tests, (ii) carers or family members of people with MS to support them in completing tests, (iii) clinicians who need to implement the screening in clinical services and help interpret the results of the tests to plan future care, (iv) commissioners who fund clinical services within the NHS, amongst others. Using qualitative methods, we believed we would be able to identify stakeholders' perceptions of the assessments that are acceptable and feasible to be administered in routine UK healthcare systems.

5.4 Methods

5.4.1 Ethical Considerations

Ethical approval was obtained from the University of Nottingham's Division of Psychiatry and Applied Psychology's Research Ethics Committee (Ethics reference number: 0233) and the University of Nottingham's Faculty of Medicine and Health Sciences Research Ethics Committee (Ethics reference Number: 263-1903). Ethics approval letters can be found in Appendix 8.

As described in the recruitment strategy, we used personal networks and a Patient and Public Involvement (PPI) group to recruit participants for both categories. To address implied coercion, potential participants were assured that they were under no obligation to take part, even if they knew a member of the research team.

Written or verbal informed consent was obtained, and a Participant Information Sheet (PIS) was given to all participants. We were aware that discussing sensitive issues could lead to distress, and so we signposted the participant to contact their GP or neurology care team in this case. The PIS also detailed how to contact the researcher, research supervisor or ethics committee if they had any questions or complaints.

All data collected were held in password protected files, including the database of participants, the anonymisation key, and all audio recordings and transcripts. All participants were given a pseudonym based on their participant characteristic (e.g., occupation therapists were ascribed OT 01). All identifying details were removed from the transcripts. A signed written agreement with the professional transcription service was put in place with the service to ensure confidentiality of data.

5.4.2 **Design**

Interviews and a focus group were conducted to gather insights regarding aspects of cognitive screening in MS. Both methods were used to enhance the richness of data (Lambert & Loiselle, 2008).

5.4.3 **Data collection**

Data were collected as part of two research projects; the current PhD; and for an NIHR-funded programme, Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS). NEuRoMS aims to develop and evaluate a cognitive screening pathway and rehabilitation for people with MS. Qualitative data in this phase were collected by four researchers, who interviewed participants about both cognitive screening and cognitive rehabilitation. Only data collected regarding cognitive screening is used in this thesis, and the complete data is used as a NEuRoMS output. Through NEuRoMS, a logic model and preliminary pathway was devised prior to data collection (Appendix 1) and parts of this were used in the current study to guide discussion. The proposed logic model was developed using a variety of sources, including literature reviews, input from PPI groups and clinician experience.

5.4.4 **Participants**

Based on recommendations of determining sample size for qualitative research (Vasileiou et al., 2018), we determined that a sample size of 40-50 participants across the two participant categories previously mentioned, would be sufficient to address our aims. However, we reviewed the sample size on an iterative basis to ensure data saturation, and to ensure we did not collect more data than necessary. Saturation was operationalised as the point where there were no new data being generated which supported ongoing theme development or creating the possibility for a new theme. The wider research team and I regularly reviewed the data to determine whether data collection had reached this point, and whether the analysis would benefit from further interviews.

We used a purposive convenience sampling strategy to recruit participants. All participants invited to participate if they were 18 years old or older, could communicate in English, and were able to give informed consent to participate. People with MS had to have a diagnosis of MS; family members had to have a relative or care for a person with MS, and charity volunteers had to be affiliated with MS charities such as the MS Society, Overcoming MS, or the MS Trust. These participants were recruited through PPI networks, online advertisements (using relevant communities and forums on Twitter, Reddit, and Facebook), contacts at MS charities, and by word of mouth. The participant advert used is given in Appendix 2.

Clinicians, identified through professional networks of the research team, representing different healthcare professionals working in MS, were invited to participate. Commissioners were invited to participate if they were a service commissioner working within the Clinical Commissioning Groups that fund NHS services.

5.4.5 **Materials**

A brief video was produced which showed digitised administrations of three commonly used screening assessments: a symbol substitution test, a Stroop test, and a paced addition test. This was shown to participants during the interviews. We also provided participants with a document showing screenshots of the tests to facilitate the discussion (Appendix 7). This document also contained a word list learning test, a verbal fluency test, a trail making task, and the first ten items of the patient form of the Multiple Sclerosis Neuropsychological Screening questionnaire (MSNQ, Benedict et al., 2003). The seven measures were selected because they: (i) are commonly applied both as individual assessments and within batteries such as BICAMS and the BRB-N (Benedict et al., 2002; Langdon et al., 2012); (ii) demonstrate different modalities (electronic, written and verbal); (iii) and provide examples of both subjective and objective measures. Clinicians and commissioners were also shown the preliminary logic model and asked for feedback.

Two semi-structured interview schedules, one for people with MS, family members and charity volunteers, and one for clinicians and commissioners, were developed (Appendix 5) using the preliminary logic model and input from a PPI group. The interview schedule was amended on an iterative basis if the interviewers found that some questions required clarification or further areas needed to be investigated. The interview and focus group schedules consisted of both general and specific questions, and probing questions to clarify responses or to ask further questions if needed.

5.4.6 **Procedure**

Upon receiving interest in participating in the study, we emailed participants copies of all the interview materials including the participant information (Appendix 3), consent forms (Appendix 4), a link to the video of computerised assessments, and details of the assessments and the proposed cognitive screening pathway (Appendix 1 and Appendix 7). Participants were asked to review all materials before their interview.

One focus group was run for people with MS and was co-facilitated by three researchers and a PPI partner. After introductions and providing some background on the study, participants were shown the materials and encouraged to participate in the focus group. Researchers guided the focus group using the same interview schedule as used for the interviews.

Interviews were held at participants' homes (people with MS), place of work (clinicians), or over the phone. The focus group was held at a convenient time and place, and participants were offered breaks. At the beginning of the interview or focus group, consent was obtained, and interviewers explained the research aims and procedure, and what would happen during the interview or focus group. Demographic information was collected for each person with MS, family member and charity volunteer. All participants we approached agreed to take part, and no participants dropped out of the study. We showed the brief

video and assessment document to participants at an appropriate time during the interview.

Participants who had MS, family members and charity volunteers were offered a £20 shopping voucher for their participation. Travel expenses were offered for all participants who attended an interview or the focus group in person.

All participants were asked about their perceptions of implementing a routine cognitive screening pathway. We asked about a proposed situation whereby people with MS would be sent a link to complete the cognitive screening at home prior to their appointment with a clinician, so that their results would be available to be discussed at their appointment, and the necessary referral to neuropsychological rehabilitation could take place if needed. In the interviews we explored the understanding behind the pathway, the feasibility, and how it will fulfil the unmet needs.

5.4.7 **Transcription and data analysis**

Audio recordings of the interviews and focus group were sent to a professional transcription service for verbatim transcription. All transcripts were anonymised, imported into NVivo software (version 12), and analysed using framework analysis, a hierarchical, matrix-based method which aims to systematically approach the data with the predetermined objectives of the logic model (Pope et al., 2000). The logic model was used to develop an initial thematic framework for analysis, and through an iterative process, this was refined to capture the themes derived from the data.

Two researchers (myself and a postdoctoral researcher) then immersed ourselves in the data from transcripts and any field notes collected, to become familiar with the data, and gain an overview of the recurring patterns within the data. Framework analysis was chosen for analysis as a method of systematically approaching the data with the predetermined objectives of the logic model (Pope et al., 2000). We developed iterations of the framework, with feedback from a third

researcher who was familiar with the logic model and methodology. The framework uses a **Context-Mechanisms-Outcomes** configuration of a cognitive assessment programme based on realist evaluation (Pawson & Tilley, 1997). **Context** referred to the prevalence of cognitive problems in MS, and the need for the identification of these problems in routine clinical care, and early enough for further cognitive assessment and/or the provision of rehabilitation. **Mechanisms** refer to the resources needed for cognitive assessment to be successfully implemented in clinical practice. The **outcomes** include the results and next stages of cognitive assessment, and the perceived benefits of the programme. These are used as overarching themes, with themes and subthemes within these.

Once the framework was finalised, data from the transcripts were indexed using NVivo, using each thematic aspect of the developed framework. These indexed data were then charted within a matrix which mapped the data onto the framework and could be interpreted. We then wrote summaries to correspond to each data point for each participant's transcript. After each researcher coded a small number of interviews each, we met again to discuss the effectiveness of the framework for mapping the data, and on this basis no further amendments were made to the theoretical framework.

Once summaries for each case and each theme were written, these were distilled further through overall summaries for the two participant groups: (i) people with MS, carers, and charity volunteers with MS, and (ii) clinicians and healthcare commissioners. The initial results in the form of these summaries mapped onto the pathway were presented to both the NEuRoMS programme management group, and a PPI group for feedback and a sense check of the final framework and interpretations of the data.

5.5 Results

5.5.1 Sample characteristics

Forty-nine people participated in the interviews and focus group. The demographics of the people with MS, family members and charity volunteers are given in Table 12. Participants have been assigned a code which is given in Table E in the appendix together with individual level demographic and information about MS characteristics. Seventeen clinicians were interviewed: four each of MS nurse specialists, neurologists, occupational therapists (OTs), and neuropsychologists, one neuro-physiotherapist, and three commissioners. The focus group lasted 125 minutes and all participants contributed to the discussion. All interviews lasted an average of 52 minutes.

Table 12: Demographics of people with MS, family members and charity volunteers

Interviews with people with MS (n=15)	Age	*Mean*	48
		Range	36-67
	Time Since Diagnoses	*Mean*	12.93
		Range	2-41
	Gender	*Women*	12
		Men	3
	Ethnicity	*White*	14
		Mixed/Multiple Ethnic Groups	1
	Highest Education	*GCSE*	4
		A Level	2
		Degree	5
		Higher Degree	4
	Employment	*Not employed*	4
		Employed Full Time	4
		Employed Part Time	3
		Retired	4
	Type of MS	*Relapsing-remitting MS*	9
		Primary Progressive MS	1
		Secondary Progressive MS	4
		Unknown	1
Charity Volunteer Interviews (n=4)	Age	*Mean*	42.25
		Range	36-53
	Time Since Diagnoses	*Mean*	8.75
		Range	2-18
	Gender	*Female*	3

		Male	1
	Ethnicity	*White*	3
		Mixed/Multiple Ethnic Groups	1
	Highest Education	*Degree*	3
		Higher Degree	1
	Employment	*Voluntary Part Time*	2
		Employed Part Time and Voluntary Part Time	2
	Type of MS	*Relapsing-remitting MS*	3
		Primary Progressive MS	1
	Age	*Mean*	54.2
		Range	21-68
	Time Since Family Member's Diagnosis	*Mean*	15.8
		Range	3-28
	Gender	*Female*	1
		Male	4
	Ethnicity	*White*	5
Family Member Interviews (n=5)	Highest Education	*A Level*	1
		Degree	3
		Higher Degree	1
	Employment	*Education Full Time*	1
		Employed Full Time	1
		Employed Part Time	1
		Voluntary Part Time	1
		Retired	1
	Family Member's Type of MS	*Relapsing-remitting MS*	3
		Secondary Progressive MS	1
		Unknown	1
	Age	*Mean*	50
		Range	38-57
	Time Since Diagnoses	*Mean*	12
		Range	4-26
	Gender	*Female*	4
		Male	1
Focus Group with people with MS (n=5)	Ethnicity	*White*	5
	Highest Education	*GCSE*	3
		A Level	1
		Higher Degree	2
	Employment	*Not employed*	4
		Employed Full Time	4
		Employed Part Time	3
		Retired	4
	Type of MS	*Relapsing-remitting MS*	4
		Secondary Progressive MS	1

5.5.2 **Context**

5.5.2.1 *Cognitive problems in MS*

All participants recognised cognitive problems as a prevalent symptom for many people with MS, which impact upon daily life, employment, and social life. Memory, processing speed, decision making, concentration, and communication were all highlighted as affected cognition domains for people with MS. Clinicians also acknowledged that cognition, fatigue, and quality of life are interlinked in MS. Despite cognitive impairment being perceived as an important symptom of MS, both people with MS and clinicians reported that discussion or assessment of cognitive symptoms does not routinely occur in clinical care, as other symptoms take priority. Some neurologists interviewed did ask people with MS about their cognition in annual appointments, but due to the competition for time in routine appointments, other symptoms often took priority over 'invisible' symptoms such as cognitive impairment.

People with MS were not always directly told about cognitive problems during their appointments. They often reported their physical symptoms take precedence over invisible symptoms and thought that cognitive problems were often not taken as seriously as other problems.

> *"[clinicians] just think it's mild. It's mild in comparison with dementia but actually it's not really – often it is mild but the effects that mild memory loss can have on your life – and there are things I'm always forgetting, you know, it drives me mad. It really annoys me"* (Person with MS F 03)[2].

Consequently, some participants thought that younger or newly diagnosed people may be unaware that cognitive problems could be an MS symptom, and that some people could be in denial of their problems by "*burying their head in the sand, just not willing to accept*" (Volunteer 01). A charity volunteer acknowledged that, although cognitive issues

[2] People with MS have been coded to indicate if they participated in an interview (I) or the focus group (F)

had not affected them, it would still be beneficial to be informed by their clinician about potential cognitive problems: *"it might be quite handy to actually hear it from somebody clinically saying, you know, you need to do something about this"* (Volunteer 01). Some people with MS thought that their neurologist "*doesn't think [cognitive problems] actually even exist*" and that cognitive problems were not "*taken seriously*" (Person with MS I 11).

5.5.2.2 *Current identification of cognitive problems*

Neurologists and MS nurses largely perceived addressing cognitive problems the responsibility of neuropsychologists and OTs, and referrals to primary and secondary care were being prioritised for those with severe or advanced problems, or when cognitive problems were having an adverse impact on employment or education. One nurse reported that they "*don't actually formally screen in clinic and we send them off to neuro-psychology to be screened*" (MS Nurse 01).

Neurologists reported that they were able to refer people with MS to psychologists or neuropsychologists if they report cognitive problems that are disruptive, complex, or severe.

> *"It could be people who are functioning pretty well but perhaps struggling a little bit in the workplace for example through to people who are more at the severe end and anything in between" (Physiotherapist 04)*

Clinicians observed that there are few neuropsychologists available for neurology patients, and referral "*is quite clumsy, it can take quite a while"* (Neurologist 02). Neurologists stated that referrals can come about through joint decisions within multidisciplinary teams, "*I'd probably refer them into the MDT and then let the discussions there guide the best approach*" (Neurologist 04), and referrals can also be made to OTs and neurorehabilitation services if available. OTs and neuropsychologists emphasised that assessments should encompass the domains of processing speed, working memory, executive function,

and attention and clinicians find that patients who are referred do show impairment on objective cognitive assessments.

People with MS also reported that the availability of support for cognitive problems is mixed but generally limited, and there is often a lack of information about what is available, with one person with MS stating that they *"don't know whether there's anything that they can advise you on*" (Person with MS I 06), or there is nothing available and "*either it's not picked up, you're not given any guidance or support*" (Person with MS I 18). One person with MS "*had to fight to see a neuropsychologist*" (Person with MS I 18), and another had asked to be referred for help and felt frustrated by the process.

> *"I feel quite angry that I have to do all that and no-one's doing anything for me… to go through all that for no gain is – and I know that there's not going to be anything positive at the end of it anyway"* (Person with MS I 11)

Some people with MS reported that they received cognitive assessments through charity programmes, research, occupational health, but only one person had a referral from their neurologist. Participants who had experience of cognitive assessments stated that they could be abstract or removed from everyday life, and some could induce panic, and be long and tiring, especially given their fatigue, stating they are "*a very intense and probably quite a long assessment, you know, given that your MS is there and can be, I don't know, you know, you're more impaired, aren't you, than just somebody going in normally*" (Person with MS I 20) and one participant "*came away feeling a bit disappointed in [herself]"* (Volunteer 05). A charity volunteer thought that although they did not think cognitive issues affected them, it would still be beneficial to be informed by their clinician: "*I think this might be quite handy to actually hear it from somebody clinically saying, you know, you need to do something about this*" (Volunteer 01).

OTs and psychologists reported using the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS), BICAMS,

SDMT, MoCA, Addenbrooke's Cognitive Examination III (ACE-III), Rivermead Behavioural Memory Test (RBMT) and/or Test of Everyday Attention (TEA) if required, but these were not routinely administered for all people with MS. The neuro-physiotherapist reported using the SDMT as part of a larger battery of assessments and used it to aid discussion with the people with MS about their perception of cognitive changes.

5.5.2.3 *Related factors*

Clinicians working in MS acknowledged that factors, such as mood and medication, having an impact on cognition need to be managed prior to referrals. Most clinicians and people with MS identified that cognitive impairment is interlinked with other symptoms and aspects of individuals' lives. Both clinicians and people with MS expressed views about the importance of holistic care for MS where component parts affect one another and identified the need to maintain both mental and physical aspects. Fatigue was a significant symptom that was frequently described as both having an impact on, and being affected by cognitive impairment. One participant stated that they "*get irritated by [cognitive impairment] and the more tired I am the more irritable I get"* (Person with MS I 16). Some people with MS emphasised factors such as age and hormonal changes impacting their cognition and expressed uncertainty as to whether cognitive impairment was due to their MS or other factors and *"it could be due to lack of sleep, could be due to pregnancy, it could be due to MS, you don't know*" (Person with MS I 15).

Another important factor related to cognition was that of mood, "*if somebody's particularly low in mood or particularly high, it's going to have a huge impact on their cognition*" (OT 06). Participants also discussed the impact of cognitive impairment on their quality of life, including personal relationships, family life and employment.

> *"There are days when it affects every level really on a relationship basis as well, you know, you want to be an equal*

with your partner but there's some days I just feel absolutely like I can't have a conversation" (Person with MS I 20).

5.5.3 **Mechanisms**

5.5.3.1 *Cognitive assessment*

Most people with MS thought that cognitive assessments proposed were acceptable and would complete them if required. However, one person with MS strongly opposed completing the cognitive tasks, expressing a negative reaction towards the tasks, but said they would complete questionnaires. Another thought questionnaires would be less intimidating as "*you feel like it's not being done to you, you have got some control over what's actually happening*" (Person with MS I 11). Clinicians suggested that subjective measures can add further insight into cognitive impairment, and the perceived difficulties of people with MS. However, there were concerns about recall periods, and whether questionnaires were nuanced enough, with one person with MS thinking they might be "*too subjective, a bit too fluffy… unless I've got a space for text which gives nuance, these are worthless"* (Person with MS F 02). One person with MS thought that "*when you go through a questionnaire, you don't really want to revisit the symptoms that you feel and the changes that have come about in you as a person*" (Person with MS I 04). Similarly, both patient- and proxy-reported measures were considered acceptable: "*it would be quite interesting getting both of you, the family member and yourself and see what the difference is*" (Person with MS I 04), but one clinician had concerns that family members may not want to fill in proxy-reported measures.

Some participants thought that symbol substitution tests were relevant and had ecological validity. One person with MS reported that it was "*like looking at a spreadsheet isn't it! [Laughs] Making sure your data is correct*" (Person with MS I 09), although conversely, one person with MS felt that the task would benefit from extra context about the test and would benefit from a "*simple explanation of 'this is what this test is*

looking for"" (Person with MS I 11). This was a common theme from participants who were shown the measures:

> *"I think it's really, really important to provide people that information because otherwise you don't know if people just look at it and go 'oh it's just another one of these things that I have to fill in and then never hear back about'"* (Person with MS I 13).

Many people with MS found the paced addition style tasks difficult, fatiguing and potentially anxiety inducing. One participant reported they would "*probably hate to do that one*" (Person with MS I 02), others were also unwilling to complete it, stating "*I'm not doing that. I've more important things in life, that's what I'd think*" (Person with MS F 01) and "*I wouldn't even enjoy doing it, so why would I bother"* (Person with MS I 07). Some people with MS indicated that they would still try to complete the task, despite struggling with it and believed that their premorbid mathematical ability may have more of an impact than their cognitive impairment "*the reason the score's so low is because I just can't do it!"* (Person with MS F 02). A neuropsychologist agreed that "*the resounding feedback for the PASAT with patients is a 'no!'"* (Psychologist 05).

5.5.3.2 *Resources needed*

5.5.3.2.1 Time

People with MS thought that cognitive assessments should be brief enough to prevent fatigue and a lapse in concentration, and 2-5 minutes of screening would be acceptable, but a few people with MS felt tests could be up to 60 minutes long if they needed to be. Stakeholders thought that brief screening would encourage people with MS to complete screening at home and ensure they were not too tiring:

> *"These tests are requiring a great deal of commitment and input and, as I said, I have to be very careful about how I spend my energy and I would probably end up at the end of the day*

feeling a little hard done by because I'd done that, and I wanted to do something else" (*Person with MS I* 07).

5.5.3.2.2 Staff

The lack of time during appointments with clinicians was also discussed when asked about who should feedback results to people with MS. Some people with MS thought that their neurologist would be the best person, especially if the results indicated cognitive decline, to afford an opportunity for people with MS to ask questions and discuss the results in-person. One carer stated that "*it's very important this trust element, that is vital, but you can only create this trust if the professional has got the correct information"* (Family Member 09). Some people with MS stated a preference for their results to be given by an MS nurse, because they had a better relationship and longer appointment times with them compared with their neurologist.

Clinicians observed that the discussion would have to be brief to fit into existing appointments with neurologists or MS nurses. MS nurses reported not having enough time in their current workload to contend with cognitive problems, and some viewed this as the job of other healthcare professionals, such as OTs, while they saw their remit focusing on disease modifying therapies and managing physical symptoms. One OT thought "*that wasn't the most effective use of [OT's] time*" (OT 02). Neurologists thought that other healthcare professionals such as physiotherapists, psychologists or OTs, or volunteers could be trained to administer the screening. All participant groups thought that volunteers could play a part in the screening, including providing help and support to complete the assessments. People with MS felt it was important that the person administering the cognitive tests should understand MS and be able to explain the rationale and provide support for completion.

Clinicians also thought that administrative staff could be involved and trained to administer the screening pathway. Commissioners and neurologists felt that the screening programme could increase the workload for administrative staff:

"You could provisionally have a trained receptionist to give a tablet, and say can you do this, and then can you give the tablet to me please, so I can give it to the next patient. That of course, you know, implies some cooperation and willingness of the receptionist to facilitate" (Neurologist 01).

5.5.3.2.3 Mode of delivery

Provision of a digital cognitive assessment would require it to operate with current NHS systems, which clinicians and one commissioner thought may be challenging. Some clinicians reported that they have no infrastructure for embedding new processes, including the ability to email patients, and this may pose difficulties administering the screening.

All stakeholders felt that digital administration of cognitive tests was an acceptable option, and digitisation may fit in well with other electronic systems such as patient records. However, there were concerns from clinicians that it may lead to non-completion, "*there will be a big chunk of people that don't do it. But that is just in relation to general kind of self-management, I suppose, and how well… kind of engaged people are with that"* (OT 01). People with MS and their family members thought that other people with MS may have challenges with electronic completion, especially older individuals, and those lacking confidence with technology or not having the appropriate equipment. However, none of those interviewed thought they themselves would have difficulty with the proposed technology.

Digital versions may overcome obstacles, such as anxiety induced by in-person administrations. One volunteer said, "*I'd be a bit worried; I'd be a bit panicky, and then I'll be sitting there thinking I can't think… in a face-to-face situation I think it could be quite intense*" (Volunteer 04). Electronic versions may also overcome dexterity issues. One volunteer thought they would have difficulties with pen-and-paper assessments, which could lead to inaccurate scoring, as "*pressing a screen is a lot easier than holding a pen for me… I would struggle more… I would definitely be slower*" (Volunteer 01).

Clinicians raised concerns that some tests, such as figure copying, word list learning, and verbal fluency tests would be challenging, if not impossible to digitise, and the change in modality would require new normative data to be collected for assessments. One neurologist emphasised that handheld administrations need further work by "*looking to see how the tests perform in mobile versus tablet and what kind of people use one or the other or how it compares to the two*" (Neurologist 03).

More difficulty was ascribed to using a keyboard and mouse rather than touchscreen devices, which people with MS reported to be more familiar with, and felt "*it's most convenient, it's easier*" (Person with MS I 13). Clinicians agreed that tablets would be easier to use in clinic than a laptop or desktop computer but restricting assessment to this format may preclude people from self-administration, however, traditional verbal or written tests could be made available.

5.5.3.2.4 Location of testing

MS nurses thought in-clinic testing would be impractical. Waiting rooms were viewed as inappropriate for assessment – both people with MS and clinicians perceived them as too busy, loud, and unsuitable to complete tests: *"just think chaos*" (Volunteer 05). Screening during appointments was also perceived as inappropriate due to time constraints. One participant was very resistant to spending any more time in clinic than they had to, saying that before and after their appointments they are *"angry and all it does is give me a hard time dealing with my illness"* (Person with MS I 07).

Most people with MS preferred the option of completing the assessment at home, and emphasised the potential benefits of this. These included saving time and being more comfortable at home, thus providing a better reflection of their "true" cognitive ability. An OT agreed with this, saying "*we probably would get more honest information if people had time to give it at home, and bring the results with them*" (OT 02).

Participants agreed that it was important to still have the option of completing the screening in clinic, if people with MS could not complete it at home. There were also perceived advantages to completing in clinic, including having "*all assessments and results in hand and patients don't forget about it"* (MS Nurse 05). The results of the screening were felt to potentially be anxiety inducing and people with MS may want reassurance about their results as soon as possible:

> *"If they feel that they're a bit isolated and they're not going to know for a while what the results are, if they – I guess it's a bit different than doing it kind of face-to-face and getting reassurance straight away of you" (*OT 01*).*

For successful completion of screening in individuals' homes, there needs to be the infrastructure to send a weblink and information to the patient. In many services, appointment letters are generated from a central department and incorporating a link could be challenging. Clinicians reported that their service sent text message reminders of appointments in which links could be sent, however the timing of letter and text messages vary between trusts and departments. People with MS stated a preference for receiving the link to the cognitive assessment in an email for the easy access. Some people with MS reported receiving letters months in advance of their appointment, and only noting details such as time and location, and so extra detail of the screening programme may not be noted or forgotten if it is sent too far in advance.

All participant groups noted that some people with MS may not have the ability or access to the technology to complete electronic screening at home, and both support and the ability to complete screening in clinic if necessary is important. A commissioner raised the point that information concerning routine screening is essential, and a programme could need to have a contingency if people could not complete the screening on their own devices at home. One neurologist reiterated that tablets could be made available in waiting rooms for

screening in clinic. There was also the need for support to be made available if any people with MS had difficult with the technology, "*as long as there's a support that I can access and say, 'I can't do this' or 'I don't understand it', as long as that's behind it, that's great. Otherwise, I would probably get frustrated if that wasn't there*" (Person with MS I 04).

There are challenges of administering the cognitive screening in appointments and in clinic waiting rooms. Separate rooms are generally unavailable for most clinics, and any available desktop computers were thought to be inadequate. Clinicians thought that tablets would work best for in clinic administration, "*tablets will work better because it's more portable, so if they are waiting, they could do it while they're waiting*" (Physiotherapist 01).

5.5.3.3 *Role of 'change agents'*

Commissioning was seen as important in the feasibility of the cognitive screening pathway, with one commissioner commenting that "*that's a bit about mapping… what areas do have commissioned services"* (Commissioner 01), as services need to be present currently to cope with the increased referrals to neuropsychology. Clinicians had concerns about the time necessary for the pathway, one MS Nurse thought "*It does sound feasible. It does. But as long as it's not taking fifteen, twenty minutes. Otherwise, it just runs over*" (MS Nurse 07) and that it would have to fit into existing appointment times with patients, otherwise it would not be possible to implement due to time pressures. People with MS thought that the pathway would be worthwhile and feasible, and from their perspective it "*sounds like it's worth following up to find out what's happening*" (Person with MS I 04).

Implementation of programmes may vary across different trusts and services, and commissioning could be an important factor in the success of a screening programme across multiple services. Commissioning decisions were reported to be driven by both NICE guidelines, NHS funding strategies, and strong evidence, including evidence of return on investments. Commissioners stated the need for

required evidence of tangible outcomes, such as employment, and a reduction in service use "*would be more attractive to commissioners if it had the preventative focus*" (Commissioner 03).

> *"You would have to be very clear about what that means, I mean, is this going to stop people ending up having unnecessary attendances or admissions into hospital, or is it going to just keep people more able to live at home for longer on their own… what they'll be looking for is to see improved care with less money"* (Commissioner 03)

Early intervention was viewed positively as it could lead to better outcomes and a reduction in service use "*if you intervene early, it might reduce the numbers that deteriorate*" (Commissioner 02). It may be necessary for some staff to be embedded in community services, to promote a multidisciplinary service. This could ensure that "*you've got opportunity for so much more joint working and communication"* (Psychologist 03).

5.5.3.4 *Perceptions of a proposed cognitive assessment pathway*

The proposed cognitive screening programme was well received by all participants, who thought there would be myriad benefits on an individual and organisational level. Clinicians thought that this pathway would add structure to a system that does not have it currently, and one commissioner thought "*it's the way forward*" (Commissioner 01).

Patient stakeholders noted that "*there's not really anything in place at the moment*" (Charity Volunteer 01) and so having a routine cognitive screening programme would be helpful, and people with MS understood the importance of screening for cognitive problems before they became evident "*We might not realise that we've got problems, you know, because we live with it every day. It's become a way of life, you know, it's the norm for us*" (Person with MS F 02).

One person with MS thought that it would be important to explain what the pathway was, why people are being asked to complete assessments; "*it's not really explained why you're being asked these*

questions and what's the end goal of… but what's the point of it, what benefit are you going to get" (Person with MS I 16) and how help can be accessed if needed "*if you can answer from there that this is why you need to do it and if they need any help, this is the way of getting help"* (Person with MS I 05).

A charity volunteer living with MS said that "*this cognitive thing really frightens me so, you know, it's a huge thing and I haven't been able to get support really anywhere*" (Volunteer 05). One neurologist stated that *"[neurologists] don't go looking for problems that aren't volunteered"* (Neurologist 03) and so questions about cognitive impairment occur once they have developed, rather than being identified early.

One psychologist thought that the tools available to clinicians were often inappropriate for people with MS because "*you're looking at processing speed and executive functions and screens notoriously… don't capture that*" (Psychologist 05). They reiterated that generic cognitive assessments were often inadequate for capturing the cognitive problems faced by people with MS.

People with MS said that they saw the value in routine screening to identify any problems and monitor them regularly.

> *"You don't normally get to see anybody… So, for it to be picked up it would be good because I know it's a long time since I had any tests done and I do know I have got worse since then"* (Person with MS 03).

A carer felt that their family member with MS did not receive satisfactory follow up after they underwent cognitive screening: "*very little seems to happen after that… is that [they are] getting worse and worse and there doesn't seem to be any support there*" (Family Member 03). They felt that having an explanation for why people with MS were being asked to complete the assessments, and what the potential outcomes and actions would be would increase engagement with a

cognitive screening programme; "*She'd have a go at it if she felt it was useful for you or for her treatment*" (Family Member 10).

5.5.3.5 *Beliefs and understanding of cognitive screening*

Clinicians agreed that cognition symptoms were an important part of management of MS, "*the cognition, the sort of memory, thinking, is significant"* (OT 04), and care of patients should be holistic. Participants thought that cognitive impairment should be addressed as part of routine care offered to people with MS to manage MS symptoms as a whole, "*these assessments should just be part of the whole offer, so it's yes it's about walking and it's about lifting and the same, it's about memory and understanding*" (Commissioner 02). It was seen as important to reiterate the benefit of the programme with people with MS to promote engagement. Psychologists and occupational therapists in particular saw the benefit of early identification and management of cognitive symptoms, "*as clinicians within a team we've said right from day one what we want is to capture people early on and provide rehab and some self-management as early as we can, because it delays people developing complex needs*" (Psychologist 04).

However, attitudes of which symptoms take precedence were common. MS nurses showed concern that cognition would be a lower priority than other symptoms; "*bladder and bowel, it can dictate, and it can be horrendous for people. It can be absolutely horrendous. So they are, in a way, they have to be dealt with*" (MS Nurse 07). Nurses also expressed worry that there would be insufficient time to discuss cognition in routine appointments, and cognitive screening could "*open up a can of worms… it would be taking away from that clinic appointment*" (MS Nurse 07). A neurologist felt that it may be difficult to change clinicians' beliefs regarding cognitive screening and "*there will be some clinicians in the country that will not engage, whatever*" (Neurologist 01).

People with MS thought that focusing on cognition and self-help strategies would empower them, help them to understand their

symptoms and provide some assurance that these symptoms can be eased. There was a belief that the programme would be helpful and worthwhile, and it is therefore unsurprising that people with MS were open and willing to participate in the screening programme. But it was important to provide a purposeful follow up to the screening

> *"As long as I know what the end goal is and it's really worthwhile doing, I haven't got a problem with it at all, but if it's something oh do something for the sake of it, I object to that"* (Person with MS I 16)

5.5.3.6 *Barriers and enablers to implementation of the programme*

One clinician thought that the pathway could be positive for both nursing staff and patients, however it would need to be planned carefully, with estimates of the time and workload resources to ensure that staff are motivated. Clinicians emphasised that there may be some difficulty in convincing nurses that it was worthwhile of the change:

> *"Nurses being nurses is it's about trying to encourage people to do things differently and that's one of the big things, so they're used to doing things in a set way and [the programme] is something else"* (Commissioner 02).

Increasing staff engagement may require a shift away from cognitive problems being the role of psychologists and physical issues (for example, medication, bowel, and bladder) towards the remit of MS nurses. A neurologist thought that without evidence that a cognitive screening programme is beneficial for patients, motivation to engage with it could be low, "*some of them will say that my priority list, treating this, you know, is worth my time more than screening for cognition if I have not seen the benefit"* (Neurologist 01).

Participants from all groups raised issues of practicalities of the cognitive screening, and the administration of the programme. There were concerns about the licencing or cost of proprietary assessments, practice effects of the assessments, accessibility for different formats, languages barriers and confidentiality. People with MS also had some

reservations as to how their data would be used, and whether the screening would be tiring to complete. One person with MS highlighted that people may want reassurance as some may be concerned that the Department of Work and Pensions may use results of cognitive assessments as evidence of the individual's disability.

Clinicians spoke of the organisational constraints. Screening programmes could require services to be adapted so they are equipped to deal with cognitive problems. There is a need not just to be able to administer the screening programme and discuss problems, but then have resources to address these.

There was an emphasis that people with MS and clinicians being aware of the assessments, having pertinent information as to why they were being asked to participate, and are told what will happen next:

> "*Say like we'll do this again in three months/six months/year and then we can actually compare and see, you know, is there improvement/decline, anything like that, I think people would be more receptive to that" (Volunteer 01).*

Clinicians agreed that this information would help people with MS engage with the programme.

People with MS found that the pathway was acceptable, and it may help with coping with symptoms of MS, but people commented that some may not want to accept they have cognitive problems and may have difficulty coping. An OT reported that historically there is low engagement with cognitive screening assessment, but this could be mitigated.

> *"We've got to develop the skills and use our language and use our input right from the beginning, with these patients, that actually you know, the responsibility for managing your MS is you. Part of it is filling in some of these screening tools that we've got for you know" (OT 04)*

5.5.4 **Outcomes**

5.5.4.1 *Results of cognitive assessments*

Clinicians stated that ideally results would be automatically embedded within NHS databases, however if this is not possible, results could be emailed across in a format compatible with NHS system (for example, PDF, Excel spreadsheet or a Word document), to be uploaded to the database manually. There was some concern that results may got lost or fail to be sent to the correct person to administer the results, and few clinical teams have a central inbox for results to be sent to. A psychologist stated the need for meaningful scoring and benchmarks against premorbid functioning, however one physiotherapist stated that cognitive screening could open conversations about cognitive functioning, "*if somebody's already highly functional and they score 55 and then they start telling you that there are changes*" (Physiotherapist 01).

All participant groups thought it was important for the results to be given in-person, by a trained healthcare professional. People with MS stated that the results could help them understand their MS better. People with MS and their family members reported that results should facilitate discussions about cognitive problems

> *"I'd want to know exactly how I've done and would it highlight any deterioration. And I like to hit things head on and I like to understand what's happening. Now not everybody wants that… I like to understand exactly what's going on inside my body and my brain"* (Volunteer 03).

Results should also focus on severity and progression since their previous assessment, and people with MS should have the opportunity to discuss the results with their clinician. Clinicians reporting the results to patients was considered to potentially be upsetting and arduous; some people with MS thought worsening results could set back patients' progress and clinicians should bear in mind the sensitivity of such discussions.

Clinicians wanted the results to be clarified in terms of the domains of cognitive impairment, and being provided with an interpretation of the results (such as a categorisation into severity), and a raw score. Clinicians also reported wanting to see patients' results of related symptoms, such as fatigue and depression, to facilitate interpretation of the cognitive assessment results in the context of these, "*what's causing what, so is it the cognitive decline that's causing the depression or the depression that's having an effect on the cognitive decline"* (MS Nurse 01).

Clinicians reported that having the results of screening tests would be useful to facilitate discussions and increase understanding of cognition both between clinicians, for example in multi-disciplinary teams, and between the clinician and the patient.

5.5.4.2 *Referral pathway*

For the purposes of referral, the results that clinicians receive would need to factor in reported mood and fatigue, and self-reported cognitive function, and the objective cognitive screening results. This would enable clinicians to discuss the best course of action with their patient and make a referral to the most appropriate service.

> *"People can fly through, you know, people who are pre morbidly pretty bright and fly through screens without too much of a problem, but actually they're still experiencing some kind of cognitive decline or difficulty, so I just wanted to make sure that you're capturing that group"* (Psychologist 04).

Psychologists expressed concern that risk that only people with MS with 'severe' scores would be referred to neuropsychological services, when in some instances, the 'mild' or 'moderate' cognitively impaired individuals could potentially stand to benefit from more specialist input also.

5.5.4.3 *Perceived outcomes of the programme*

The proposed cognitive assessment programme was well received by all participants, who thought there would myriad benefits on an individual and organisational level.

Commissioners showed enthusiasm for the proposed programme, stating that it could lead to "*increased knowledge, better treatment, decision making, clear referral*" (Commissioner 02). Clinicians reiterated that patients would benefit, and through an increase in knowledge and understanding of cognitive problems in people with MS could reduce care costs and time. There could be a beneficial impact on commissioning decisions by informing healthcare commissioners of the patient and clinical outcomes, and lead to MS services operating more effectively. If clinicians knew what cognitive problems people with MS were experiencing,

> *"it'll be fantastic if it is digital because I think then you will come out of it with a real product that will be really bespoke and fantastic for people with MS, it could change their lives and the lives of clinicians. So, you know, it'd make everything better when it comes to quality of care."* (MS Nurse 06).

The implementation of assessing patients' cognition at all stages of the progression of their MS could lead to the earlier provision of cognitive rehabilitation, which in turn could promote self-management, enabling patients to live well for longer.

> *"it's a really good idea to get in early to provide any sort of rehab and kind of saying that from a self-management perspective, if you get in early provide people with the skills and the tools, they should be able to manage by themselves for as long as possible, or for longer than they might if they didn't have that"* (Psychologist 04).

People with MS and family members also perceived that the programme would raise awareness of cognitive problems in MS and had the potential to improve their care and lead to better understanding

of symptoms of MS, "*it can only be seen as progress really and any kind of support covering a different area of MS is always a good thing*" (Volunteer 03).

Participants from all groups acknowledged that other symptoms such as mood and fatigue are associated with cognitive impairment and stated that it was appropriate to also assess and manage these symptoms at the same time as cognition. Clinicians argued that other support would be needed for other apparent symptoms, alongside the cognitive symptoms, *"what are we going to do with the ones who score purely cognitively but also score high for anxiety and depression*" (Neurologist 03). One occupational therapist reported that as part of the MS cognitive rehabilitation offered by their service:

> *"[clinicians] talk about the impact that that has on cognition, which is why we tend to target fatigue first, to try and see if we can reduce levels of fatigue, to then influence any other domains" (OT 06).*

Clinicians highlighted that a potential challenge of the programme may be the increase in workload and provision of services. It would be likely that more cognitively impaired people with MS would be identified through the screening, especially those with mild and moderate cognitive problems, which may "*increase the nursing workload quite significantly*" (MS Nurse 01). This could also impact commissioning as "*services are receiving more and more referrals because people have done an online screening*" (Commissioner M01), which services may not be able to effectively manage.

5.6 Discussion

All participants showed awareness that cognitive impairment was a symptom of MS, and most people with MS reported experiences of cognitive impairment. The impact of cognitive impairment was varied, but many people with MS explained how it had affected their employment, relationships, the management of their MS, and quality of life. People with MS perceived that all aspects of their memory

(including working, prospective and episodic memory), attention and concentration, processing speed, and language skills were affected. Clinicians agreed that these cognitive domains were altered. It was also clear that cognitive symptoms were closely related to mental health issues and fatigue.

Like Mortensen et al. (2020), we found that routine cognitive screening was seen as valuable for both people with MS and clinicians. Both objective and subjective cognitive assessments were viewed as appropriate. However, despite the prevalence and perceived importance of cognitive impairment, both clinicians and people with MS in our study did not routinely discuss this during consultations, and referrals to specialist services were infrequent. Time constraints of appointments, and physical symptoms (such as bowel and bladder problems) took precedence in appointments over cognitive problems. Very few people with MS had received a referral or assessment for cognitive problems. Clinicians also reported a lack of staffing and resources, such as private rooms in MS clinics, making administration of cognitive assessments challenging. Clinicians' disinclination to discuss cognitive problems in routine appointments could be because of the lack of appropriate cognitive assessments or management clinics/services available to refer their patients to. Clinicians, such as neurologists and MS nurses, were thought to be key drivers in whether a pathway would be commissioned.

