



City Research Online

City St George's, University of London

Citation: Newman, M. L. (2022). Examining the drawing process used to complete the Rey-Osterrieth Complex Figure Task in individuals with mild traumatic brain injury. (Unpublished Doctoral thesis, City, University of London)

This is the accepted version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://city-ref-test.eprints-hosting.org/id/eprint/30158/>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).

**Examining the drawing process used to
complete the Rey-Osterrieth Complex Figure
Task in individuals with mild traumatic brain
injury.**

Michelle Lyndsey Newman

*Department of Psychology
City, University of London*

Thesis submitted in fulfilment of the requirements for the
degree of Doctor of Philosophy

September 2022

Declaration

I, Michelle Lyndsey Newman confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.



Abstract

Mild traumatic brain injury (mTBI) accounts for 65-90% of all brain injuries. Whilst the majority make a full recovery within three months, 2-24% report persisting symptoms beyond this window, sometimes for years. This may result in a diagnosis of Post-Concussion Syndrome (PCS). However, a long-standing debate questions the cause of these symptoms, arguing there is a lack of evidence of tissue damage or cognitive deficit. Symptoms are attributed to other factors such as low intelligence, mental health or malingering. Supporters of PCS having an organic basis have highlighted limitations in the tools used to measure cognition and behaviour. New neuroimaging technology illustrates organic damage occurs at a substructural and metabolic level not identified in earlier techniques. Further, the research design and statistical analysis has been criticised. Research has indicated that researchers and clinicians should be focusing on *how* individuals with mTBI completed tasks, not on *what* they achieve. The Rey-Osterrieth Complex Figure tasks offers a solution to observe how individuals perceive and problem-solve complex visual information.

A mixed-methods approach within a clinical neuropsychological framework was used to examine whether observations of drawing process used to complete the Rey-Osterrieth Complex Figure task can support a diagnosis of PCS. Drawing process scores from both non-mTBI and mTBI populations were used to validate a novel scoring system of four constructs, and to identify if drawing process measurements were predictors of post-concussive symptoms and behavioural challenges. Thematic analysis was also used to analyse examine in greater detail the challenges experienced following mTBI, and to consider how cognitive performance correlates with subjective reports of these challenges.

Results indicated observations of drawing process were strong predictors of subjective reporting of persisting symptoms and behavioural challenges post-mTBI. Examination of drawing process constructs alongside other neuropsychological measures indicated that reduced inhibitory control and cognitive flexibility underpins persisting symptoms and, in particular, cognitive fatigue. The theme of Effort and Energy indicated that fatigue was a significant long-term challenge following mTBI, supporting theoretical models that increased cognitive effort underpins cognitive fatigue. Analysis showed that mental health was also a strong predictor of symptoms and challenges. However, the nature of how this is experienced by individuals with mTBI illustrated different sources for these psychological challenges – concluding that mental health is a moderator of cognitive dysfunction following mTBI, rather than the cause.

Table of contents

Declaration	2
Abstract	3
Table of contents	4
Index of tables	10
Index of figures	12
Acknowledgements	13
Abbreviations	14
 Chapter One – Introduction to Mild Traumatic Brain Injury	 15
1.1 Chapter premise	15
1.2 Acquired brain injury	15
1.3 Traumatic brain injury	16
1.4 Mild traumatic brain injury	17
1.5 Concussion	18
1.6 Epidemiology and prognosis of mTBI	19
1.7 Pathological processes of mTBI	20
1.8 Post-concussion syndrome and the symptoms of mTBI	21
1.9 Outcomes of mTBI and PCS	23
1.10 Neuroimaging of mTBI	23
1.11 Cognitive deficits following mTBI	24
1.12 Sensitivity of neuropsychological measures – are we measuring the right thing?	26
1.13 The research question	27
 Chapter Two – Methodology and Methods	 29
2.1 Chapter premise	29
2.2 Cognitive neuropsychological approach	29
2.3 Research design and methods	30
2.4 Participant recruitment	31
2.5 Participant inclusion and exclusion criteria	31
2.5.1 <i>Inclusion criteria</i>	32
2.5.2 <i>Exclusion criteria</i>	32
2.5.3 <i>Mild traumatic brain injury classification and time since injury</i>	36
2.6 Protocol	36