The existing situation suggests that the implementation of routine cognitive screening would need to take these resource limitations into consideration. Participants agreed that self-administered electronic screening completed in patients' homes would mitigate some of these obstacles.

A screening programme would need to be managed carefully to ensure sustainability within the UK healthcare system. Participants were clear that a cognitive screening programme should be administered with a rationale as to why people with MS were being asked to engage with it, and what would happen after the screening in

terms of getting results, and potential cognitive treatment or management that could be offered.

It is critical to establish which forms of assessments people with MS would find acceptable to complete, and clinicians to administer. The paced addition style task was the only assessment that people with MS viewed negatively. This was echoed by clinicians, who reported that participants often did not want to complete this test, and is consistent with research findings (Cortés-Martínez et al., 2019). Other assessments, such as symbol substitution style tests, were more widely accepted, and an enthusiasm for self-administration is also reflected in other studies (Patel et al., 2019). People with MS showed an understanding of the rationale of the assessments and thought they reflected relevant cognitive domains they had difficulty with but would not be so challenging that they would not want to undergo the test. Some people may decline assessments they are asked to complete outright, and an assessment programme needs to take this into account. Noncompletion objective assessments could lead to cognitive impairments not being identified, and not receiving appropriate care if people with MS are experiencing cognitive problems. This suggests that clinicians should have other methods of identifying impairment, such as self- or proxy-report questionnaires, and discussions with their patients. Despite their limitations, self-report questionnaires, such as the MSNQ-P, may be less intimidating than objective assessments of cognition. Some people with MS also thought that people may feel anxious about discussing cognitive problems, and may not want to examine this sensitive topic, either in their medical appointments or in conversations with researchers.

Electronic self-administration at home was acceptable to all participants and reflects recommendations for brief screening in MS (Kalb et al., 2018). Self-administration could protect clinic time, and for people with MS, reduce the time spent at the hospital. However, care should be taken to ensure that a lack of technology, or an unwillingness or inability to complete the assessments, does not preclude people with

MS in receiving treatment for cognitive problems. The development of an electronic, self-administered cognitive screening assessment would require careful consideration and evidence of its validity, and new normative data for interpretation (Bauer et al., 2012). UK healthcare is moving towards a digital future, and the role of the patient is important to ensure that the technology meets their needs (Topol, 2020). It is likely that due to changes in neurology clinics during COVID-19, and for the foreseeable future, telemedicine and digital assessments will be adopted more frequently (Mummery & Kipps, 2020). Traditionally, the administration of neuropsychological assessments is restricted by test publishers to those with specified neuropsychological expertise and experience (Puente et al., 2006). However, digitisation can automate the administration and scoring of screening assessments, and widen administration to include MS nurses, and other clinic staff. This could overcome some of the resource constraints that were reported by clinicians in coordinating and assisting patients to complete cognitive screening assessments. This will also be facilitated by other aspects of digitisation such as sending a link to patients, and directing digital results to clinicians, which could form a part of routine administration by staff. However, formal neuropsychological assessments should continue to be administered, scored, and interpreted only by those trained to do this.

These conclusions are limited by some methodological limitations. The overarching research programme of NEuRoMS encompasses a proposed pathway of both cognitive screening and management of cognitive problems provided in routine care within the UK NHS. This chapter has focused on the cognitive screening element, but it is important to note that participants were aware that screening forms a part of a pathway that could offer cognitive rehabilitation as opposed to screening presented as a discrete undertaking. In addition, the focus of the pathway's implementation within the UK healthcare system and we do not know how this would translate to different countries, and/or different systems.

We presented seven different assessments that are frequently used in MS; however, this is not the full range of assessments in-use. We recruited people with MS using an opportunistic sample from our PPI network, promotion of the research through MS charities, advertisements placed on social media sites and word of mouth, and clinicians from our professional networks. This, like in most qualitative research, is not (meant to be) a representative sample. In our PIS, we stated that interviews lasted for up to an hour, and this could be unappealing, or not feasible for people with MS with severe symptoms, both physically and cognitively. We attempted to allay this concern by offering to conduct interviews at participants' homes, or remotely via video conferencing or telephone. Similarly, we intended to recruit a greater sample of both clinician and people with MS to conduct several focus groups with people with MS, and clinicians. We found that participants, particularly clinicians, preferred interviews, which had more flexibility in terms of time and place, than a focus group. The focus group had a smaller number of people with MS (five) than we wanted to recruit. However, this study benefited from a mixture of interviews and a focus group, a large sample, and a range of stakeholders.

5.7 Conclusions

The results of this interview study show that brief cognitive screening with simple assessments, such as symbol substitution tests, for routine administration is both feasible and acceptable for people with MS and clinicians. However cognitive assessments should be managed carefully. Patient preference indicates self-administration and electronic assessments completed on their own handheld device could overcome barriers such as limited clinic resources. However, there are challenges with this approach, including adapting tests successfully to this modality, and these need to be considered carefully.

The next stages of the research explore the development of a digital cognitive assessment platform by determining its usability, as the

first step to ascertaining its validity. Using Human Computer Interaction approaches to place the user experience at the heart of cognitive screening technology development, the next chapter uses cognitive 'think-aloud' interviews, based on Technology Acceptance Model. This aims to establish whether people with MS are willing and able to engage with a cognitive screening platform developed on the findings in the current chapter. I then propose a further study to collect normative data and determine the psychometric properties of digital assessments.

Chapter 6 Usability Testing of a Cognitive Assessment Platform for Use with People with Multiple Sclerosis

6.1 Chapter overview

This chapter presents a cognitive interview study that used a modified think-aloud approach to assess the usability and acceptability of an assessment platform hosting a digital symbol substitution task. The introduction presents a background on digital cognitive assessments developed specifically for people with MS, outlines issues related to converting assessments to a digital format and provides rationale for the chosen methodology of this study.

The digital symbol substitution task developed used findings from previous chapters in this thesis. Interviews with people with MS recorded any problems experienced and reported by participants, which were then used to inform the next iteration of the platform. I describe the results of these interviews including technological problems of the platform, the context of the assessment, and any barriers or enablers to completion.

6.2 Introduction

Previous chapters have examined how the Symbol Digit Modalities Test (SDMT) can be interpreted, and its relationship with other measures. The interview study found that it is acceptable for both people with MS and clinicians, and a qualitative inquiry found that the SDMT was one of the least demanding and fatiguing of the three tests in the Brief International Cognitive Assessment for MS (BICAMS; Mortensen et al., 2020).

The study described in Chapter 5 found that providing digital assessments can have a number of benefits, including being able to provide assessment to a greater number of patients, in a more accessible manner, and in a more cost-effective method than traditional administrations that require purchasing materials per participant (Bauer et al., 2012). Both general recommendations of digital health

management such as the Topol Review (Topol, 2020) and MS-specific guidance on cognitive assessment (Kalb et al., 2018) have emphasised the potential for technology to have a positive impact on MS clinical care. In 2006, a paper promoting a common framework for cognitive assessment stated that the "*future of cognitive assessment lies in the integration of computerized tests into clinics and centers treating MS patients*" (Achiron & Barak, 2006, p. 49). However, fifteen years later, this has not come to fruition. The COVID-19 pandemic pushed necessary and rapid changes in telemedicine, such as video conferencing for routine appointments. The successes of some of these initiatives have opened a discussion as to why telehealth was not implemented before this crisis, and how it can be implemented proactively to the benefit of patients and clinicians (Smith et al., 2020). The interview study in Chapter 5 demonstrated that both people with MS and clinicians found that completing digital administrations of cognitive assessment on personal mobile devices to be acceptable. This would also fit into the limited resources of MS clinics, such as time, staffing and privacy. There have been digital translations of assessments for use in MS research and clinical care, and these have highlighted some of the difficulties associated with this development process.

6.2.1 **Digital cognitive assessments for MS**

Assessments that have been digitised for use with people with MS have largely used technician-administered computer assessments (Wojcik et al., 2019). There has been a development of a digitised iPad tablet version of the BICAMS to be deployed in-clinic, the iCAMS (Beier, Alschuler, et al., 2019). Finding difficulties in licencing the second edition of the California Verbal Learning Test for use in the battery, the authors replaced this test with the Rey's Auditory Verbal Learning Test. Whilst they found strong and significant correlations between the pen-and-paper and tablet administrations, the tablet version is more for the technician convenience. The SDMT is still conducted orally, with the technician using the tablet to record correct

responses. However, there is still an advantage that the scores are automated and require less manual interpretation.

Self-administered cognitive assessments that can be completed remotely, such as on tablets or smartphones, have been found to be acceptable to people with MS (Patel et al., 2019) and have several benefits (Settle et al., 2015). Whilst tablet-based assessments are being developed, they are mostly restricted to certain devices such as iPads (e.g., Middleton et al., 2020; Rao et al., 2017; van Dongen et al., 2020), which may preclude people with MS from completing assessments if clinics or people with MS do not have a specific device.

Another consideration is the input mechanism for a digital test. An attempt to digitise a verbal version of the SDMT using voice recognition reported that participants including people with MS had a preference for completing the computer-based verbal SDMT over the technician-administered verbal SDMT (Patel et al., 2019). However, when another study gave participants the option to complete a mobile version of the SDMT verbally or by inputting their responses manually, very few participants chose the verbal version (Pham et al., 2021). The authors of this study also noted that the automated voice recognition often failed for people with MS with dysarthria, a common symptom in MS (Noffs et al., 2018). There are also technical implications for a verbal input, requiring much more technical development and testing. Digital developments of cognitive assessments have more considerations than deploying traditional pen-and-paper assessments.

6.2.2 **Digitisation of cognitive assessments**

A systematic review identifying barriers to telemedicine drew attention to computer literacy and access to technology, age and education of patients, and concerns about privacy and confidentiality (Kruse et al., 2018). Extrapolating from trends in technology (Office of National Statistics, 2020), many of these issues (such as access to technology) will become redundant. A scoping review of digital cognitive assessments used in ageing found that using mobile assessments

increased the rate of completion, even in ageing populations with less confidence in technology (Koo & Vizer, 2019). Digitisation can enable automatic integration of results into electronic health records and instantaneous scoring and interpretation of results. However, there are barriers to overcome with digitisation of assessments.

One of the proposed methodologies for overcoming some of the technological challenges is due consideration for how the original task can be deployed on a digital device (for example, if aspects of the task can be adapted to optimise its digital administration), and robust testing of the user interface (Germine et al., 2019). This can include changing the instructions from written to visual or video examples and include practice tasks so participants can familiarise themselves with the interface.

Another barrier identified by Kruse et al. (2018) was that of application design. If developers of digital health implementation do not examine insight into the user experience and impact on patients and clinicians, engagement and uptake may suffer. A recent review from a Human Computer Interaction (HCI) perspective emphasised the importance of an iterative design process with user feedback to entrench user experience within patient-facing digital health initiatives (Blandford, 2019).

6.2.2.1 *Technology acceptance and usability testing*

Developers of digital health initiatives have encouraged user experience to be built into patient-facing technology during the early stages of development, and developers can look to HCI research to ensure usability and acceptance (Blandford, 2018). For neuropsychological assessments, clinicians and patients perceiving the benefits, and their acceptance of the technology may be critical in effective implementation.

The Technology Acceptance Model (TAM; Davis et al., 1989) states the importance of user acceptability in intention to use a particular technology. This model asserts the value of users being able

to perceive the effectiveness of the technology, and the ease of using it. Ensuring a user-centred design can increase technology uptake and lead to a successful implementation upon deployment. A guide to digital health interventions again exemplifies the role of the end-user in its development (Bradbury et al., 2014). Although this is focused on digital interventions, much of the methodological process can be applied to digital assessment.

Blandford (2019) emphasised that successful digital health developments need to be underpinned by three elements:

1. A robust understanding and application of the scientific principles to the technology.
2. The technology to be conceptually meaningful, engaging, and accessible to the user.
3. The technology must not cause undue burden or emotional stress to the users.

The proposed development cycle can be used to conceptualise the stages of development.

The first of these elements has been fulfilled through the work detailed in this thesis, and the involvement of experts in neuropsychology and technology development for psychology as part of the National Institute for Health Research funded programme Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS). The next two elements are explored in the research described in this chapter.

There are a multitude of user-centred design approaches that can be used to evaluate digital health technology. Methodologies such as usage data, diary studies, and surveys of users and clinicians (all of which are described in Kuniavsky et al., 2012), can add insight as to the acceptability of technology. However, to capture the perceived usefulness and perceived ease of use (as described by TAM), and enable an iterative development of the technology, a qualitative ‘think-aloud’ approach was adopted in the current study.

6.2.3 **Development of a cognitive assessment platform**

The SDMT is a licenced pen-and-paper test, and the digital development of a conceptually similar symbol digit substitution task is considered an original test (Bauer et al., 2012) and is therefore subject to further validation. The development of this test described in this thesis has been given the name Symbol Substitution Test (SST) to distinguish it from the SDMT. The SST was developed and integrated into the platform as: i. the test it is based on, the SDMT, is considered a valid and reliable cognitive screening task for use in MS (Benedict, et al., 2017); ii. in our interview study, both people with MS and clinicians reported that it was acceptable and iii. it is straightforward to implement on handheld mobile devices.

The SST has been deployed on the Discrete and Ellipse-based Response Capture System (Ellerby et al., 2019; DECSYS); a webapp platform that was designed through a collaboration of MindTech and the Digital Research Service of the University of Nottingham. It has been developed to host several neuropsychological tasks that can be completed remotely by people with MS as part of their routine care.

The first step in evaluating the SST is ascertaining its usability and feasibility for use as a self-administered, device-agnostic test. This will substantiate the findings of the qualitative inquiry as to the acceptance of smartphone deployment of assessments. Previous research, consensus recommendations, and the systematic review and primary qualitative research described in Chapters 2 and 5 contributed towards the first step of this development cycle, through identifying the problem and planning the next stages. This chapter explores the next phases of the lifecycle: design, evaluation, and testing. This chapter explores the next phases of the lifecycle: design, evaluation, and testing. The methodology used to evaluate and test uses a qualitative cognitive interviewing approach to ascertain the usability, feasibility, and acceptability of the digital assessment and platform, as recommended as the first step for digital and mobile health developments (Bradbury et al., 2014). Using this approach, which

draws from both health and HCI research, this chapter places the views and experience of people with MS at the core of the digitised assessment development.

6.3 Aims

There were two phases, each with different aims:

Phase I:

1. To corroborate the findings of the interview study in Chapter 5 that mobile deployment and symbol substitution tasks were suitable for cognitive assessment
2. To devise an interview schedule and pilot the cognitive interview technique.

Phase II:

1. To explore perceptions of the usability, accessibility, and acceptability of a symbol substitution style task (the SST), and a platform to host this test.

6.4 Methods

6.4.1 Ethical considerations

Ethical approval was obtained through University of Nottingham's School of Medicine ethics committee (Reference number: 372-1908). All participants gave their verbal consent and were provided with the Participant Information Sheet (PIS).

Data uploaded and transcribed by the Automated Transcription Service (ATS) were held securely and deleted after 90 days. All files associated with the project, including audio files, transcripts, demographic sheets, and recruitment database with password protection. Any identifying details were removed from the transcripts, and participants were assigned a pseudonym (e.g., Person with MS 01)

6.4.2 **Design**

There were two phases of interviews in this study. Both used a cognitive interviewing ‘think-aloud’ approach to explore the usability, and enablers and barriers of accessing and completing digital cognitive assessments.

6.4.3 **Participants**

6.4.3.1 *Phase I – Pilot interviews*

Six people with MS were recruited from the University of Nottingham’s MS PPI group. Inclusion criteria required participants to be aged 18 or over, could communicate in English, and could give verbal consent.

6.4.3.2 *Phase II – Cognitive interviews*

We recruited eight people with MS to participate in the cognitive interviews. Inclusion criteria required participants to be aged 18 or over, could communicate in English, had access to video conferencing software on a computer or device, and a separate handheld, touchscreen device (a smartphone or tablet) to access DECSYS, and could give verbal consent. The sample size was reviewed iteratively, and we stopped recruiting once we believed that we had reached data adequacy (Vasileiou et al., 2018) for this stage of the development , with no new issues being raised.

We used a purposive convenience strategy to recruit participants. Participants were invited to take part through advertisements in newsletters and on social media (including Facebook, and Twitter), and approached the University of Nottingham’s MS PPI group for participants.

6.4.4 **Materials**

6.4.4.1 *Symbol Substitution Task*

Two iterations of the SST were used. In Phase I, participants completed a prototype test that was developed close to the original SDMT (shown in Figure 5). This was developed in Unity in the C#

programming language and were deployed on a laptop, large tablet and Samsung S9+ smartphone. This prototype showed a key of symbols and numbers from 1-9, and a grid of 90 symbols. Participants had to type or tap the number associated with each symbol.

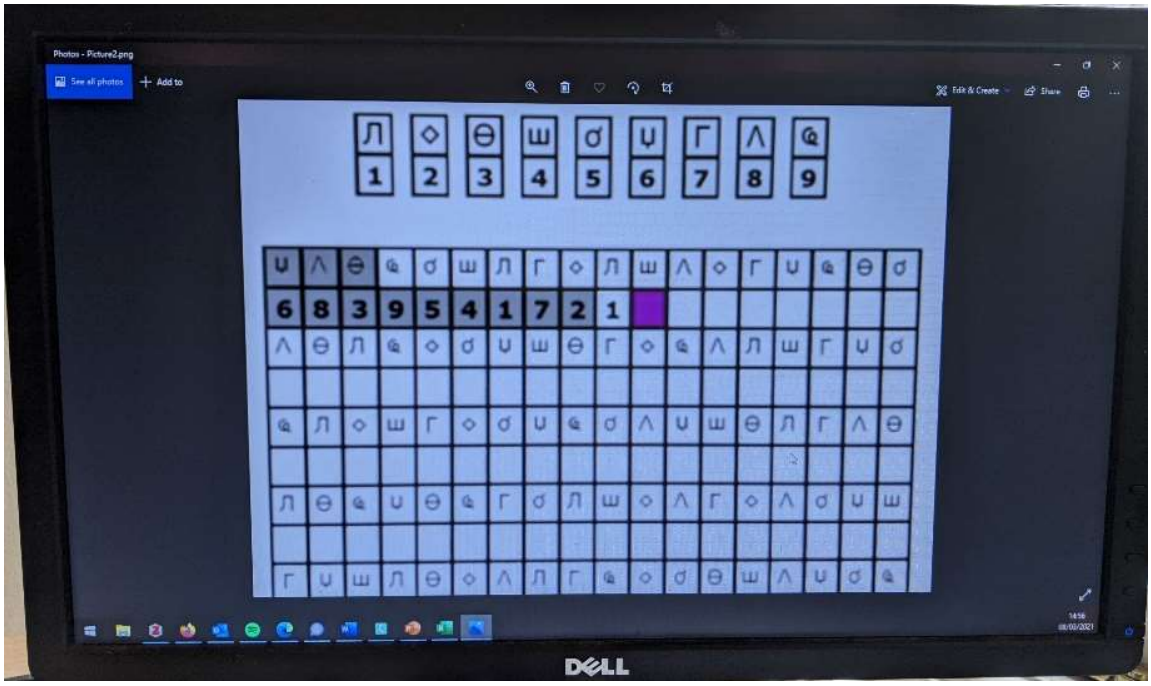


Figure 5: Prototype of a symbol substitution task deployed on a PC

The next iteration of the SST used in Phase II was developed based on a similar task developed for an iPad administration (Middleton et al., 2020), however a number of changes were made through feedback from the wider research team, and to optimise the task for smaller devices. The iPad deployment showed participants the next stimuli they would be presented with below the current stimuli, designed to mimic showing future stimuli in the original SDMT test. However, concerns were raised during informal testing within the research team that the presentation of both the current and next symbol was confusing.

Instructions for the task were written based on the original SDMT instructions and digital deployments and reviewed by PPI members to ensure clarity. The instructions were displayed across two screens, with a visual example of the SST on each. Screengrabs of the instructions

as to how they appear on two different mobile devices are shown in Figure 7 and Figure 8 in Appendix 14.

The task used in this study showed participants a key which assigns a symbol to numbers 1-9. Participants are presented with a Unicode symbol in the middle of the screen, and they are asked to match the symbol with a number using the key and tap the correct number from a number pad at the bottom of the screen. Once they have tapped a number (correctly or incorrectly) for the symbol, a new symbol appears immediately. Participants are given a 20 second practice task, and then 90 seconds for the scored task.

6.4.4.2 *DECSYS Platform*

The DECSYS platform (Ellerby et al., 2019) was developed as an open-source survey tool by the Digital Research Team at the University of Nottingham. Although initially designed to use Likert scale responses, its functionality was expanded to add cognitive tasks. It was developed in JavaScript and HTML5 and runs as a native web application. The DECSYS platform has been developed on an iterative basis, from early prototypes that could be deployed on a full range of devices, to the final platform on participants' own handheld devices.

The platform consists of several elements; a login screen to input a user ID, an introduction page, demographic questions, task instructions and examples, a practice task, the full task, and a completion confirmation page. As it is a webapp, it can be completed on a computer, tablet, or smartphone and it only requires an internet connection through a browser to access.

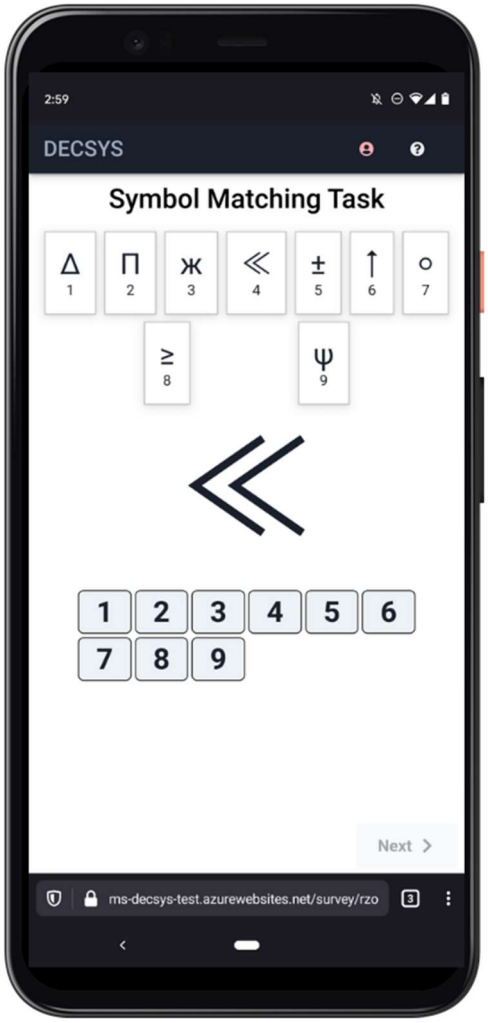


Figure 6: DECSYS mobile platform showing the Symbol Substitution Task

6.4.5 **Procedure**

6.4.5.1 *Phase I – Pilot interviews*

Participants in the pilot interviews were invited to attend an in-person interview at the University of Nottingham. After reading the PIS (Appendix 10) and giving their consent, the researcher provided the participant with a phone, tablet, and/or a laptop with which to complete a prototype of the SST on. Participants were encouraged to verbalise any thoughts they had as they underwent the task.

6.4.5.2 *Phase II – Cognitive interviews*

Once participants expressed interest in taking part in the research, researchers sent the PIS (Appendix 10) and ensured their eligibility to participate. Due to COVID-19 restrictions on in-person data collection,

all interviews were conducted through a video conference hosted on Microsoft Teams. Before the interview, we ensured that participants had access to a handheld mobile device with access to the internet that they could use during the interview. Participants were sent a link to join an online meeting through email and a link to the DECSYS platform.

During the interview, participants were first asked for their consent using the verbal consent form (Appendix 11). We confirmed that participants were able to use the link to the platform on their own handheld device, and if not, the researcher resent the link through email or the preferred means for the participant to access it. We asked participants demographic questions, including their age, gender, type of MS and how long they had been diagnosed.

We followed a semi-structured interview schedule (Appendix 12), and this was modified on an iterative basis, with probes used to clarify any answers. We began the interview by asking about the participant's experience and comfort with technology, and if they had any problems with their manual dexterity or vision, and if this had any impact on their access to technology. The participant then accessed the cognitive assessments using their personal mobile device, and we encouraged them to talk through their experience, and express any impressions they had about the platform as they went through it. Once the participants had finished the task, we asked further questions about their experience and thoughts of completing the task and using the platform.

Participants were offered a £20 shopping voucher for their participation.

6.4.6 **Patient and public involvement**

Members of the PPI group at the University of Nottingham provided input on all stages of the usability interviews, including reviewing the proposed methodology, and all participant-facing materials.

6.4.7 **Cognitive interviewing procedure**

'Think-aloud' protocols in cognitive interviewing typically involve participants concurrently verbalising their thoughts as they use a technology, to identify issues encountered by users and comprehend how participants are completing tasks (Cooke, 2010). Concurrent think-aloud verbalisation can be an unnatural experience, and may impact the user's experience of using the technology under investigation, and in such cases a hybrid model of concurrent verbalisation and retrospective questioning may be more appropriate (Alhadreti & Mayhew, 2018). In Phase I, we used a concurrent approach, however we used the hybrid model in Phase II, with probing questions whilst participants navigated the technology to facilitate the interview.

6.4.8 **Transcription and data analysis**

The pilot interviews were audio recorded, and cognitive interviews were audio and video recorded through Microsoft Teams. All recordings were converted to MP3 files and an initial transcription was conducted through the University of Nottingham's ATS. These transcripts were read carefully in comparison with the audio recording, and errors were manually corrected. The text files were imported into NVivo software (version 12) for analysis. Minor changes that could be amended on an iterative basis were captured using a change log table.

An *a priori* framework with which to analyse the data was constructed using O'Connor et al.'s (2016) Digital Health Engagement Model (DIEGO). This uses four overall themes: i. Personal agency and motivation, ii. Personal life and values, iii. Engagement and recruitment approach, and iv. Quality of the digital health intervention.

6.5 Results

6.5.1 **Pilot interviews**

Six people with MS recruited from a PPI group attended pilot interviews held at the University of Nottingham. We did not record demographic information for these participants. Participants confirmed

that they preferred cognitive assessments deployed on mobile devices to computer administrations. Participants also stated that they would prefer to complete assessments in their own homes, on their personal devices.

Whilst the cognitive interview procedure was mostly suitable for the usability testing, we found that participants were verbalising the responses they were choosing for the tasks (for example, saying aloud the numbers that matched the symbol) rather than their thought process whilst completing the task. We reiterated to the participant that they were not being assessed, and we incorporated concurrent probing questions throughout each part of the platform, and retrospective questions after they had finished interacting with the platform. On the basis of these results, we modified the 'think-aloud' approach to incorporate more probing questions, and revised the interview schedule.

6.5.2 **Cognitive interviews**

Eight people with MS took part in the formal cognitive interviews. Their demographic information is given in Table 13. Interviews lasted 30 minutes on average.

The framework analysis was amended on an iterative basis to analyse the data based on three of the four DIEGO themes (i. Personal agency and motivation, ii. Personal life and values, and iv. Quality of the digital health intervention). The theme of "engagement and recruitment" was not relevant to this analysis. The names of the themes were changed for greater relevance to the data, and reordered, and are presented as: i. Personal context, ii. Quality of the digital assessment platform iii. Enablers and barriers. Summaries for each participant under these themes and subthemes were written, followed by overall summaries.

Table 13: Sample characteristics information of people with MS interviewed

Age (in years)	*Mean*	50.63
	Range	36-62
Years Since Diagnosis	*Mean*	12.94
	Range	2.5-20
Gender	*Women*	6
	Men	2
Ethnicity	*White*	7
	Asian/Asian British	1
Highest Education	*GCSE*	2
	A Level	1
	Degree	1
	Higher Degree	4
Employment	*Employed Full Time*	2
	Employed Part Time	3
	Retired	2
	Voluntary Part-Time	1
MS phenotype	*Relapsing-remitting MS*	4
	Primary Progressive MS	1
	Secondary Progressive MS	3

6.5.3 **Personal context**

This theme covers the personal context surrounding accessing the platform and completing the task. This includes digital literacy and technology access, MS-related problems that may hinder technology use, and privacy and security.

Seven participants used their smartphones to access the platform (four using an Apple iPhone, and three using an Android phone), and one participant used an Apple iPad. All participants were comfortable with using their mobile device and technology in general, assuming that "*most people are going to be using this out with fairly savvy in the fact that they are using device at all*" (P04). One participant said that they were *"so non-techie*" (P08) although they did not encounter any difficulties.

No participants had problems with vision that were not corrected by glasses or contact lenses, but some people said that they have previously had MS-related visual problems, such as optic neuritis which may affect how they experienced the platform. Although half of the

participants reported difficulties with their dexterity that affected their use of their device, none used aids or adaptive technology to mitigate these problems. One person with difficulty with their manual dexterity thought that they were slower at responding to the stimuli during the SST due to their dexterity, rather than their information processing speed, saying that their "*brain goes faster than [their] fingers"* (P08).

Cognitive problems were an issue for most participants, and these could be worse on some days, at certain times, and with distractions, being unable to "*concentrate on two things at once, even something quite minor*" (P01) and expressed concern these may affect their results. Participants thought that the provision of additional context would be useful in the email or at the beginning of the access to the platform, to ensure that they would complete the assessment in a suitable environment. Participants stated that they would ensure that their environment was appropriate, for example quiet and free of distractions, before accessing the platform. One participant emphasised that they would "*sit down out of the way… I'll do it when it's quiet and I'm hydrated*" (P01).

6.5.4 **Quality of the digital assessment platform**

Quality of the platform and test included issues related to the usability, including ease of accessing and navigating the platform, the instructions and completing the task, and the look and feel of the platform. This theme is described through subthemes related to the distinct stages of the platform, and the easy of using the platform.

6.5.4.1 *Accessing and navigating the DECSYS platform*

One participant did not have access to their email account on the device they were using, and therefore required brief guidance on how to access their email on their tablet. All other participants were able to access the link from the email sent by the researcher with ease. When asked to enter a unique identifier in the form of an adjective and animal (e.g., SmellyRabbit), participants were able to type their ID in easily,

and they viewed this configuration of words positively as it was easier to parse and type than random characters.

During stages of the interview, two participants highlighted problems with their devices while completing the task. One participant's smartphone turned off multiple times due to low battery or a faulty device, and the another received multiple emails and text messages during the task. The participants said that "*during the course... I'm going to see random emails to be popping up slightly open*" (P04). This distracted them from the task, despite not opening the messages.

Some of the participants reported that they were using their device differently during the interview. For example, one participant had their device on the table in front of them but recognised that they would normally hold their smartphone in one hand.

6.5.4.2 *Look and feel of the platform*

Participants felt neutral about the look of the platform, saying that "*it wasn't amazing… it was just fine*." (P01). Two participants thought that it could be visually more interesting, with colour, but also understood that the platform should be as accessible as possible. Overall, the platform looked and functioned well, and participants were able to navigate through the platform. Most of the participants noted an issue with the layout of the screen and the location of the 'next' button that will be fixed in the next iteration of the platform development and detailed in the change log in Table 14.

6.5.4.3 *Instructions provided*

Whilst mostly perceived as straightforward and clear, participants raised some issues with the instructions that may hinder completion of the task. The most significant of these was the lack of instructions as to whether respondents should complete the task as quickly or as accurately, and an uncertainty about how frequently the symbols appeared. Participants differed on their perceptions as to speed vs. accuracy, some thinking "*it has to be as accurate as possible"* (P05), and some viewing it being scored as "*the faster you are going"* (P08).

One participant showed a misinterpretation about how often the symbols would appear on the screen:

> *"I also don't understand why there's a disparity between the 20 second practice time and the 90 second completion time. Unless you want me to do more than one task as you want me to match more than one symbol"* (*P03*).

With regards to the two screens, and two examples of answers, participants though these were clear, saying that with the second example it "*helps reiterate what you're doing and also making you prepared for the practice*" (P02), and highlights that the symbol key changes. However, a couple of issues of distractions were raised: one participant ignored the text preceeding the example, looking "*straight at the symbols and numbers, as opposed to reading the text as to what I should actually be doing*" (P02), and another finding the red text and highlights on the example "*a little bit too much to me"* (P08). We asked participants if a video explaining the instructions would be effective, but participants thought the text and images were clear enough, and allowed them to understand at their own time, with one participant saying "*a video won't be helpful because then you'll continue continuously be pausing the video… with still pictures like this, I can take my time*" (P05).

6.5.4.4 *Symbol Substitution Task*

Participants found that, overall, the SST was straightforward, easy to complete, and took an acceptable amount of time to complete. The practice task was viewed positively as it helped participants to orient themselves to the task, finding that "*it's quite useful having that practice*" (P02). The length of the practice and real tasks (20 second and 90 seconds respectively) were seen as reasonable and "*probably quite a good length time*" (P01) saying "*it's not too long to have to concentrate for*" (P03).

Once the task was complete, we asked participants what they perceived the task to be measuring, and they expressed a range of

different responses. Participants thought that the test assessed both cognitive and physical problems. One participant thought it was evaluating their visual acuity and manual dexterity, and other participants perceived it to be assessing their short-term memory, recognition of symbols and remembering them when looking to the number key, processing speed, and their concentration. Two participants thought that the task was not particularly relevant to their MS symptoms. Participants reported that they responded to the symbols quicker as the task went on, but one participant was frustrated that their vision and dexterity slowed them down. This participant expressed a preference for answering verbally to record their responses, a "*voice-activated kind of exercise*" (P08) to overcome their manual dexterity difficulties.

6.5.5 **Barriers and enablers to use**

This theme explores the perceived benefits that participants identified, and any potential barriers and enablers to navigating the platform, completing the assessment, or the assessment programme more generally. This included awareness and understanding of the technology.

All participants thought that the assessment would be acceptable to complete as part of their clinical care, and it was reasonable to complete the assessment on regular basis (for example, every 12 months), linked with their neurology appointments, and that "*people [will] get used to it being part of the check-up*" (P02). Most participants saw the benefit in repeated assessment, with one participant saying that "*it doesn't matter how you perform relative to anybody else, it's how you perform relative to yourself*" (P04).

Overall, participants thought that the platform and task was straightforward and clear, and perceived a benefit in using their own mobile device to access it, one saying "*given a choice I would go for the digital version every time*" (P03). Some of these benefits included being able to complete it in their own home and time, linked to their neurology

appointments, and not having to spend more time in clinics than necessary. One participant emphasised that "*it's a stress going to hospital*" (P07) and remote, self-administered assessment could mitigate this.

Participants thought that the SST engaged their minds in diverse ways, for example:

> *"It would get me thinking on a completely different level from outside of what I do… It is something quite removed from your life. It's something quite different. It's a different task, I never have thought of doing to get my brain thinking. Thinking at a pace."* (P05)

Two participants enjoyed the task, stating "*I love things like this*" (P08), but others thought that "*it was it was just sort of okay, really, I think that was alright. That was enough and it didn't make me feel stupid*" (P01). After completing the task, one participant reported that they had "*little bit of anxiety*" (P05) prior to completion, however felt relieved at how simple the task was. In contrast, one participant thought that the task might be too easy and not challenging enough, saying stating that they were "*not convinced that being asked to do something that you can do is testing you*" (P03). Similarly, although the SST lasts 90 seconds, not having a visible countdown was welcomed, as awareness of a timed task was problematic, with one participant saying that they "*have a problem with being timed on a test. That makes [them] so much more anxious*" (P05).

There were different opinions about being able to retry the task, and participants wanted more guidance as to if they could retry if they were disturbed during the task or did something wrong. One participant expressed anxiety saying that "*most of my thinking is not what happens when I do it right, but what happens when I do it wrong*" (P03) which may be an intrusive thought if they could not retry the task if needed. Similarly, participants would like more context and reassurance to encourage them to complete the assessment, with one participant

stating that "*you as a team can set that encouragement*" (P05). This should include an explanation about how the results will be used, and what help may be available for any detected cognitive impairment. One participant thought that other individuals with MS may deny they have problems and therefore may not wish to engage with the assessment. One participant thought that there should be reassurance as to privacy concerns that people with MS may have, both with technological privacy (such as storing cookies) and their data being reported to third parties. The participant said that "*someone with Relapsing Remitting MS might be just a bit more reluctant to share data if they if they think it's going to leak out somehow*" (P03), such as to their employers, insurance companies, or the Driver and Vehicle Licensing Agency.

Table 14: Change log of amendments made due to minor usability issues

Category	Issue Raised	Changes
Instructions	In the first visual example of the test, participants are shown the symbol for '2', but shown the answer number '3'.	Amended to the correct answer.
	Participants wanted more context in the instructions, as they were unsure about completing the task as quickly or as accurately as possible.	Instructions to be amended to emphasise that the task should be completed as quickly and as accurately as possible.
	Participants did not necessarily read the instructions but looked at images first.	A first page will be added with brief written instructions, followed by the images showing examples on the following screen.
	Participants looked at examples and did not read the text preceding the images.	Have the text instructions on one screen, followed by visual example on the next screen.
Symbol Substitution Task	The arrow symbol was shown as the first symbol, and participants thought it was drawing their attention to something rather than being a stimulus.	The arrow symbol will be replaced with another Unicode symbol.
	Ability to retake task needs more context or to not be an option.	Text to be added to explain that participants could complete the task once again if they were interrupted.
Look and feel	Participant found that the symbol response number pad was pushed off the screen.	To be resolved by the technology team to ensure all stimuli fit on a single screen
	'Next' button was too far at the bottom of the screen, creating too much white space.	The 'next' symbol will be moved further up the screen to reduce blank space.

6.6 Discussion

The cognitive interview study described in this chapter identified usability issues experienced by people with MS and asserted that the platform is mostly acceptable. As a result of the findings of this study, changes can be made to the platform, including amending the instructions to clarify how the task will function, and providing more context.

A number of usability issues were identified: some necessitate small, technical changes to be made to the platform (as described in Table) which were communicated to the wider research team and technology developers to be amended after the cognitive interviews. However, some raised wider issues involved in cognitive assessments deployed on individuals' own devices. One of these was experienced by two participants in this study and involved device failures and interruptions. Whilst people with MS can be directed to ensure that their device is charged and to minimise external interruptions, the web application requires connection to mobile internet or Wi-Fi, and therefore it is likely some respondents would experience interruptions of emails, messages, or phone calls, which may be confounding to the timed nature of the task. There are some approaches to mitigating this, for example by allowing people with MS to retry the assessment if they were interrupted or advising them how to report the issue to be given an in-clinic re-test if required.

A pertinent finding of the qualitative interviews described in Chapter 5 was that participants want assurance of the purpose of the assessments to make their participation meaningful. This was reiterated in the current cognitive interviews: people with MS wanted more context than was currently provided as to the purpose of the assessment, and how clinicians will receive and process the results, and privacy of the platform and their results. To assuage these concerns, participants thought more information should be provided, either in the initial email or letter inviting them to complete the assessment, or upon accessing the platform. This includes when and where they should complete the

assessment, and information about what will happen to their results, how it could inform their clinical care, and reassurance about who will see it and how it will be kept private. This is lacking in the current iteration of the platform, demonstrated through some of the participants' differing opinions as to what the purpose of the task was, and how useful it may be to them. For example, some participants thought the SST was a task of visual acuity, or manual dexterity, rather than cognition.

Adapting tests may also require changes to be made to the instructions. Whilst the instructions for a self-administered, digital assessment will diverge from the original task's, these can be optimised for the modality and information that needs to be provided. Whilst participants thought the instructions were mostly straightforward, some made suggestions, including providing text on a separate screen before the visual examples to ensure that the text is read.

Participants preferred straightforward, simple tasks to ensure they are accessible and engaging. As previously found, more complex tasks can discourage completion, especially with people with MS with more severe cognitive impairment (Cortés-Martínez et al., 2019). Therefore, the cognitive interviews showed that participants were mostly comfortable with the simplicity of the SST, and the straight-forward nature of it was acceptable.

Studies have described how people with MS are concerned about how the time of day, stress or general fatigue would impact on the results of cognitive assessments (Mortensen et al., 2020), which participants in the current interviews corroborated. Self-administered digital assessments that can be completed remotely can allow people with MS to choose the best time and environment to complete the task, which was perceived as a positive feature in the interviews.

In the interview study in Chapter 5, participants stated a preference for completing assessments on their own mobile devices, and therefore we restricted access of the platform to mobile devices in

Phase II of the current study. Participants' completing assessments on their personal device has the benefit of an existing familiarity with the device. An overview of digital neuropsychology literature has highlighted some of the challenges with touchscreen devices for deploying assessments, including latency in user responses which could lead to participants recording a slower response than their actual responses (Germine et al., 2019). There is a trade-off between difficulties such as these, and the benefits of allowing people with MS to use their own device. Individual studies may allow the use of a single, common device, but this is not feasible in a remote, self-administered assessment in routine healthcare. Further research may examine the impact of device variability on scores, and this can be ascertained through normative data collection (described in Chapter 7).

Using an adapted DIEGO framework of factors that affect engagement with digital health (O'Connor et al., 2016), we found that people with MS felt motivated to engage with the technology, and an understanding of why they may be asked to complete an assessment on the DECSYS platform. Whilst some participants expressed greater enthusiasm for the SST and the platform than others, no participants reported a negative experience that could potentially inhibit future participation with the platform. Similarly, no technical issues were identified that would hinder the use of the platform in a significant way. Future iterations can ensure that the task is conceptually meaningful and provide greater reassurance to allay any anxiety respondents may experience. The results of the cognitive interviews provide evidence in line with Blandford's (2019) elements of successful digital health technologies, ensuring that the platform is engaging, accessible, and does not cause undue burden to the end user.

Technological problems identified in the change log in Table can be used to inform the next iteration of the platform, and problems around the lack of context and information can be improved in the wording of instructions on the platform, and communications in clinical care with people with MS. Whilst some of the problems highlighted,

such as device failures and interruptions are more difficult to overcome, these can be managed, for example, by allowing individuals to report these interruptions and allow them to retry the task or complete it in-clinic completion on provided devices.

There were some limitations around the methodology of this study that should be considered. We found that in both our pilot interviews with PPI members and the cognitive interviews, participants showed a tendency to focus on completing the timed task, rather than talking about their experience throughout. This could be due to participants wishing to complete the task well in the presence of a researcher. Therefore, during the eight cognitive interviews, we chose to modify the 'think-aloud' protocol to a hybrid model to facilitate the process, and as such, the researcher had a more active participation in the interviews.

One PPI participant also expressed difficulty in communicating her thoughts and word finding due to her MS-related cognitive impairment. We amended the interview schedule and 'think-aloud' approach to overcome some of these difficulties. Guidance has stated that greater interviewer involvement through probing questions is more useful than continuing in silence in think-aloud studies (McDonald et al., 2016), and indeed we found that offering encouragement and asking questions throughout each stage of the assessment generated more insight from participants. Especially with a move towards video conferencing platforms to restrict in-person research, it would be worthwhile to consider how think-aloud methods can be structured to encourage verbalisation in future, especially in populations that may have difficulty with word-finding.

Due to COVID-19 restrictions on in-person interviews, we conducted all Phase II cognitive interviews over video conferencing. This suggested that all participants had a basic understanding of IT systems and participation may have excluded those who have less experience, who may also have difficulties with the cognitive screening technology. Upon completion of the first two video conference interviews, it became apparent that participants did not necessarily

need to use a desktop or laptop computer and could use a device (e.g., a tablet to conduct the interview), whilst using an alternate device (e.g., a smartphone) to complete the screening. However, this still assumed access to multiple devices. It is also possible that the people with MS who participated in this research had an interest in technology development, and indeed two participants described their strong technological backgrounds. Therefore, insights into how participants less comfortable with technology may be missing from this analysis. This is particularly true for older people with MS, who may not have access to their own technology, and might be aided by a family member to complete cognitive assessments. Once it is feasible to recruit and collect data in-person with participants, it may be necessary to conduct further interviews with participants who are less confident with technology to ensure that the platform is accessible to them.

A further limitation with remote interviews was the inability to watch how participants interacted with the technology and relying on verbal descriptions of their interactions. During the first interviews conducted through Microsoft Teams, we asked participants to show the interviewer the device's display by holding their smartphone or tablet to the webcam or posing their webcam down. This presented difficulties as the screens were too small to view fully, and there are difficulties viewing screens through webcams due to the power frequency, causing flickering. Through discussions within the research team, we decided to rely on descriptions from the participants rather than trying to view how they were using the technology. With the hybrid model of cognitive interviewing integrating concurrent questioning, we found that participants were able to describe the process of interacting with the platform well. Despite these limitations, we collected important user feedback that will inform both the next iteration of the platform, and future technology development for self-administered cognitive assessment platforms deployed on individuals' own devices.

6.7 Conclusions

The qualitative interviews described in Chapter 5 found that people with MS were willing to complete mobile assessments, and the cognitive interviews confirmed this conclusion. O'Connor et al. (2016) emphasised that this indication of engagement provides evidence for which digital health approaches are successful. As such, this platform and test can be used with confidence that people with MS perceive the value and are motivated to undergo routine assessment. However, the SST is still considered a new test (Bauer et al., 2012), and further evidence of its psychometric properties is needed. Its utility is also limited without normative data with which to compare scores against. The next chapter details a protocol to determine the validity and reliability of the developed cognitive screening technology, and aims to collect normative data from control participants."

Chapter 7 Protocol for the Validation of a Digital Cognitive Assessment for Use in MS

7.1 Chapter overview

This chapter presents a protocol that was devised and granted ethical approval in February 2020. However due to delays in the technological development of the assessment platform and tests caused by the COVID-19 pandemic, this research could not be conducted. The utility of the cognitive assessment platform described in earlier chapters could be furthered from the inclusion of an additional task alongside the symbol digit substitution test, a Stroop test. The two-task battery requires evidence of validity and reliability, as well as the collection of new normative data. This chapter provides justification for this approach, the methodology, and analysis proposed for this study. This chapter is presented in future tense, as it is a planned protocol of work

7.2 Introduction

Previous research on cognitive assessment for people with MS has largely used assessments which required a trained technician to administer in clinic, even for electronic versions (Wojcik et al., 2019). Proof of concept work (Patel et al., 2019) has reported that both people with MS and neurologically healthy participants have shown a preference for a self-administered version to the traditional written oral administrations of the Symbol Digit Modalities Test (SDMT). As described in Chapters 5 and 6, people with MS confirmed that a digital, self-administered symbol substitution task (named the SST) was acceptable and feasible to complete. However, this is a single test that has limited capability of assessing the range of cognitive impairment. The inclusion of other brief tests can broaden the assessment to add further insight to cognitive dysfunction in MS.

An inclusion criteria was formulated to choose additional tasks to include in the assessment battery. Choosing tests that fulfil these criteria may support implementation within clinical practice by confirming tests are relevant, simple to administer and complete, and can thus prevent burden on both people with MS and clinicians. The criteria are as follows:

1. Relevant to the condition and cognitive impairment present in MS.
2. Adaptable to a small-screen digital modality. The task needs to be clearly displayed on a small screen, and easy to respond to, to prevent manual errors.
3. Acceptable to patients and clinicians. To avoid undue anxiety and ensure willingness to complete assessment, tasks should have been evaluated for acceptability and usability. This will ensure that the assessment is straight-forward to complete without the intervention of a technician.
4. Automated scoring to prevent the need for a technician

For the inclusion of a second task to the battery, the Stroop test was chosen as it meets the four criteria. The Stroop task was presented to participants (as described in Chapter 5), and both clinicians and people with MS felt that it was an acceptable assessment for use in MS. Additionally, the Stroop task has also been recommended as a sensitive measure for assessing complex attention in people with MS, and research has suggested that administering the Stroop test alongside the SDMT may make cognitive screening efficient at detecting subtle changes in impairment in individuals with RMMS (Prakash et al., 2008). During the interviews as part of the first stage of the usability testing (presented in Chapter 6), Patient and Public Involvement (PPI) group members completing the Stroop task reported that it was acceptable, usable on mobile devices, and some respondents reported preferring the task to the SST. On the basis of this evidence, the Stroop test has been integrated into the Discrete and Ellipse-based Response Capture System (DECSYS) platform. This

chapter details the background and use of the Stroop in MS, and the research needed to assess the validity and reliability of the SST and the Stroop test together as a multidomain assessment.

7.2.1 **Inclusion of the Stroop task**

Stroop tasks were first developed in 1935 and have been used frequently in psychological research. In the original administration, there were two control trials – word reading (reading colour words in a black text) and naming of colours (naming colours of squares or symbols), and an experimental interference trial of colour-word in which participants read the colour of the words, as opposed to the text of the word which are incongruent (Stroop, 1935). The Stroop effect is the interference of what the word says (e.g., blue) on respondents' naming the colour of the word (e.g., red ink). There are several different administrations and versions of the Stroop test, differing on the assessment procedures, stimuli used (types, amounts, sizes, presentation), scoring, and therefore their validity and reliability. Some of these different versions are the Comalli and Kaplan Stroop (Comalli et al., 1962); Golden Stroop (Golden, 1978), Dodrill Stroop (Dodrill, 1978), the Victoria Stroop test (Regard, 1981), and the Trenerry Stroop (Trenerry et al., 1989).

The Stroop test has been frequently used in MS research, accounting for 5% of all neuropsychological tests, as reported in Chapter 2. It also forms a part of some commonly used neuropsychological batteries such as the Delis Kaplan Executive Function System (Delis et al., 2001). A meta-analysis showed that the Stroop test had a higher effect size (g=−0.79) than other tests measuring attention, such as the Paced Auditory Serial Addition Test (PASAT), suggesting that it may be more sensitive, and the authors recommended its inclusion in a cognitive assessment battery for people with MS (Prakash et al., 2008).

One of the benefits of the Stroop test is that it can be easily modified to a digital administration, which can improve the

measurement of response speed and accuracy with automated scoring. A recent systematic review found that digital administrations of the Stroop were largely valid and reliable, although such administrations are used in research without stating which version of the test has been used and the psychometric properties of the test, and therefore extrapolation is contentious (Din & Tat Meng, 2019). A smartphone native application version of the Stroop, the EncephalApp, has been developed for use in Minimal Hepatic Encephalopathy (MHE; Bajaj et al., 2013), and consists of two trials. In the first trial, there are multiple runs in which 10 coloured hash symbols (####) appear and participants are required to categorise them based on their colour by pressing one of three colour names at the bottom of the screen. The second trial uses incongruent coloured names of colours. Validation of the test found good specificity (81%), and test-retest reliability (r=0.941, p<0.0001; Bajaj et al., 2015). Although these results are promising, the integration of a similar version of the Stroop into the DECSYS platform still requires further evidence due to potential discrepancies in formats.

7.2.2 **Psychometric validation**

Guidance states the necessity to consider newly developed digital cognitive assessments as new assessments, due to the difference in administration to a device, and changes such as the presentation of stimuli and response modality (Bauer et al., 2012). This includes evidence of validity (such as equivalence to their traditional administrations) and the collection of new normative data to establish benchmarks for clinical scores.

Studies have shown that there is good validity and reliability between traditional verbal administrations and newly developed digitised adaptations with people with MS. Convergent validity of symbol substitution tests deployed on mobile devices have strong, positive correlations (r>0.7) with results on the written and oral SDMT, confirming its convergent validity (Middleton et al., 2020; Pham et al., 2020; Rao et al., 2017; van Dongen et al., 2020). Similarly, test-retest reliability of digital symbol substitution tests is consistently high (r>0.8)

over a period of (Middleton et al., 2020; Pham et al., 2020; Rao et al., 2017). We would therefore expect that the mobile tasks in the current study would have similarly good validity and reliability. Test-retest reliability will use an interval of 14 days based on recommendations (Polit, 2014) and similar studies (e.g., Lam et al., 2021; Pham et al., 2021).

7.3 Aims

There are three sub-studies in the proposed research, and the aims of each are as follows:

Study I: Ascertain the validity of the mobile, self-administrations of the SST and Stroop. This includes equivalence to their traditional verbal administrations and construct validity.

Study II: Collect UK-based normative data for both tasks to enable regression-based analyses to provide benchmarks for clinical scores.

Study III: Examine the test-retest reliability of the SST and the Stroop after a period of 14 days.

7.4 Methods

7.4.1 Ethical considerations

The Division of Psychiatry and Applied Psychology at the University of Nottingham granted ethical approval for the three sub-studies, and all study materials on 27/04/20 (REC reference DPAP-2020-1517-3; Appendix 15).

A relevant Participant Information Sheet (PIS) will be provided to all participants in all studies, which will explain the background and rationale of the research, why they are being invited to take part, what will happen to their data, and provide contact details for the researcher should they wish to be in touch after their participation or have any concerns about the study. Participants will be informed that taking part is their choice, and they are free to withdraw before or during the assessments. All participants will be given sufficient time to decide if

they wish to take part and required to give their consent, either verbally or digitally to proceed with the study.

Participants taking part in Studies II and III will be given a unique identifier to match their data. These will be kept in a password protected file, and once matched, the data will be fully anonymised.

7.4.2 **Participants**

We will recruit two groups of participants: those with MS and those without MS. Neurologically healthy participants will be eligible to take part in all studies if they are aged 18 years old or over, able to provide consent, can communicate in English, have their own mobile device (smartphone or tablet) and are willing to use it in the study. In Study I, they will require access to a computer and a webcam. Participants with MS will be eligible to take part if they fulfil the above inclusion criteria and have a diagnosis of MS. Neurologically healthy participants will be excluded from taking part if they have any neurological conditions or other pre-existing conditions that may impact cognition, and people with MS will be excluded if they have any co-morbid neurological conditions that may impact cognition.

7.4.3 **Materials**

7.4.3.1 *Symbol Substitution Task*

The digital adaptation of the SDMT, the SST, (the development of which has been described in Chapter 6) is one of the two tests hosted on the DECSYS platform. This test is modelled on both the traditional written version of the SDMT (Smith, 1982) and digital adaptations that have been developed (e.g. Middleton et al., 2020; Rao et al., 2017). Responses are scored by the number of correct responses within the 90-second period, and the number of incorrect responses is also recorded.

7.4.3.2 *Oral administration of the Symbol Digital Modalities Test*

In Study I, a researcher will administer the oral SDMT to the two groups of participants to assess the equivalence between this test and the conceptually similar SST. The oral administration is recommended

and more frequently used in MS than the written version, due to removing the motor element of completing the test (Rao, 1990). During the test, respondents view the response sheet, which displays a key of symbols at the top of the page, and a full page of symbols. They have 90 seconds to say aloud the number associated with the symbols using the key. Responses are scored on the number of correct responses.

The oral SDMT was chosen because it is more feasible to conduct remotely during COVID-19 restrictions on research that requires in-person contact. A remote, oral administration of the SDMT has been found to be strongly correlated with in-person administration (Eilam-Stock et al., 2021). In this validation study, participants accessed the test materials through a website, and completed the test via a telephone call with a technician. Therefore, this is the administration that will be used in the current study to administer the SDMT remotely. A link to a website with a static image of the SDMT response sheet will be sent to participants, who will be asked to access it during the test session conducted on their computer.

7.4.3.3 *Digital Stroop*

The digital development of a Stroop task will be closely aligned with existing digital applications of the Stroop (for example, Bajaj et al., 2015), which have been optimised to work on mobile devices but still retain aspects of original Stroop tests that have been used to identify cognitive impairment in people with MS.

The digital Stroop will consist of two trials: a neutral colour naming trial and an incongruent word trial, each consisting of 100 stimuli. The number of incorrect responses made for each trial, and the times taken for each trial will be recorded. The scoring of the two trials within the digital Stroop will be calculated as follows:

a. Number of incorrect responses on the incongruent word trial compared to incorrect responses on the colour naming trial.
b. Time taken to respond to all stimuli on the incongruent trial – time taken to complete all stimuli on the neutral colour trial.

7.4.3.4 *Golden Stroop*

In Study I, a researcher will administer the verbal administration of the Golden version of the Stroop task (Golden, 1978), which will be used to examine the equivalence to the digital Stroop. The Golden Stroop uses three trials presented in the same order for each administration, and each trial consists of 100 items that participants must respond to by reading the word, or the colour aloud to the researcher. All items in a trial are presented on a single side of page. The trials are as follows:

1. A neutral word naming trial with names of colours in black ink.

2. A neutral colour naming trial with coloured Xs.

3. An incongruent colour naming trial consisting of words of colours presented in incongruent colours.

The stimuli will be presented by the researcher, and participants say the words (Trial 1 and 3) or the colour (Trial 2) of the stimuli aloud. Incorrect responses are corrected by the technicians, and participants are scored on how many items they can complete within 45 seconds.

7.4.4 **Study I**

7.4.4.1 *Design*

This study will use a within-subjects, counterbalanced and randomised design to control for practice effects. This will determine the equivalence of the SST to previously validated written administration of these tests with neurologically healthy participants.

7.4.4.2 *Participants*

A sample size of 66 neurologically healthy participants has been chosen based on a previous study examining the equivalence of neuropsychological assessments for use with people with MS (van Dongen et al., 2020) and a sample size calculation using a minimum Intraclass Correlation coefficient of 0.3 and a power of 80% (Cohen, 2013). A pilot sample of 10 people with MS (n=10) will be recruited to

demonstrate whether these two participant samples are comparable in examining equivalence between the two administrations.

7.4.4.3 *Procedure*

Upon interest in taking part, participants will be sent the PIS. During the test session, a researcher will briefly explain what the participant will be asked to do as part of the study, and participants will be asked several questions based on the eligibility criteria to ensure they are eligible to take part. The researcher will take verbal consent and ask the participant a series of questions to collect demographic data.

The participant will then be individually randomised to either Condition 1 or Condition 2. As the order of the tests hosted on the assessment platform will remain the same, the same order will be used for both administrations. Participants in Condition 1 (*n*=38) will complete the oral SDMT, and Stroop followed by the digital SST and Stroop. Participants in Condition 2 (*n*=38) will complete the digital tasks followed by the oral tasks.

The link to the test materials and the digital platform will be sent by email, and all participants will be asked to use their computer to access the test materials, and their handheld devices to access the DECSYS platform.

Participants in Condition 1 will access the test material website through their computer and complete the SDMT, followed by the Stroop. The researcher will present the cards from the Golden Stroop described in Section 7.4.3.4, and participants read the words or colours aloud. Once participants have finished the verbal tasks, they will then be asked to pick up their handheld device and access the platform using a unique identifier, and complete both digital tests.

Participants in Condition 2 will be asked to complete the digital tasks first, followed by the verbal tests.

7.4.5 **Study II**

7.4.5.1 *Design*

Study II is a cross-sectional study of cognitively unimpaired, neurologically healthy participants.

7.4.5.2 *Participants*

Approximately 200 cognitively unimpaired neurologically healthy participants in the UK will take part in Study II. Previous psychometric literature has identified that a sample size of 200 is the minimum for normative data collection (Nunally & Bernstein, 1994), and similar research has used similar sample sizes (for example, *n*=236 in van Dongen et al., 2020). Samples that are representative of the gender and age of the population of the UK will be recruited, and the demographics of the sample will be regularly reviewed, so that specific groups can be targeted if necessary. Educational attainment will also be collected as part of the demographics, and reviewed to ensure that a range of years of education is collected. If when reviewed the sample is not representative of a range of demographics, further recruitment will target these in particular, for example by approaching and advertising to specific groups and in communities that are underrepresented.

7.4.5.3 *Procedure*

Participants will be able to directly access the tests on their mobile device using a weblink from recruitment materials or the researcher. Participants will be given the PIS upon indicating their interest to the researcher, or upon access to the system, and asked to read and confirm their consent. They will then answer questions about the device they are using and their demographic information (including their age, gender, highest educational attainment, ethnicity and socioeconomic status), before moving on to the tests. On completion of these, participants will be asked if they wish to take part in Study III after 14-days. If interested, participants will provide their email address to receive further information.

7.4.6 **Study III**

7.4.6.1 *Design*

Study III is a repeated-measures, within-subjects, test-retest reliability study.

7.4.6.2 *Participants*

This study uses participants from Study II who consented to participating in this study will be. All participants will be invited, with the aim to include 27 participants who express their interest, allowing for a randomisation of participants, and to account for any attrition of those who are not able or willing to participate in Study III. This sample size is based on the correlation between the two administrations to be >0.6, using a power of 80% and a significance level of 0.01 (Cohen, 2013).

7.4.6.3 *Procedure*

A researcher will send a reminder email one week prior to the 14-day timepoint to participants to confirm they are still able and willing to take part in the study and are able to complete the tests on the specified day. At 14 days after their first assessment, they will be emailed the link to the assessment platform, and a unique identifier and asked to complete it that day using their own mobile device.

7.5 Patient and public involvement

PPI members have provided their feedback and ‘sense-checked’ the findings of the qualitative study in Chapter 5, as well as confirming that the Stroop task was acceptable to include as a further assessment. A PPI group which included people with MS and their family members were asked to review all participant-facing documents to ensure they were clear and easy to understand. PPI members confirmed the appropriateness of research materials, methodology, and recruitment strategies within each study, and thought that the research would have acceptable potential outcomes for people with MS.

7.6 Recruitment

All participants will be invited to take part in the research using fliers and information posted on social media (including Facebook and Twitter) and snowballing techniques. We will also approach MS charities and organisations to publicise the research. Additionally, we will approach the PPI research network of the University of Nottingham to contact people with MS and their family members to participate.

For the normative data collection, we will recruit a sample that is representative of the gender, age, education, ethnicity, and socioeconomic status of the population of the UK, and this will be regularly reviewed to target specific groups if necessary.

7.7 Analyses

The analyses proposed in this section have been guided by previous validation studies, and literature on psychometric evaluations. An overview of psychometric considerations for neuropsychological assessments is provided in the Introduction chapter, in section 1.4.2. All analyses will be conducted using IBM's Statistical Package for the Social Sciences (SPSS) 26th edition. Statistical significance will be set at $p<0.05$.

The data will be presented as frequencies (gender, ethnicity, socioeconomic status, and device used) and descriptive statistics (age, educational attainment, raw scores on each test). The raw scores for all tasks are described in the materials section 7.4.3. Histograms and normal probability plots, and results of Kolmogorov-Smirnov tests will assess the normality of the distribution of test scores and will inform which statistical tests will be used.

7.7.1 **Convergent validity**

Equivalence provides an indication of convergent validity by determining if the scores on the digital tests are analogous to those on the verbal tests, and a high correlation between the two versions will evidence that they are measuring the same construct. As the digital

assessments use different stimuli presented randomly, this will mitigate practice effects by producing alternate versions. Using the tasks' raw scores from both neurologically healthy participants and people with MS groups, a Pearson's product-moment or Spearman's rho correlation will measure the association between the digital and verbal modalities. Convergent validity can be confirmed through a strong positive correlation (r>0.7) between the two scores (Cohen, 2013). A systematic review reported strong correlations between traditional and digital administrations of processing speed and attention (r>0.7; Wojcik et al., 2019).

7.7.1.1 *Construct Validity*

The construct validity will examine the relationship between scores on a test measuring a similar construct that we would expect to be correlated. The SST and Stroop tests are hypothesised to have a positive relationship, due to both measuring facets of cognition. Previous studies assessing the associations between Stroop and symbol substitution tests have found positive correlations between the two (r=0.3-0.5; Hughes et al., 2013; Morrow, 2013; Periáñez et al., 2021). Again, Pearson's product-moment correlation or Spearman's rho will measure the correlation between the two test scores collected as part of Studies I and II. A medium positive correlation (r>0.4) indicates construct validity.

7.7.2 **Cross modality differences**

An analysis of the mean difference in scores between the types of devices used will be examined. Participants in Study II will have the choice of a tablet or a smartphone to complete the assessments. Descriptive statistics of how many participants in Study II choose to use each different device will be reported, and the difference in scores on both tasks will be examined using a one-way ANOVA. Significant differences between scores and device used indicates that the device type may need to be adjusted for in comparison with norming data.

7.7.3 **Reliability**

7.7.3.1 *Test-retest reliability*

Test-retest reliability indicates the stability of the assessment over time in a sample of participants. Depending on the normality of data, a Pearson or Spearman's correlation coefficient of scores on the digital tasks between the two time periods will indicate if the tests have achieved test-retest reliability by reaching a medium to strong positive association (r>0.6; Cohen, 2013), demonstrated in similar research with associations of r>0.7 in digital assessments (Wojcik et al., 2019).

7.7.3.2 *Internal consistency*

The correlation between scores on each digital test can assess internal consistency. A minimum Cronbach's alpha value of 0.70 will suggest that the battery of the two tests is internally consistent (Cronbach, 1951; Nunally & Bernstein, 1994).

7.7.4 **Normative data analysis**

Regression-based normative data have been found to be more sensitive in detecting cognitive impairment in MS than discrete approaches (Burggraaff et al., 2017; Parmenter et al., 2010), and therefore will be used in this analysis.

A regression analysis will examine the predictive value of any of the demographic variables on scores on the Stroop and SST. Entering the variables of age, gender, education, socioeconomic status, and device used will assess the relative contribution of each variable on each test score, and devise formulae that can adjust for these variables if they are found to significantly predict test performance, as described in Fellows and Schmitter-Edgecombe (2019). Based on previous research, we would expect age and educational attainment to be associated with performance on cognitive assessments (van Dongen et al., 2020). The directions of any relationships will be evaluated using scatterplots. Regression-based corrections based on significant relationships can be applied to the data before raw scores are converted to z-scores using means and SDs. These corrections can be

implemented in the DECSYS platform to provide adjusted scores and interpretations for clinical scores.

7.8 Discussion

This chapter describes a protocol of work to be completed as the next phase of development of the assessment platform. It was devised and was granted ethical approval in early 2020 to be conducted in-person sessions. An amendment based on the data collection described in this chapter was then submitted to consider COVID-19 restrictions on in-person contact. Due to the delays in the technology development of the platform, it was not feasible to conduct this study.

There are potential limitations of the proposed research that must be acknowledged; administering traditional verbal tests over video call could be problematic. Whilst the SDMT has been validated for remote administration through participants accessing the test sheet through a website (e.g., Eilam-Stock et al., 2021), such evidence for the Golden Stroop is not available. A meta-analysis of remote cognitive assessment over video conferencing found small, insignificant effect sizes between remote and in-person administrations (Brearly et al., 2017). However, the included studies did not involve any version of a Stroop tests, and results of visually mediated tests were more likely to differ from their in-person administrations.

The APA has provided guidance on teleneuropsychological assessments in light of the pandemic (American Psychological Association, 2020), but acknowledges its limitations. These include technological impediments, such as poor Wi-Fi connections. Indeed, when using time-limited tasks (as described in this protocol), a slow internet connection could impact results of the verbal tests, and any technological delays should be observed, and their impact assessed during the data collection.

Interruptions in the technology development due to COVID-19 caused further delays in the usability study described in Chapter 6. The usability study was conducted in December 2020 and January and

2021; the results of which were used to inform further development of the DECSYS platform and SST, however the Stroop task is still under development as of September 2021. Due to these delays, this proposed study was not feasible to conduct during my period of PhD registration. As part of the NIHR-funded programme, the Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS), the validation and collection of normative data will begin in Winter 2021 by another researcher.

Chapter 8 Discussion

8.1 Summary of the research

The overall aim of this research was to explore, critically appraise, and develop an assessment for detecting and monitoring cognitive impairment for people with MS. There have been multiple recommendations that have acknowledged the practical difficulties that may hinder uptake of assessment with people with MS in clinical practice (Kalb et al., 2018; Sumowski et al., 2018). These have advocated for the implementation of routine brief screening of cognitive problems in MS clinical practice, followed by a full neuropsychological examination if needed. Consensus approaches to assessment and screening have not consider the preferences and practical limitations in using cognitive assessment.

The systematic review described in Chapter 2 established the breadth of the assessments used in MS research. The most popular individual test, the Symbol Digit Modalities Test (SDMT), is also recommended as a minimal screening assessment in MS (Kalb et al., 2018). The next chapters focused on the utility of the SDMT for use with people with MS. The scoping review in Chapter 3 determined the different methodologies employed for interpreting results of this individual cognitive test. This review highlights the diverse approaches by summarising the most common methods of categorising cognitive function, with examples of these. I also highlight some of the potential opportunities that digital administration and scoring of the SDMT.

I then examined the association of processing speed (as measured by the SDMT) with other variables in Chapter 4. Using data from the Cognitive Rehabilitation for Attention and Memory in Multiple Sclerosis (CRAMMS) Randomised Controlled Trial, I determined the domain structure of the self-report form of the Multiple Sclerosis Neuropsychological Questionnaire (MSNQ-P), to ascertain whether different domains could predict processing speed deficits. During the development of the MSNQ-P, an *a priori* factor structure was used with

the intention of a unidimensional structure. However, studies have found mixed results with regards to its relationship to objective measures of cognition (e.g. Benedict et al., 2004; Migliore et al., 2021), and similarly to its latent factor structure implying a multidimensional scale, which may be of greater utility in detecting cognitive impairment (Konstantinopoulou et al., 2018; Migliore et al., 2021). The results of the principal components analysis showed that the MSNQ-P may consist of five components relating to various aspects of cognition. Two of these components, related to complex attention and prospective memory, were significantly, negatively correlated with SDMT scores in the whole sample of people with MS. This suggests that a component-based short form of the MSNQ-P could be administered as an alternative method of assessing cognition of people with MS who may not want to complete an objective assessment. There is also potential for the individual domains to be used to explore subjective experiences of other aspects of cognition.

Having established the attributes of the SDMT, I proceeded to conduct two qualitative studies determining how cognitive assessment would operate in clinical practice. The interviews and a focus group with people with MS, family members, clinicians and healthcare commissioners confirmed that the SDMT was an acceptable task to administer in a clinical context within the NHS from the perspectives of both people with MS and clinicians. Participants showed a preference for a digital assessment that can be completed remotely and deployed on individual's own mobile devices. This would fit into existing resources in NHS clinical care, and potentially increase the uptake in assessment, ensuring it could be routinely provided to all people with MS. This guided the development of a digital platform for hosting cognitive tasks in collaboration with MindTech and the Digital Research Service at the University of Nottingham.

Once the platform was developed, I determined the technical usability of this test and the platform it would be hosted on, as well as identifying other mechanisms of use, including enablers and barriers.

Establishing that these factors are frequently not assessed in the development of user-facing technology, and these studies highlighted how important it is to integrate the end user in the development process. These interviews with people with MS highlighted technical issues that needed to be addressed, confirmed that the platform was usable and accessible on a range of mobile devices, and evidenced that the platform and a digitised symbol substitution task was acceptable and well received.

The next phase of the development would be quantitative data collection. As described in the protocol of research in Chapter 7, an additional task, a version of the Stroop Colour and Word Task, could be added to the platform to expand its capacity for identifying cognitive impairment. The research proposed in Chapter 7 would evaluate the validity and reliability of these tests and sets out the importance of normative data to interpret scores. Due to the COVID-19-related delays in the technology development, it was not possible to conduct this research as part of this PhD, however the task development, validation, and normative data collection is being completed as part of Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS).

8.2 General discussion

Despite consensus recommendations for cognitive assessment in MS, it is not widespread (Klein et al., 2019). The interviews with people with MS and clinicians described in Chapter 5 confirmed that cognition is not discussed or assessed routinely but would be a valuable addition to clinical care. Within the NHS, lack of resources hinders clinicians' ability to implement assessment. The lack of time, staff, specific expertise, the private space to conduct an assessment, and financial constraints all contribute to this issue. There have been mixed results regarding the value that a full cognitive assessment (including chart reviews, interviews and other methodologies alongside cognitive tests and batteries) adds to clinical practice (Donders, 2020). Therefore, brief screening offered to all people with MS in their routine care, followed by

a full neuropsychological battery if required, can optimise limited resources, while providing input to more patients. Whilst research and recommendations emphasise the importance of routine assessment for people with MS (Foley et al., 2012; Hansen & Lautenbacher, 2017), they often do not consider the practical difficulties. For example, the Brief International Cognitive Assessment for MS (BICAMS) promotes the use of three separate tests, all of which are appropriate for assessing cognition in MS (Corfield & Langdon, 2018). However, accessing these tests requires clinicians to approach three different test publishers to purchase three different sets of test materials, whilst ensuring that the technician administering these tests understands the procedures of each, and how to score and interpret the results. One of the tests within the BICAMS, the SDMT is emerging as a gold-standard, single assessment to screen for cognitive impairment in people with MS (Kalb et al., 2018; Schependom et al., 2014), however is still subject to some of the same barriers in its implementation in clinical practice due to being a proprietary test.

Another potential barrier is a reluctance to complete objective cognitive tests. Cortés-Martínez et al. (2019) highlighted that people with MS with greater cognitive impairment were less willing to complete the PASAT, however in the qualitative interviews some participants expressed a dislike of *all* objective cognitive tests. Therefore, provision of an alternative method of assessment may ensure greater inclusion of people with MS. Previous research has recommended the MSNQ-P to be administered alongside the SDMT (Benedict et al., 2008), however its use has been contentious, often showing a greater association with mood than cognition measured by objective tests (Benedict et al., 2003; Benedict et al., 2004). The secondary data analysis described in Chapter 4 used a Principal Components Analysis to determine a potential factor structure, and showed promise in a reduced eight-item, or a two-subscale version of this questionnaire. Questionnaires such as the MSNQ-P are simple to digitise, and do not require the stringent evidence of equivalence that digital objective cognitive tests are subject

to. This analysis highlights the opportunity of domain-based self-report cognitive assessments, but further development and validation is needed.

In focusing on the feasibility and how assessment can fit into the existing structure of clinical care, the research in this thesis has identified factors to the implementation of brief screening. The barriers and enablers identified in Chapter 5 were considered and put at the centre of the model of assessment described in this thesis. This person-centred approach provides a framework that can inform clinical practice, through a proof of concept that digitised assessments, administered remotely, on a patient's own devices enables routine monitoring.

8.2.1 **How the research relates and adds to the literature**

While there are recommendations for routine cognitive assessment in MS (Kalb et al., 2018), and many digital assessments are emerging (e.g., Lam et al., 2021; Middleton et al., 2020; Pham et al., 2021 amongst others), there has been limited research into patient and clinician preference in cognitive assessments in MS. Mortensen et al. (2020) explored attitudes to cognitive assessment. This was a small sample of people with MS (*n*=12) who had already completed the BICAMS, and so only the tests considered were the SDMT, the California Verbal Learning Test (CVLT-II), and the Brief Visuospatial Memory Test (BVMT-R), in their pen-and-paper versions. The qualitative study described in Chapter 5 supports this research. This study explored more cognitive assessments, in different administrative modalities, and with more stakeholders (including family members, clinicians and commissioners, and people with MS). The results of this study showed a need for assessments to operate within limited resources, and a clear preference for digitised assessment that can be completed in individuals' own homes. Moore et al. (2017) found good validity, reliability, and equivalence between self-administered, mobile assessments deployed on smartphones and Personal Digital Assistants in a systematic review, demonstrating the feasibility of using such

administrations in clinical research. The research described in this research supports these findings in the MS population.

One of the limitations of previous research that has developed digitised assessments for use in MS is the lack of user input. A digital assessment, the Multiple Screener, detailed some feedback from people with MS regarding the clarity and feasibility of the technology, and made adjustments on this basis (van Dongen et al., 2020). This study used a questionnaire to record difficulties with using the technology rather than a qualitative approach such as cognitive interviewing. The current research goes beyond this by ensuring the usability and acceptability of a cognitive assessment platform from a HCI perspective.

A narrative review (Marziniak et al., 2018) focused on broad communication technologies currently used for people with MS, and has applicability to more specific digital technologies. This review emphasised that digital health initiatives should be secure, unintrusive, and carefully implemented, which was echoed in the findings of both qualitative studies described in Chapter 5 and Chapter 6. People with MS reported a need for reassurance about the safety of their data, and the outcomes of the assessment. Both this review and the current research showed that telehealth is largely accepted and could lead to greater satisfaction of clinical care.

Whilst the COVID-19 pandemic has encouraged the development of teleneuropsychological assessments, these have often been developed without end user or expert input. A review of available self-administered cognitive assessment uncovered 39 applications (apps) and web-based tools that can be easily accessed through Google or distribution services such as Google Play (Charalambous et al., 2020). During the year in which this review took place, 12 of these apps had been removed, and the psychometric details could only be ascertained for seven of the assessments left, despite attempts to contact the authors. Only two of the assessments offered respondents a method to follow up their results; the remaining 19 had no further

guidance or support to their users if they were concerned about their results. Given how ubiquitous and accessible these apps are to anyone with internet access, their quality and safety are questionable. There is a need for formal evaluation, to evidence their psychometric properties and suitability for assessing cognitive function. The qualitative studies, and the protocol described in Chapter 7 attempt to provide a formalised approach to the digital development of a multidomain platform of cognitive assessments.

The usability testing discussed in Chapter 6 used Human Computer Interaction (HCI) methodology to formally assess usability and perceptions of the SST under development. This is the first time this approach has been applied to digitised cognitive assessments in this manner. This usability research identified aspects of the platform and assessment important to the people with MS being interviewed.

8.3 Limitations of the research

There are limitations to the research presented in this thesis. Firstly, this research concentrated on the setting of how cognitive assessment would work in clinical practice within the UK-based NHS, and therefore its applicability to different healthcare settings and countries may be limited. However, the overall findings of the potential for the developed digital cognitive assessment platform could be investigated in different settings.

Having identified the SDMT as the most frequently used measure in MS research, this thesis concentrated on its use as an individual test. The current SST uses a manual touch response input. In the originally developed SDMT, the verbal is preferred in MS to reduce any impact of dexterity impairment (Benedict et al., 2017), which may still be a problem with the digital SST. Voice recognition technology could be implemented, but slowed speech through dysarthria may also contribute significantly to low neuropsychological performance on verbal tasks (Arnett et al., 2008).

Whilst some research has suggested the SDMT is not just a measure of information processing speed, it also may involve aspects of memory and lexical access speed, its specificity may be problematic (Sandry et al., 2021). Similarly, recent approaches using domain-based interpretation have suggested that information processing speed may not be the most impaired cognitive domain in MS (De Meo et al., 2021; Leavitt et al., 2018). Both studies emphasise that memory dysfunction may be more common in people with MS and may be the first domain of cognitive impairment to become impaired. This challenges recent recommendations that information processing speed should be at the forefront of neuropsychological assessment for people with MS (Kalb et al., 2018; Sumowski et al., 2018). Whilst I have described the integration of an additional test, the Stroop task, and its validation, a memory task may further increase the capabilities and overcome some of the specificity issues suggested of the SDMT (Sandry et al., 2021).

The usability testing in Chapter 6 demonstrated some of the difficulties with developing a device-agnostic technology. For example, some participants' devices did not present the task optimally (for example, by pushing response options off the screen), which may negatively impact their performance and falsely detect impairment. The purpose of the tasks and platform is not diagnostic or comprehensive, and its use in clinical practice would need to be managed carefully to ensure that false positives are identified. Although most participants in both qualitative studies reported being receptive to completing cognitive assessments digitally, it is possible that some people with MS would not have the capability or willingness to complete the tests on mobile devices. This may especially apply to those who have greater disability or longer duration of MS, are older, or may not have the devices or familiarity with technology, who may not have been willing to participate in the interviews. The pilot usability interviews were conducted at the university due to the availability of a range of equipment (a tablet, a smartphone and a laptop). However, the second phase of the usability cognitive interviews with participants' personal devices were completed

remotely over video conferencing due to COVID-19 constraints on in-person data collection. This precluded those who did not have access to at least two devices; one of which to communicate with the researcher, and the other on which to complete the test. Therefore, people with MS who lacked the equipment were precluded from participating. Several of the participants in these interviews had experience and interest in technology and user design, which could have prompted their interest to take part.

The usability testing involved the DECSYS platform, using a single digital test based on the SDMT. With the integration of more tests in future, including the planned Stroop task, further cognitive interviews would need to be conducted to ensure that these digital tests are acceptable to people with MS. Whilst this research was not able to conduct a quantitative validation, or examine the clinical utility of the screening assessment, I am hopeful that the initial findings are able to have benefits going forward. Providing an acceptable cognitive assessment may promote discussion between people with MS and their clinicians, and open the opportunity for further assessment and intervention, where indicated.

8.4 Recommendations for future direction

With the continued development of technology that can be applied to cognitive assessment, the self-administered, mobile device approach has many potential directions. Moving towards an evidence-based consensus could be a principal factor in ensuring uptake. The current research has highlighted the importance of patient and clinician acceptability, how to implement a screening programme in clinical practice, considerations needed for digitisation, and the further research required for these. The development of a common framework for cognitive assessment in neurological conditions could be helpful in establishing how developments should proceed.

Whilst the protocol detailed further steps that are needed to validate the newly developed assessments, there would still need to be

further work involved in assessing its clinical utility. Previously determined cut-off scores cannot be used between the different modalities, and these need to be ascertained for new tests. Future research using the digital tests with people with MS could also administer a more comprehensive, traditional battery, such as the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS), to categorise participants as cognitively impaired or preserved. Examining the potential cut-offs for the digital tests that correctly identify cognitively impaired or preserved people with MS through assessment of their sensitivity and specificity can allow further utility of the scores. This could indicate severity of cognitive impairment which could be used to triage people with MS for appropriate treatment. Correspondingly, using approaches discussed in Chapter 3 could add further insight into scoring and interpretation of the tests.

There is further potential to expand the utility of the digital assessment platform. A memory task could be included in future iterations of the platform, as memory is seen as a predominantly impaired cognitive domain in MS (as described in section 1.3.2). However, there is a difficulty associated with digitising memory tests, and was therefore precluded from the next iteration described in the protocol Chapter 7. Whilst a figure recognition task was originally considered for inclusion in the DECSYS platform, recognition tends to largely be intact in MS (Calabrese, 2006). People with MS show more deficits in verbal and non-verbal recall (Zakzanis, 2000), however tasks often involve replicating figures (such as in the BVMT-R) or recalling word lists (such as in the CVLT-II), and these would these present distinct obstacles for digitisation of self-administrated tests. I established the requirement for automated scoring in the current development of the platform, Automated scoring can provide greater accuracy of scores, together with different metrics that cannot typically be captured by a technician. However, the process of developing these technologies is an emerging field (Birchmeier & Studer, 2019), and considerably more complex for inputs such as drawing or speech

recognition. This is particularly challenging for a device-agnostic approach that would allow respondents to use their own tablets or smartphones.