2.6.1 <i>Estimated intelligence</i>	37
2.6.2 <i>Neuropsychological assessments</i>	38
2.6.3 <i>Mental health inventories</i>	41
2.6.4 <i>Symptoms and behavioural challenges</i>	42
2.7 Participant history	44
2.7.1 <i>Demographic details</i>	44
2.8 Traumatic brain injury history	45
2.9 Variations in administration of protocol	46
2.9.1 <i>Covid-19 pandemic adaptations</i>	46
2.10 Participant incentives	49
2.11 Concluding comments	50
 Chapter Three – The Rey-Osterrieth Complex Figure Task	 51
3.1 Chapter premise	51
3.2 The Rey-Osterrieth complex figure task	51
3.3 The original system for measuring accuracy	53
3.4 Development of the Rey-Osterrieth system for observing drawing process	54
3.5 Revisions to the Rey-Osterrieth scoring system	55
3.6 Psychometric properties and selection of scoring systems	59
3.6.1 <i>Inter-rater and intra-rater reliability of Osterrieth accuracy scoring</i>	59
3.6.2 <i>Inter-rater and intra-rater reliability of other accuracy scoring systems</i>	60
3.6.3 <i>Psychometric analyses of qualitative scoring systems</i>	60
3.7 Copy and recall phases	61
3.7.1 <i>Number of phases</i>	61
3.7.2 <i>Length of recall delay</i>	62
3.8 Alternate figures and group design	63
3.9 Demographic factors	65
3.9.1 <i>Age</i>	65
3.9.2 <i>Intellect & education</i>	66
3.9.3 <i>Biological Sex</i>	67
3.10 Qualitative characteristics scoring	67
3.11 Construct of Wholeness	73
3.12 Construct of Order	74
3.13 Construct of Continuation	76
3.14 Construct of Strategy	78
3.15 Concluding comments	79

Chapter Four – Use of digital technology in neuropsychological assessment	81
4.1 Chapter premise	81
4.2 Use of a digital pen to record the drawing process used to complete the Rey-Osterrieth Complex Figure Task	81
4.2.1 <i>Traditional methods of recording the drawing process</i>	81
4.2.2 <i>Digitisation of the administration of the RCF</i>	82
4.2.3 <i>Empirical evidence for the use of digital pens in neuropsychological assessment</i>	84
4.2.4 <i>Aims and hypotheses</i>	85
4.2.5 <i>Study design, group allocation and participants</i>	85
4.2.6 <i>Equipment and procedure</i>	86
4.2.7 <i>Results</i>	87
4.2.7.1 ANOVA analysis	87
4.2.7.2 Correlation coefficient analysis	88
4.2.7.3 Intraclass correlation coefficient	89
4.2.7.4 Bland-Altman plots of agreement	89
4.2.7.5 Usability and comfort	89
4.2.8 <i>Study conclusions</i>	92
4.3 Video-based teleneuropsychology	94
4.3.1 <i>History of teleneuropsychology</i>	
4.3.2 <i>Empirical evidence for the use of video-based teleneuropsychology</i>	95
4.3.3 <i>Aims and hypotheses</i>	101
4.3.4 <i>Conditions and allocations</i>	101
4.3.5 <i>Participants</i>	101
4.3.6 <i>Procedure and online adaptations</i>	102
4.3.7 <i>RCF Results</i>	102
4.3.7.1 ANOVA analysis	103
4.3.7.2 Correlation coefficient analysis	103
4.3.7.3 Intraclass correlation coefficient	104
4.3.7.4 Bland-Altman plots of agreement	105
4.3.8 <i>Cognitive screen results</i>	105
4.3.8.1 Group differences analysis	105
4.3.8.2 Correlation coefficient analysis	109
4.3.8.3 Intraclass correlation analysis	109
4.3.8.4 Bland-Altman plots of agreement	110
4.3.9 <i>Study conclusions</i>	114