Advances in artificial intelligence (AI) which improve automated speech recognition could enable the integration of word list learning and verbal fluency tests to increase the capability of what can be assessed. One digitisation of the SDMT using speech recognition of oral responses had promising results with a small sample (Patel et al., 2019), but recognition of words over numbers may be more difficult to implement (for example, there are a greater amount of words possible for verbal fluency than digit options for the SDMT). Additionally, preferences of digitised verbal input from people with MS should be sought – one study found people with MS overwhelmingly chose touchscreen over verbal responses in a mobile SDMT (Pham et al., 2021).

Additionally, there have been some developments using digitised drawing tasks on smartphones. One digital development of a symbol substitution task uses symbol drawing, instead of writing the number associated with the symbol (Allen et al., 2021). This smartphone-based digital test uses a grid of five dots on which respondents draw a geometric pattern with their finger, similar to pattern locks for mobile devices. This could be used for figure drawing tasks in future but would need to be assessed in terms of its acceptance through qualitative research, and equivalence and validation against similar pen-and-paper tests.

Re-evaluating the profile of cognitive impairment in MS could also lead to greater insight into the range of domains affected. A recommendation emerging for a different neurological condition, epilepsy, supports the use of the DSM-5 diagnostic divisions for cognitive function alongside expert opinion for a common cognitive assessment framework (Norman et al., 2021). This is analogous to the number of expert consensus recommendations for cognitive assessment in MS, however, raises an interesting approach to use of

the common framework of the DSM-5's domains to describe cognitive profiles in neurological conditions. As explored in the introductory chapter, the homogenous cognitive profile of MS is contentious, driven by different assessments, interpretations, and inconsistent conclusions. This is particularly apparent when considering which domains are overlooked in multidomain batteries. For example, more evidence is emerging regarding deficits in social cognition (Bora et al., 2016; Cotter et al., 2016) and prospective memory deficits in MS (Rouleau et al., 2018), but tests of these domains are mostly absent from recommendations, batteries, and research. Using a top-down approach of a common framework of cognitive profiling, more comprehensive cognitive assessment can construct a more thorough model of impairment. This in turn could standardise how cognition is evaluated in MS, and involve previously neglected domains, such as prospective and autobiographical memory, and social cognition. The findings of Chapter 4 and the potential of a domain-based self-report questionnaire could form the basis of the development of a new psychometric tool.

There is a potential for research to go beyond adapting tests and begin to develop tests that are optimised for the digital format. A cognitive assessment that uses AI in a visual animal categorisation task has shown remarkable promise for use in MS (Khaligh-Razavi et al., 2020). The authors have reported excellent psychometric properties, including high sensitivity and specificity, paired with advantages over traditional assessments including no bias through practice, education, or language. Pitteri et al.'s (2020) tablet-based game of attentional load shows similar psychometric properties. Gamification of tests could reduce anxiety that some respondents may feel about neuropsychological assessment by appearing less threatening to people with MS and may increase engagement with such assessments. Whilst these two assessments are deployed on tablets, which may significantly reduce the applicability in clinical practice, device-agnostic developments and validation could mitigate these challenges.

There are opportunities to integrate digitised assessment into research and clinical practice, including different measurement techniques, different conditions, and the potential for interdisciplinary approaches. One of the obstacles to digitisation of neuropsychological assessment is that of resources. Development is expensive, time-consuming, and requires expertise from patient, clinical, neuropsychological and software development perspectives. There is the risk that platforms will become obsolete with newer devices, and futureproofing is an important consideration. For example, tests deployed on iPads may only function on one model.

Digital assessments also necessitate considerable data collection for validation and normative comparisons. A solution to these barriers is a centralised resource of assessments. There are initiatives, for example the Mobile Toolbox initiative, funded by the US National Institute on Aging, is developing a collection of validated cognitive assessments that can be used on smartphones and have associated normative data (Gershon et al., 2021). There is a possibility that cognitive assessments could fall under the scope of a medical device, and therefore be subject to further regulations, which needs to be considered if the assessment is being used for diagnostic purposes.

Mobile administration also opens the possibility of frequent testing, to explore what cognition looks like over the course of everyday life and explore variability of daily factors (such as fatigue). However, this could be overly onerous on people with MS and lead to attrition as experienced in a study of older adults (Valdes et al., 2016). A recent review highlighted the advantage that mobile testing has with regards to the extra data that can be collected to include context within cognitive assessment (Weizenbaum et al., 2020). There may be internal and external factors that have an impact on cognitive performance, and using mobile devices can attempt to measure these, such as using the microphone to measure background noise, or GPS data to capture where an individual completed the assessment. There are clear ethical

and privacy concerns with this approach, but self-reported contextual questions could be integrated into platforms to assess these.

8.5 Conclusions

This research has identified the need for routine screening of cognitive impairment for people with MS, and explored how this could be implemented within NHS MS clinics. Determining the breadth of cognitive assessments used in MS research, and demonstrating the variability in methods used to score and interpret one of the most popular tests, the SDMT, emphasised the inconsistency in current research. However, these were the first steps in consolidating the prevalent consensus recommendations with a more formal appraisal and evaluation of how cognitive assessments could be implemented into clinical care. The quantitative and qualitative inquiries have examined both subjective and objective aspects of cognitive assessments further, and explored the potential for these different assessments to be administered to people with MS.

There are many potential future directions for research. Further development of the platform could integrate other tests and provide evidence as to the assessments' clinical utility, as described in Chapter 7. With a digital deployment, scoring mechanisms can be built into the digital platform, allowing for comparisons of clinical data with the collected normative data and provide meaningful interpretation of scores.

The research detailed in this thesis describes the first stages required to integrate cognitive screening into clinical care, in an efficient, acceptable, and efficacious manner. In Chapter 5, participants recognised the value of being able to provide cognitive screening for all people with MS, which could lead to cognitive problems being identified earlier, and support such as rehabilitation being provided when it could be most effective. However, this is a complex process, which requires further stages to implement and ascertain the feasibility in clinical practice. As part of the NEuRoMS programme (www.neuroms.org), a

pathway of routine cognitive screening and neuropsychological rehabilitation will be evaluated within UK clinical care. This will build upon the research described in this thesis.

Principally, the provision of a brief, screening platform that is easy to access and use, and feasible to implement in clinical practice for all people with MS, can aid in the identification of cognitive impairment. This could be used as an endpoint to evidence the impact that rehabilitation may have in treating cognitive problems.

References

Achiron, A., & Barak, Y. (2006). Cognitive changes in early MS: A call for a common framework. *Journal of the Neurological Sciences*, *245*(1), 47–51. https://doi.org/10.1016/j.jns.2005.05.019

Alhadreti, O., & Mayhew, P. (2018). *Rethinking thinking aloud: A comparison of three think-aloud protocols*. CHI Conference on Human Factors in Computing Systems. https://doi.org/10.1145/3173574.3173618

Allen, C. M., Mowry, E., Tintore, M., & Evangelou, N. (2020). Prognostication and contemporary management of clinically isolated syndrome. *Journal of Neurology, Neurosurgery, and Psychiatry*. https://doi.org/10.1136/jnnp-2020-323087

Allen, L., Baker, E., Thorp, E., Evans, M., Thorpe, D., Granger, K., Cormack, F., Barnett, J., & Cashdollar, N. (2021). Validation and comparability of smartphone-based digit symbol substitution task with written version. *The International Society For CNS Clinical Trials and Methodology*. 17th Annual Scientific Meeting.

Amato, M. P., Prestipino, E., & Bellinvia, A. (2019). Identifying risk factors for cognitive issues in multiple sclerosis. *Expert Review of Neurotherapeutics*, *19*(4), 333–347. https://doi.org/10.1080/14737175.2019.1590199

Amato, M. P., & Zipoli, V. (2003). Clinical management of cognitive impairment in multiple sclerosis: A review of current evidence. *International MS Journal*, *10*(3), 72–83.

Amato, M. P., Zipoli, V., & Portaccio, E. (2006). Multiple sclerosis-related cognitive changes: A review of cross-sectional and longitudinal studies. *Journal of the Neurological Sciences*, *245*(1), 41–46. https://doi.org/10.1016/j.jns.2005.08.019

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*.

American Psychological Association. (2020). *Guidance on psychological tele-assessment during the COVID-19 crisis*. https://www.apaservices.org/practice/reimbursement/health-codes/testing/tele-assessment-covid-19

Anens, E., Emtner, M., Zetterberg, L., & Hellström, K. (2014). Physical activity in subjects with multiple sclerosis with focus on gender differences: A survey. *BMC Neurology*, *14*(1), 47. https://doi.org/10.1186/1471-2377-14-47

Anstey, K. J., Matters, B., Brown, A. K., & Lord, S. R. (2001). Normative data on neuropsychological tests for very old adults living in retirement villages and hostels. *The Clinical Neuropsychologist*, *14*(3), 309–317. https://doi.org/10.1076/1385-4046(200008)14:3;1-p;ft309

Arnett, P. A., Smith, M. M., Barwick, F. H., Benedict, R. H. B., & Ahlstrom, B. P. (2008). Oralmotor slowing in multiple sclerosis: Relationship to neuropsychological tasks requiring an oral response. *Journal of the International Neuropsychological Society*, *14*(3), 454–462. https://doi.org/10.1017/S1355617708080508

Atkinson, T. M., Ryan, J. P., Kryza, M., & Charette, L. M. (2011). Using versions of the trail making test as alternate forms. *The Clinical Neuropsychologist*, *25*(7), 1193–1206. https://doi.org/10.1080/13854046.2011.589410

Aupperle, R. L., Beatty, W. W., deNAP Shelton, F., & Gontkovsky, S. T. (2002). Three screening batteries to detect cognitive impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *8*(5), 382–389. https://doi.org/10.1191/1352458502ms832oa

Aupperle, R. L., Beatty, W. W., Shelton, N., & Gontkovsky, S. T. (2002). Three screening batteries to detect cognitive impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *8*(5), 382–389. https://doi.org/10.1191/1352458502ms832oa

Baddeley, A. (2012). Working Memory: Theories, Models, and Controversies. *Annual Review of Psychology*, *63*(1), 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Bagert, B., Camplair, P., & Bourdette, D. (2002). Cognitive Dysfunction in Multiple Sclerosis. *CNS Drugs*, *16*(7), 445–455. https://doi.org/10.2165/00023210-200216070-00002

Baird, J. F., Cederberg, K. L. J., Sikes, E. M., Jeng, B., Sasaki, J. E., Sandroff, B. M., & Motl, R. W. (2019). Changes in cognitive performance with age in adults with multiple sclerosis. *Cognitive and Behavioral Neurology : Official Journal of the Society for Behavioral and Cognitive Neurology*, *32*(3), 201–207. https://doi.org/10.1097/WNN.0000000000000200

Bajaj, J. S., Heuman, D. M., Sterling, R. K., Sanyal, A. J., Siddiqui, M., Matherly, S., Luketic, V., Stravitz, R. T., Fuchs, M., Thacker, L. R., Gilles, H., White, M. B., Unser, A., Hovermale, J., Gavis, E., Noble, N. A., & Wade, J. B. (2015). Validation of encephalapp, smartphone-based Stroop test, for the diagnosis of covert hepatic encephalopathy. *Clinical Gastroenterology and Hepatology*, *13*(10), 1828-1835.e1. https://doi.org/10.1016/j.cgh.2014.05.011

Bajaj, J. S., Thacker, L. R., Heuman, D. M., Fuchs, M., Sterling, R. K., Sanyal, A. J., Puri, P., Siddiqui, M. S., Stravitz, R. T., Bouneva, I., Luketic, V., Noble, N., White, M. B., Monteith, P., Unser, A., & Wade, J. B. (2013). The Stroop smartphone application is a short and valid method to screen for minimal hepatic encephalopathy. *Hepatology*, *58*(3), 1122–1132. https://doi.org/10.1002/hep.26309

Bakirtzis, C., Ioannidis, P., Messinis, L., Nasios, G., Konstantinopoulou, E., Papathanasopoulos, P., & Grigoriadis, N. (2018). The rationale for monitoring cognitive function in multiple sclerosis: Practical issues for clinicians. *The Open Neurology Journal*, *12*, 31–40. https://doi.org/10.2174/1874205X01812010031

Barcellos, L. F., Horton, M., Shao, X., Bellesis, K. H., Chinn, T., Waubant, E., Bakshi, N., Marcus, J., Benedict, R. H., & Schaefer, C. (2020). A validation study for remote testing of cognitive function in multiple sclerosis. *Multiple Sclerosis Journal*. https://doi.org/10.1177/1352458520937385

Bardi, C. A., Hamby, S. L., & Wilkins, J. W. (1995). Stability of several brief neuropsychological tests in an HIV+ longitudinal sample. *Archives of Clinical Neuropsychology*, *10*(4), 295–295. https://doi.org/10.1093/arclin/10.4.295

Barker-Collo, S. L. (2005). Within session practice effects on the PASAT in clients with multiple sclerosis. *Archives of Clinical Neuropsychology*, *20*(2), 145–152. https://doi.org/10.1016/j.acn.2004.03.007

Barry, D., Bates, M. E., & Labouvie, E. (2008). FAS and CFL forms of verbal fluency differ in difficulty: A meta-analytic study. *Applied Neuropsychology*, *15*(2), 97–106. https://doi.org/10.1080/09084280802083863

Basso, M. R., Beason-Hazen, S., Lynn, J., Rammohan, K., & Bornstein, R. A. (1996). Screening for cognitive dysfunction in multiple sclerosis. *Archives of Neurology*, *53*(10), 980–984. https://doi.org/10.1001/archneur.1996.00550100046014

Bauer, R. M., Iverson, G. L., Cernich, A. N., Binder, L. M., Ruff, R. M., & Naugle, R. I. (2012). Computerized neuropsychological assessment devices: Joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. *Archives of Clinical Neuropsychology*, *27*(3), 362–373. https://doi.org/10.1093/arclin/acs027

Beatty, W. W., & Goodkin, D. E. (1990). Screening for cognitive impairment in multiple sclerosis. An evaluation of the Mini-Mental State Examination. *Archives of Neurology*, *47*(3), 297–301.

Beier, M., Alschuler, K., Amtmann, D., Hughes, A., Madathil, R., & Ehde, D. (2019). iCAMS: Assessing the reliability of a BICAMS tablet application. *International Journal of MS Care*. https://doi.org/10.7224/1537-2073.2018-108

Beier, M., Hughes, A. J., Williams, M. W., & Gromisch, E. S. (2019). Brief and cost-effective tool for assessing verbal learning in multiple sclerosis: Comparison of the Rey Auditory Verbal Learning Test (RAVLT) to the California Verbal Learning Test – II (CVLT-II). *Journal of the Neurological Sciences*, *400*, 104–109. https://doi.org/10.1016/j.jns.2019.03.016

Benedict, R. B. H. (1997). *Brief Visuospatial Memory Test—Revised*. Psychological Assessment Resources, Inc.

Benedict, R. B. H., Amato, M. P., DeLuca, J., & Geurts, J. J. G. (2020). Cognitive impairment in multiple sclerosis: Clinical management, MRI, and therapeutic avenues. *The Lancet Neurology*, *19*(10), 860–871. https://doi.org/10.1016/S1474-4422(20)30277-5

Benedict, R. B. H., Cookfair, D., Gavett, R., Gunther, M., Munschauer, F., Garg, N., & Weinstock-Guttman, B. (2006). Validity of the Minimal Assessment of Cognitive Function in Multiple Sclerosis (MACFIMS). *Journal of the International Neuropsychological Society*, *12*(4), 549–558.

Benedict, R. B. H., Duquin, J., Jurgensen, S., Rudick, R., Feitcher, J., Munschauer, F., Panzara, M., & Weinstock-Guttman, B. (2008). Repeated assessment of neuropsychological deficits in multiple sclerosis using the Symbol Digit Modalities Test and the MS

Neuropsychological Screening Questionnaire. *Multiple Sclerosis Journal*, *14*(7), 940–946. https://doi.org/10.1177/1352458508090923

Benedict, R. B. H., Fischer, J. S., Archibald, C. J., Arnett, P. A., Beatty, W. W., Bobholz, J., Chelune, G. J., Fisk, J. D., Langdon, D. W., & Caruso, L. (2002). Minimal neuropsychological assessment of MS patients: A consensus approach. *The Clinical Neuropsychologist*, *16*(3), 381–397.

Benedict, R. B. H., Munschauer, F., Linn, R., Miller, C., Murphy, E., Foley, F., & Jacobs, L. (2003). Screening for multiple sclerosis cognitive impairment using a self-administered 15-item questionnaire. *Multiple Sclerosis Journal*, *9*(1), 95–101. https://doi.org/10.1191/1352458503ms861oa

Benedict, R. B. H., Wahlig, E., Bakshi, R., Fishman, I., Munschauer, F., Zivadinov, R., & Weinstock-Guttman, B. (2005). Predicting quality of life in multiple sclerosis: Accounting for physical disability, fatigue, cognition, mood disorder, personality, and behavior change. *Journal of the Neurological Sciences*, *231*(1–2), 29–34. https://doi.org/10.1016/j.jns.2004.12.009

Benedict, R. H. B., Cox, D., Thompson, L. L., Foley, F., Weinstock-Guttman, B., & Munschauer, F. (2004). Reliable screening for neuropsychological impairment in multiple sclerosis. *Multiple Sclerosis Journal*, *10*(6), 675–678. https://doi.org/10.1191/1352458504ms1098oa

Benedict, R. H. B., Deluca, J., Enzinger, C., Geurts, J. J. G., Krupp, L. B., & Rao, S. M. (2017). Neuropsychology of multiple sclerosis: Looking back and moving forward. *Journal of the International Neuropsychological Society*, *23*(9–10), 832–842. https://doi.org/10.1017/s1355617717000959

Benedict, R. H. B., DeLuca, J., Phillips, G., LaRocca, N., Hudson, L. D., & Rudick, R. (2017). Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Multiple Sclerosis Journal*, *23*(5), 721–733. https://doi.org/10.1177/1352458517690821

Benedict, R. H. B., Wahlig, E., Bakshi, R., Fishman, I., Munschauer, F., Zivadinov, R., & Weinstock-Guttman, B. (2005). Predicting quality of life in multiple sclerosis: Accounting for physical disability, fatigue, cognition, mood disorder, personality, and behavior change. *Journal of the Neurological Sciences*, *231*(1–2), 29–34. https://doi.org/10.1016/j.jns.2004.12.009

Benedict, R. H., DeLuca, J., Phillips, G., LaRocca, N., Hudson, L. D., & Rudick, R. (2017). Validity of the Symbol Digit Modalities Test as a cognition performance outcome measure for multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *23*(5), 721–733. https://doi.org/10.1177/1352458517690821

Benton, A. L., deS, K., & Sivan, A. B. (1994). *Multilingual Aphasia Examination*. Psychological Assessment Resources, Inc.

Bever, C. T., Grattan, L., Panitch, H. S., & Johnson, K. P. (1995). The Brief Repeatable Battery of Neuropsychological Tests for

multiple sclerosis: A preliminary serial study. *Multiple Sclerosis*, *1*(3), 165–169. https://doi.org/10.1177/135245859500100306

Bilder, R. M., Postal, K. S., Barisa, M., Aase, D. M., Cullum, C. M., Gillaspy, S. R., Harder, L., Kanter, G., Lanca, M., Lechuga, D. M., Morgan, J. M., Most, R., Puente, A. E., Salinas, C. M., & Woodhouse, J. (2020). Inter organizational practice committee recommendations/guidance for teleneuropsychology in response to the COVID-19 pandemic. *Archives of Clinical Neuropsychology*, *35*(6), 647–659. https://doi.org/10.1093/arclin/acaa046

Birchmeier, M. E., & Studer, T. (2019). Automated rating of multiple sclerosis test results using a convolutional neural network. *Studies in Health Technology and Informatics*, *259*, 105–108. https://doi.org/10.3233/978-1-61499-961-4-105

Blandford, A. (2018). Lessons from working with researchers and practitioners in healthcare. *Interactions*, *26*(1), 72–75. https://doi.org/10.1145/3292023

Blandford, A. (2019). HCI for health and wellbeing: Challenges and opportunities. *International Journal of Human-Computer Studies*, *131*, 41–51. https://doi.org/10.1016/j.ijhcs.2019.06.007

Bodling, A. M., Denney, D. R., & Lynch, S. G. (2009). Cognitive aging in patients with multiple sclerosis: A cross-sectional analysis of speeded processing. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *24*(8), 761–767. https://doi.org/10.1093/arclin/acp076

Bora, E., Özakbaş, S., Velakoulis, D., & Walterfang, M. (2016). Social cognition in multiple sclerosis: A meta-analysis. *Neuropsychology Review*, *26*(2), 160–172. https://doi.org/10.1007/s11065-016-9320-6

Boringa, J. B., Lazeron, R. H., Reuling, I. E., Adèr, H. J., Pfennings, L. E., Lindeboom, J., de Sonneville, L. M., Kalkers, N. F., & Polman, C. H. (2001). The Brief Repeatable Battery of Neuropsychological Tests: Normative values allow application in multiple sclerosis clinical practice. *Multiple Sclerosis Journal*, *7*(4), 263–267. https://doi.org/10.1177/135245850100700409

Bowler, R., Sudia, S., Mergler, D., Harrison, R., & Cone, J. (1992). Comparison of Digit Symbol and Symbol Digit Modalities Tests for assessing neurotoxic exposure. *The Clinical Neuropsychologist*, *6*(1), 103–104. https://doi.org/10.1080/13854049208404123

Bradbury, K., Watts, S., Arden-Close, E., Yardley, L., & Lewith, G. (2014). Developing digital interventions: A methodological guide. *Evidence-Based Complementary and Alternative Medicine*, *2014*, 561320. https://doi.org/10.1155/2014/561320

Brandt, J., & Benedict, R. (2001). *Hopkins Verbal Learning Test-Revised Professional manual.* Psychological Assessment Resources.

Brearly, T. W., Shura, R. D., Martindale, S. L., Lazowski, R. A., Luxton, D. D., Shenal, B. V., & Rowland, J. A. (2017). Neuropsychological test administration by videoconference: A

systematic review and meta-analysis. *Neuropsychology Review*, *27*(2), 174–186. https://doi.org/10.1007/s11065-017-9349-1

Bruce, J., Hancock, L., Arnett, P., & Lynch, S. (2010). Treatment adherence in multiple sclerosis: Association with emotional status, personality, and cognition. *Journal of Behavioral Medicine*, *33*, 219–227. https://doi.org/10.1007/s10865-010-9247-y

Burgess, P. W., Alderman, N., Evans, J., Emslie, H., & Wilson, B. A. (1998). The ecological validity of tests of executive function. *Journal of the International Neuropsychological Society*, *4*(6), 547–558.

Burggraaff, J., Knol, D. L., & Uitdehaag, B. M. J. (2017). Regression-based norms for the symbol digit modalities test in the Dutch population: Improving detection of cognitive impairment in multiple sclerosis? *European Neurology*, *77*(5–6), 246–252. https://doi.org/10.1159/000464405

Calabrese, P. (2006). Neuropsychology of multiple sclerosis—An overview. *Journal of Neurology*, *253 Suppl 1*, I10-15. https://doi.org/10.1007/s00415-006-1103-1

Campbell, J., Rashid, W., Cercignani, M., & Langdon, D. (2017). Cognitive impairment among patients with multiple sclerosis: Associations with employment and quality of life. *Postgraduate Medical Journal*, *93*(1097), 143–147. https://doi.org/10.1136/postgradmedj-2016-134071

Casaletto, K. B., & Heaton, R. K. (2017). Neuropsychological Assessment: Past and Future. *Journal of the International Neuropsychological*, *23*(9–10), 778–790. https://doi.org/10.1017/S1355617717001060

Charalambous, A. P., Pye, A., Yeung, W. K., Leroi, I., Neil, M., Thodi, C., & Dawes, P. (2020). Tools for app- and web-based self-testing of cognitive impairment: Systematic search and evaluation. *Journal of Medical Internet Research*, *22*(1), e14551. https://doi.org/10.2196/14551

Chiaravalloti, N., & Deluca, J. (2009). Cognitive impairment in multiple sclerosis. *Lancet Neurology*, *7*, 1139–1151. https://doi.org/10.1016/S1474-4422(08)70259-X

Clemens, L., & Langdon, D. (2018). How does cognition relate to employment in multiple sclerosis? A systematic review. *Multiple Sclerosis and Related Disorders*, *26*, 183–191. https://doi.org/10.1016/j.msard.2018.09.018

Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences*. Routledge.

Comalli, P. E., Wapner, S., & Werner, H. (1962). Interference effects of Stroop color-word test in childhood, adulthood, and aging. *The Journal of Genetic Psychology*, *100*, 47–53. https://doi.org/10.1080/00221325.1962.10533572

Compston, A., & Coles, A. (2008). Multiple sclerosis. *Lancet*, *372*(9648), 1502–1517. https://doi.org/10.1016/S0140-6736(08)61620-7

Cook, D. A., & Beckman, T. J. (2006). Current concepts in validity and reliability for psychometric instruments: Theory and application. *The American Journal of Medicine*, *119*(2), 166.e7-16. https://doi.org/10.1016/j.amjmed.2005.10.036

Cooke, L. (2010). Assessing concurrent think-aloud protocol as a usability test method: A technical communication approach. *IEEE Transactions on Professional Communication*, *53*(3), 202–215. https://doi.org/10.1109/TPC.2010.2052859

Corfield, F., & Langdon, D. (2018). A systematic review and meta-analysis of the Brief Cognitive Assessment for Multiple Sclerosis (BICAMS). *Neurology and Therapy*, *7*(2), 287–306. https://doi.org/10.1007/s40120-018-0102-3

Cortés-Martínez, A., Matias-Guiu, J. A., Pytel, V., Montero, P., Moreno-Ramos, T., & Matías-Guiu, J. (2019). What is the meaning of PASAT rejection in multiple sclerosis? *Acta Neurologica Scandinavica*, *139*(6), 559–562. https://doi.org/10.1111/ane.13090

Costa, S. L., Genova, H. M., DeLuca, J., & Chiaravalloti, N. D. (2017). Information processing speed in multiple sclerosis: Past, present, and future. *Multiple Sclerosis Journal*, *23*(6), 772–789. https://doi.org/10.1177/1352458516645869

Cotter, J., Firth, J., Enzinger, C., Kontopantelis, E., Yung, A. R., Elliott, R., & Drake, R. J. (2016). Social cognition in multiple sclerosis: A systematic review and meta-analysis. *Neurology*, *87*(16), 1727–1736. https://doi.org/10.1212/WNL.0000000000003236

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297–334. https://doi.org/10.1007/BF02310555

Cullen, B., O'Neill, B., Evans, J. J., Coen, R. F., & Lawlor, B. A. (2007). A review of screening tests for cognitive impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, *78*(8), 790–799. https://doi.org/10.1136/jnnp.2006.095414

Dagenais, E., Rouleau, I., Demers, M., Jobin, C., Roger, E., Chamelian, L., & Duquette, P. (2013). Value of the MoCA test as a screening instrument in multiple sclerosis. *Canadian Journal of Neurological Sciences*, *40*(3), 410–415.

Damasceno, A., Pimentel-Silva, L. R., Damasceno, B. P., & Cendes, F. (2020). Cognitive trajectories in relapsing–remitting multiple sclerosis: A longitudinal 6-year study. *Multiple Sclerosis Journal*, *26*(13), 1740–1751. https://doi.org/10.1177/1352458519878685

Davis, D. H., Creavin, S. T., Yip, J. L., Noel-Storr, A. H., Brayne, C., & Cullum, S. (2015). Montreal Cognitive Assessment for the diagnosis of Alzheimer's disease and other dementias. *Cochrane Database of Systematic Reviews.*, *10*, Cd010775. https://doi.org/10.1002/14651858.CD010775.pub2

Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*, 982–1003. https://doi.org/10.1287/mnsc.35.8.982

de Caneda, M. A. G., Cuervo, D. L. M., Marinho, N. E., & de Vecino, M. C. A. (2018). The Reliability of the Brief Visuospatial Memory Test—Revised in Brazilian multiple sclerosis patients. *Dementia & Neuropsychologia*, *12*(2), 205–211. https://doi.org/10.1590/1980-57642018dn12-020014

De Meo, E., Portaccio, E., Giorgio, A., Ruano, L., Goretti, B., Niccolai, C., Patti, F., Chisari, C. G., Gallo, P., Grossi, P., Ghezzi, A., Roscio, M., Mattioli, F., Stampatori, C., Simone, M., Viterbo, R. G., Bonacchi, R., Rocca, M. A., De Stefano, N., … Amato, M. P. (2021). Identifying the distinct cognitive phenotypes in multiple sclerosis. *JAMA Neurology*. https://doi.org/10.1001/jamaneurol.2020.4920

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System: Examiners manual*. Psychological Corporation.

Delis, D. C., Kramer, J. H., Kaplan, E., & Holdnack, J. (2004). Reliability and validity of the Delis-Kaplan Executive Function System: An update. *Journal of the International Neuropsychological Society*, *10*, 301–303. https://doi.org/10.1017/s1355617704102191

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *Manual for the California Verbal Learning Test, (CVLT-II).* The Psychological Corporation.

Deluca, J. (2008). Information processing speed: How fast, how slow, and how come? In *Information processing speed in clinical populations.* (pp. 265–273). Taylor & Francis.

DeLuca, J., Chelune, G. J., Tulsky, D. S., Lengenfelder, J., & Chiaravalloti, N. D. (2004). Is speed of processing or working memory the primary information processing deficit in multiple sclerosis? *Journal of Clinical and Experimental Neuropsychology*, *26*(4), 550–562. https://doi.org/10.1080/13803390490496641

DeLuca, J., Chiaravalloti, N. D., & Sandroff, B. M. (2020). Treatment and management of cognitive dysfunction in patients with multiple sclerosis. *Nature Reviews Neurology*, *16*(6), 319–332. https://doi.org/10.1038/s41582-020-0355-1

DeLuca, J., Gaudino, E. A., Diamond, B. J., Christodoulou, C., & Engel, R. A. (1998). Acquisition and storage deficits in multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology*, *20*(3), 376–390. https://doi.org/10.1076/jcen.20.3.376.819

Denney, D. R., & Lynch, S. G. (2009). The impact of multiple sclerosis on patients' performance on the Stroop Test: Processing speed versus interference. *Journal of the International Neuropsychological Society*, *15*(3), 451–458. https://doi.org/10.1017/s1355617709090730

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Dickinson, M. D., & Hiscock, M. (2011). The Flynn effect in neuropsychological assessment. *Applied Neuropsychology*, *18*(2), 136–142. https://doi.org/10.1080/09084282.2010.547785

Diehr, M. C., Heaton, R. K., Miller, W., & Grant, I. (1998). The Paced Auditory Serial Addition Task (PASAT): Norms for age, education, and ethnicity. *Assessment*, *5*(4), 375–387. https://doi.org/10.1177/107319119800500407

Din, N., C., & Tat Meng, E., C. (2019). Computerized Stroop tests: A review. *Journal of Psychology & Psychotherapy*, *9*(1), 5.

Dodrill, C. B. (1978). A neuropsychological battery for epilepsy. *Epilepsia*, *19*(6), 611–623. https://doi.org/10.1111/j.1528-1157.1978.tb05041.x

Donders, J. (2020). The incremental value of neuropsychological assessment: A critical review. *The Clinical Neuropsychologist*, *34*(1), 56–87. https://doi.org/10.1080/13854046.2019.1575471

Drake, A., Weinstock-Guttman, B., Morrow, S. A., Hojnacki, D., Munschauer, F., & Benedict, R. (2010). Psychometrics and normative data for the Multiple Sclerosis Functional Composite: Replacing the PASAT with the Symbol Digit Modalities Test. *Multiple Sclerosis Journal*, *16*(2), 228–237. https://doi.org/10.1177/1352458509354552

Drew, M., Tippett, L. J., Starkey, N. J., & Isler, R. B. (2008). Executive dysfunction and cognitive impairment in a large community-based sample with Multiple Sclerosis from New Zealand: A descriptive study. *Archives of Clinical Neuropsychology*, *23*(1), 1–19. https://doi.org/10.1016/j.acn.2007.09.005

Eijlers, A. J. C., Van Geest, Q., Dekker, I., Steenwijk, M. D., Meijer, K. A., Hulst, H. E., Barkhof, F., Uitdehaag, B. M. J., Schoonheim,

M. M., & Geurts, J. J. G. (2018). Predicting cognitive decline in multiple sclerosis: A 5-year follow-up study. *Brain*. https://doi.org/10.1093/brain/awy202

Eilam-Stock, T., Best, P., Sherman, K., Shaw, M. T., Ventura, J., Krupp, L. B., & Charvet, L. E. (2021). An interview-based assessment of the experience of cognitive impairment in multiple sclerosis: The cognitive assessment interview (CAI). *Frontiers in Neurology*, *12*, 23. https://doi.org/10.3389/fneur.2021.637895

Eilam-Stock, T., Shaw, M. T., Sherman, K., Krupp, L. B., & Charvet, L. E. (2021). Remote administration of the Symbol Digit Modalities Test to individuals with multiple sclerosis is reliable: A short report. *Multiple Sclerosis Journal - Experimental, Translational and Clinical*, *7*(1), 2055217321994853. https://doi.org/10.1177/2055217321994853

Ellerby, Z., McCulloch, J., Young, J., & Wagner, C. (2019). DECSYS – Discrete and Ellipse-based response Capture SYStem. *2019 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1–6. https://doi.org/10.1109/FUZZ-IEEE.2019.8858996

Elwick, H., Smith, L., Mhizha-Murira, J. R., Topcu, G., Leighton, P., Drummond, A., Evangelou, N., & Nair, R. das. (2021). Cognitive assessment in multiple sclerosis clinical care: A qualitative evaluation of stakeholder perceptions and preferences. *Neuropsychological Rehabilitation*, 1–19. https://doi.org/10.1080/09602011.2021.1899942

Elwick, H., Topcu, G., Allen, C. M., Drummond, A., Evangelou, N., & Nair, R. das. (2021). Cognitive measures used in adults with multiple sclerosis: A systematic review. *Neuropsychological Rehabilitation*, 1–18. https://doi.org/10.1080/09602011.2021.1936080

Ernst, A. (2020). Autobiographical memory and future thinking impairments in multiple sclerosis: Cognitive and neural mechanisms, functional impact and rehabilitation. *Annals of Physical and Rehabilitation Medicine*, *63*(2), 159–166. https://doi.org/10.1016/j.rehab.2019.06.006

Ernst, A., Blanc, F., Voltzenlogel, V., Seze, J. de, Chauvin, B., & Manning, L. (2013). Autobiographical memory in multiple sclerosis patients: Assessment and cognitive facilitation. *Neuropsychological Rehabilitation*, *23*(2), 161–181. https://doi.org/10.1080/09602011.2012.724355

Feenaughty, L., Guo, L.-Y., Weinstock-Guttman, B., Ray, M., Benedict, R. H. B., & Tjaden, K. (2021). Impact of cognitive impairment and dysarthria on spoken language in multiple sclerosis. *Journal of the International Neuropsychological Society*, 1–11. https://doi.org/10.1017/S1355617720001113

Feenaughty, L., Tjaden, K., Weinstock-Guttman, B., & Benedict, R. H. B. (2018). Separate and combined influence of cognitive impairment and dysarthria on functional communication in multiple sclerosis. *American Journal of Speech-Language*

Pathology, *27*(3), 1051–1065. https://doi.org/10.1044/2018_ajslp-17-0174

Feinstein, A. (2006). Mood disorders in multiple sclerosis and the effects on cognition. *Journal of the Neurological Sciences*, *245*(1), 63–66. https://doi.org/10.1016/j.jns.2005.08.020

Feinstein, A., Magalhaes, S., Richard, J.-F., Audet, B., & Moore, C. (2014). The link between multiple sclerosis and depression. *Nature Reviews Neurology*, *10*(9), 507–517. https://doi.org/10.1038/nrneurol.2014.139

Fellows, R. P., & Schmitter-Edgecombe, M. (2019). Symbol Digit Modalities Test: Regression-based normative data and clinical utility. *Archives of Clinical Neuropsychology*, *35*(1), 105–115. https://doi.org/10.1093/arclin/acz020

Ferreira, M. (2010). Cognitive deficits in multiple sclerosis A systematic review. *Arquivos de Neuro-Psiquiatria*, *68*, 632–641. https://doi.org/10.1590/S0004-282X2010000400029

Filser, M., Schreiber, H., Pöttgen, J., Ullrich, S., Lang, M., & Penner, I. K. (2018). The Brief International Cognitive Assessment in Multiple Sclerosis (BICAMS): Results from the German validation study. *Journal of Neurology*, *265*(11), 2587–2593. https://doi.org/10.1007/s00415-018-9034-1

Fischer, M., Kunkel, A., Bublak, P., Faiss, J. H., Hoffmann, F., Sailer, M., Schwab, M., Zettl, U. K., & Köhler, W. (2014). How reliable is the classification of cognitive impairment across different criteria in early and late stages of multiple sclerosis? *Journal of the*

Neurological Sciences, *343*(1), 91–99. https://doi.org/10.1016/j.jns.2014.05.042

Fisk, J. D., Pontefract, A., Ritvo, P. G., Archibald, C. J., & Murray, T. J. (1994). The impact of fatigue on patients with multiple sclerosis. *The Canadian Journal of Neurological Sciences*, *21*(1), 9–14.

Foley, F. W., Benedict, R. H. B., Gromisch, E. S., & DeLuca, J. (2012). The need for screening, assessment, and treatment for cognitive dysfunction in multiple sclerosis: Results of a multidisciplinary CMCS consensus conference. *International Journal of MS Care*, *14*(2), 58–64. https://doi.org/10.7224/1537-2073-14.2.58

Forn, C., Belenguer, A., Parcet-Ibars, M. A., & Avila, C. (2008). Information-processing speed is the primary deficit underlying the poor performance of multiple sclerosis patients in the Paced Auditory Serial Addition Test (PASAT). *Journal of Clinical and Experimental Neuropsychology*, *30*(7), 789–796. https://doi.org/10.1080/13803390701779560

Franzen, M. D., Tishelman, A. C., Sharp, B. H., & Friedman, A. G. (1987). An investigation of the test-retest reliability of the Stroop Color-Word Test across two intervals. *Archives of Clinical Neuropsychology*, *2*(3), 265–272. https://doi.org/10.1016/0887-6177(87)90014-X

Freitas, S., Batista, S., Afonso, A. C., Simoes, M. R., de Sousa, L., Cunha, L., & Santana, I. (2016). The Montreal Cognitive Assessment (MoCA) as a screening test for cognitive

dysfunction in multiple sclerosis. *Applied Neuropsychology*, *25*(1), 57–70. https://doi.org/10.1080/23279095.2016.1243108

Gaines, J. J., Gavett, R. A., Lynch, J. J., Bakshi, R., & Benedict, R. H. B. (2008). New Error type and Recall Consistency Indices for the Brief Visuospatial Memory Test – Revised: Performance in Healthy Adults and Multiple Sclerosis Patients. *The Clinical Neuropsychologist*, *22*(5), 851–863. https://doi.org/10.1080/13854040701565216

Gallagher, C., & Burke, T. (2007). Age, gender and IQ effects on the Rey-Osterrieth Complex Figure Test. *British Journal of Clinical Psychology*, *46*, 35–45.

Genova, H. M., Lancaster, K., Lengenfelder, J., Bober, C. P., DeLuca, J., & Chiaravalloti, N. D. (2020). Relationship between social cognition and fatigue, depressive symptoms, and anxiety in multiple sclerosis. *Journal of Neuropsychology*, *14*(2), 213–225. https://doi.org/10.1111/jnp.12185

Germine, L., Reinecke, K., & Chaytor, N. S. (2019). Digital neuropsychology: Challenges and opportunities at the intersection of science and software. *The Clinical Neuropsychologist*, *33*(2), 271–286. https://doi.org/10.1080/13854046.2018.1535662

Gershon, R. C., Weiner, M. F., & Rentz, D. (2021). *Technology | Mobile Toolbox*. Mobile Toolbox: Digital Cognition. https://www.mobiletoolbox.org

Gholamzad, M., Ebtekar, M., Ardestani, M. S., Azimi, M., Mahmodi, Z., Mousavi, M. J., & Aslani, S. (2019). A comprehensive review on the treatment approaches of multiple sclerosis: Currently and in the future. *Inflammation Research*, *68*(1), 25–38. https://doi.org/10.1007/s00011-018-1185-0

Glanz, B. I., Healy, B. C., Rintell, D. J., Jaffin, S. K., Bakshi, R., & Weiner, H. L. (2010). The association between cognitive impairment and quality of life in patients with early multiple sclerosis. *Journal of the Neurological Sciences*, *290*(1–2), 75–79. https://doi.org/10.1016/j.jns.2009.11.004

Golan, D., Wilken, J., Doniger, G. M., Fratto, T., Kane, R., Srinivasan, J., Zarif, M., Bumstead, B., Buhse, M., Fafard, L., Topalli, I., & Gudesblatt, M. (2019). Validity of a multi-domain computerized cognitive assessment battery for patients with multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *30*, 154–162. https://doi.org/10.1016/j.msard.2019.01.051

Golden, C. J. (1978). *Stroop Color and Word Test: A manual for clinical and experimental uses.* Stoelting Company.

Goverover, Y., & DeLuca, J. (2015). Actual reality: Using the Internet to assess everyday functioning after traumatic brain injury. *Brain Injury*, *29*(6), 715–721. https://doi.org/10.3109/02699052.2015.1004744

Graf, P., & Uttl, B. (2001). Prospective memory: A new focus for research. *Consciousness and Cognition*, *10*(4), 437–450. https://doi.org/10.1006/ccog.2001.0504

Gromisch, E. S., Portnoy, J. G., & Foley, F. W. (2018). Comparison of the abbreviated minimal assessment of cognitive function in multiple sclerosis (aMACFIMS) and the brief international cognitive assessment for multiple sclerosis (BICAMS). *Journal of the Neurological Sciences*, *388*, 70–75. https://doi.org/10.1016/j.jns.2018.03.012

Gromisch, E. S., Turner, A. P., Haselkorn, J. K., Lo, A. C., & Agresta, T. (2020). Mobile health (mHealth) usage, barriers, and technological considerations in persons with multiple sclerosis: A literature review. *JAMIA Open*, *ooaa067*. https://doi.org/10.1093/jamiaopen/ooaa067

Grothe, M., Opolka, M., Berneiser, J., & Dressel, A. (2021). Testing social cognition in multiple sclerosis: Difference between emotion recognition and theory of mind and its influence on quality of life. *Brain and Behavior*, *11*(1), e01925. https://doi.org/10.1002/brb3.1925

Hadjimichael, O., Vollmer, T., Oleen-Burkey, M., & North American Research Committee on Multiple Sclerosis. (2008). Fatigue characteristics in multiple sclerosis: The North American Research Committee on Multiple Sclerosis (NARCOMS) survey. *Health and Quality of Life Outcomes*, *6*, 100. https://doi.org/10.1186/1477-7525-6-100

Hancock, L., Hermann, B., Schoonheim, M., Hetzel, S., Brochet, B., & Deluca, J. (2022). Comparing Diagnostic Criteria for the Diagnosis of Neurocognitive Disorders in Multiple Sclerosis.

Multiple Sclerosis and Related Disorders, *58*, 103479. https://doi.org/10.1016/j.msard.2021.103479

Hansen, S., & Lautenbacher, S. (2017). Neuropsychological assessment in multiple sclerosis. *Zeitschrift Für Neuropsychologie*, *28*(2), 117–148. https://doi.org/10.1024/1016-264X/a000197

Hartelius, L., Runmarker, B., & Andersen, O. (2000). Prevalence and characteristics of dysarthria in a multiple-sclerosis incidence cohort: Relation to neurological data. *Folia Phoniatrica et Logopaedica*, *52*(4), 160–177. https://doi.org/10.1159/000021531

Healy, B. C., Barker, L., Bakshi, R., Benedict, R. H. B., Gonzalez, C. T., Chitnis, T., Weiner, H. L., & Glanz, B. I. (2021). Trajectories of Symbol Digit Modalities Test performance in individuals with multiple sclerosis. *Multiple Sclerosis Journal*, *27*(4), 593–602. https://doi.org/10.1177/1352458520913439

Henry, J. D., & Beatty, W. W. (2006). Verbal fluency deficits in multiple sclerosis. *Neuropsychologia*, *44*(7), 1166–1174. https://doi.org/10.1016/j.neuropsychologia.2005.10.006

Herdman, M., Gudex, C., Lloyd, A., Janssen, M., Kind, P., Parkin, D., Bonsel, G., & Badia, X. (2011). Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *20*(10), 1727–1736. https://doi.org/10.1007/s11136-011-9903-x

Hobart, J. C., Riazi, A., Lamping, D. L., Fitzpatrick, R., & Thompson, A. J. (2004). Improving the evaluation of therapeutic interventions in multiple sclerosis: Development of a patient-based measure of outcome. *Health Technology Assessment*, *8*(9), iii, 1–48. https://doi.org/10.3310/hta8090

Hobart, J., Lamping, D., Fitzpatrick, R., Riazi, A., & Thompson, A. (2001). The Multiple Sclerosis Impact Scale (MSIS-29): A new patient-based outcome measure. *Brain: A Journal of Neurology*, *124*(Pt 5), 962–973. https://doi.org/10.1093/brain/124.5.962

Hughes, A. J., Bhattarai, J., Paul, S., & Beier, M. (2019). Depressive symptoms and fatigue as predictors of objective-subjective discrepancies in cognitive function in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *30*, 192–197. https://doi.org/10.1016/j.msard.2019.01.055

Hughes, A. J., Denney, D. R., Owens, E. M., & Lynch, S. G. (2013). Procedural variations in the stroop and the symbol digit modalities test: Impact on patients with multiple sclerosis. *Archives of Clinical Neuropsychology*, *28*(5), 452–462. https://doi.org/10.1093/arclin/act041

Jones, K. H., Ford, D. V., Jones, P. A., John, A., Middleton, R. M., Lockhart-Jones, H., Peng, J., Osborne, L. A., & Noble, J. G. (2013). The physical and psychological impact of multiple sclerosis using the MSIS-29 via the web portal of the UK MS Register. *PLOS ONE*, *8*(1), e55422. https://doi.org/10.1371/journal.pone.0055422

Kalb, R., Beier, M., Benedict, R., Charvet, L., Costello, K., Feinstein, A., Gingold, J., Goverover, Y., Halper, J., Harris, C., Kostich, L., Krupp, L., Lathi, E., LaRocca, N., Thrower, B., & DeLuca, J. (2018). Recommendations for cognitive screening and management in multiple sclerosis care. *Multiple Sclerosis Journal*, *24*(13), 1665–1680. https://doi.org/10.1177/1352458518803785

Kalbe, E., Calabrese, P., Fengler, S., & Kessler, J. (2013). Demtect, PANDA, EASY, and MUSIC: Cognitive screening tools with age correction and weighting of subtests according to their sensitivity and specificity. *Journal of Alzheimer's Disease : JAD*, *34*. https://doi.org/10.3233/JAD-122128

Kalmar, J. H., Gaudino, E. A., Moore, N. B., Halper, J., & Deluca, J. (2008). The relationship between cognitive deficits and everyday functional activities in multiple sclerosis. *Neuropsychology*, *22*(4), 442–449. https://doi.org/10.1037/0894-4105.22.4.442

Kesselring, J., & Beer, S. (2005). Symptomatic therapy and neurorehabilitation in multiple sclerosis. *The Lancet Neurology*, *4*(10), 643–652. https://doi.org/10.1016/S1474-4422(05)70193-9

Khaligh-Razavi, S.-M., Sadeghi, M., Khanbagi, M., Kalafatis, C., & Nabavi, S. M. (2020). A self-administered, artificial intelligence (AI) platform for cognitive assessment in multiple sclerosis (MS). *BMC Neurology*, *20*(1), 193. https://doi.org/10.1186/s12883-020-01736-x

Kinsinger, S. W., Lattie, E., & Mohr, D. C. (2010). Relationship between depression, fatigue, subjective cognitive impairment, and objective neuropsychological functioning in patients with multiple sclerosis. *Neuropsychology*, *24*(5), 573–580. https://doi.org/10.1037/a0019222

Klein, O. A., Nair, R. das, Ablewhite, J., & Drummond, A. (2019). Assessment and management of cognitive problems in people with multiple sclerosis: A National Survey of Clinical Practice. *International Journal of Clinical Practice*, *73*(3), e13300. https://doi.org/10.1111/ijcp.13300

Kobelt, G., Berg, J., Lindgren, P., Fredrikson, S., & Jönsson, B. (2006). Costs and quality of life of patients with multiple sclerosis in Europe. *Journal of Neurology, Neurosurgery & Psychiatry*, *77*(8), 918–926. https://doi.org/10.1136/jnnp.2006.090365

Konstantinopoulou, E., Ioannidis, P., Bakirtzis, C., Giantzi, V., Afrantou, T., Parissis, D., Messinis, L., Nasios, G., & Grigoriadis, N. (2018). Estimating everyday neuropsychological functioning in multiple sclerosis: Reliability and validity of the Greek Multiple Sclerosis Neuropsychological Questionnaire. *Multiple Sclerosis International*, *2018*, 6301535. https://doi.org/10.1155/2018/6301535

Koo, B. M., & Vizer, L. M. (2019). Mobile technology for cognitive assessment of older adults: A scoping review. *Innovation in Aging*, *3*(1), igy038. https://doi.org/10.1093/geroni/igy038

Korakas, N., & Tsolaki, M. (2016). Cognitive impairment in multiple sclerosis: A review of neuropsychological assessments. *Cognitive and Behavioral Neurology*, *29*(2), 55–67. https://doi.org/10.1097/wnn.0000000000000097

Kos, D., Kerckhofs, E., Nagels, G., D'hooghe, M. B., & Ilsbroukx, S. (2008). Origin of fatigue in multiple sclerosis: Review of the literature. *Neurorehabilitation and Neural Repair*, *22*(1), 91–100. https://doi.org/10.1177/1545968306298934

Kratz, A. L., Braley, T. J., Foxen-Craft, E., Scott, E., Murphy, J. F., & Murphy, S. L. (2017). How do pain, fatigue, depressive, and cognitive symptoms relate to well-being and social and physical functioning in the daily lives of individuals with multiple sclerosis? *Archives of Physical Medicine and Rehabilitation*, *98*(11), 2160–2166. https://doi.org/10.1016/j.apmr.2017.07.004

Krupp, L. (2006). Fatigue is intrinsic to multiple sclerosis (MS) and is the most commonly reported symptom of the disease. *Multiple Sclerosis Journal*, *12*(4), 367–368. https://doi.org/10.1191/135248506ms1373ed

Krupp, L. B., Alvarez, L. A., LaRocca, N. G., & Scheinberg, L. C. (1988). Fatigue in multiple sclerosis. *Archives of Neurology*, *45*(4), 435–437.

Kruse, C., Karem, P., Shifflett, K., Vegi, L., Ravi, K., & Brooks, M. (2018). Evaluating barriers to adopting telemedicine worldwide: A systematic review. *Journal of Telemedicine and Telecare*, *24*(1), 4–12. https://doi.org/10.1177/1357633X16674087

Kujala, P., Portin, R., & Ruutiainen, J. (1997). The progress of cognitive decline in multiple sclerosis. A controlled 3-year follow-up. *Brain: A Journal of Neurology*, *120 (Pt 2)*, 289–297.

Kuniavsky, M., Goodman, E., & Moed, A. (2012). *Observing the User Experience: A Practitioner's Guide to User Research* (2nd Revised edition). Morgan Kaufmann.

Kurtzke, J. F. (1983). Rating neurologic impairment in multiple sclerosis: An expanded disability status scale (EDSS). *Neurology*, *33*(11), 1444–1444. https://doi.org/10.1212/WNL.33.11.1444

Lalkhen, A. G., & McCluskey, A. (2008). Clinical tests: Sensitivity and specificity. *Continuing Education in Anaesthesia Critical Care & Pain*, *8*(6), 221–223. https://doi.org/10.1093/bjaceaccp/mkn041

Lam, K., van Oirschot, P., den Teuling, B., Hulst, H., de Jong, B., Uitdehaag, B., de Groot, V., & Killestein, J. (2021). Reliability, construct and concurrent validity of a smartphone-based cognition test in multiple sclerosis. *Multiple Sclerosis Journal*, 13524585211018104. https://doi.org/10.1177/13524585211018103

Landrø, N. I., Celius, E. G., & Sletvold, H. (2004). Depressive symptoms account for deficient information processing speed but not for impaired working memory in early phase multiple sclerosis (MS). *Journal of the Neurological Sciences*, *217*(2), 211–216. https://doi.org/10.1016/j.jns.2003.10.012

Langdon, D. W., Amato, M. P., Boringa, J., Brochet, B., Foley, F., Fredrikson, S., Hämäläinen, P., Hartung, H.-P., Krupp, L., Penner, I. K., Reder, A. T., & Benedict, R. (2012). Recommendations for a Brief International Cognitive Assessment for Multiple Sclerosis (BICAMS). *Multiple Sclerosis*, *18*(6), 891–898. https://doi.org/10.1177/1352458511431076

Larrabee, G. J., & Curtiss, G. (1995). Construct validity of various verbal and visual memory tests. *Journal of Clinical and Experimental Neuropsychology*, *17*(4), 536–547. https://doi.org/10.1080/01688639508405144

Lau, A., Chan, C. C. H., & Keung, Y. H. (1998). Visual perception and hand function in persons with multiple sclerosis. *Occupational Therapy International*, *5*(3), 194–205. https://doi.org/10.1002/oti.75

Leavitt, V. M., Tosto, G., & Riley, C. S. (2018). Cognitive phenotypes in multiple sclerosis. *Journal of Neurology*, *265*(3), 562–566. https://doi.org/10.1007/s00415-018-8747-5

Lensch, E., Matzke, M., Petereit, H.-F., Scherer, P., Schramm, S., & Calabrese, P. (2006). Identification and management of cognitive disorders in multiple sclerosis. *Journal of Neurology*, *253*(1), i29–i31. https://doi.org/10.1007/s00415-006-1107-x

Lin, S.-J., Lam, J., Beveridge, S., Vavasour, I., Traboulsee, A., Li, D. K. B., MacKay, A., McKeown, M., & Kosaka, B. (2017). Cognitive performance in subjects with multiple sclerosis is robustly influenced by gender in canonical-correlation analysis. *The*

Journal of Neuropsychiatry and Clinical Neurosciences, *29*(2), 119–127. https://doi.org/10.1176/appi.neuropsych.16040083

Lincoln, N., bradshaw, L. E., Constantinescu, C. S., Day, F., Drummond, A., Fitzsimmons, D., Harris, S., Montgomery, A. A., & das Nair, R. (2020). Group cognitive rehabilitation to reduce the psychological impact of multiple sclerosis on quality of life: The CRAMMS RCT. *Health Technology Assessment*, *24*(4). https://doi.org/10.3310/hta24040

Locke, D. E. C., Stonnington, C. M., Thomas, M. L., & Caselli, R. J. (2011). Correlates of quitting the Paced Auditory Serial Addition Test in cognitively normal older adults participating in a study of normal cognitive aging. *Journal of Clinical and Experimental Neuropsychology*, *33*(8), 937–943. https://doi.org/10.1080/13803395.2011.578571

Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., Wolinsky, J. S., Balcer, L. J., Banwell, B., Barkhof, F., Bebo, B., Calabresi, P. A., Clanet, M., Comi, G., Fox, R. J., Freedman, M. S., Goodman, A. D., Inglese, M., Kappos, L., … Polman, C. H. (2014). Defining the clinical course of multiple sclerosis: The 2013 revisions. *Neurology*, *83*(3), 278–286. https://doi.org/10.1212/WNL.0000000000000560

Mackenzie, I. S., Morant, S. V., Bloomfield, G. A., MacDonald, T. M., & O'Riordan, J. (2014). Incidence and prevalence of multiple sclerosis in the UK 1990-2010: A descriptive study in the General Practice Research Database. *Journal of Neurology,*

Neurosurgery & Psychiatry, *85*(1), 76–84. https://doi.org/10.1136/jnnp-2013-305450

Macleod, D., & Prior, M. (1996). Attention deficits in adolescents with ADHD and other clinical groups. *Child Neuropsychology*, *2*(1), 1–10. https://doi.org/10.1080/09297049608401345

Marrie, R. A., & Horwitz, R. I. (2010). Emerging effects of comorbidities on multiple sclerosis. *The Lancet Neurology*, *9*(8), 820–828. https://doi.org/10.1016/S1474-4422(10)70135-6

Marrie, R. A., Reingold, S., Cohen, J., Stuve, O., Trojano, M., Sorensen, P. S., Cutter, G., & Reider, N. (2015). The incidence and prevalence of psychiatric disorders in multiple sclerosis: A systematic review. *Multiple Sclerosis Journal*, *21*(3), 305–317. https://doi.org/10.1177/1352458514564487

Marziniak, M., Brichetto, G., Feys, P., Meyding-Lamadé, U., Vernon, K., & Meuth, S. G. (2018). The use of digital and remote communication technologies as a tool for multiple sclerosis management: Narrative review. *JMIR Rehabilitation and Assistive Technologies*, *5*(1), e7805. https://doi.org/10.2196/rehab.7805

Matarazzo, J. D. (1978). *Wechsler's Measurement and Appraisal of Adult Intelligence* (5th edition). Oxford University Press Inc.

Mazancieux, A., Souchay, C., Casez, O., & Moulin, C. J. A. (2019). Metacognition and self-awareness in Multiple Sclerosis. *Cortex*, *111*, 238–255. https://doi.org/10.1016/j.cortex.2018.11.012

Mc Auliffe, A., & Hynes, S. (2019). The impact of cognitive functioning on daily occupations for people with multiple sclerosis: A qualitative study. *The Open Journal of Occupational Therapy*, *7*(3), 1–12. https://doi.org/10.15453/2168-6408.1579

McDonald, S., Zhao, T., & Edwards, H. M. (2016). Look who's talking: Evaluating the utility of interventions during an interactive think-aloud. *Interacting with Computers*, *28*(3), 387–403. https://doi.org/10.1093/iwc/iwv014

McGuigan, C., & Hutchinson, M. (2004). The Multiple Sclerosis Impact Scale (MSIS-29) is a reliable and sensitive measure. *Journal of Neurology, Neurosurgery & Psychiatry*, *75*(2), 266–269. https://doi.org/10.1136/jnnp.2003.016899

McNicholas, N., O'Connell, K., Yap, S. M., Killeen, R. P., Hutchinson, M., & McGuigan, C. (2018). Cognitive dysfunction in early multiple sclerosis: A review. *QJM: An International Journal of Medicine*, *111*(6), 359–364. https://doi.org/10.1093/qjmed/hcx070

Middleton, L. S., Denney, D. R., Lynch, S. G., & Parmenter, B. (2006). The relationship between perceived and objective cognitive functioning in multiple sclerosis. *Archives of Clinical Neuropsychology*, *21*(5), 487–494. https://doi.org/10.1016/j.acn.2006.06.008

Middleton, R. M., Pearson, O. R., Ingram, G., Craig, E. M., Rodgers, W. J., Downing-Wood, H., Hill, J., Tuite-Dalton, K., Roberts, C., Watson, L., Ford, D. V., Nicholas, R., & UK MS Register

Research Group. (2020). A rapid electronic cognitive assessment measure for multiple sclerosis: Validation of cognitive reaction, an electronic version of the Symbol Digit Modalities Test. *Journal of Medical Internet Research*, *22*(9), e18234. https://doi.org/10.2196/18234

Migliore, S., Landi, D., Proietti, F., D'Aurizio, G., Squitieri, F., Mataluni, G., Nicoletti, C. G., Curcio, G., & Marfia, G. A. (2021). Validity of the Italian Multiple Sclerosis Neuropsychological Screening Questionnaire. *Neurological Sciences*. https://doi.org/10.1007/s10072-021-05141-1

Miller, D., Barkhof, F., Montalban, X., Thompson, A., & Filippi, M. (2005). Clinically isolated syndromes suggestive of multiple sclerosis, Part I: Natural history, pathogenesis, diagnosis, and prognosis. *The Lancet Neurology*, *4*(5), 281–288. https://doi.org/10.1016/S1474-4422(05)70071-5

Miller, J. B., & Barr, W. B. (2017). The technology crisis in neuropsychology. *Archives of Clinical Neuropsychology*, *32*(5), 541–554. https://doi.org/10.1093/arclin/acx050

Mills, R., Young, C., Nicholas, R., Pallant, J., & Tennant, A. (2009). Rasch analysis of the Fatigue Severity Scale in multiple sclerosis. *Multiple Sclerosis Journal*, *15*(1), 81–87. https://doi.org/10.1177/1352458508096215

Mitrushina, M. N., Boone, K. B., Razani, L. J., & D'Elia, L. F. (2005). *Handbook of Normative Data for Neuropsychological Assessment* (Second Edition). Oxford University Press.

Moher, D. ; L., A. ;. Tetzlaff, J. ;. Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*, *6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Moore, R. C., Swendsen, J., & Depp, C. A. (2017). Applications for self-administered mobile cognitive assessments in clinical research: A systematic review. *International Journal of Methods in Psychiatric Research*, *26*(4). https://doi.org/10.1002/mpr.1562

Morrow, S. A. (2013). Normative data for the Stroop Color Word test for a North American population. *Canadian Journal of Neurological Sciences*, *40*(6), 842–847. https://doi.org/10.1017/S0317167100015997

Morrow, S. A., O'Connor, P., Polman, C., Goodman, A., Kappos, L., Lublin, F., Rudick, R., Jurgensen, S., Paes, D., Forrestal, F., & Benedict, R. (2010). Evaluation of the Symbol Digit Modalities Test (SDMT) and MS Neuropsychological Screening Questionnaire (MSNQ) in natalizumab-treated MS patients over 48 weeks. *Multiple Sclerosis Journal*, *16*(11), 1385–1392. https://doi.org/10.1177/1352458510378021

Morrow, S., Weinstock-Guttman, B., Munschauer, F., Hojnacki, D., & Benedict, R. (2009). Subjective fatigue is not associated with cognitive impairment in multiple sclerosis: Cross-sectional and longitudinal analysis. *Multiple Sclerosis Journal*, *15*(8), 998–1005. https://doi.org/10.1177/1352458509106213

Mortensen, G. L., Theódórsdóttir, Á., Sejbæk, T., & Illes, Z. (2020). Patient attitudes to routine cognitive testing in multiple sclerosis. *Patient Preference and Adherence*, *14*, 693–704. https://doi.org/10.2147/PPA.S245623

MS Society. (2020). *MS Prevalence Report*. MS Society. https://www.mssociety.org.uk/what-we-do/our-work/our-evidence/ms-in-the-uk

Mummery, C. J., & Kipps, C. M. (2020). UK neurology response to the COVID-19 crisis. *Clinical Medicine*, *20*(3), 266–269. https://doi.org/10.7861/clinmed.2020-0159

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

National Institute for Health and Care Excellence. (2019). *Multiple sclerosis in adults: Management*. NICE. https://www.nice.org.uk/guidance/cg186/chapter/Context

Nauta, I. M., Balk, L. J., Sonder, J. M., Hulst, H. E., Uitdehaag, B. M., Fasotti, L., & Jong, B. A. de. (2018). The clinical value of the patient-reported multiple sclerosis neuropsychological screening questionnaire: *Multiple Sclerosis Journal*. https://doi.org/10.1177/1352458518777295

Newland, P., Starkweather, A., & Sorenson, M. (2016). Central fatigue in multiple sclerosis: A review of the literature. *The Journal of Spinal Cord Medicine*, *39*(4), 386–399. https://doi.org/10.1080/10790268.2016.1168587

Noffs, G., Perera, T., Kolbe, S. C., Shanahan, C. J., Boonstra, F. M. C., Evans, A., Butzkueven, H., van der Walt, A., & Vogel, A. P. (2018). What speech can tell us: A systematic review of dysarthria characteristics in Multiple Sclerosis. *Autoimmunity Reviews*, *17*(12), 1202–1209. https://doi.org/10.1016/j.autrev.2018.06.010

Norman, M., Wilson, S. J., Baxendale, S., Barr, W., Block, C., Busch, R. M., Fernandez, A., Hessen, E., Loring, D. W., McDonald, C. R., & Hermann, B. P. (2021). Addressing neuropsychological diagnostics in adults with epilepsy: Introducing the International Classification of Cognitive Disorders in Epilepsy: The IC CODE Initiative. *Epilepsia Open*, *6*(2), 266–275. https://doi.org/10.1002/epi4.12478

Nucci, M., Mapelli, D., & Mondini, S. (2011). Cognitive Reserve Index questionnaire (CRIq): A new instrument for measuring cognitive reserve. *Aging Clinical and Experimental Research*, *24*. https://doi.org/10.3275/7800

Nunally, J. C., & Bernstein, I. H. (1994). *Psychometric theory (3rd ed.).* McGraw-Hill, Inc.

O'Brien, A., Gaudino-Goering, E., Shawaryn, M., Komaroff, E., Moore, N. B., & DeLuca, J. (2007). Relationship of the Multiple Sclerosis

Neuropsychological Questionnaire (MSNQ) to functional, emotional, and neuropsychological outcomes. *Archives of Clinical Neuropsychology*, *22*(8), 933–948. https://doi.org/10.1016/j.acn.2007.07.002

O'Connor, S., Hanlon, P., O'Donnell, C. A., Garcia, S., Glanville, J., & Mair, F. S. (2016). Understanding factors affecting patient and public engagement and recruitment to digital health interventions: A systematic review of qualitative studies. *BMC Medical Informatics and Decision Making*, *16*(1), 120. https://doi.org/10.1186/s12911-016-0359-3

Office of National Statistics. (2020). *Internet access—Households and individuals*. https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/datasets/internetaccesshouseholdsandindividuals95confidenceintervals

Oreja-Guevara, C., Ayuso Blanco, T., Brieva Ruiz, L., Hernández Pérez, M. Á., Meca-Lallana, V., & Ramió-Torrentà, L. (2019). Cognitive dysfunctions and assessments in multiple sclerosis. *Frontiers in Neurology*, *10*. https://doi.org/10.3389/fneur.2019.00581

Ostrosky-Solís, F., & Lozano, A. (2006). Digit Span: Effect of education and culture. *International Journal of Psychology*, *41*(5), 333–341. https://doi.org/10.1080/00207590500345724

Parmenter, B. A., Testa, S. M., Schretlen, D. J., Weinstock-Guttman, B., & Benedict, R. (2010). The utility of regression-based norms

in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). *Journal of the International Neuropsychological Society*, *16*(1), 6–16. https://doi.org/10.1017/S1355617709990750

Parmenter, B. A., Weinstock-Guttman, B., Garg, N., Munschauer, F., & Benedict, R. H. B. (2007). Screening for cognitive impairment in multiple sclerosis using the Symbol digit Modalities Test. *Multiple Sclerosis*, *13*(1), 52–57. https://doi.org/10.1177/1352458506070750

Patel, V. P., Shen, L., Rose, J., & Feinstein, A. (2019). Taking the tester out of the SDMT: A proof of concept fully automated approach to assessing processing speed in people with MS. *Multiple Sclerosis Journal*, *25*(11), 1506–1513. https://doi.org/10.1177/1352458518792772

Pawson, R., & Tilley, N. (1997). *Realistic Evaluation*. Sage.

Periáñez, J. A., Lubrini, G., García-Gutiérrez, A., & Ríos-Lago, M. (2021). Construct validity of the stroop color-word test: Influence of speed of visual search, verbal fluency, working memory, cognitive flexibility, and conflict monitoring. *Archives of Clinical Neuropsychology*, *36*(1), 99–111. https://doi.org/10.1093/arclin/acaa034

Pham, L., Harris, T., Varosanec, M., Morgan, V., Kosa, P., & Bielekova, B. (2021). Smartphone-based symbol-digit modalities test reliably captures brain damage in multiple sclerosis. *Digital*

Medicine, *4*(1), 1–13. https://doi.org/10.1038/s41746-021-00401-y

Pinkston, J. B., Kablinger, A., & Alekseeva, N. (2007). Multiple Sclerosis and Behavior. In *International Review of Neurobiology* (Vol. 79, pp. 323–339). Academic Press. https://doi.org/10.1016/S0074-7742(07)79014-1

Pitteri, M., Dapor, C., Ziccardi, S., Guandalini, M., Meggiato, R., & Calabrese, M. (2020). Visual-attentional load unveils slowed processing speed in multiple sclerosis patients: A pilot study with a tablet-based videogame. *Brain Sciences*, *10*(11), 871. https://doi.org/10.3390/brainsci10110871

Pitteri, M., Ziccardi, S., Dapor, C., Guandalini, M., & Calabrese, M. (2019). Lost in classification: Lower cognitive functioning in apparently cognitive normal newly diagnosed RRMSpatients. *Brain Sciences*, *9*(11). https://doi.org/10.3390/brainsci9110321

Polit, D. F. (2014). Getting serious about test–retest reliability: A critique of retest research and some recommendations. *Quality of Life Research*, *23*(6), 1713–1720. https://doi.org/10.1007/s11136-014-0632-9

Polman, C. H., & Rudick, R. A. (2010). The Multiple Sclerosis Functional Composite: A clinically meaningful measure of disability. *Neurology*, *74 Suppl 3*, S8-15. https://doi.org/10.1212/WNL.0b013e3181dbb571

Pope, C., Ziebland, S., & Mays, N. (2000). Analysing qualitative data. *BMJ : British Medical Journal*, *320*(7227), 114–116.

Powell, D. J. H., Liossi, C., Schlotz, W., & Moss-Morris, R. (2017). Tracking daily fatigue fluctuations in multiple sclerosis: Ecological momentary assessment provides unique insights. *Journal of Behavioral Medicine*, *40*(5), 772–783. https://doi.org/10.1007/s10865-017-9840-4

Prakash, R. S., Snook, E. M., Lewis, J. M., Motl, R. W., & Kramer, A. F. (2008). Cognitive impairments in relapsing-remitting multiple sclerosis: A meta-analysis. *Multiple Sclerosis Journal*, *14*(9), 1250–1261. https://doi.org/10.1177/1352458508095004

Puente, A. E., Adams, R., Barr, W. B., Bush, S. S., NAN Policy and Planning Committee, Ruff, R. M., Barth, J. T., Broshek, D., Koffler, S. P., Reynolds, C., Silver, C. H., & Tröster, A. I. (2006). The use, education, training and supervision of neuropsychological test technicians (psychometrists) in clinical practice. *Archives of Clinical Neuropsychology*, *21*(8), 837–839. https://doi.org/10.1016/j.acn.2006.08.011

Rabin, L. A., Barr, W. B., & Burton, L. A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, *20*(1), 33–65. https://doi.org/10.1016/j.acn.2004.02.005

Rao, S. M. (1990). *A manual for the Brief Repeatable Battery of neuropsychological tests in multiple sclerosis*. Medical College of Wisconsin.

Rao, S. M., Leo, G. J., Bernardin, L., & Unverzagt, F. (1991). Cognitive dysfunction in multiple sclerosis. I. Frequency, patterns, and prediction. *Neurology*, *41*(5), 685–691.

Rao, S. M., Losinski, G., Mourany, L., Schindler, D., Mamone, B., Reece, C., Kemeny, D., Narayanan, S., Miller, D. M., Bethoux, F., Bermel, R. A., Rudick, R., & Alberts, J. (2017). Processing speed test: Validation of a self-administered, iPad®-based tool for screening cognitive dysfunction in a clinic setting. *Multiple Sclerosis Journal*, *23*(14), 1929–1937. https://doi.org/10.1177/1352458516688955

Rao, S. M., Martin, A. L., Huelin, R., Wissinger, E., Khankhel, Z., Kim, E., & Fahrbach, K. (2014). Correlations between MRI and information processing speed in MS: A meta-analysis. *Multiple Sclerosis International*, *2014*, 975803. https://doi.org/10.1155/2014/975803

Regard, M. (1981). Cognitive rigidity and flexibility. A neuropsychological study. *PhD Dissertation, University of Victoria*. https://ci.nii.ac.jp/naid/10016808834/

Renauld, S., Mohamed-Saïd, L., & Macoir, J. (2016). Language disorders in multiple sclerosis: A systematic review. *Multiple Sclerosis and Related Disorders*, *10*, 103–111. https://doi.org/10.1016/j.msard.2016.09.005

Renner, A., Baetge, S. J., Filser, M., Ullrich, S., Lassek, C., & Penner, I. K. (2020). Characterizing cognitive deficits and potential predictors in multiple sclerosis: A large nationwide study applying

Brief International Cognitive Assessment for Multiple Sclerosis in standard clinical care. *Journal of Neuropsychology*, *14*(3), 347–369. https://doi.org/10.1111/jnp.12202

Rey, A. (1941). L'examen psychologie dans les cas d'encephalopathie traumatique. *Archives de Psychologie.*, *28*, 286–340.

Rimkus, C. de M., Avolio, I. M. B., Miotto, E. C., Pereira, S. A., Mendes, M. F., Callegaro, D., & Leite, C. da C. (2018). The protective effects of high-education levels on cognition in different stages of multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *22*, 41–48. https://doi.org/10.1016/j.msard.2018.03.001

Romero, K., Shammi, P., & Feinstein, A. (2015). Neurologists accuracy in predicting cognitive impairment in multiple sclerosis. *Multiple Sclerosis and Related Disorders*, *4*(4), 291–295. https://doi.org/10.1016/j.msard.2015.05.009

Rosti-Otajärvi, E., Ruutiainen, J., Huhtala, H., & Hämäläinen, P. (2014). Relationship between subjective and objective cognitive performance in multiple sclerosis. *Acta Neurologica Scandinavica*, *130*(5), 319–327. https://doi.org/10.1111/ane.12238

Rouleau, I., Dagenais, E., Tremblay, A., Demers, M., Roger, É., Jobin, C., & Duquette, P. (2018). Prospective memory impairment in multiple sclerosis: A review. *The Clinical Neuropsychologist*, *32*(5), 922–936. https://doi.org/10.1080/13854046.2017.1361473

Ruano, L., Branco, M., Severo, M., Sousa, A., Castelo, J., Araújo, I., Pais, J., Cerqueira, J., Amato, M. P., Lunet, N., & Cruz, V. T.

(2020). Tracking cognitive impairment in multiple sclerosis using the Brain on Track test: A validation study. *Neurological Sciences: Official Journal of the Italian Neurological Society and of the Italian Society of Clinical Neurophysiology*, *41*(1), 183–191. https://doi.org/10.1007/s10072-019-04088-8

Ruano, L., Portaccio, E., Goretti, B., Niccolai, C., Severo, M., Patti, F., Cilia, S., Gallo, P., Grossi, P., Ghezzi, A., Roscio, M., Mattioli, F., Stampatori, C., Trojano, M., Viterbo, R. G., & Amato, M. P. (2017). Age and disability drive cognitive impairment in multiple sclerosis across disease subtypes. *Multiple Sclerosis Journal*, *23*(9), 1258–1267. https://doi.org/10.1177/1352458516674367

Rudick, R. A., Cutter, G., & Reingold, S. (2002). The Multiple Sclerosis Functional Composite: A new clinical outcome measure for multiple sclerosis trials. *Multiple Sclerosis Journal*, *8*(5), 359–365. https://doi.org/10.1191/1352458502ms845oa

Ruff, R. M., Light, R. H., Parker, S. B., & Levin, H. S. (1996). Benton Controlled Oral Word Association Test: Reliability and updated norms. *Archives of Clinical Neuropsychology*, *11*(4), 329–338.

Ryan, J. J., Gontkovsky, S. T., Kreiner, D. S., & Tree, H. A. (2012). Wechsler Adult Intelligence Scale-Fourth Edition performance in relapsing-remitting multiple sclerosis. *Journal of Clinical and Experimental Neuropsychology.*, *34*(6), 571–579. https://doi.org/10.1080/13803395.2012.666229

Salthouse, T. A. (2011). What cognitive abilities are involved in trail-making performance? *Intelligence*, *39*(4), 222–232. https://doi.org/10.1016/j.intell.2011.03.001

Sampson, H. (1958). Serial addition as a function of stimulus duration and pacing. *Canadian Journal of Psychology*, *12*(3), 179–183. https://doi.org/10.1037/h0083750

Sandry, J., Simonet, D. V., Brandstadter, R., Krieger, S., Sand, I. K., Graney, R. A., Buchanan, A. V., Lall, S., & Sumowski, J. F. (2021). The Symbol Digit Modalities Test (SDMT) is sensitive but non-specific in MS: Lexical access speed, memory, and information processing speed independently contribute to sdmt performance. *Multiple Sclerosis and Related Disorders*, 102950. https://doi.org/10.1016/j.msard.2021.102950

Santangelo, G., Altieri, M., Enzinger, C., Gallo, A., & Trojano, L. (2019). Cognitive reserve and neuropsychological performance in multiple sclerosis: A meta-analysis. *Neuropsychology*, *33*(3), 379–390. https://doi.org/10.1037/neu0000520

Scarrabelotti, M., & Carroll, M. (1999). Memory dissociation and metamemory in multiple sclerosis. *Neuropsychologia*, *37*(12), 1335–1350. https://doi.org/10.1016/S0028-3932(99)00042-1

Schäffler, N., Schönberg, P., Stephan, J., Stellmann, J.-P., Gold, S. M., & Heesen, C. (2013). Comparison of patient-reported outcome measures in multiple sclerosis. *Acta Neurologica Scandinavica*, *128*(2), 114–121. https://doi.org/10.1111/ane.12083

Schependom, J. V., D'hooghe, M. B., Cleynhens, K., D'hooge, M., Haelewyck, M. C., Keyser, J. D., & Nagels, G. (2014). The Symbol Digit Modalities Test as sentinel test for cognitive impairment in multiple sclerosis. *European Journal of Neurology*, *21*(9), 1219-e72. https://doi.org/10.1111/ene.12463

Schoonheim, M., Popescu, V., Rueda Lopes, F., Wiebenga, O., Vrenken, H., Douw, L., Polman, C., Geurts, J., & Barkhof, F. (2012). Subcortical atrophy and cognition: Sex effects in multiple sclerosis. *Neurology*, *79*, 1754–1761. https://doi.org/10.1212/WNL.0b013e3182703f46

Sepulcre, J., Vanotti, S., Hernández, R., Sandoval, G., Cáceres, F., Garcea, O., & Villoslada, P. (2006). Cognitive impairment in patients with multiple sclerosis using the Brief Repeatable Battery-Neuropsychology test. *Multiple Sclerosis Journal*, *12*(2), 187–195. https://doi.org/10.1191/1352458506ms1258oa

Settle, J. R., Robinson, S. A., Kane, R., Maloni, H. W., & Wallin, M. T. (2015). Remote cognitive assessments for patients with multiple sclerosis: A feasibility study. *Multiple Sclerosis*, *21*(8), 1072–1079. https://doi.org/10.1177/1352458514559296

Sheridan, L. K., Fitzgerald, H. E., Adams, K. M., Nigg, J. T., Martel, M. M., Puttler, L. I., Wong, M. M., & Zucker, R. A. (2006). Normative Symbol Digit Modalities Test performance in a community-based sample. *Archives of Clinical Neuropsychology*, *21*(1), 23–28. https://doi.org/10.1016/j.acn.2005.07.003

Smith, A. (1982). *Symbol Digit Modalities Test (SDMT). Manual (Revised)*. Western Psychological Services.

Smith, A. C., Thomas, E., Snoswell, C. L., Haydon, H., Mehrotra, A., Clemensen, J., & Caffery, L. J. (2020). Telehealth for global emergencies: Implications for coronavirus disease 2019 (COVID-19). *Journal of Telemedicine and Telecare*, *26*(5), 309–313. https://doi.org/10.1177/1357633X20916567

Sonder, J. M., Burggraaff, J., Knol, D. L., Polman, C. H., & Uitdehaag, B. M. (2014). Comparing long-term results of PASAT and SDMT scores in relation to neuropsychological testing in multiple sclerosis. *Multiple Sclerosis Journal*, *20*(4), 481–488. https://doi.org/10.1177/1352458513501570

Sonder, J. M., Mokkink, L. B., van der Linden, F. A. H., Polman, C. H., & Uitdehaag, B. M. J. (2012). Validation and interpretation of the Dutch version of the Multiple Sclerosis Neuropsychological Screening Questionnaire. *Journal of the Neurological Sciences*, *320*(1), 91–96. https://doi.org/10.1016/j.jns.2012.06.024

Spedo, C., Foss, M., Pereira, D., Conde, R., Motta, T., Marques, V., & Barreira, A. (2014). The Wechsler Memory Scale-fourth edition (WMS-IV) on memory evaluation of multiple sclerosis patients. *Multiple Sclerosis Journal*, *20*, 67–284.

Stegen, S., Stepanov, I., Cookfair, D., Schwartz, E., Hojnacki, D., Weinstock-Guttman, B., & Benedict, R. (2009). Validity of the California Verbal Learning Test-II in multiple sclerosis. *The*

Clinical Neuropsychologist, *24*(2), 189–202. https://doi.org/10.1080/13854040903266910

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society: JINS*, *8*(3), 448–460.

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary*. Oxford University Press.

Strober, L. B., DeLuca, J., Benedict, R., Jacobs, A., Cohen, J. A., Chiaravalloti, N., Hudson, L. D., Rudick, R. A., & LaRocca, N. G. (2018). Symbol Digit Modalities Test: A valid clinical trial endpoint for measuring cognition in multiple sclerosis. *Multiple Sclerosis Journal*, 1352458518808204. https://doi.org/10.1177/1352458518808204

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. https://doi.org/10.1037/h0054651

Sullivan, M. J., Edgley, K., & Dehoux, E. (1990). A survey of multiple sclerosis: I. Perceived cognitive problems and compensatory strategy use. *Canadian Journal of Rehabilitation*, *4*(2), 99–105.

Sumowski, J. F., Benedict, R., Enzinger, C., Filippi, M., Geurts, J. J., Hamalainen, P., Hulst, H., Inglese, M., Leavitt, V. M., Rocca, M. A., Rosti-Otajarvi, E. M., & Rao, S. (2018). Cognition in multiple sclerosis. *Neurology*, *90*(6), 278–288. https://doi.org/10.1212/WNL.0000000000004977

Sumowski, J. F., & Leavitt, V. M. (2013). Cognitive reserve in multiple sclerosis. *Multiple Sclerosis Journal*, *19*(9), 1122–1127. https://doi.org/10.1177/1352458513498834

The Multiple Sclerosis International Federation. (2020). *Atlas of MS* (Third edition).

Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., Correale, J., Fazekas, F., Filippi, M., Freedman, M. S., Fujihara, K., Galetta, S. L., Hartung, H. P., Kappos, L., Lublin, F. D., Marrie, R. A., Miller, A. E., Miller, D. H., Montalban, X., … Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *The Lancet Neurology*, *17*(2), 162–173. https://doi.org/10.1016/S1474-4422(17)30470-2

Thornton, A. E., & Raz, N. (1997). Memory impairment in multiple sclerosis: A quantitative review. *Neuropsychology*, *11*(3), 357–366. https://doi.org/10.1037/0894-4105.11.3.357

Tobin, W. O. (2019). Management of multiple sclerosis symptoms and comorbidities. *Continuum*, *25*(3), 753–772. https://doi.org/10.1212/CON.0000000000000732

Tombaugh, T. N. (2006). A comprehensive review of the Paced Auditory Serial Addition Test (PASAT). *Archives of Clinical Neuropsychology*, *21*(1), 53–76. https://doi.org/10.1016/j.acn.2005.07.006

Topol, E. (2020). *The Topol Review*. https://topol.hee.nhs.uk/

Trahan, L., Stuebing, K. K., Hiscock, M. K., & Fletcher, J. M. (2014). The Flynn Effect: A Meta-analysis. *Psychological Bulletin*, *140*(5), 1332–1360. https://doi.org/10.1037/a0037173

Trenerry, M. R., Crosson, B., DeBoe, J., & Leber, W. R. (1989). *Stroop Neuropsychological Screening Test manual*. Psychological Assessment Resources.

Valdes, E. G., Sadeq, N. A., Harrison Bush, A. L., Morgan, D., & Andel, R. (2016). Regular cognitive self-monitoring in community-dwelling older adults using an internet-based tool. *Journal of Clinical and Experimental Neuropsychology*, *38*(9), 1026–1037. https://doi.org/10.1080/13803395.2016.1186155

Van der Elst, W., Molenberghs, G., Van Boxtel, M. P. J., & Jolles, J. (2013). Establishing normative data for repeated cognitive assessment: A comparison of different statistical methods. *Behavior Research Methods*, *45*(4), 1073–1086. https://doi.org/10.3758/s13428-012-0305-y

van Dongen, L. van, Westerik, B., Hiele, K. van der, Visser, L. H., Schoonheim, M. M., Douw, L., Twisk, J. W. R., Jong, B. A. de, Geurts, J. J. G., & Hulst, H. E. (2020). Introducing Multiple Screener: An unsupervised digital screening tool for cognitive deficits in MS. *Multiple Sclerosis and Related Disorders*, *38*. https://doi.org/10.1016/j.msard.2019.101479

Van Oirschot, P., Heerings, M., Wendrich, K., Den Teuling, B., Martens, M. B., & Jongen, P. J. (2020). Symbol Digit Modalities Test variant in a smartphone app for persons with multiple sclerosis:

Validation study. *JMIR MHealth and UHealth*, *8*(10), e18160. https://doi.org/10.2196/18160

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: Systematic analysis of qualitative health research over a 15-year period. *BMC Medical Research Methodology*, *18*(1), 148. https://doi.org/10.1186/s12874-018-0594-7

Ventura, J., Reise, S. P., Keefe, R. S. E., Baade, L. E., Gold, J. M., Green, M. F., Kern, R. S., Mesholam-Gately, R., Nuechterlein, K. H., Seidman, L. J., & Bilder, R. M. (2010). The Cognitive Assessment Interview (CAI): Development and validation of an empirically derived, brief interview-based measure of cognition. *Schizophrenia Research*, *121*(1–3), 24–31. https://doi.org/10.1016/j.schres.2010.04.016

Visser, L. A., Louapre, C., Uyl-de Groot, C. A., & Redekop, W. K. (2021). Health-related quality of life of multiple sclerosis patients: A European multi-country study. *Archives of Public Health*, *79*(1), 39. https://doi.org/10.1186/s13690-021-00561-z

Vleugels, L., Lafosse, C., van Nunen, A., Nachtergaele, S., Ketelaer, P., Charlier, M., & Vandenbussche, E. (2000). Visuoperceptual impairment in multiple sclerosis patients diagnosed with neuropsychological tasks. *Multiple Sclerosis*, *6*(4), 241–254. https://doi.org/10.1177/135245850000600406

Walker, L. A. S., Cheng, A., Berard, J., Berrigan, L. I., Rees, L. M., & Freedman, M. S. (2012). Tests of information processing speed:

What do people with multiple sclerosis think about them? *International Journal of MS Care*, *14*(2), 92–99. https://doi.org/10.7224/1537-2073-14.2.92

Walton, C., King, R., Rechtman, L., Kaye, W., Leray, E., Marrie, R. A., Robertson, N., La Rocca, N., Uitdehaag, B., van der Mei, I., Wallin, M., Helme, A., Angood Napier, C., Rijke, N., & Baneke, P. (2020). Rising prevalence of multiple sclerosis worldwide: Insights from the Atlas of MS, third edition. *Multiple Sclerosis*, *26*(14), 1816–1821. https://doi.org/10.1177/1352458520970841

Weber, E., Goverover, Y., & DeLuca, J. (2019). Beyond cognitive dysfunction: Relevance of ecological validity of neuropsychological tests in multiple sclerosis. *Multiple Sclerosis Journal*, *25*(10), 1412–1419. https://doi.org/10.1177/1352458519860318

Weizenbaum, E., Torous, J., & Fulford, D. (2020). Cognition in context: Understanding the everyday predictors of cognitive performance in a new era of measurement. *JMIR MHealth and UHealth*, *8*(7), e14328. https://doi.org/10.2196/14328

Wills, S., & Leathem, J. (2004). The effects of test anxiety, age, intelligence level, and arithmetic ability on paced auditory serial addition test performance. *Applied Neuropsychology*, *11*(4), 178–185. https://doi.org/10.1207/s15324826an1104_2

Wojcik, C. M., Beier, M., Costello, K., DeLuca, J., Feinstein, A., Goverover, Y., Gudesblatt, M., Jaworski, M., Kalb, R., Kostich, L., LaRocca, N. G., Rodgers, J. D., Benedict, R., & National MS

Society Cognition Work Team. (2019). Computerized neuropsychological assessment devices in multiple sclerosis: A systematic review. *Multiple Sclerosis*, *25*(14), 1848–1869. https://doi.org/10.1177/1352458519879094

Yigit, P., Acikgoz, A., Mehdiyev, Z., Dayi, A., & Ozakbas, S. (2020). The relationship between cognition, depression, fatigue, and disability in patients with multiple sclerosis. *Irish Journal of Medical Science*. https://doi.org/10.1007/s11845-020-02377-2

Zakzanis, K. K. (2000). Distinct neurocognitive profiles in multiple sclerosis subtypes. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists*, *15*(2), 115–136.

Zec, R., Kohlrus, S., Robbs, R., & Ala, T. (2015). Effects of age, education, and gender on Trail Making Test parts A & B in non-demented older adults. *Archives of Clinical Neuropsychology*, *30*(6), 525–525. https://doi.org/10.1093/arclin/acv047.105

Zimmermann, P., & Fimm, B. (1995). *Test for Attentional Performance (TAP)*. PsyTest.

Zurawski, J., Healy, B. C., Ratajska, A., Barker, L., Glanz, B. I., & Houtchens, M. (2020). Identification of a predominant cognitive phenotype in patients with multiple sclerosis. *European Journal of Neurology*, *27*(6), 1083–1088. https://doi.org/10.1111/ene.14186

Tables

Table A: Data extraction table

	First Author	Study Title	Year	Location of study	Intervention/treatment/par	Who completed the measure of						Cognitive Test		Use of Cognitive Test						Modality						Reference
						PwMS			Carers/ families	Clinician reported	Healthy Control	Other	Name of measure	Self/proxy reported	For screening	To report cognitive	As dependent	As an independe	Other	Face to face	Postal	Phone	Electronic	No Informatio	
						Number	Mean Age	Type of MS																		
1	Abad	The analysis of semantic networks in multiple	2015	Spain	N/a	36	36	No information			200		National Adult Reading Test - Revised		X									X	Abad, E., Sepulcre, J., Martinez-Lapiscina, E. H., Zubizarreta, I., Garcia-Ojalvo, J., & Villoslada, P. (2015). The analysis of semantic networks in
													VFT - animals			X				X					
2	Achiron	Modeling of Cognitive Impairment by Disease Duration in Multiple Sclerosis: A Cross-Sectional Study	2013	Israel	N/a	1500	40.5	CIS, RRMS, SPMS, PPMS					MindStreams GAB			X							X			Achiron, A., Chapman, J., Magalashvili, D., Dolev, M., Lavie, M., Bercovich, E., . . . Barak, Y. (2013). Modeling of Cognitive Impairment by Disease Duration in Multiple Sclerosis: A Cross-Sectional Study. PLoS ONE, 8 (8) (no pagination)(e71058).
3	Achiron	Superior temporal gyrus thickness correlates with cognitive performance in multiple sclerosis.	2013	Israel	N/a	20	32.2	Mild RRMS			20		MCCB (http://www.neurotrax.com)			X							X			Achiron, A., Chapman, J., Tal, S., Bercovich, E., Gil, H., & Achiron, A. (2013). Superior temporal gyrus thickness correlates with cognitive performance in multiple sclerosis. Brain Structure & Function, 218(4), 943-950.
4	Achiron	Prolonged response	2007	Israel	N/a	58	41.5	Probable,			71		NSBMS			X								X	Achiron, A., Doniger, G. M., Harel, Y.,
													Mindstreams			X							X		Appleboim-Gavish, N., Lavie, M., &
5	Achiron	Cognitive patterns and progression in multiple sclerosis: Construction and validation of percentile curves.	2005	Israel	N/a	233	40.8	RRMS					NSBMS			X				X						Achiron, A., Polliack, M., Rao, S. M., Borak, Y., Lavie, M., Appelboim, N., & Harel, Y. (2005). Cognitive patterns and progression in multiple sclerosis: Construction and validation of percentile curves. Journal of Neurology, Neurosurgery and Psychiatry, 76(5), 744-749.

Table B: Individual neuropsychological tests used in MS research with adults

Class	Examples of Tests	Description	Domains	Time Taken to Administer	Advantages	Disadvantages
ST-Substitution tests	Symbol Digit Modalities Test (SDMT (Smith, 1982)), Digit Symbol (DS), Letter Digit Substitution Test	Substitution tests consist of a coding key, which can display symbols, digits, letters, or other stimuli. Respondents use the key to match the numbers and symbols and write down or say aloud the corresponding letter/symbol to the stimuli as quickly as they are able to. It can be scored on correct responses, and time taken.	Attention, perceptual speed, motor speed, and memory	5 minutes	SDMT is one of the strongest measures of cognitive processing speed in MS patients psychometrically. Normative data is available, test re-test data is adequate and equivalent alternate versions have been developed to moderate practice effects (Benedict et al., 2017).	Age and education can have a significant effect this might be due to the change in motor response and visual acuity (Bowler et al., 1992). Gender has been shown to have an effect on score (with females usually scoring higher than males), and this variation can be pronounced (Smith, 1982).

Class	Examples of Tests	Description	Domains	Time Taken to Administer	Advantages	Disadvantages
SAT – Serial Addition Test	Paced Auditory Serial Addition Test (PASAT), Paced Visual Serial Addition Test (PVSAT)	Participants are presented with a number every few seconds and asked to add it to the proceeding number. The time between numbers and presentation of numbers (verbally or visually) vary between tests. It can be scored in different ways, including number of correct answers, time per correct response, latency of response, number of errors.	Attention/concentration, immediate memory, and information processing.	10-15 minutes	SAT is relatively simple to administer and score. There is considerably evidence of construct validity, and that the PASAT loads onto a range of domains and has good internal consistency (Macleod & Prior, 1996). Considerable norming data exists for the version used in the MSFC.	SAT was found to cause significant anxiety in participants and some will not wish to complete or participate in the test (Aupperle, et al., 2002; Cortés-Martínez et al., 2019). This measure also has significant practice effects over repeated measures (Barker-Collo, 2005). Significant effects of intelligence, ethnicity, educational level, age and mathematical ability have been found (Diehr et al., 1998).
WLL - Word List Learning	California Verbal Learning Test (CVLT; Delis et al., 2004), Hopkins Verbal Learning Test (HVLT; Brandt & Benedict, 2001), Rey's Auditory Verbal Learning	Respondents are shown a list of words and asked to recall as many as they can when the list is taken away. The type of word (e.g., nouns, semantically similar words), the time between being given the words and repeating them, and the presentation of words can vary between tests. Some tests may ask participants to point out the words they remember seeing previously on a new list of words. There are brief forms that use	Memory	15-50 minutes, depending on length of list and time to recall	Normative data exists for many of the versions of the tests. WLL tests have been shown to have good reliability and validity. The scoring system can give a wealth of information related to memory. A shorter version can be used, which may be more suitable for severe cognitive impairment. Shown to be one of the most sensitive measures of memory deficit and found to be valid in MS (Stegen et al., 2009).	The longer version especially can be quite demanding of the participant. There is a lack of data on practice effects.

Class	Examples of Tests	Description	Domains	Time Taken to Administer	Advantages	Disadvantages
	Test (RAVLT; Rey, 1941)	fewer words and are suitable for severe cognitive impairment.				
VFT - Verbal Fluency Test	Controlled Oral Word Association Test (COWAT), or Word List Generation (WLG), FAS test,	Patients are asked to name as many words as possible beginning with a certain letter (usually F, A and S), or category (e.g., animals or professions) and asked to name as many words as possible in a specified time frame. Scored based on how many words the participant can name.	Verbal fluency	5-10 minutes	Research has suggested VFTs can be sensitive to mild cognitive impairment. Normative data adjusted for age, education, and gender is available. Word generation tests have high internal consistency and good test-retest reliability (Ruff et al., 1996).	Age and education can have an effect on the results, and this test can be subject to practice effects (Barry et al., 2008).
ST - Stroop Test (Stroop, 1935)	Color and Word Test	The Stroop Test can have multiple formats, but the commonly, there are two trials. Participants are first presented with words of colours, written in a different colour ink (such as the word 'green written in red ink), and asked to say the word (i.e., green). In the second trial, participants are presented with words of colours, written in a different colour ink (such as the word 'red' written in	Attention, processing speed, cognitive flexibility and working memory	5-10 min	The Stroop Test has been used extensively in MS patients, and seen to be a valid measure of processing speed (Denney & Lynch, 2009). Norming data exists but can vary between administrations. Parallel versions can be created and administrated with ease.	No individual standardised version exists, and variations can include the number and format of stimuli, the number of tasks, how it is scored. Research suggests that the Stroop test is not valid as a measure of selective attention in MS patients (Denney & Lynch, 2009). Practice effects on the test have been found (Franzen et al., 1987). Poorer performance is correlated with depression, and can be influenced by visual acuity (Anstey et al., 2001).

Class	Examples of Tests	Description	Domains	Time Taken to Administer	Advantages	Disadvantages
		green ink), and the patient must say the word (i.e., green). Scoring can vary and can include correct responses given in the period; to the time taken to complete the trials; and number of errors.				
TMT - Trail Making Test	Colour Trails, Oral Trails, Making. Trails Test, Letter Number Sequencing	The participant is asked to connect a series of stimuli as quickly as they can. To start, the stimuli are numbers (1, 2, 3) and they are connected in sequential order. In the second part, the participant alternates between numbers and letters in sequential order (e.g., 1, A, 2, B). This test is scored by time taken to complete. Variants exist – including a Colour Trails Test and verbal administrations (such as Letter Number Sequencing).	Processing speed and executive function	5-10 minutes	Different versions can be administered to mitigate practice effects (Atkinson et al., 2011) . Norming data is available, stratified by age and education levels. It is thought to be sensitive to neurological impairment and has good construct validity (Salthouse, 2011).	Motor function may have an effect on performance. Education and age have a significant effect on speed of completion, and gender has a small effect (Zec et al., 2015) Reliability in clinical populations, especially on Part B, has been found inadequate (Bardi et al., 1995).

Class	Examples of Tests	Description	Domains	Time Taken to Administer	Advantages	Disadvantages
DS - Digit Span	Digit Span, Forward Span, Backwards Span,	Sequences of numbers are presented, beginning with 2 digits and 2 trials, and gradually increasing length, and participants are asked to recall the sequences. Digit span can be measured for forward and backward recall of sequences. The test ends when it cannot be accurately recalled, or the longest length list is reached. There are many variations of the Digit Span test, such as visual or auditory presentation, electronic or human presentation; length of lists and speed of presentation.	Working Memory, Attention	5-10 minutes	DS tests have good construct validity is and load on multiple domains of cognition(Larrabee & Curtiss, 1995). Digit Span tests have adequate internal consistency and test re-test reliability. Norming data exists.	There are concerns that when the test is administrated by a human, the administration can influence performance of participants due to emphasis or speed of reading the numbers. Effects of IQ have been found (Ostrosky-Solís & Lozano, 2006).
CST -Card Sorting Test	Wisconsin Card Sorting Task (WCST) Modified Card Sorting, Weigl Test	In CSTs, cards are presented with a variety of different coloured shapes (e.g., 1 red star, 2 green stars, 3 yellow crosses, 4 blue circles). Participants are then given a pack cards with similar designs, and they are asked to match these cards with the key cards. Feedback is given as to if the match is correct or not, and participants have to deduce the rules for sorting	Executive function	12-20 minutes	CSTs are correlated with behavioural and cognitive deficits as reported by an informant and so is thought to have ecological validity (Burgess et al., 1998). It is a relatively simple task.	Scoring of CST can be confusing for the administrator, and errors are likely when scoring. Reliability has been reportedly mixed, especially in healthy populations, and warning has been given in analysing data on an individual basis (Strauss et al., 2006).

Class	Examples of Tests	Description	Domains	Time Taken to Administer	Advantages	Disadvantages
		the cards. The rule for sorting is changed intermittently.				
VR – Visual Reproduction Test	ROCFT - Rey-Osterrieth Complex Figure Test (ROCFT; Rey, 1941) Brief Visuospatial Memory Test (BVMT; Benedict, 1997)), Figural Learning Tests, Taylor Complex Figure (TCF)	Visual Reproduction tests involve presenting figures, and asking participants to copy the figure, and then recall and reproduce the figures from memory. Some tests use a recognition trial, in which participants identify figures they have previously seen. The conditions can be immediate or delayed. There are multiple ways to score visual reproduction tests, including accuracy of drawings,	Visual memory	10-45 minutes	VR tests are sensitive and specific in people with MS (Anstey et al., 2001) and has high internal consistency and correlates with other measures of cognition in people with MS (de Caneda et al., 2018).	These tests rely on visual acuity, so may not be appropriate for participants with visual difficulties. Results can be affected by gender, age and IQ (Gallagher & Burke, 2007). Scoring can be complex and vary between administrations and could reduce interrater reliability.

Table C: Multidomain neuropsychological assessments used in multiple sclerosis research

Abbreviation	Name	Description	Therapeutic classification/ population	Domains	Time Taken to Administer	Advantages	Disadvantages
BRB-N	Brief Repeatable Battery of Neuropsychological Tests	Designed to measure for cognitive impairment in MS sensitively, and consists of SRT, 10/36 Spatial Recall Test (SPART), SDMT, PASAT and the Word List Generation (WLG).	Multiple sclerosis	Verbal memory and delayed recall, visuospatial memory and delayed recall, sustained attention and concentration and semantic retrieval.	30 minutes	A sensitive measure of impairment in MS and can be used for screening in early stage MS (Rao et al., 1991) Parallel versions exist for test re-test ability and to overcome practice effects. Can be shortened to reduce administration time and participant burden as some of the tests are similar and measure similar constructs.	Studies have shown that the difficulty is low, and patterns could be easily verbalised. Differences in scores were found in the parallel versions, which could not be explained. Age has been found to be an important and significant factor on many of the subtests, and gender and education level can influence the test scores (Boringa et al., 2001).
MMSE	Mini-Mental State Examination	The MMSE is a screen of cognitive impairment, developed especially to screen in an elderly population. It contains items pertaining to orientation, attention and calculation, language and 3-word recall. Consists of the following subscales: orientation of time and place, 3-word recall, arithmetic, spelling backwards, naming test, repetition of sentence, following commands, write a sentence, figure copying.	Developed for geriatric population	Orientation, memory, language, visuospatial abilities	10 minutes	Norming data available for score ranges, test-retest versions available for multiple testing of patients. Relatively quick, and easy to administer and score, does not require equipment.	Not recommended in early stages of cognitive impairment. The MMSE is not sensitive enough to detect precisely defined cognitive impairment in people with MS (Beatty & Goodkin, 1990). Research has found that the MMSE did not detect large changes of cognitive impairment in MS (Kujala et al., 1997).

Abbreviation	Name	Description	Therapeutic classification/ population	Domains	Time Taken to Administer	Advantages	Disadvantages
WAIS	Wechsler Adult Intelligence Scale	Consists of 10 subtests: Block Design, Similarities, Digit Span, Matric Reasoning, Vocabulary, Arithmetic, Symbol Search, Visual Puzzles, Information, Coding, Letter Number Sequencing, Figure Weights, Comprehension, Cancellation and Picture Completion.	N/A	Verbal comprehension, perceptual reasoning, working memory and processing speed. The results of the subtests can be scored into a full score of general intellectual ability, or into functioning of specific domains.	75 minutes	Comprehensive test that spans a range of domains that are impaired in people with MS and research has found the WAIS to be sensitive and specific in identifying cognitive impairment in RRMS (Ryan et al., 2012).	Large time burden for test administration. The WAIS has changed considerably over time, so comparing results is contentious.
MACFIMS	Minimal assessment of Cognitive Function in MS	MACFIMS was devised through an expert panel of neurologists, neuropsychologists, and psychologists. It consists of COWAT, JOL, CVLT-II, BVMT-R, PASAT, SDMT, and the Sorting Test from the Delis-Kaplan Executive Function System (D-KEFS). An abbreviated version exists which reduces the administration time and adds a verbal fluency and executive function measures.	Multiple Sclerosis	Information processing speed, learning and memory, executive function, visual-spatial processing, and word retrieval	90 minutes	Research has suggested that the MACFIMS has good sensitivity and specificity as a whole and it is a good measure of impairment in MS patients and has good validity. MACFIMs discriminated between employed and unemployed patients (Benedict et al., 2006). Norming data adjusted for age, sex, and education is available.	At 90 minutes, this is one of the longest batteries and may induce fatigue effects in patients. The length of the measure may also mean unacceptability to clinicians and participants. The BICAMS has been found to have a higher sensitivity compared to the aMACFIMS (Gromisch et al., 2018).

Abbreviation	Name	Description	Therapeutic classification/ population	Domains	Time Taken to Administer	Advantages	Disadvantages
WMS, WMS-III, WMS-IV	Wechsler Memory Scale, Wechsler Memory Scale Fourth Edition	The WMS is designed to be a measure of different memory functions. It consists of the following subtests: Information and Orientation, Logical Memory, Facial recognition, Family Pictures, Word list learning, Letter-Number Sequencing, Spatial Span, Mental Control, Digit Span, Cognitive Status Exam, Verbal Paired Associates, Visual Reproduction, substitution test, Spatial Addition.	N/a	Auditory Memory, Visual Memory Index, Visual Working Memory Index, Immediate Memory Index, Delayed Memory Index	30-50 minutes	Scoring software exists and interrater reliability is good. Internal consistency is good. Test re-test reliability is adequate in many of the subtests, and it is appropriate for evaluating memory in people with MS (Spedo et al., 2014). Norming data is extensive.	Scoring and administration could be overly complex for many settings. The WMS only measures aspects of memory, and not all aspects of cognitive impairment. The expense and administration burden may be a barrier to use in research and clinician use in people with MS.
BICAMS	Brief International Cognitive Assessment for Multiple Sclerosis	BICAMS was developed as a practical and brief assessment of CI in people with MS, and it intends to be internationally acceptable. It consists of the SDMT, CVLT-II and the BVMT-R.	Multiple Sclerosis	Mental processing speed and memory	15 minutes	Norms exist for a wide range of countries. It is easy to administer and does not require extensive training. All tests within the BICAMS were chosen to target cognitive domains that are apparent in MS and selected as they have good psychometric properties, feasibility and acceptability (Langdon et al., 2012).	Lack of assessment of executive functions and a short test means that it may not assess the full range of cognitive impairment.

Abbreviation	Name	Description	Therapeutic classification/ population	Domains	Time Taken to Administer	Advantages	Disadvantages
MOCA	Montreal Cognitive Assessment	Developed to be a brief screening tool for mild cognitive impairment, consists of a number of tasks: Word List Learning, Clock drawing task, Trail Making Task, verbal fluency, sentence repetition, drawing of a 3d cube, digit span, orientation, serial 7s, similarity between 2 nouns, target detection with tapping.	Developed for detecting mild cognitive impairment	Visuospatial, executive functions, attention, concentration, working memory, language, orientation	10 minutes	Excellent sensitivity for mild cognitive impairment in people with MS (Freitas et al., 2016). Parallel versions exist and electronic version enables processing speed metrics and improves the reliability of testing and retesting, and improved scoring. Good construct validity and factors load onto speed of processing, learning and delayed recall, and correlates well with the MSNQ-I even with patients with mild functional disability (Dagenais et al., 2013).	A systematic review found that the MOCA can produce false positives in other conditions, and care should be taken if utilising as a diagnostic tool (Davis et al., 2015) and further validation work should be undertaken.
TAP	Test for Attentional Performance	Computerised test consisting of 14 subtests: Alertness, Covert Attention Shift, Cross-Modal Integration, Divided Attention, Eye Movement, Flexibility, Go/NoGo, Incompatibility, Sustained Attention, Vigilance, Visual Field/Neglect, Visual Scanning, Working Memory	N/A	Attention and working memory	100 minutes	Computerised scoring can mean reduce interrater error and increased precision of results. Good reliability, and construct validity (Zimmermann & Fimm, 1995).	PC version only, which may be a barrier to use. Limited domains and lengthy administration time.

Table D: Extracted studies as part of the scoping review

	Study Properties				Administration			Norming and scoring				Interpretation			
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
1.	Beier et al. (2017). Proposed cut scores for tests of the Brief International Cognitive Assessment of Multiple Sclerosis (BICAMS). Journal of the Neurological Sciences, 381, 110–116. https://doi.org/10.1016/j.jns.2017.08.019	USA	703 people with MS (626 completed SDMT)	To derive cut-off scores	Battery - BICAMS	Verbal in-person administration	Total correct responses said aloud in 90s	Previous normative data used	Normative data taken from SDMT manual	None	Impaired/ Non-impaired	Total raw scores for each cognitive test were converted to z-scores using normative data. Participants were then categorised into “impaired” vs. “unimpaired,” using two SD cut-offs. Optimal cut-off scores were determined using ROC.	Previous research utilising the same SDs	Using a score below 1.5SD of the norm 32% of participants were categorised as impaired, and 2SDs below categorised 22% of participants as impaired.

	Study Properties			Administration			Norming and scoring				Interpretation		
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
												For 1.5SDs, a cut-off score of 44 was determined as the best score balancing sensitivity and specificity. For 2SDs, a cut-off score of 38 was given.		
2.	Deloire et al. (2006). How to detect cognitive dysfunction at early stages of multiple sclerosis? Multiple Sclerosis, 12(4), 445–452.	France	58 people with RRMS, 44 control participants	To evaluate psychometric performance of individual tests within the BRB-N	Battery - BRB-N	BRB-N verbal in-person version	No data reported	Collected as part of study	Control participant norming data collected as part of study	Matched for age, gender, and education	Impaired/ Non-impaired	Participants were considered "cognitively impaired" if they performed worse than the fifth percentile	Not reported	50.88% of participants scored less than the fifth percentile of matched control participants on the SDMT. 59.7% of people with MS were cognitively

			Study Properties		Administration			Norming and scoring				Interpretation			
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
	https://doi.org/10.1191/1352458506ms1289oa											of control participants on any test, adjusted for age, gender and education.		impaired, scoring less than the fifth percentile of control participants on two or more tests of BRB-N. 87.7% scored less than the fifth percentile of control data on one test, e.g., SDMT. 13.6% control participants were considered cognitively impaired on a single test.
3.	De Meo E, Portaccio E, Giorgio A, et al. Identifying the Distinct Cognitive Phenotypes	Italy	1212 clinically stable people with MS, 196 healthy control	Used as a part of a battery to formulate cognitive phenotypes	Battery - BRB-N	BRB-N verbal in-person version	No data reported	Previous normative data used	Italian norms collected in 2006 were used	Adjusted for age, gender, and education	Five cognitive phenotypes were identified using all tests. Significantl	Total raw scores for each cognitive test were converted to z-scores	Previous norming study, and "current knowledge about test interpretation".	19.4% people with MS had no cognitive impairment. 29.9% of people with MS had mild

		Study Properties		Administration			Norming and scoring			Interpretation			
Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
in Multiple Sclerosis. JAMA Neurology. Published online January 4, 2021. doi:10.1001/jamaneurol.2020.4920		participants								y poorer performance on the SDMT (compared to control participants) contributed to the four bolded phenotypes: **1. No cognitive impairment** **2. Mild verbal memory/semantic fluency impairment** **3. Mild multidomain impairment**	using normative data, performance below <1.5SD was used as a threshold to distinguish between performance from control participants. Latent profile analysis was used to identify cognitive phenotypes based on individual test performance.		verbal memory/semantic fluency impairment. 19.5% of people with MS had mild multidomain impairment. 13.8% had severe executive function/attention impairment. 17.5% had severe multidomain impairment.

			Study Properties		Administration			Norming and scoring			Interpretation				
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
											4. Severe executive function/attention impairment **5. Severe multidomain impairment**			
4.	Golan et al. (2019). Validity of a multi-domain computerized cognitive assessment battery for patients with multiple sclerosis. Multiple Sclerosis and Related Disorders, 30, 154–162. https://doi.org/10.1016/j.	USA	81 people with MS, 15 control participants	To screen participants into study, and as part of a gold standard neuropsychological battery	Individual to screen participants into study, and as a battery (MACFIMS)	MACFIMS verbal in-person version	No data reported	Previously collected norms	Sheridan's (2006) community-based sample	Age	Stratification for screening into study (no labels); Impaired/ Non-impaired	For stratification in the study, the following cut-offs were used: - 1SD -1 to -1.5SD -1.51 to 2.SD >2SD SDMT and PASAT scores were grouped	Original manual for the SDMT	In the initial screening, an equal number of people with MS in each subgroup were chosen for the sample (n=20) using the stratification. Using the 'gold standard' tests, 72% were categorised as cognitively impaired, and 28% were

			Study Properties		Administration			Norming and scoring				Interpretation				
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category		
	msard.2019.01.051											into a 'processing speed and working memory' domain and an average of both was reported as a z-score. This average domain z-score was considered cognitively impaired at -1.5 or lower.		cognitively normal.		
5.	Leavitt et al (2018). Cognitive phenotypes in multiple sclerosis. Journal of Neurology,	USA	128 RRMS	As a measure of processing speed to categorise cognitive impairment compared	Combined with Stroop test to form a measure of processing speed	Verbal in-person administration	No data reported	Previous norms	MS-COG norms were used and taken from personal communication with Benedict.	Age and education	Not impaired - Memory impaired - Processing speed and	A z-score of -1SDs or lower on both SDMT and Stroop was considere	Authors acknowledge that this score is different from the suggested z-score of -	Not impaired (56.3%) Memory impaired (18.8%). Processing speed and memory		

		Study Properties			**Administration**			**Norming and scoring**				**Interpretation**			
	Reference	**Location**	**Sample**	**Use of SDMT**	**Individual administration or part of a battery?**	**Format of Test**	**Scoring**	**Norms**	**Details of norming data, including any demographic**	**Normative adjustments made**	**Classification categories**	**SDMT Cut-off**	**Justification for cut-off**	**Classification criteria, and % of people with MS in each category**
	265(3), 562–566. https://doi.org/10.1007/s00415-018-8747-5			with a battery of NP tests					Norms were adjusted for age and education.		memory impaired - Processing speed impaired	d to be processing speed impaired	1.5. The z-score of -1 was chosen as the sample had a mean pre-morbid IQ of 110.	impaired (17.2%). Processing speed impaired (7.8%).
6.	López-Góngora et al (2015). A one-year follow-up study of the Symbol Digit Modalities Test (SDMT) and the Paced Auditory Serial Addition Test (PASAT) in relapsing-remitting multiple sclerosis: An appraisal of comparative longitudinal	Spain	237 people with MS, 57 control participants	Comparison of scores on the SDMT to PASAT	Battery - BRB-N	BRB-N verbal in-person version	No data reported	Previous norms	Normative data taken from previous study of the Spanish BRB-N	None	Impaired/ Non-impaired	Cognitive impairment was defined as ≤1.5 standard deviations below control data. A cut-off score of 49 was thus chosen.	The cut-off was chosen using acceptable ROC.	33.8% were classified as cognitively impaired.

			Study Properties		Administration			Norming and scoring				Interpretation			
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
	sensitivity. BMC Neurology, 15, 40. https://doi.org/10.1186/s12883-015-0296-2													
7.	Parmenter et al (2010). The utility of regression-based norms in interpreting the minimal assessment of cognitive function in multiple sclerosis (MACFIMS). Journal of the International Neuropsychological Society: JINS, 16(1), 6–16. https://doi.or	USA	395 people with MS, 100 healthy control participants	Establish and compare norming data using regression-based model	Battery - MACFIMS	MACFIMS verbal in-person version	No data reported	Previous norms	Rao's (1991) norms, compared against norm data collected as part of study.	Adjusted for age, sex and education using regression model.	Impaired/ Non-impaired	T scores: T >35 intact T ≤ 35 impaired	Age, gender, and education adjusted T-scores were determined for people with MS based on control data collected. Predicted scores were subtracted from actual scores, and this was divided by the SD of the control scores. These t-	33.92% impaired on SDMT using norms. 54.68% impaired on SDMT using regression-based norms.

			Study Properties		Administration			Norming and scoring				Interpretation			
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
	g/10.1017/S1355617709990750												scores were then compared against published norms.	
8.	Parmenter BA, Weinstock-Guttman B, Garg N, Munschauer F, Benedict RHB. Screening for cognitive impairment in multiple sclerosis using the Symbol digit Modalities Test. Multiple Sclerosis. 2007;13(1):52-57. doi:10.1177/1352458506070750	USA	100 people with MS, 50 healthy control participants	Assess performance of SDMT against a full NP battery	Individual	Verbal in-person administration using Rao's administration	No data reported	Collected	Comparisons with control data collected as part of study	None	People with MS were categorised into cognitively impaired vs cognitively normal.	A z-score using -0.5 or -1.5 or lower were used to examine different cut-offs. 3 cut-offs were considered, ≤35, ≤55 and ≤67 based on ROC. ≤55 was found to be the score that optimally classified people with MS	No data reported	55 were considered impaired (55%) at a cut-off using -0.5SD as a z-score.

	Study Properties			Administration			Norming and scoring				Interpretation		
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
												with reasonable accuracy.		
9.	Schependom et al (2014). The Symbol Digit Modalities Test as sentinel test for cognitive impairment in multiple sclerosis. European Journal of Neurology, 21(9), 1219-e72. https://doi.org/10.1111/ene.12463	Belgium	359 people with MS	Assess performance of SDMT against a full NP battery	Individual	No data reported	No data reported	Previous norms	Rao's (1991) norms	None	No classification	Failure on SDMT was defined by scoring in the 5th percentile of Rao's (1991) norms. A cut-off for the SDMT was a score of 40.	Sensitivity was examined using comparison to a comprehensive battery (NSBMS). Participants were considered cognitive impaired if they failed two or more tests on the battery. Using the cut-off of 0.6-0.65, the SDMT had a sensitivity of 0.9 as determined	No data reported

	Study Properties				Administration			Norming and scoring				Interpretation			
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
													by participants categorised as CI.	
10	Xiang et al (2019). Single Scan Quantitative Gradient Recalled Echo MRI for Evaluation of Tissue Damage in Lesions and Normal Appearing Gray and White Matter in Multiple Sclerosis. Journal of Magnetic Resonance Imaging. 49, 487–498. https://doi.org/10.1002/jmri.26218	USA	44 people with MS completed SDMT (and 19 control participants who did not)	Correlation between WM lesions and SDMT scores	Individual	No data reported	No data reported	N/a	N/a	None	No classification	None used	N/a	No classification, results reported were mean, SD and range for RRMS, PPMS and SPMS. Correlation was used to compare results of SDMT.

	Study Properties				Administration			Norming and scoring				Interpretation		
	Reference	Location	Sample	Use of SDMT	Individual administration or part of a battery?	Format of Test	Scoring	Norms	Details of norming data, including any demographic	Normative adjustments made	Classification categories	SDMT Cut-off	Justification for cut-off	Classification criteria, and % of people with MS in each category
11	Burggraaff, et al (2017). Regression-Based Norms for the Symbol Digit Modalities Test in the Dutch Population: Improving Detection of Cognitive Impairment in Multiple Sclerosis? European Neurology, 77(5–6), 246–252. https://doi.org/10.1159/000464405	Netherlands	157 people with MS, 96 healthy control participants	Psychometric evaluation of regression-based normative data for SDMT amongst people with MS	Individual	BRB-N verbal in-person version	No data reported	Previous norms	Control data was matched to Dutch normative data from Boringa et al (2001) by age and gender	Regression model - Age and gender	Impaired/ Non-impaired	z-score≤-2 SD below control normative data were considered cognitively impaired	SDMT scores were converted to z-scores, and a regression model was applied. Age was found to be a predictor of SDMT score.	Norms categorised 10% of people with MS as impaired, regression-based norms categorised 20% as impaired

List of abbreviations:
BICAMS - Brief International Cognitive Assessment for MS
BRB-N - Brief Repeatable Battery of Neuropsychological Tests
MACFIMS - Minimal Assessment of Cognitive Function in Multiple Sclerosis

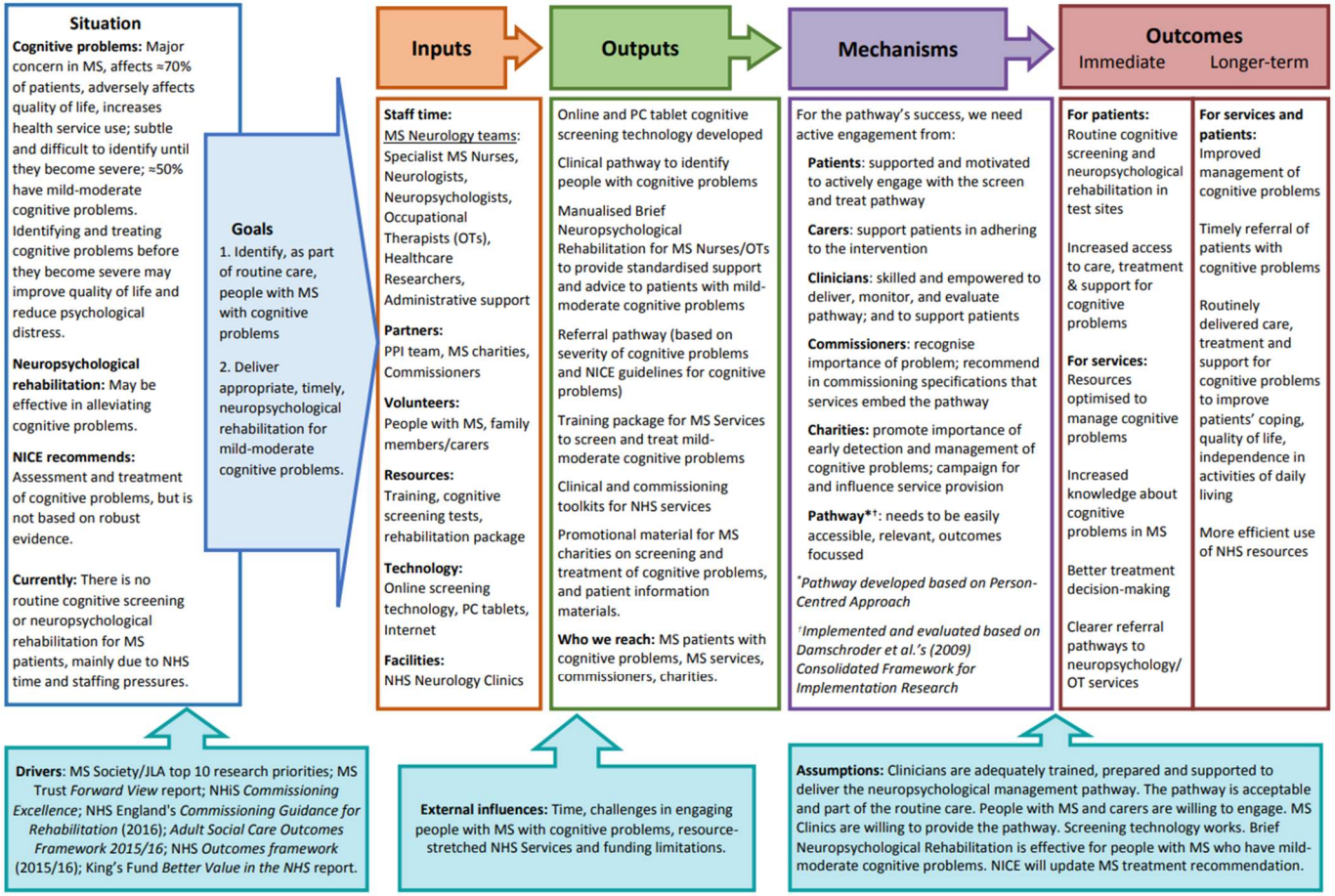
NP - Neuropsychological
PASAT – Paced Auditory Serial Addition Test
SDMT – Symbol Digit Modalities Test
NSBMS - Neuropsychological Screening Battery For MS
PPMS - Primary Progressive Multiple Sclerosis
ROC - Receiver Operating Characteristic
RRMS - Relapsing-Remitting Multiple Sclerosis
SPMS - Secondary Progressive Multiple Sclerosis
WM - White Matter

Table E: Demographics and MS characteristics of participants with MS

	Code Used	Age	Gender	Ethnicity	Highest Education	Employment	How long had MS (years)	Type of MS
Interview Participants	Person with MS I 01	43	Female	White	Degree	Not employed	4	RRMS
	Person with MS I 02	54	Female	White	A-Level	Not employed	14	RRMS
	Person with MS I 03	62	Male	White	Degree	Retired	41	RRMS
	Person with MS I 04	51	Female	White	GCSE	Employed Full Time	16	SPMS
	Person with MS I 05	39	Female	White	GNVQ	Employed Part Time	2	RRMS
	Person with MS I 06	67	Female	White	A Level	Retired	14	Mild MS
	Person with MS I 07	60	Female	White	GCSE	Not employed	11	PPMS
	Person with MS I 08	41	Female	White	Degree	Not employed	19	SPMS
	Person with MS I 09	37	Female	White	Higher Degree	Employed Part Time	9	RRMS
	Person with MS I 10	47	Male	White	HNC	Employed Full Time	10	SPMS
	Person with MS I 11	47	Female	Mixed/ Multiple ethnic groups	Higher Degree	Retired	28	SPMS
	Person with MS I 12	52	Female	White	GCSE	Employed Part Time	3	RRMS
	Person with MS I 13	39	Female	White	Higher Degree	Employed Full Time	6	RRMS

	Code Used	Age	Gender	Ethnicity	Highest Education	Employment	How long had MS (years)	Type of MS
Focus Group Participants	Person with MS I 14	36	Female	White	Degree	Employed Full Time	9	RRMS
	Person with MS I 15	47	Male	White	Higher Degree	Retired	8	RRMS
	Person with MS F 01	51	Female	White	Higher Degree	Employed Part Time	6	RRMS
	Person with MS F 02	57	Female	White	GCSE	Employed Full Time	4	RRMS
	Person with MS F 03	49	Female	White	GCSE	Employed Full Time	14	RRMS
	Person with MS F 04	38	Male	White	Higher Degree	Employed Part Time	10	RRMS
	Person with MS F 05	53	Female	White	BTEC	Employed Part Time	26	SPMS

Appendix 1. Cognitive Screening Pathway and Rehabilitation Logic Model



Appendix 2. Qualitative Interviews – Adverts for Participants

University of Nottingham
UK | CHINA | MALAYSIA

Living with MS?