4.4 Chapter Discussion	117
Chapter Five – Drawing process in healthy adults	121
5.1 Chapter premise	121
5.2 Participants	121
5.3 Protocol and scoring procedures	122
5.4 Accuracy scoring population data	123
5.5 Wholeness construct population data	125
5.6 Order construct population data	127
5.7 Continuity construct population data	131
5.8 Strategy population data	133
5.8.1 <i>Principal component analysis</i>	133
5.8.2 <i>Strategy construct population data</i>	137
5.9 Time to completion	138
5.10 Handedness	139
5.11 Demographic factors	139
5.12 Correlation between RCF measures and other cognitive measures	139
5.13 Chapter discussion	140
Chapter Six – Cognitive, Psychological and Demographic predictors of post-concussive symptoms	147
6.1 Chapter premise	147
6.2 Introduction	147
6.3 Study design	151
6.4 Description of participants	154
6.5 Brain imaging report	154
6.6 Procedures	155
6.7 Outcome measure descriptives and inferentials	155
6.8 Cognitive predictor variable descriptives and inferentials	156
6.9 Psychological predictor variable descriptives and inferentials	157
6.10 Variable correlation coefficient analysis	157
6.10.1 <i>Predictor correlations with outcome measures</i>	158
6.10.2 <i>Correlation coefficients between predictor variables</i>	159
6.11 Predictor variable screening	163
6.12 Regression analysis assumptions	163
6.12.1 <i>Variable characteristics</i>	163

6.12.2 <i>Independence of observations</i>	163
6.12.3 <i>Linearity</i>	164
6.12.4 <i>Homoscedasticity</i>	164
6.12.5 <i>Multicollinearity</i>	164
6.12.6 <i>Outliers</i>	165
6.12.7 <i>Normality of residuals</i>	165
6.13 Multiple liner regression analysis	165
6.13.1 <i>Rivermead post-concussion symptoms questionnaire</i>	165
6.13.2 <i>Mental fatigue scale</i>	167
6.13.3 <i>Dysexecutive Questionnaire</i>	168
6.13.4 <i>Cognitive Failures Questionnaire</i>	169
6.14 Chapter discussion	170
 Chapter Seven – Living with post-concussive symptoms	 178
7.1 Chapter premise	178
7.2 Introduction	178
7.3 Participants	181
7.4 Procedure	181
7.4.1 <i>Reflexive positioning</i>	183
7.5 Thematic analysis	183
7.5.1 <i>Reconciliation and validation of recovery status</i>	185
7.5.2 <i>Tangible symptoms</i>	189
7.5.3 <i>Effort and energy</i>	191
7.5.4 <i>Living with mTBI within a social world</i>	193
7.5.5 <i>Psychology of living with mTBI</i>	195
7.6 Quantitative post-concussive symptoms	197
7.7 Chapter discussion	199
 Chapter Eight - Discussion	 205
8.1 Chapter premise	205
8.2 The Research question – a reminder	205
8.3 How to measure drawing process used for the Rey-Osterrieth Complex Figure	209
8.4 Increasing functionality of the Rey-Osterrieth Complex Figure through the use of digital recording devices	210
8.5 Adapting to Covid-19 – extending literature on video-based teleneuropsychology	212

8.6 Binary versus spectrum: why the rectangle is not a good indicator of cognitive deficit	215
8.7 Examining drawing performance of the Rey-Osterrieth Complex Figure following mTBI	217
8.8 The patient report – understanding the nature of long-term post-concussive symptoms	219
8.9 Answering the research question	221
8.10 Limitations	224
8.11 Impact and future research	227
8.12 Concluding summary	228
Appendices	230
References	250