Experienced problems with your memory, or thinking, paying attention and concentrating?

We are inviting people with MS who have experienced problems with thinking, concentrating, or their memory; and their families to share their experience of how these are assessed.

We would like your opinions and experience so we can develop a pathway for identifying cognitive problems in people with MS.

You will be part of a group discussion with other people with MS and their relatives, or have an individual interview if you prefer.

The group discussions will take no longer than 1.5 hours, will be located at the University of Nottingham and refreshments will be provided. An interview will last for 45 minutes and can be carried out on campus, via telephone or Skype, or at your home. Any travel expenses incurred will be paid.



If you are interested or would like more information, please contact Hannah Elwick using the email address or QR code below:

hannah.elwick@nottingham.ac.uk

This study has been funded by the MS Society. Ethical approval has been obtained from the University of Nottingham. The information you provide will be collected anonymously. Your name will not be recorded and anything that might identify you will be removed.

Help us with research into cognitive problems in multiple sclerosis

- **Do you have MS or care for someone with MS?**
- **Are you aged 18-69?**
- **Do you want to help research aimed at improving the everyday lives of people living with MS?**

Cognition refers to people's ability to pay attention, remember things, multi-task and problem solve.

You are being invited to share your views and experiences to help develop:

- a cognitive screening pathway to assess all people with MS for cognitive problems
- a brief rehabilitation intervention programme for people with 'mild to moderate' cognitive problems

We would like you to take part in a phone or face-to-face **interview / discussion group.**

We are hosting discussion groups on 26th July and 9th August.

You will be offered travel expenses for attending an interview in person.

If you are interested in taking part, please get in touch.

If you would like to know more information about the study, please contact:

Name: Hannah Elwick
Email: Hannah.elwick@nottingham.ac.uk
Telephone: 07340615551

Appendix 3. Qualitative Interviews – Participant Information Sheets

PARTICIPANT INFORMATION SHEET: Interview for people with MS
Research Ethics Reference: 263-1903
Final Version 1.1 Date: 11.04.2019

Title of Study: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS): Qualitative study for developing a clinical management pathway (Phase 1 Work Package 1a)
Study ID: NEuRoMS Qualitative phase WP1a

We would like to invite you to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. One of our team will go through the information sheet with you and answer any questions you have. Please take time to read this carefully and discuss it with others if you wish. Ask us anything that is not clear.

What is the purpose of the research?
This study aims to improve our ability to detect and manage cognitive problems (such as problems with memory, attention, and problem solving) in people with multiple sclerosis (MS). These cognitive problems can have a negative impact on people's mood, ability to work, and quality of life. 'Neuropsychological rehabilitation' is a type of rehabilitation that may help deal with cognitive problems (e.g., by teaching people how to better concentrate or remember things). However, cognitive problems are not routinely checked (screened) for in people with MS. Early recognition through screening can help with rehabilitation activities before problems become severe. The current study aims to develop a cognitive screening pathway to routinely test all people with MS for cognitive problems, and develop a brief rehabilitation programme for people with 'mild to moderate' cognitive problems.

Why have I been invited to take part?
You have been invited to take part in this research either because you have been living with MS and you are aged 18-69 years.

We are inviting up to 15 participants like you to take part in one-to-one interviews.

Do I have to take part?
No. It is up to you to decide if you want to take part in this research. We will describe the study and go through this information sheet with you to answer any questions you may have. If you agree to participate, we will ask you to sign a consent form and will give you a copy to keep. However, you would still be free to withdraw from the study at any time, without giving a reason and without any negative consequences, by telling the researchers of your decision. This would not affect your medical care or legal rights.

What will happen to me if I take part?
You will be invited to take part in a one-to-one interview to share your views and/or experiences on cognitive screening and rehabilitation in MS.

A researcher will contact you to go over the information sheet, explain the procedures, and answer any questions that you have about the study. Interviews can be done face-to-face, over the phone or via Skype, whichever you prefer.

If you agree to take part in the study, the researcher will agree the interview date and time with you. If you are doing an interview face-to-face, the researcher will ask you to complete a consent form before the interview starts. If you are doing the interview over the phone, the consent form will be completed with you over the 'phone. As part of the interview, you will be shown some examples of cognitive screening tools (these are paper-pencil tests or computerised tests). If you have chosen to do your interview over the 'phone or Skype then the screening tool examples will be provided before the interview happens, either by post or through a link to online versions.

The interview will last up to 60 minutes. During the interview, you will be asked for your views on our proposals on how to screen for and manage cognitive problems. You will also be asked for your views on both the format and the content of the screening tools, and the impact you think they might have on a person with MS to use this. You will be asked about what you believe is the usual care that people with MS currently receive relating to cognitive problems. This is to help us better understand what happens in routine practice and any gaps in treatment. Then you will be shown the outline of a cognitive rehabilitation intervention and asked for your opinions on how well you think this might work for people with MS, how it might be improved, and what possible barriers might exist for delivering this intervention within the NHS.

We will be audio recording the interviews so that we do not miss any important points. The recordings will be sent to an external company for transcription (meaning that they will type out the interview word for word). The company will have a duty of confidentiality to you as a research participant and they will do their best to meet this duty. We will remove any names of people (including your own name) or services so that you cannot be identified in anything that is said.

What are the screening tools and the rehabilitation intervention?

The **screening tools** will be provided to you as a link/attachment before the time of the interview to allow you to familiarise yourself with them. These are typically questionnaires and puzzles.
The **rehabilitation intervention** will be developed to help improve cognitive problems. You will be presented with some preliminary ideas about this intervention and how it might be delivered.

Are there any risks in taking part?
When talking about experiences and issues associated with cognitive problems, you may feel upset. This may also cause you some concerns about your own cognitive abilities. If this happens during the interview, the researcher will be there to talk to you and address any concerns. You will have their contact information in case you feel upset after the session for reasons associated with participation. You may also choose to leave the interview if you wish.
A possible disadvantage is that it may inconvenience you to attend an interview in person or be available at a specific time during the day. We will arrange the interview at a time and date that is suitable for you and if you prefer, it can be conducted over the phone.

Are there any benefits in taking part?
There will be no direct benefit to you from taking part in this research but your contribution may help people with MS in the future by contributing to the development of a cognitive screening and rehabilitation programme.

Will my time/travel costs be reimbursed?
You will receive a £20 gift voucher as a thank you for your time. Travel expenses will be offered for you to attend the interview if travel is involved.

What happens to the data provided?
The research data will be stored confidentially to ensure your privacy. You will be assigned a unique identification number (for example P01/A for participant 1 recruited from centre A) that will be used instead of your name on all forms used to save data. We will save all audio recordings and research data using your unique identification number so that none of the data will have your real name or other individual identifiers. Your name and any other information about you will not be disclosed outside the study centre.

Access to the information will be limited to the research team and any relevant regulatory authorities. Computer held data including the study database will be held securely and password protected. All data will be stored separately on secure university networks. Access will be restricted by user identifiers and passwords (encrypted using a one way encryption method). The research team will have access to personal and research data, and a transcriber will have access to some research data. Where a transcription service is used, an appropriate confidentiality contract will be in place between the university-approved provider of the transcription service.

The research will be undertaken by research staff at the University of Nottingham. Data will be analysed and stored by the University of Nottingham. All research data and records will be stored for a minimum of 7 years after publication or public release of the work of the research. Your optional consent for the use of your anonymised data in future research is requested as part of the Participant Consent Form. Such usage in future research would have to be approved by investigators at the University of Nottingham before anonymised data is released.

With your permission, we would like to use fully anonymised direct quotes in research publications.
We would also like your permission to use anonymised data in future studies, and to share our research data (e.g. in online databases) with other researchers in other Universities and organisations both inside and outside the European Union. This would be used for research in health and social care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research) and to understand the bigger picture in particular areas of research. All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public. Data sharing in this way is usually anonymised (so that you could not be identified).

What will happen if I don't want to carry on with the study?
Even after you have signed the consent form, you are free to withdraw from the study at any time without giving any reason, and without your legal rights being affected. Any personal data will be destroyed.

If you decide to leave the study, we will not contact you further. The information you would have provided us so far cannot be erased and this information may still be used in the analysis because your information may have already been used in some analyses and may still be used in the final study analyses. To safeguard your rights, we will use the minimum personally-identifiable information possible.

Who will know that I am taking part in this research?
All information collected about you during this research would be kept strictly confidential. Any audio recordings and electronic data will be anonymised with a code as detailed above. A password-protected list of code numbers and participant names will be stored separately from the audio recordings. All such data are kept on password-protected databases sitting on a restricted access computer system and any paper information (such as your consent

form, contact details and any research questionnaires) would be stored safely in lockable cabinets in a swipe-card secured building and would only be accessed by the research team.

Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security) and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights we will use the minimum personally – identifiable information possible.

You can find out more about how we use your information and to read our privacy notice at:

https://www.nottingham.ac.uk/utilities/privacy.aspx/

Designated individuals of the University of Nottingham may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines.

With your consent, we will keep your personal information on a secure database in order to contact you for future studies.

Anything you say during an interview will be kept confidential, unless you reveal something of concern that may put yourself or anyone else at risk. It will then be necessary to report to the appropriate persons.

What will happen to the results of the research?
This research will help us to develop a screening pathway for use in NHS MS services and a brief rehabilitation programme to support people with multiple sclerosis who have mild to moderate cognitive problems. The findings will be published in scientific journals. The study findings will also be presented to local patient or service provider groups, at academic conferences and to relevant charities. All your personal data will be removed in order to preserve your anonymity and you will not be identifiable in any publications/presentations. A summary of the study findings will be made available through the NEuRoMS web page for you to access, in due course.

Who has reviewed this study?
All research involving people is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Faculty of Medicine and Health Sciences Research Ethics Committee (Reference number: FMHS 263-1903).

Who is organising and funding the research?
This research is being organised by the University of Nottingham and is being funded by the National Institute for Health Research (NIHR).

What if something goes wrong?
If you have a concern about any aspect of this project, please speak to the researcher Hannah Elwick or the Chief Investigator Professor Roshan dasNair who will do their best to answer your query. The researchers should acknowledge your concern within 10 working days and give you an indication of how he/she intends to deal with it. If you remain unhappy and wish to complain formally, you can do this by contacting the FMHS Research Ethics Committee Administrator, c/o The University of Nottingham, Faculty PVC Office, B Floor, Medical School, Queen's Medical Centre Campus, Nottingham University Hospitals, Nottingham, NG7 2UH. E-mail: FMHS-ResearchEthics@nottingham.ac.uk

Contact Details
If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Researcher:
Hannah Elwick
Institute of Mental Health, School of Medicine
Email: hannah.elwick@nottingham.ac.uk

Chief Investigator:
Professor Roshan das Nair
Institute of Mental Health, School of Medicine
Phone: +44 (0)115 823 0589
Email: roshan.dasnair@nottingham.ac.uk

PARTICIPANT INFORMATION SHEET: Interview for Charity volunteers
Research Ethics Reference: 263-1903
Final Version 1.1 Date: 04.03.2019

Title of Study: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS): Qualitative study for developing a clinical management pathway (Phase 1 Work Package 1a)
Study ID: NEuRoMS Qualitative phase WP1a

We would like to invite you to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. One of our team will go through the information sheet with you and answer any questions you have. Please take time to read this carefully and discuss it with others if you wish. Ask us anything that is not clear.

What is the purpose of the research?
This study aims to improve our ability to detect and manage cognitive problems (such as problems with memory, attention, and problem solving) in people with multiple sclerosis (MS). These cognitive problems can have a negative impact on people's mood, ability to work, and quality of life. 'Neuropsychological rehabilitation' is a type of rehabilitation that may help deal with cognitive problems (e.g., by teaching people how to better concentrate or remember things). However, cognitive problems are not routinely checked (screened) for in people with MS. Early recognition through screening can help with rehabilitation activities before problems become severe. The current study aims to develop a cognitive screening pathway to routinely test all people with MS for cognitive problems, and develop a brief rehabilitation programme for people with 'mild to moderate' cognitive problems.

Why have I been invited to take part?
You have been invited to take part in this research because your voluntary work has given you experience of supporting people living with MS.

We are inviting up to five participants like you to take part in one-to-one interviews.

Do I have to take part?
No. It is up to you to decide if you want to take part in this research. We will describe the study and go through this information sheet with you to answer any questions you may have. If you agree to participate, we will ask you to sign a consent form and will give you a copy to keep. However, you would still be free to withdraw from the study at any time, without giving a reason and without any negative consequences, by telling the researchers of your decision. This would not affect your legal rights.

What will happen to me if I take part?
You will be invited to take part in a one-to-one interview to share your views and experiences on cognitive screening and rehabilitation in MS.

A researcher will contact you to go over the information sheet, explain the procedures, and answer any questions that you have about the study. Interviews can be done face-to-face, over the phone or via Skype, whichever you prefer.

If you agree to take part in the study, the researcher will agree the interview date and time with you. If you are doing an interview face-to-face, the researcher will ask you to complete a consent form before the interview starts. If you are doing the interview over the phone, the consent form will be completed with you over the 'phone. As part of the interview, you will be shown some examples of cognitive screening tools (these are paper-pencil tests or computerised tests). If you have chosen to do your interview over the 'phone or Skype then the screening tool examples will be provided before the interview happens, either by post or through a link to online versions.
The interview will last up to 60 minutes. During the interview, you will be asked for your views on our proposals on how to screen for and manage cognitive problems. You will also be asked for your views on both the format and the content of the screening tools, and the impact you think they might have on a person with MS to use this. You will be asked about what you believe is the usual care that people

with MS currently receive relating to cognitive problems. This is to help us better understand what happens in routine practice and any gaps in treatment. Then you will be shown the outline of a cognitive rehabilitation intervention and asked for your opinions on how well you think this might work for people with MS, how it might be improved, and what possible barriers might exist for delivering this intervention within the NHS.

We will be audio recording the interviews so that we do not miss any important points. The recordings will be sent to an external company for transcription (meaning that they will type out the interview word for word). The company will have a duty of confidentiality to you as a research participant and they will do their best to meet this duty. We will remove any names of people (including your own name) or services so that you cannot be identified in anything that is said.

What are the screening tools and the rehabilitation intervention?

The **screening tools** will be provided to you as a link/attachment before the time of the interview to allow you to familiarise yourself with them. These are typically questionnaires and puzzles.
The **rehabilitation intervention** will be developed to help improve cognitive problems. You will be presented with some preliminary ideas about this intervention and how it might be delivered.

Are there any risks in taking part?
We do not anticipate any risks associated with taking part in the interview.
A possible disadvantage is that it may inconvenience you to attend an interview in person or be available at a specific time during the day. We will arrange the interview at a time and date that is suitable for you and if you prefer, it can be conducted over the phone or Skype.

Are there any benefits in taking part?
There will be no direct benefit to you from taking part in this research but your contribution may help people with MS in the future by contributing to the development of a cognitive screening and rehabilitation programme.

Will my time/travel costs be reimbursed?
Travel expenses will be offered for you to attend the interview if travel is involved.

What happens to the data provided?
The research data will be stored confidentially to ensure your privacy. You will be assigned a unique identification number (for example P01/A for participant 1 recruited from centre A) that will be used instead of your name on all forms used to save data. We will save all audio recordings and research data using your unique identification number so that none of the data will have your real name or other individual identifiers. Your name and any other information about you will not be disclosed outside the study centre.

Access to the information will be limited to the research team and any relevant regulatory authorities. Computer held data including the study database will be held securely and password protected. All data will be stored separately on secure university networks. Access will be restricted by user identifiers and passwords (encrypted using a one way encryption method). The research team will have access to personal and research data, and a transcriber will have access to some research data. Where a transcription service is used, an appropriate confidentiality contract will be in place between the university-approved provider of the transcription service

The research will be undertaken by research staff at the University of Nottingham. Data will be analysed and stored by the University of Nottingham. All research data and records will be stored for a minimum of 7 years after publication or public release of the work of the research. Your optional consent for the use of your anonymised data in future research is requested as part of the Participant Consent Form. Such usage in future research would have to be approved by investigators at the University of Nottingham before anonymised data is released.

With your permission, we would like to use fully anonymised direct quotes in research publications.

We would also like your permission to use anonymised data in future studies, and to share our research data (e.g. in online databases) with other researchers in other Universities and organisations both inside

and outside the European Union. This would be used for research in health and social care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research) and to understand the bigger picture in particular areas of research. All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public. Data sharing in this way is usually anonymised (so that you could not be identified).

What will happen if I don't want to carry on with the study?
Even after you have signed the consent form, you are free to withdraw from the study at any time without giving any reason, and without your legal rights being affected. Any personal data will be destroyed.

If you decide to leave the study, we will not contact you further. The information you would have provided us so far cannot be erased and this information may still be used in the analysis because your information may have already been used in some analyses and may still be used in the final study analyses. To safeguard your rights, we will use the minimum personally-identifiable information possible.

Who will know that I am taking part in this research?
All information collected about you during this research would be kept strictly confidential. Any audio recordings and electronic data will be anonymised with a code as detailed above. A password-protected list of code numbers and participant names will be stored separately from the audio recordings. All such data are kept on password-protected databases sitting on a restricted access computer system and any paper information (such as your consent form, contact details and any research questionnaires) would be stored safely in lockable cabinets in a swipe-card secured building and would only be accessed by the research team.

Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security) and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights we will use the minimum personally – identifiable information possible.

You can find out more about how we use your information and to read our privacy notice at:

https://www.nottingham.ac.uk/utilities/privacy.aspx/

Designated individuals of the University of Nottingham may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines.

With your consent, we will keep your personal information on a secure database in order to contact you for future studies.

Anything you say during an interview/focus group will be kept confidential, unless you reveal something of concern that may put yourself or anyone else at risk. It will then be necessary to report to the appropriate persons.

What will happen to the results of the research?
This research will help us to develop a screening pathway for use in NHS MS services and a brief rehabilitation programme to support people with MS who have mild to moderate cognitive problems. The findings will be published in scientific journals. The study findings will also be presented to local patient or service provider groups, at academic conferences and to relevant charities. All your personal data will be removed in order to preserve your anonymity and you will not be identifiable in any publications/presentations. A summary of the study findings will be made available through the NEuRoMS web page for you to access, in due course.

Who has reviewed this study?
All research involving people is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Faculty of Medicine and Health Sciences Research Ethics Committee (Reference number: FMHS 263-1903).

Who is organising and funding the research?
This research is being organised by the University of Nottingham and is being funded by the National Institute for Health Research (NIHR).

What if something goes wrong?
If you have a concern about any aspect of this project, please speak to the researcher Hannah Elwick or the Chief Investigators Professor Roshan dasNair who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how he/she intends to deal with it. If you remain unhappy and wish to complain formally, you can do this by contacting the FMHS Research Ethics Committee Administrator, c/o The University of Nottingham, Faculty PVC Office, B Floor, Medical School, Queen's Medical Centre Campus, Nottingham University Hospitals, Nottingham, NG7 2UH. E-mail: FMHS-ResearchEthics@nottingham.ac.uk

Contact Details
If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Researcher:
Hannah Elwick
Institute of Mental Health, School of Medicine
Email: hannah.elwick@nottingham.ac.uk

Chief Investigator:
Professor Roshan das Nair
Institute of Mental Health, School of Medicine
Phone: +44 (0)115 823 0589
Email: roshan.dasnair@nottingham.ac.uk

PARTICIPANT INFORMATION SHEET: Focus Group for people with MS
Research Ethics Reference: 263-1903
Final Version 1.1 Date: 04.03.2019

Title of Study: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS): Qualitative study for developing a clinical management pathway (Phase 1 Work Package 1a)
Study ID: NEuRoMS Qualitative phase WP1a

We would like to invite you to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. One of our team will go through the information sheet with you and answer any questions you have. Please take time to read this carefully and discuss it with others if you wish. Ask us anything that is not clear.

What is the purpose of the research?
This study aims to improve our ability to detect and manage cognitive problems (such as problems with memory, attention, and problem solving) in people with multiple sclerosis (MS). These cognitive problems can have a negative impact on people's mood, ability to work, and quality of life. 'Neuropsychological rehabilitation' is a type of rehabilitation that may help deal with cognitive problems (e.g., by teaching people how to better concentrate or remember things). However, cognitive problems are not routinely checked (screened) for in people with MS. Early recognition through screening can help with rehabilitation activities before problems become severe. The current study aims to develop a cognitive screening pathway to routinely test all people with MS for cognitive problems, and develop a brief rehabilitation programme for people with 'mild to moderate' cognitive problems.

Why have I been invited to take part?
You have been invited to take part in this research because you have been living with MS and are aged 18-69 years.

We are inviting up to 10 participants like you to take part in one focus group.

Do I have to take part?
No. It is up to you to decide if you want to take part in this research. We will describe the study and go through this information sheet with you to answer any questions you may have. If you agree to participate, we will ask you to sign a consent form and will give you a copy to keep. However, you would still be free to withdraw from the study at any time, without giving a reason and without any negative consequences, by telling the researchers of your decision. This would not affect your medical care or legal rights.

What will happen to me if I take part?
You will be invited to take part in a group discussion (called a focus group) with other people who have a MS diagnosis to share your views and experiences on cognitive screening and rehabilitation in MS. A researcher will contact you to go over the information sheet, explain the procedures, and answer any questions that you have about the study. If you agree to take part in the study, the researcher will discuss and agree the date of the focus group that you can join. You will be given a consent form to sign and return to the researcher before the focus group takes place.

The focus group will happen on one day, in a private room and last up to 4 hours, with regular breaks, and refreshments in between. Upon arrival we will talk you through the study procedures and give you chance to ask any questions.

During the focus group you will be presented with our proposals on how to screen for and manage cognitive problems, and you will be asked to comment on and offer feedback on our proposals. You will also be shown some screening tools in various formats. You will be asked for your views on both the format and the content of the screening tools (these are paper-pencil tests or computerised tests), and the impact you think they might have on a person with MS to use this. You will be asked about what you believe is the usual care that people with MS currently receive relating to cognitive problems. This is to help us better understand what happens in routine practice and any gaps in treatment. Then you will be shown the outline of a cognitive rehabilitation intervention and asked for your opinions on how

well you think this might work for people with MS, how it might be improved, and what possible barriers might exist for delivering this intervention within the NHS.

We will be audio-recording these groups so that we do not miss any important points. The recordings will be sent to an external company for transcription (meaning that they will type out the focus group word for word). The company will have a duty of confidentiality to you as a research participant and they will do their best to meet this duty. We will remove any names of people (including your own name) or services so that you cannot be identified in anything that is said.

What are the screening tools and the rehabilitation intervention?
The **screening tools** will be provided to you as a link/attachment before the time of the focus group to allow you to familiarise yourself with them. These are typically questionnaires and puzzles.
The **rehabilitation intervention** will be developed to help improve cognitive problems. You will be presented with some preliminary ideas about this intervention and how it might be delivered.

Are there any risks in taking part?
When talking about experiences and issues associated with cognitive problems, you may feel upset. This may also cause you some concerns about your own cognitive abilities. If this happens during the focus group, the researcher will be there to talk to you and address any concerns. You will have their contact information in case you feel upset after the session for reasons associated with participation. You may also choose to leave the session if you wish.
A possible disadvantage is that it may inconvenience you to attend a focus group in person or be available at a specific time during the day.

Are there any benefits in taking part?
There will be no direct benefit to you from taking part in this research but your contribution may help people with MS in the future by contributing to the development of a cognitive screening and rehabilitation programme.

Will my time/travel costs be reimbursed?
You will receive a £20 gift voucher as a thank you for your time and you will be offered travel expenses to attend the focus group.

What happens to the data provided?
The research data will be stored confidentially to ensure your privacy. You will be assigned a unique identification number (for example P01/A for participant 1 recruited from centre A) that will be used instead of your name on all forms used to save data. We will save all audio recordings and research data using your unique identification number so that none of the data will have your real name or other individual identifiers. Your name and any other information about you will not be disclosed outside the study centre.

Access to the information will be limited to the research team and any relevant regulatory authorities. Computer held data including the study database will be held securely and password protected. All data will be stored separately on secure university networks. Access will be restricted by user identifiers and passwords (encrypted using a one way encryption method). The research team will have access to personal and research data, and a transcriber will have access to some research data. Where a transcription service is used, an appropriate confidentiality contract will be in place between the university-approved provider of the transcription service.

The research will be undertaken by research staff at the University of Nottingham. Data will be analysed and stored by the University of Nottingham. All research data and records will be stored for a minimum of 7 years after publication or public release of the work of the research. Your optional consent for the use of your anonymised data in future research is requested as part of the Participant Consent Form. Such usage in future research would have to be approved by investigators at the University of Nottingham before anonymised data is released.

With your permission, we would like to use fully anonymised direct quotes in research publications. We would also like your permission to use anonymised data in future studies, and to share our research data (e.g. in online databases) with other researchers in other Universities and organisations both inside and outside the European Union. This would be used for research in health and social

care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research) and to understand the bigger picture in particular areas of research. All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public. Data sharing in this way is usually anonymised (so that you could not be identified).

What will happen if I don't want to carry on with the study?
Even after you have signed the consent form, you are free to withdraw from the study at any time without giving any reason, and without your legal rights being affected. Any personal data will be destroyed.

If you decide to leave the study, we will not contact you further. The information you would have provided us so far cannot be erased and this information may still be used in the analysis because your information may have already been used in some analyses and may still be used in the final study analyses. To safeguard your rights, we will use the minimum personally-identifiable information possible.

Who will know that I am taking part in this research?
All information collected about you during this research would be kept strictly confidential. Any audio recordings and electronic data will be anonymised with a code as detailed above. A password-protected list of code numbers and participant names will be stored separately from the audio recordings. All such data are kept on password-protected databases sitting on a restricted access computer system and any paper information (such as your consent form, contact details and any research questionnaires) would be stored safely in lockable cabinets in a swipe-card secured building and would only be accessed by the research team.

Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security) and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights we will use the minimum personally – identifiable information possible.

You can find out more about how we use your information and to read our privacy notice at:

https://www.nottingham.ac.uk/utilities/privacy.aspx/

Designated individuals of the University of Nottingham may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines.

With your consent, we will keep your personal information on a secure database in order to contact you for future studies.

Anything you say during the focus group discussion will be kept confidential, unless you reveal something of concern that may put yourself or anyone else at risk. It will then be necessary to report to the appropriate persons.

What will happen to the results of the research?
This research will help us to develop a screening pathway for use in NHS MS services and a brief rehabilitation programme to support people with MS who have mild to moderate cognitive problems. The findings will be published in scientific journals. The study findings will also be presented to local patient or service provider groups, at academic conferences and to relevant charities. All your personal data will be removed in order to preserve your anonymity and you will not be identifiable in any publications/presentations. A summary of the study findings will be made available through the NEuRoMS web page for you to access, in due course.

Who has reviewed this study?
All research involving people is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable

opinion by the Faculty of Medicine and Health Sciences Research Ethics Committee (Reference number: FMHS 263-1903).

Who is organising and funding the research?
This research is being organised by the University of Nottingham and is being funded by the National Institute for Health Research (NIHR).

What if something goes wrong?
If you have a concern about any aspect of this project, please speak to the researcher Hannah Elwick or the Chief Investigators Professor Roshan dasNair who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how he/she intends to deal with it. If you remain unhappy and wish to complain formally, you can do this by contacting the FMHS Research Ethics Committee Administrator, c/o The University of Nottingham, Faculty PVC Office, B Floor, Medical School, Queen's Medical Centre Campus, Nottingham University Hospitals, Nottingham, NG7 2UH. E-mail: FMHS-ResearchEthics@nottingham.ac.uk

Contact Details
If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Researcher:
Hannah Elwick
Institute of Mental Health, School of Medicine
Email: hannah.elwick@nottingham.ac.uk

Chief Investigator:
Professor Roshan das Nair
Institute of Mental Health, School of Medicine
Phone: +44 (0)115 823 0589
Email: roshan.dasnair@nottingham.ac.uk

PARTICIPANT INFORMATION SHEET: Interview for clinicians and commissioners
Research Ethics Reference: 263-1903
Final Version 1.0 Date: 04/03/2019

Title of Study: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS): Qualitative study for developing a clinical management pathway (Phase 1 Work Package 1a)
Study ID: NEuRoMS Qualitative phase WP1a

We would like to invite you to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. One of our team will go through the information sheet with you and answer any questions you have. Please take time to read this carefully and discuss it with others if you wish. Ask us anything that is not clear.

What is the purpose of the research?
This study aims to improve our ability to detect and manage cognitive problems (such as problems with memory, attention, and problem solving) in people with multiple sclerosis (MS). These cognitive problems can have a negative impact on people's mood, ability to work, and quality of life. 'Neuropsychological rehabilitation' is a type of rehabilitation that may help deal with cognitive problems (e.g., by teaching people how to better concentrate or remember things). However, cognitive problems are not routinely checked (screened) for in people with MS. Early recognition through screening can help with rehabilitation activities before problems become severe. The current study aims to develop a cognitive screening pathway to routinely test all people with MS for cognitive problems, and develop a brief rehabilitation programme for people with 'mild to moderate' cognitive problems.

Why have I been invited to take part?
You have been invited to take part in this research because you have experience of providing or commissioning services for people with MS.

We are inviting up to 15 participants like you to take part in one-to-one interviews.

Do I have to take part?
No. It is up to you to decide if you want to take part in this research. We will describe the study and go through this information sheet with you to answer any questions you may have. If you agree to participate, we will ask you to sign a consent form and will give you a copy to keep. However, you would still be free to withdraw from the study at any time, without giving a reason and without any negative consequences, by telling the researchers of your decision. This would not affect your legal rights.

What will happen to me if I take part?
You will be invited to take part in a one-to-one interview to share your views and experiences on cognitive screening and rehabilitation in MS.

A researcher will contact you to go over the information sheet, explain the procedures, and answer any questions that you have about the study. Interviews can be done face-to-face, over the phone or via Skype, whichever you prefer.

If you agree to take part in the study, the researcher will agree the interview date and time with you. If you are doing an interview face-to-face, the researcher will ask you to complete a consent form before the interview starts. If you are doing the interview over the phone, the consent form will be completed with you over the 'phone. As part of the interview, you will be shown some examples of cognitive screening tools (these are paper-pencil tests or computerised tests). If you have chosen to do your interview over the 'phone or Skype then the screening tool examples will be provided before the interview happens, either by post or through a link to online versions.

The interview will last up to 60 minutes. During the interview, you will be asked for your views on our proposals on how to screen for and manage cognitive problems. You will also be asked for your views on both the format and the content of the screening tools, and the impact you think they might have on a person with MS to use this. You will be asked about what you believe is the usual care that people

with MS currently receive relating to cognitive problems. This is to help us better understand what happens in routine practice and any gaps in treatment. Then you will be shown the outline of a cognitive rehabilitation intervention and asked for your opinions on how well you think this might work for people with MS, how it might be improved, and what possible barriers might exist for delivering this intervention within the NHS.

We will be audio recording the interviews so that we do not miss any important points. The audiotapes will be sent to an external company for transcription. The company will have a duty of confidentiality to you as a research participant and they will do their best to meet this duty. We will remove any names of people (including your own name) or services so that you remain anonymous.

What are the screening tools and the rehabilitation intervention?
The **screening tools** will be provided to you as a link/attachment before the time of the interview to allow you to familiarise yourself with them. These are typically questionnaires and puzzles.

The **rehabilitation intervention** will be developed to help improve cognitive problems. You will be presented with some preliminary ideas about this intervention and how it might be delivered.

Are there any risks in taking part?
We do not anticipate any risks associated with taking part in the interview.
A possible disadvantage is that it may inconvenience you to attend an interview in person or be available at a specific time during the day. We will arrange the interview at a time and date that is suitable for you and if you prefer, it can be conducted over the phone or Skype.

Are there any benefits in taking part?
There will be no direct benefit to you from taking part in this research but your contribution may help people with MS in the future by contributing to the development of a cognitive screening and rehabilitation programme.

Will my time/travel costs be reimbursed?
Travel expenses will be offered for you to attend the interview if travel is involved.

What happens to the data provided?
The research data will be stored confidentially to ensure your privacy. You will be assigned a unique identification number (for example P01/A for participant 1 recruited from centre A) that will be used instead of your name on all forms used to save data. We will save all audio recordings and research data using your unique identification number so that none of the data will have your real name or other individual identifiers. Your name and any other information about you will not be disclosed outside the study centre.

Access to the information will be limited to the research team and any relevant regulatory authorities. Computer held data including the study database will be held securely and password protected. All data will be stored separately on secure university networks. Access will be restricted by user identifiers and passwords (encrypted using a one way encryption method). The research team will have access to personal and research data, and a transcriber will have access to some research data. Where a transcription service is used, an appropriate confidentiality contract will be in place between the university-approved provider of the transcription service.

The research will be undertaken by research staff at the University of Nottingham. Data will be analysed and stored by the University of Nottingham. All research data and records will be stored for a minimum of 7 years after publication or public release of the work of the research. Your optional consent for the use of your anonymised data in future research is requested as part of the Participant Consent Form. Such usage in future research would have to be approved by investigators at the University of Nottingham before anonymised data is released.

With your permission, we would like to use fully anonymised direct quotes in research publications.
We would also like your permission to use anonymised data in future studies, and to share our research data (e.g. in online databases) with other researchers in other Universities and organisations both inside and outside the European Union. This would be used for research in health and social care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research)

and to understand the bigger picture in particular areas of research. All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public. Data sharing in this way is usually anonymised (so that you could not be identified).

What will happen if I don't want to carry on with the study?
Even after you have signed the consent form, you are free to withdraw from the study at any time without giving any reason and without your legal rights being affected. Any personal data will be destroyed.

If you decide to leave the study, we will not contact you further. The information you would have provided us so far cannot be erased and this information may still be used in the analysis because your information may have already been used in some analyses and may still be used in the final study analyses. To safeguard your rights, we will use the minimum personally-identifiable information possible.

Who will know that I am taking part in this research?
All information collected about you during this research would be kept strictly confidential. Any audio recordings and electronic data will be anonymised with a code as detailed above. A password-protected list of code numbers and participant names will be stored separately from the audio recordings. All such data are kept on password-protected databases sitting on a restricted access computer system and any paper information (such as your consent form, contact details and any research questionnaires) would be stored safely in lockable cabinets in a swipe-card secured building and would only be accessed by the research team.

Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security) and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights we will use the minimum personally – identifiable information possible.

You can find out more about how we use your information and to read our privacy notice at:

https://www.nottingham.ac.uk/utilities/privacy.aspx/

Designated individuals of the University of Nottingham may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines.

With your consent, we will keep your personal information on a secure database in order to contact you for future studies.

Anything you say during an interview will be kept confidential, unless you reveal something of concern that may put yourself or anyone else at risk. It will then be necessary to report to the appropriate persons.

What will happen to the results of the research?
This research will help us to develop a screening pathway for use in NHS MS services and a brief rehabilitation programme to support people with MS who have mild to moderate cognitive problems. The findings will be published in scientific journals. The study findings will also be presented to local patient or service provider groups, at academic conferences and to relevant charities. All your personal data will be removed in order to preserve your anonymity and you will not be identifiable in any publications/presentations. A summary of the study findings will be made available through the NEuRoMS web page for you to access, in due course.

Who has reviewed this study?
All research involving people is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Faculty of Medicine and Health Sciences Research Ethics Committee (Reference number: FMHS 263-1903).

Who is organising and funding the research?

This research is being organised by the University of Nottingham and is being funded by the National Institute for Health Research (NIHR).

What if something goes wrong?
If you have a concern about any aspect of this project, please speak to the researcher Jacqueline Mhizha-Murira or the Chief Investigator Professor Roshan das Nair who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how he/she intends to deal with it. If you remain unhappy and wish to complain formally, you can do this by contacting the FMHS Research Ethics Committee Administrator, c/o The University of Nottingham, Faculty PVC Office, B Floor, Medical School, Queen's Medical Centre Campus, Nottingham University Hospitals, Nottingham, NG7 2UH. E-mail: FMHS-ResearchEthics@nottingham.ac.uk

Contact Details
If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Researcher:
Dr Jacqueline Mhizha-Murira
Institute of Mental Health, School of Medicine
Phone: +44 (0)115 8231463
Email: jacqueline.mhizha-murira2@nottingham.ac.uk

Chief Investigator:
Professor Roshan das Nair
Institute of Mental Health, School of Medicine
Phone: +44 (0)115 823 0589
Email: roshan.dasnair@nottingham.ac.uk

Appendix 4. Qualitative Interviews – Consent Forms

University of Nottingham
UK | CHINA | MALAYSIA

Faculty of Psychiatry & Applied Psychology
School of Medicine
University of Nottingham
B Floor, Institute of Mental Health
Innovation Park
Triumph Road
Nottingham, NG7 2TU

Participant Consent Form: Interview
Final version 2.0: 12.07.2019

Title of Study: Qualitative Study to Develop a Cognitive Screening and Management Pathway for People with MS

REC ref: 263-1903

Names of Researchers: Roshan das Nair (Chief Investigator), Shirley Thomas, Paul Leighton, Jennifer Martin, Stephen Timmons, Deborah Fitzsimmons, Gogem Topcu, Kyle Harrington, Hannah Elwick

Name of Participant: **Please initial box**

1. I confirm that I have read and understand the information sheet for the above study which is attached and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I understand that should I withdraw after the interview has taken place then the information collected so far cannot be erased and that this information may still be used in the study analysis. ☐
4. I understand that relevant sections of my data collected in the study may be looked at by the research group and by other responsible individuals for monitoring and audit purposes. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐
5. I understand that what I say during the interview will be kept confidential unless I reveal something of concern that may put myself or someone else at any risk. It will then be necessary to report this to the appropriate persons. ☐
6. I understand that information about me recorded during the study will be made anonymous before it is stored. It will be uploaded into a secure database on a computer kept in a secure place. Data will be kept for 7 years after the study has ended and then destroyed. ☐
7. I agree that my research data may be stored and used in possible future research during and after 7 years, and shared with other researchers including those working outside the University. ☐
8. **Optional:** I agree that my personal data (telephone number, email, address) will be kept up to 12 months after the end of the study so that I can be contacted about the findings of the study. ☐
9. I agree to take part in the above study. ☐

_____	_____	_____
Name of Participant	Date	Signature
_____	_____	_____
Name of Person taking consent	Date	Signature

2 copies: 1 for participant, 1 for the project notes.