Index of Tables

Reference	Title	Page
1.1	The Mayo classification system of traumatic brain injury severity	17
2.1	Neuropsychological screen protocol, including protocol section, item name and function measured	37
3.1	Objects of the Rey-Osterrieth Complex Figure task, reproduced from Spreen & Strauss (2006)	54
3.2	Scoring of accuracy for the Rey-Osterrieth Complex Figure Task. Reproduced from Spreen & Strauss (2006).	54
3.3	Accuracy based scoring systems of the Rey-Osterrieth Complex Figure Task.	55
3.4	Drawing process scoring systems of the Rey-Osterrieth Complex Figure Task	57
3.5	Drawing process taxonomy review	70
3.6	Drawing process systems organised by construct	72
3.7	Objects included in Wholeness construct scoring	74
4.1	Descriptive data of participants by condition: sex, age, estimated IQ	86
4.2	Mean and standard deviations of RCF scores for each phase, split by condition (Normal versus Digital)	87
4.3	Usability, comfort and perception of performance using a digital pen (%)	90
4.4	Mean and standard deviation of age, estimated IQ and years of education by group	102
4.5	Mean and standard deviations of RCF scores for each phase, split by condition (in-person versus videoconference)	103
4.6	Mean, standard deviation and significance for each test, by condition (in-person versus videoconference)	107
4.7	Correlation coefficients for neuropsychological tests	109
4.8	Intraclass correlation coefficients between in-person and videoconference conditions	110
4.9	Unstandardized beta and significant values to test proportional bias between in-person and videoconference conditions	114
5.1	Percentage of elements drawn as whole units	127
5.2	Principal components analysis of RCF Strategy, with rotation variations	135
5.3	Principal components analysis with equimax rotation, restricted to four factors	136
5.4	RCF Strategy clusters with underlying components	136
6.1	Predicting variables included within regression analysis	153
6.2	Group mean and standard deviations for the RPCSQ, MFS, CFS and DEX	156
6.3	Mean, standard deviation, and independent t-test analysis for cognitive predictor variables, split by group	157
6.4	Mean, standard deviation, and independent t-test analysis for psychological predictors, split by group	157
6.5	Pearson's correlation coefficients and significance between outcome measures	158
6.6	Pearson's correlation coefficients and significance between outcome measures and psychological predictor variables	158
6.7	Pearson's correlation coefficients and significance between outcome measures and cognitive predictor variables	159
6.8	Pearson's correlation coefficients and significance between demographic and cognitive outcomes	160
6.9	Mean, standard deviation and independent t-test analysis for predictor variables, split by sex	161
6.10	Pearson's correlation coefficients and significance between psychological and cognitive variables	161
6.11	Pearson's correlation coefficients and significance between RCF measures	163
6.12	Multiple linear regression model with RPCSQ as the outcome measure, in order from most influential to least influential variable	166
6.13	Multiple linear regression model with MFS as the outcome measure, in order from most influential to least influential variable	167

Reference	Title	Page
6.14	Multiple linear regression model with MFS as the outcome measure, in order from most influential to least influential variable	168
6.15	Multiple linear regression model with CFS as the outcome measure, in order from most influential to least influential variable	169
7.1	Initial, final and sub-themes of thematic analysis	184
7.2	Percentage of experience and severity of post-concussive symptoms in individuals with mTBI at least three-months post injury (N=15)	197
7.3	Fisher's Exact test, examining proportional differences of recovery status by experience for individual RPCSQ symptoms	198

Index of Figures

Reference	Title	Page
3.1	The Rey-Osterrieth Complex Figure	52
3.2	Rey-Osterrieth Complex Figure with numbered elements. Image credit: Martin & Velay, 2010	54
3.3	Timeline and family tree of 19 scoring system methodologies of qualitative features of the drawing process for the Rey-Osterrieth Complex Figure Task	58
3.4	Score points of Continuation construct	77
4.1	Example of significant distortion. Image credit: Cherrier et al, 1999	82
4.2	Example of a haptic stylus device	83
4.3	The Neo Smartpen and Neo Paper	86
4.4	Camera and tripod set up	86
4.5	Scatterplot comparing copy accuracy scores of a normal versus digital pen. Line of perfect correlation indicated.	88
4.6	Bland-Altman plot for Copy phase (digital – normal)	91
4.7	Bland-Altman plot for Immediate recall phase (digital - normal)	91
4.8	Bland-Altman plot for Delayed recall phase (digital – normal)	91
4.9	Illustration of literature search of the term 'neuropsychology'	97
4.11	Scatterplot comparing accuracy scores of in-person versus videoconference conditions. Line of perfect correlation indicated.	104
4.12	Bland-Altman plot for Copy phase (in-person – videoconference)	106
4.13	Bland-Altman plot for Immediate recall phase (in-person – videoconference)	106
4.14	Bland-Altman plot for Delayed recall phase (in-person – videoconference)	106
4.15	Box and Whisker plot for Delayed HVLT recall	108
4.16	Box and Whisker plot for Stroop task	108
4.17	Box and Whisker plot for TMT-B	108
4.18	Bland-Altman plot for HVLT Learning	111
4.19	Bland-Altman plot for HVLT recognition	111
4.20	Bland-Altman plot for HVLT delayed recall	111
4.21	Bland-Altman plot for FAS verbal fluency	111
4.22	Bland-Altman plot for Animal semantic fluency	112
4.23	Bland-Altman plot for Letter-number sequencing	112
4.24	Bland-Altman plot for Category switching	112
4.25	Bland-Altman plot for Stroop	112
4.26	Bland-Altman plot for TMT-A	113
4.27	Bland-Altman plot for TMA-B	113
5.1	Normative Rey-Osterrieth Complex Figure copy accuracy scores	123
5.2	Normative Rey-Osterrieth Complex Figure immediate recall accuracy scores	124
5.3	Normative Rey-Osterrieth Complex Figure delayed recall accuracy scores	124
5.4	Normative scores for Adjusted RCF Wholeness	126
5.5	Mean and range of order position, in ranked order by mean, from earliest to latest	128
5.6	Normative scores for RCF Order (core elements only)	129
5.7	Frequency of elements completed within the first five completed elements	130
5.8	Normative scores for RCF Continuation	131
5.9	Points of continuation drawn through by all participants (blue) and by less than 50% of participants (red)	132
5.10	Normative scores for RCF Strategy	138
6.1	RPCSQ, MFS, CFS & DEX ranges, split by group	156