MS Cognitive Screen & Manage Participant consent form Interview Final version 2.0, 12/07/2019

Participant Oral Consent Form (MSSocCog)
(Final version 2.0, 12/07/2019)

Title of Study: A Qualitative Study to Develop a Cognitive Screening and Management Pathway for People with MS
REC ref: 263-1903

Name of Chief Investigator: Prof Roshan das Nair
Co-Investigator(s): Dr Nikos Evangelou, Dr Shirley Thomas, Dr Paul Leighton, Dr Jennifer Martin, Dr Gogem Topcu, Prof Deborah Fitzsimmons, Prof Stephen Timmons, Kyle Harrington, Hannah Elwick

Participant Name:

Project Explained: **Yes / No***

Phone/Skype conversation either: recorded or Notes Taken*:

Date and time:

***circle as appropriate**

1. Participant confirmed to have: read and understood the information sheet version number 2.0 dated 12.07.2019 for the above study and had the opportunity to ask questions: **Yes / No***
2. Participant confirmed understanding of their rights as participants: **Yes / No***
3. Participant confirmed understanding of what will happen to their data should they withdraw: **Yes / No***
4. Participant confirmed understanding that their data may be looked at by other responsible individuals for monitoring and audit purposes: **Yes / No***
5. Participant agreed to be recorded during the interview: **Yes / No***
6. Participant agreed to quotes which would not identify them: **Yes / No***
7. Participant confirmed understanding that everything they say during the interview will be kept confidential unless they reveal something of concern that may put someone else at any risk and that it will then be necessary to report this to the appropriate person: **Yes / No***
8. Participant confirmed understanding of how data collected will be protected and handled: **Yes / No***
9. Participant agreed that the anonymised information gathered can be stored for possible use in future studies during and after 7 years, and shared with other researchers including those working outside the University: **Yes / No***
10. **Optional:** Participant agreed to be contacted up to 12 months after the end of the current study for other research studies: **Yes / No***
11. Participant agreed to take part in the study: **Yes / No***

_____ _____ _____
Name of Person taking consent Date Signature

[A copy of this to be sent to the participant upon request]

This study/project is funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (project reference RP-PG--0218-20002). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

University of Nottingham
UK | CHINA | MALAYSIA

NEuRoMS
Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis

Faculty of Psychiatry & Applied Psychology
School of Medicine
University of Nottingham
B Floor, Institute of Mental Health
Innovation Park
Triumph Road
Nottingham, NG7 2TU

Participant Consent Form: Focus Group
Final version 1.1: 04.03.2019

Title of Study: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS)

REC ref: 263-1903

Names of Researchers: Roshan das Nair (Chief Investigator), Shirley Thomas, Paul Leighton, Jennifer Martin, Stephen Timmons, Deborah Fitzsimmons, Gogem Topcu, Kyle Harrington, Hannah Elwick

Name of Participant: **Please initial box**

1. I confirm that I have read and understand the information sheet for the above study which is attached and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I understand that should I withdraw after the focus group has taken place then the information collected so far cannot be erased and that this information may still be used in the study analysis. ☐
4. I understand that relevant sections of my data collected in the study may be looked at by the research group and by other responsible individuals for monitoring and audit purposes. I give permission for these individuals to have access to these records and to collect, store, analyse and publish information obtained from my participation in this study. I understand that my personal details will be kept confidential. ☐
5. I understand that what I say during the focus group will be kept confidential unless I reveal something of concern that may put myself or someone else at any risk. It will then be necessary to report this to the appropriate persons. ☐
6. I understand that information about me recorded during the study will be made anonymous before it is stored. It will be uploaded into a secure database on a computer kept in a secure place. Data will be kept for 7 years after the study has ended and then destroyed. ☐
7. I agree that my anonymised research data may be stored and used in possible future research during and after 7 years, and shared with other researchers including those working outside the University. ☐
8. **Optional:** I agree that my personal data (telephone number, email, address) will be kept up to 12 months after the end of the study so that I can be contacted about the findings of the study. ☐
9. I agree to take part in the above study. ☐

Name of Participant	Date	Signature
Name of Person taking consent	Date	Signature

2 copies: 1 for participant, 1 for the project notes.

This study/project is funded by the National Institute for Health Research (NIHR) under its Programme Grants for Applied Research Programme (project reference RP-PG--0218-20002). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

FUNDED BY NIHR | National Institute for Health Research

Appendix 5. Qualitative Interviews – Semi-Structured Interview Schedules

Interview Schedule
Note: As per the methodology, the semi-structured interview is only a guide, and is necessarily an iterative process (i.e., it gets modified slightly after each interview, based on the information that participants provide).

All participants
Objectives
Nuanced feedback on the various proposed tests, intervention and pathways.

General order
Introduction:
Explain purpose of the interview
Explain confidentiality and disclosure policy.
Explain how long the interview is going to take- 30-60min approximately. Remind participant that the interview will be audio-taped. Also that the interviewer might take notes.
Warm up
Ask general questions to break the ice and establish rapport.
Carry out the interview
From easy, more general to in-depth questions
Cool off
Straight-forward questions to relax the interviewee
Thank them for their participation; goodbye

General questions
Participants will be asked to provide feedback on:
The proposed model for the screening and management pathway
Possible screening tools and completion formats (online, PC, tablet), location (at home, clinics)
The intervention content, format and delivery for NHS contexts
The commissioning of screening and intervention (clinician/commissioner specific)
The functionality/appropriateness of proposed tests, pathways and outcomes
Usual care for people with cognitive problems

Example questions
Using the preliminary logic model (presented to participants in paper format or on a slide) of the proposed cognitive screening and management pathway, can you make any changes to it, based on how you would like the pathway to work?
What do you see are the main barriers and facilitators to the delivery of this pathway? What can be done about the barriers?
What are your comments about the specific contents of the intervention? What's missing? What's not needed?

Specific questions
People with MS, family members and clinicians
Are the following outcomes acceptable?
Impact of cognitive problems
Cognitive impairment
Quality of life
Mood
Function
Self-efficacy
Resource-use questionnaire
Work-related issues
Improved DMT adherence
What outcomes are we not tapping?
Usual care related to cognitive assessment and intervention

Commissioners
What aspects of the services they would consider commissioning
Who would be best placed to commission and provide such services
Usual care related to cognitive assessment and intervention

Prompts
Probing questions (Rubin and Rubin, 1995) will be used to transition the conversation from general to specific. That is:
Requests for extension- e.g. 'Can you tell me more about....?', 'Is there anything else?', 'What happened then?'
Encouraging/prompting questions- e.g. 'uh huh?', 'Yes?', 'Please go on'.
Example questions- e.g. 'Can you give me an example of?'
Follow-up questions- e.g. 'What do you mean by...?', 'Would you talk a bit more about...?'

Focus group schedule

Introduction:

The co-facilitators and other team members (MindTech technology team) will present themselves to the participants. The objectives of the focus groups and ethical issues (e.g., confidentiality) will be reminded to the group. A brief discussion of the structure of the activity will be given and participants will be reminded that the session will be audio recorded.

Topics and questions:

The group will begin with a warm-up exercise, with discussions on cognitive problems in MS and participants' understanding of 'usual care' to manage these issues.

Using a digital projector, the co-facilitators will present our proposed model for the screening and management pathway. There will also be paper handouts of the model on each of the tables. We will ask participants to comment on the proposed pathway, for example, using the following set of questions:

- Using this model as a starting point, can you make any changes to it, based on how you would like the pathway to work?
- What do you see are the main barriers and facilitators to the delivery of this pathway? What can be done about the barriers?

Each group will discuss each of these questions on their own and then together with the whole group.

After these discussions, participants will be presented with proposed screening tools in a range of formats (paper, tablet). We will ask participants to comments on the content and format of the screening tools in smaller groups and then together with the whole group. Members from MindTech technology team providing technical input on the practical aspects of design and content of the proposed cognitive screening tool to be administered online.

A similar format will be followed to explore participants' perceptions on the following factors, with small group discussions leading on to discussions with the group as a whole:

- The intervention content, format and delivery for NHS contexts
- The commissioning of screening and intervention (clinician/commissioner specific)
- The functionality/appropriateness of proposed tests, pathways and outcomes

After these discussions, the session will be summarised and the participants will be given the opportunity to ask any questions. The voice recorder will then be turned off. The facilitator will then explain that the focus group discussion is now officially over and there are no more questions. The participants will then be thanked for their participation.

Appendix 6. Qualitative Interviews – Transcription Confidentiality Agreement

University of Nottingham
UK | CHINA | MALAYSIA

Faculty of Medicine & Health Sciences
School of Medicine
B Floor, Institute of Mental Health
Jubilee Campus
Innovation Park, Triumph Road
Nottingham, NG7 2TU

Confidentiality Agreement for use with Transcription Services

(Final version 1.0 date 18/02/2019)

REC ref: 263-1903

Title of Study: Ne[(Ctrl)]ological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS)

Names of Researchers: Roshan das Nair (Chief Investigator), Shirley Thomas, Paul Leighton, Jennifer Martin, Stephen Timmons, Deborah Fitzsimmons, Gogem Topcu, Kyle Harrington, Hannah Elwick

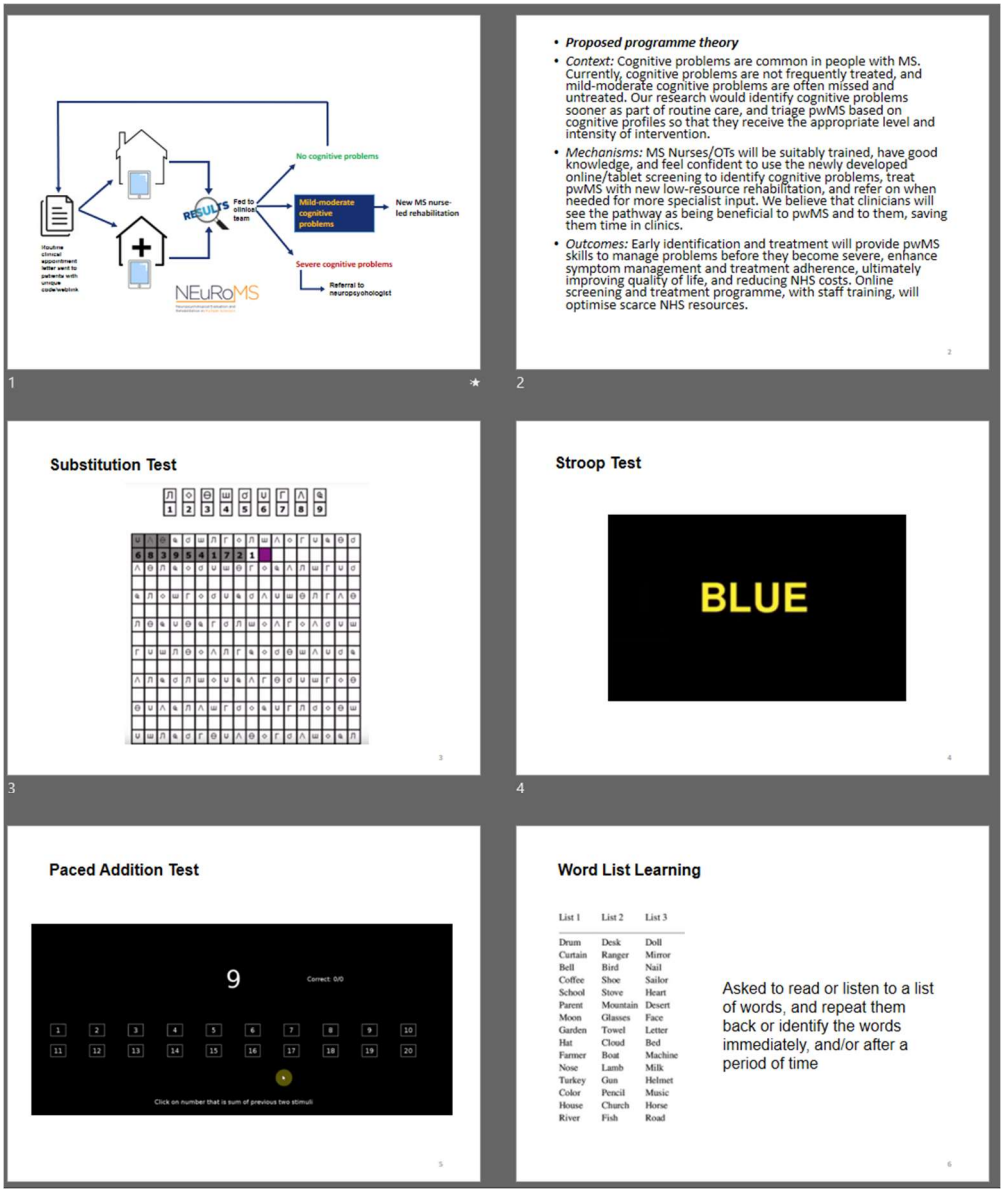
1. I, _____ transcriber, agree to maintain full confidentiality of all research data received from the research team related to this research study.
2. I will hold in strict confidence the identity of any individual that may be revealed during the transcription of interviews or in any associated documents.
3. I will not make copies of any audio-recordings, video-recordings, or other research data, unless specifically requested to do so by the researcher.
4. I will not provide the research data to any third parties without the client's consent.
5. I will store all study-related data, including audio and video recordings and transcriptions in a safe, secure location in a restricted access folder on the university network (preferred)/using encrypted storage devices (delete as appropriate) as long as they are in my possession.
6. All data provided or created for purposes of this agreement, including any back-up records, will be returned to the research team or permanently deleted. When I have received confirmation that the transcription work I performed has been satisfactorily completed, any of the research data that remains with me will be returned to the research team or destroyed, pursuant to the instructions of the research team.
7. I understand that University of Nottingham has the right to take legal action against any breach of confidentiality that occurs in my handling of the research data.

Transcriber's name (printed): _____

Transcriber's signature: _____

Date: _____

Appendix 7. Qualitative Interviews – Participant Facing Materials



Verbal Fluency tests:

"I'm going to give you a category and ask you to name all the different examples that you can think of from that category in one minute. For instance, if I said flowers, you might say rose, daisy, etc. Do you understand?"

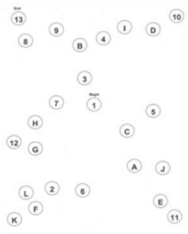
"Now go ahead and tell me all the different ANIMALS you can think of."

Can use categories of things, or letters of the alphabet

7

7

Trail Making Test



There are numbers and letters in circles on this page. Please take the pen and draw a line, alternating in order between the numbers and letters. Start at number 1, then go to the first letter, A, then go to the next number, 2, and then the next letter, B , and so on. Please try not to lift the pen as you move from one number or letter to the next. Work as quickly and accurately as you can.

8

8

Questionnaires:

Multiple Sclerosis Neuropsychological Screening Questionnaire (MSNQ)

Name: _____
Date: _____
Sex (circle one): Male / Female

INSTRUCTIONS: The following questions ask about problems that you may experience. Rate how often these problems occur **AND** how severe they are. Base your ratings on how you have been over the **last 3 months.** Please check the appropriate box.

	Very often, very disruptive (4)	Quite often, somewhat disruptive (3)	Occasionally, seldom a problem (2)	Very rarely, no problem (1)	Never, does not occur (0)
1. Are you easily distracted?	☐	☐	☐	☐	☐
2. Do you lose your thoughts while listening to somebody speak?	☐	☐	☐	☐	☐
3. Are you slow when trying to solve problems?	☐	☐	☐	☐	☐
4. Do you forget appointments?	☐	☐	☐	☐	☐
5. Do you forget what you read?	☐	☐	☐	☐	☐
6. Do you have trouble describing shows or programs recently watched?	☐	☐	☐	☐	☐
7. Do you need to have instructions repeated?	☐	☐	☐	☐	☐
8. Do you have to be reminded to do tasks?	☐	☐	☐	☐	☐
9. Do you forget errands that were planned?	☐	☐	☐	☐	☐
10. Do you have difficulty answering questions?	☐	☐	☐	☐	☐

9

9

Appendix 8. Qualitative Interviews – Ethical Approval

University of Nottingham
UK | CHINA | MALAYSIA

Email: FMHS-ResearchEthics@nottingham.ac.uk

Faculty of Medicine & Health Sciences
Research Ethics Committee
c/o Faculty PVC Office
School of Medicine Education Centre
B Floor, Medical School
Queen's Medical Centre Campus
Nottingham University Hospitals
Nottingham, NG7 2UH

3 April 2019

Professor Roshan das Nair
Professor of Clinical Psychology and Neuropsychology
Division of Psychiatry & Applied Psychology
School of Medicine, University of Nottingham
B19, Institute of Mental Health, Innovation Park
Triumph Road
Nottingham
NG7 2TU

Dear Professor das Nair

Ethics Reference No: 263-1903 **– please always quote**	
Study Title: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS): Qualitative study for developing a clinical management pathway (Phase 1 Work Package 1a)	
Chief Investigator/Supervisor: Professor Roshan das Nair, Professor of Clinical Psychology and Neuropsychology	
Other Key Investigators: Dr Nikos Evangelou, Clinical Associate Professor, Clinical Neuroscience, Dr Shirley Thomas, Rehabilitation, Ageing and Wellbeing, Dr Paul Leighton, Associate Professor or Applied Health Services Research, Dr Jennifer Martin Senior Programme Manager, NIHR MindTech MedTech Co-operative, Dr Gogem Topcu, Research Fellow, & Programme Manager (NEuROMS study), Professor Deborah Fitzsimmons Academic Director of Swansea Centre for Health Economics, Swansea University, Professor Stephen Timmons, Business School, Kyle Harrington, Research Fellow, NIHR MindTech MedTech Co-operative, Hannah Elwick, Postgraduate Research Student, Institute of Mental Health.	
Proposed Start Date: 01/05/2019	**Proposed End Date:** 31/10/2019

Thank you for submitting the above application and the following documents were received:

- FMHS REC Application form and supporting documents version 1.0: 04.03.2019

These have been reviewed and are satisfactory and the study has been given a favourable opinion.

A favourable opinion has been given on the understanding that:

1. The protocol agreed is followed and the Committee is informed of any changes using a notice of amendment form (please request a form).
2. The Chair is informed of any serious or unexpected event.
3. An End of Project Progress Report is completed and returned when the study has finished (Please request a form).

Yours sincerely



Professor Ravi Mahajan
Chair, Faculty of Medicine & Health Sciences Research Ethics Committee

University of Nottingham
UK | CHINA | MALAYSIA

DPAP Committee

20/03/2019

Supervisor: Roshan das Nair

Applicant : Hannah Elwick

Project: Project Id Cognitive Screening in Multiple Sclerosis

A favourable opinion is given to the above named study on the understanding that the applicants conduct their research as described in the above numbered application. Applicants need to adhere to all conditions under which the ethical approval has been granted and use only materials and documentation that have been approved. If any amendments to the study are required, an amendment should be submitted to the committee for approval. An end of study form will be required once the study is complete.

yours



Professor David Daley

Co-Chair of DPAP Ethics Subcommittee



Professor Amanda Griffiths

Co-Chair of DPAP Ethics Subcommittee

Appendix 9. Usability Cognitive Interviews – Advert for Participants

University of Nottingham
UK | CHINA | MALAYSIA

Living with Multiple Sclerosis? Interested in helping with research to identify problems with attention, memory and thinking?

We are inviting people with MS to help us develop ways of identifying problems with remembering things, learning new things, paying attention and concentrating.

We have developed a webapp to assess these problems electronically on mobile phones or tablets, and we would like people with MS to help us test out how user friendly the system is.

We are hoping that this webapp can be used in the NHS to help people with MS.

We would like to talk with people with MS, who are 18 and over, and have access to video chat on their computer, and a handheld touchscreen device such as a smartphone or tablet. You will be invited to take part in interviews with a researcher where you can try out the web app, and tell us what you think about it.

It shouldn't take more than 60 minutes and we are offering a £20 inconvenience shopping voucher.

If you are interested in taking part, please contact Hannah Elwick at hannah.elwick@nottingham.ac.uk



This study has been funded by the MS Society. Ethical approval has been obtained from the University of Nottingham. The information you provide will be collected anonymously. Your name will not be recorded and anything that might identify you will be removed.

Appendix 10. Usability Cognitive Interviews – Participant Information Sheet

University of Nottingham
UK | CHINA | MALAYSIA

Faculty of Medicine & Health Sciences
School of Psychiatry and Applied Psychology
B06, Institute of Mental Health Building
Jubilee Campus
7301 Wollaton Rd
Nottingham
NG8 1BB

NEuRoMS
Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis

FUNDED BY NIHR | National Institute for Health Research

PARTICIPANT INFORMATION SHEET

Research Ethics Reference: 372-1908
Version 1.3 Date: 26/11/2020

Study Title: User Testing of Cognitive Screening Tools for Multiple Sclerosis (NEuRoMS WP1b-Part 2)

Principal Investigator: Jennifer Martin, Senior Programme Lead, MindTech.
Jennifer.Martin@Nottingham.ac.uk

Lead researcher: Kyle Harrington, Research Fellow, MindTech
Kyle.Harrington@Nottingham.ac.uk

We would like to invite you to take part in a research study. Before you decide, it is important for you to understand why the research is being done and what it will involve. One of our team will go through the information sheet with you and answer any questions you have. Please take time to read this carefully and discuss it with others if you wish. Ask us anything that is not clear.

What is the purpose of the research?
Cognitive problems, which involves problems with remembering things, learning new things, paying attention and concentrating, can affect many people with Multiple Sclerosis (MS). We have developed a way of assessing people's cognitive problems online (that is, by asking people to complete a puzzle on their mobile phones or tablets). We would like to understand whether asking people with multiple sclerosis (MS) to complete cognitive assessments on their own personal devices is feasible, and to identify the ways in which people living with MS interact with their devices in order to provide design guidelines for the development of this technology. In addition, we also want to invite people living with MS to give feedback on an early working prototype of commonly used cognitive task.

Why have I been invited to take part?

You have been invited to take part in this research because you have indicated that you are interested in taking part in this research and have been diagnosed with MS and aged 18 years or over and have access to video conferencing, such as Zoom, Skype or Microsoft Teams, and your own handheld touchscreen device (a smartphone or tablet).

We will be recruiting up to 10 people like you into this study.

Do I have to take part?
No. It is up to you to decide if you want to take part in this research. We will describe the study and go through this information sheet with you to answer any questions you may have. If you agree to participate, we will ask you to sign a consent form and will give you a copy to keep. However, you would still be free to withdraw from the study at any time, without giving a reason and without any negative consequences, by advising the researchers of this decision. This would not affect your legal rights.

1. ***What will happen to me if I take part?***

The researcher will ensure that you understand the information sheet, explain the procedures, and ensure you are aware of the aims and methods of the study. If you agree to take part in the study, you will be asked to give your consent to take part verbally.

The study will involve an interview which should take approximately 30 minutes and will last no longer than 60 minutes. The researcher will arrange a convenient time for you, and send you an invitation to a video conferencing meeting and a link to the technology that we are testing in an email.

During the video chat, we will explain the research, and ask for your consent to take part. We will then ask you some questions about your MS and your use of technology, and ask you to click on the link to the platform we are testing on your own touchscreen device and we will give you a user ID so that you can access the webapp.

We will ask for your feedback on a webapp which may be used for assessments of cognitive impairment in MS. We will not be giving you a "score" or making any form of assessment about your cognition. We are simply inviting you to give your opinion and feedback on early working prototypes of the technology we have developed. You will be asked to "talk aloud" whilst you are using the prototype applications, and you will be asked in broad terms about any potential areas of frustration or difficulty you experience. We will record these interviews using the video conferencing platform.

The recordings will be sent to an external company for transcription (meaning that they will type out the interview word for word). The company will have a duty of confidentiality to you as a research participant and they will do their best to meet this duty. We will remove any names of people (including your own name) or services so that you cannot be identified in anything that is said.

2. Are there any risks in taking part?

We do not foresee any obvious risks to health and safety in participating in this study. We will ensure confidentiality and only use anonymised data in our reports. Please indicate on the consent form if you agree for your data to be used in this way.

3. Are there any benefits in taking part?

You will get the opportunity to contribute to the development of a cognitive screening technology which will hopefully be used in the NHS at some point in the future. We welcome your feedback in this area and hope that you will also find the session interesting and enjoyable.

4. Will my time/travel costs be reimbursed?

You will receive an inconvenience allowance to participate in the study (£20 of shopping vouchers)

5. What happens to the data provided?

The research data (videos and audio recordings) will be stored confidentially on password protected secure dedicated servers at the University of Nottingham, original copies stored on recording devices will be deleted. You will be given a volunteer identification number for your reference. We will keep any personal information you provide (on your consent form), separate from video or processed data, linked only by the participant identification number. We will not share any personally identifying information about you outside of the immediate research team (listed on the information sheet). Data will be stored only on University of Nottingham computers and servers, which will be password protected and encrypted at all times.

Once the data has been processed and anonymised, we will conduct analysis on this data and use the results to inform a series of design decisions. We will share our anonymised data with a variety of researchers, developers and other stakeholders. We may also publish or present the anonymised data more broadly. In addition, we may wish to use still images or video clips from the interviews for illustrative purposes, but will do so without any other personal information and only with your permission.

All research data and records will be stored for a minimum of 7 years after publication or public release of the work of the research.

We would like your permission to use anonymised data in future studies, and to share our research data (e.g. in online databases) with other researchers in other Universities and organisations both inside and outside the European Union. This would be used for research in health and social care. Sharing research data is important to allow peer scrutiny, re-use (and therefore avoiding duplication of research) and to understand the bigger picture

in particular areas of research. All personal information that could identify you will be removed or changed before information is shared with other researchers or results are made public.

Data sharing in this way will be anonymised (so that you could not be identified)

6. What will happen if I don't want to carry on with the study?

Even after you have signed the consent form, you are free to withdraw from the study at any time without giving any reason and without your legal rights being affected. Any personal data will be destroyed.

You have the right to withdraw from the study even after you have completed the interview. We will wait a week following which we will be anonymising the data for analysis purposes, and for this reason it may not be possible to withdraw once this has been done. For this reason, if you wish to withdraw once you have completed the interview, please inform us within seven days after the interview, otherwise it may not be possible to withdraw your data.

7. Who will know that I am taking part in this research?

All information collected about you during this research would be kept strictly confidential. All such data are kept on password-protected databases sitting on a restricted access computer system and servers and any paper information (such as your consent form, contact details and any research questionnaires) would be stored safely in lockable cabinets in a swipe-card secured building and would only be accessed by the research team.

Under UK Data Protection laws the University is the Data Controller (legally responsible for the data security) and the Chief Investigator of this study (named above) is the Data Custodian (manages access to the data). This means we are responsible for looking after your information and using it properly. Your rights to access, change or move your information are limited as we need to manage your information in specific ways to comply with certain laws and for the research to be reliable and accurate. To safeguard your rights, we will use the minimum personally-identifiable information possible.

You can find out more about how we use your information and to read our privacy notice at:

https://www.nottingham.ac.uk/utilities/privacy.aspx/

Designated individuals of the University of Nottingham may be given access to data for monitoring and/or audit of the study to ensure we are complying with guidelines.

With your consent, we will keep your personal information on a secure database in order to contact you for future studies.

Anything you say during an interview/focus group will be kept confidential, unless you reveal something of concern that may put yourself or anyone else at risk. It will then be necessary to report to the appropriate persons.

8. What will happen to the results of the research?

The research will be used to inform a series of design decisions by a technical development team within the University of Nottingham wishing to develop these technologies further. We also intend to publish the results and findings of this work in conferences and journal papers, and may present some of these findings in lecturers, seminars and workshops.

9. Who has reviewed this study?

All research involving people is looked at by an independent group of people, called a Research Ethics Committee, to protect your interests. This study has been reviewed and given favourable opinion by the Faculty of Medicine and Health Sciences Research Ethics Committee (Reference number: 372-1908).

10. Who is organising and funding the research?

This research is being organised by the University of Nottingham, and is being funded by the National Institute for Health Research (NIHR).

11. What if something goes wrong?

If you have a concern about any aspect of this project, please speak to the researcher Kyle Harrington or the Principal Investigator Jennifer Martin, who will do their best to answer your query. The researcher should acknowledge your concern within 10 working days and give you an indication of how he/she intends to deal with it. If you remain unhappy and wish to complain formally, you can do this by contacting the FMHS Research Ethics Committee Administrator, c/o The University of Nottingham, Faculty PVC Office, B Floor, Medical School, Queen's Medical Centre Campus, Nottingham University Hospitals, Nottingham, NG7 2UH. E-mail: FMHS-ResearchEthics@nottingham.ac.uk

12. Contact Details

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

Lead Researcher:
Kyle Harrington, Research Fellow, MindTech
B06, Institute of Mental Health Building
Jubilee Campus
7301 Wollaton Rd
Nottingham
NG8 1BB
Tel: 0115 7484214.
Email: Kyle.Harrington@Nottingham.ac.uk

Principal Investigator:
Jennifer Martin, Senior Programme Manager, MindTech
B06, Institute of Mental Health Building
Jubilee Campus
7301 Wollaton Rd
Nottingham
NG8 1BB
Tel: 0115 74 84339
Email: Jennifer.Martin@Nottingham.ac.uk

Appendix 11. Usability Cognitive Interviews – Verbal Consent Form



School of Medicine
University of Nottingham
Medical School
Nottingham
NG7 2UH

Participant Consent

Project Title: User Testing of Cognitive Screening Tools for MS
Researcher: Hannah Elwick Hannah.elwick@nottingham.ac.uk
Supervisor: Roshan das Nair roshan.dasnair@nottingham.ac.uk
Ethics Reference Number: 372-1908

- Have you read and understood the Participant Information? ☐Yes ☐No
- Do you agree to take part in an interview that will be recorded about cognitive assessments? ☐Yes ☐No
- Do you know how to contact the researcher if you have questions about this study? ☐Yes ☐No
- Do you understand that you are free to withdraw from the study without giving a reason? ☐Yes ☐No
- Do you understand that once you have been interviewed it may not be technically possible to withdraw your data? ☐Yes ☐No
- Do you give permission for your data from this study to be shared with other researchers in the future provided that your anonymity is protected? ☐Yes ☐No
- Do you understand that non-identifiable data from this study, including quotations might be used in academic research reports or publications? ☐Yes ☐No
- I confirm that I am 18 years old or over ☐Yes ☐No

If you would like a summary of the research findings please insert your email address in this text box

By ticking the button below, I indicate that I understand what the study involves, and I agree to take part.
I consent to take part in this research study ☐Yes

27 April 2020

Appendix 12. Usability Cognitive Interviews – Semi Structured Interview Schedule

Cognitive Walkthrough Script

Once participant has given consent, and researcher has asked about demographics and some preliminary question:

Please could you navigate to the link I sent you in a previous email on your smartphone or tablet? I will then give you a user ID to enter to access the tasks.

Once the participant has accessed the technology using the ID. If the participant had any difficulties accessing the platform, these will be noted.

Whilst you are using the app, I encourage you to say exactly what you are thinking as you are thinking it. The purpose of this research is to discover exactly what people think about the app whist they are using it, so any insights, impressions, or comments you have, no matter how unimportant you think they are, will be useful to us. Please just feel free to talk aloud, almost as if you are talking to yourself.

You don't need to feel under any pressure to perform well. We are not testing you, but rather, you are testing the app. So, if something doesn't work the way you expected it to, that is a problem for us to solve rather than anything you have done wrong.

During the time you are using the app, I will try not to interrupt, but you can ask for help if you do feel that you are stuck at any point. At various points I may also suggest that you try particular things, if I feel there are one or two things you might have missed. But I will seek to keep these interruptions to a minimum. I might also prompt you to "keep talking" if you stop talking.

In a moment I will start the audio and video recording. Is there anything you are unsure about or want to ask about before we begin?

Great, ok. Then let's begin".

Probing Questions

Demographic/Device/Access Questions

- Was the link easy to navigate to?
- What device and browser are you using?
- Do you have any accessibility options on your phone (e.g., large text)?
- Does this impact the platform at all? (e.g., push everything off the screen)
- Does everything fit on one page?
- Can you find the 'next' button?

Instructions

- Were the instructions clear and readable?
- What did you think about the example?
- Do you need two examples, or can you understand the task ok with one example?
- What do you think the task was asking you to do?
- Do you think you need to complete the task accurately or as fast as you can?

Practice Task

- What do you think of the time it took you to complete it?

Real Task

- Did everything fit on one page?
- Were you able to tap the buttons ok?
- What do you think of the time it took you to complete it?
- What do you think the task is assessing?
- Would more context about this be helpful?

Finishing Up

- Having done the task, has your opinion of what the task is assessing you changed?
- Can you give me your general impressions of the platform?
- Was there anything you particularly liked?
- Was there anything you particularly disliked or found difficult?
- You said that "…" could you elaborate on what you meant?
- If you were required to use this application at annually, or at semi-regular intervals, would you feel comfortable using it?
- Do you think this would be of benefit to you?
- Is there anything else you wish to say that has not been covered that you think could be of use to us?

Appendix 13. Usability Cognitive Interviews – Ethical Approval

University of Nottingham
UK | CHINA | MALAYSIA

Faculty of Medicine & Health Sciences
Research Ethics Committee
Faculty Hub
Room E41, E Floor, Medical School
Queen's Medical Centre Campus
Nottingham University Hospitals
Nottingham, NG7 2UH
Email: FMHS-ResearchEthics@nottingham.ac.uk

28 July 2020

Dr Jennifer Martin
Senior Programme Manager
NIHR MindTech MedTech Co-operative
NIHR Nottingham Biomedical Research Centre (BRC)
Institute of Mental Health
School of Medicine
Jubilee Campus
University of Nottingham
Nottingham, NG8 1BB

Dear Dr Martin

Ethics Reference No: 372-1908 **– please always quote**	
Study Title: Neuropsychological Evaluation and Rehabilitation in Multiple Sclerosis (NEuRoMS): Developing a cognitive screening pathway in Multiple Sclerosis (Phase 1 Work Package 1b parts 1 &2)	
Chief Investigator/Supervisor: Dr Jennifer Martin Senior Programme Manager, NIHR MindTech MedTech Co-operative, NIHR Nottingham BRC, Professor Roshan das Nair, Professor of Clinical Psychology and Neuropsychology, Division of Psychiatry and Applied Psychology, School of Medicine.	
Other Key Investigators: Dr Jacqueline Mhizha-Murira, Research Fellow, Division of Psychiatry and Applied Psychology, Dr Shirley Thomas, Rehabilitation, Ageing and Wellbeing, Dr Paul Leighton, Associate Professor or Applied Health Services Research, , Dr Gogem Topcu, Research Fellow, & Programme Manager (NEuROMS study), Professor Deborah Fitzsimmons Academic Director of Swansea Centre for Health Economics, Swansea University, Professor Stephen Timmons, Business School, Dr Kyle Harrington, Research Fellow, NIHR MindTech MedTech Co-operative, Hannah Elwick, Postgraduate Research Student, Institute of Mental Health.	
Proposed Start Date: Part 1 01/08/2019 **Part 2** 01/08/2019	**Proposed End Date: Part 1** 31/10/2020 **Part 2** 31/10/2020

Thank you for notifying the Committee of amendment no 1: 24.06.2020: to increase the number of participants to 20, include a time extension with a new end date of 31 October 2020, change to remote video or conference call in line with the current Covid-19 Health Crisis. The following documents were received:

- E-mail Notice of amendment no 1: dated 24.06. 2020

These have been reviewed and are satisfactory and the study amendment no 1: 24.06.2020 is approved.

Yours sincerely

Dr Bethan E Phillips, Associate Professor
Clinical, Metabolic & Molecular Physiology, Medical Sciences & Graduate Entry Medicine

Appendix 14. Usability Cognitive Interviews – DECSYS Symbol Substitution Task Instructions

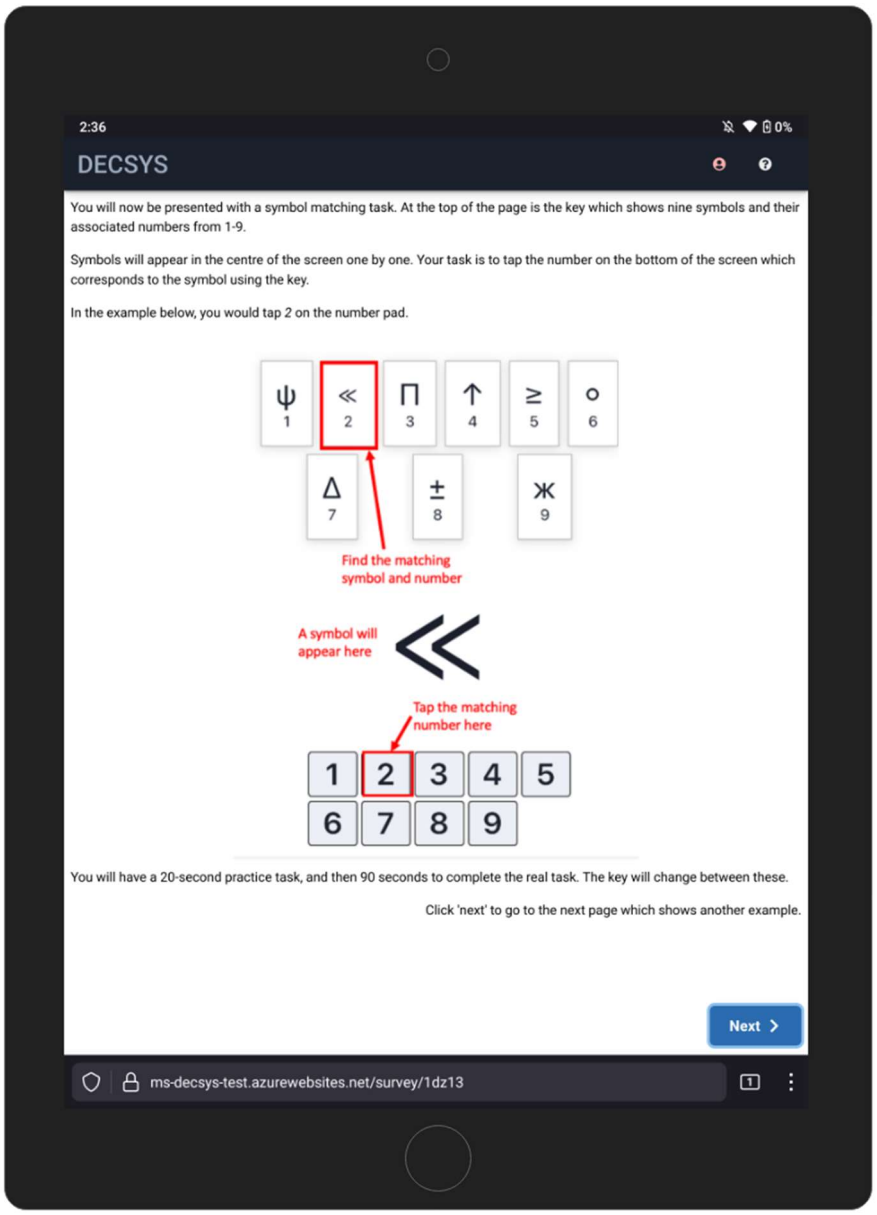


Figure 7: Symbol Substitution Task instructions – Screen 1

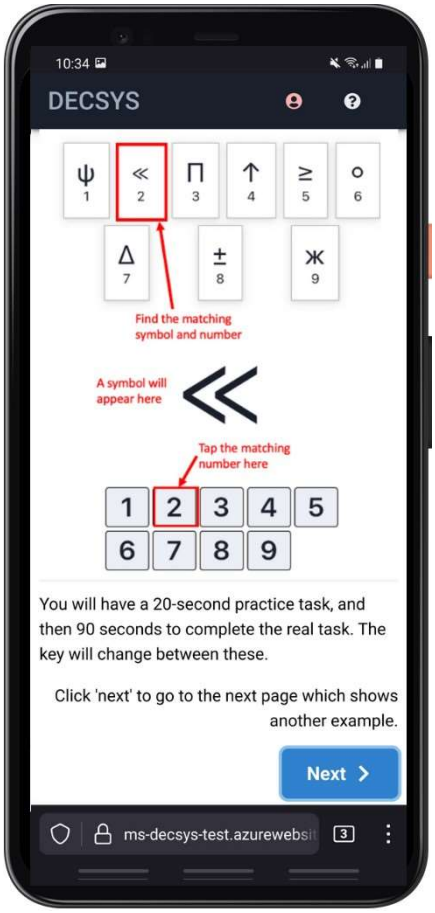


Figure 8: Symbol Substitution Task instructions - Screen 2

Appendix 15. Validation Study – Ethical Approval

University of Nottingham
UK | CHINA | MALAYSIA

DPAP Committee :27/04/2020
Supervisor: Professor Roshan Das Nair
Applicant: Hannah Elwick

Project ID 1517 - Validation of a Cognitive Screen for People with Multiple Sclerosis

Dear Hannah,

A favourable opinion is given to the above named study on the understanding that the applicants conduct their research as described in the above numbered application. Applicants need to adhere to all conditions under which the ethical approval has been granted and use only materials and documentation that have been approved.

If you need to make any any changes (for example to the date or place of data collection, or measures used), an Amendment Form should be submitted. This can be done by the Supervisor in 'Create Sub Form' in the Actions Menu on the left hand side of the page on the on-line system: Select 'Amendment Form'

Yours

David Daley

Professor David Daley

Co-Chair DPAP Ethics Subcommittee

Amanda Griffiths

Professor Amanda Griffiths

Co-Chair DPAP Ethics Subcommittee