Acknowledgments

The proverb goes *"it takes a village to raise a child"*. The same could be said for a PhD thesis. Whilst my name stands alone on the cover of this work, it would not exist without the involvement and support of others. Firstly, I must acknowledge Jamie Enoch, Karoline Arnesen and Zainab Munshi, who assisted with the dual-scoring and coding of data within this thesis. It would also be unjust of me not to acknowledge my supervisors Dr Trudi Edginton (City, University of London) and Prof. Catherine Loveday (University of Westminster). The idea for examining this topic was a result of their clinical work with patients in a neurotrauma clinic. Whilst their academic guidance and support has been invaluable, their personal support has been equally important. It has been a challenging five years since we first met, particularly with the death of a parent and a global pandemic. But I hope that we can continue to work together for many years to come.

On my first day as a PhD candidate at City, one of the lecturer's top tips was to find your peer support group. They said that they will be some of the most influential people to the success of your PhD process. This is now the advice I give those at the beginning of their own journey. Without the following people, I honestly do not believe I would have got to the end: Laura Lennuyeux-Comnene, Gemma Ryan, Sarah Donald, Jamie Enoch (again), Dr Claire Melia, Dr Holly Walton, Dr Olly Robertson, Dr Candice Whittaker, Dr Nicola McGuire, Danielle Paddock, Jenny Terry, Dr Catherine Hitch, Elizabeth Deehan, Natasha Dalton, Chris Leech, Tanya Schrader, Dr Maddi Pownall, Chris Robus, Ben Butterworth, Dr Catherine Talbot and Dr Oli Clarke. I know I have missed someone from my PsyPAG family – if I have, I apologise profusely. A huge acknowledgement and thank you goes to my husband, Stuart Newman. He has supported my academic endeavours for the entirety of our relationship – mostly in the form of doughnuts, crunchies, Fanta and white wine, but also by proof reading everything I have ever submitted. He too knows more about TBI than the average person. It was not without significant risk that I left a successful 18-year career behind. Stuart – I hope you feel the risk has paid off! To my siblings and my Dad, thanks for keeping my ego in check. Please remind me of my words next time I suggest taking on another endurance challenge.

The most important acknowledgement goes to the individuals who volunteered their time to be part of this research. Much of this was during a time when things were very uncertain for us all. I am particularly thankful to the participants who took part knowing the toll it would take on their health, which might put them out of action for hours or even days afterwards. The final note of this acknowledgement goes to two very important women who sadly are not here to see this thesis become a tangible manifestation. Mum. Nan. Your encouragement, support and belief in what I could achieve stays with me, even if I can no longer hear you say it.

List of Abbreviations

Term	Abbreviation
Acquired brain injury	ABI
Analysis of variance	ANOVA
Becks Depression Inventory – II	BDI-II
Boston Qualitative Scoring System	BQSS
Cognitive Failures Scale	CFS
Computerised tomography	CT
Diffuse Axonal Injury	DAI
Dysexecutive Questionnaire	DEX
Diffuse Tensor Imaging	DTI
Edinburgh Handedness Inventory	EHI
Controlled Oral Word Association Test – FAS	FAS
Functional magnetic resonance imaging	fMRI
Glasgow Coma Scale	GCS
Household/Support bubble member	H/SM
Hospital Anxiety and Depression Scale – Anxiety/Depression	HADS A/D
Hopkins Verbal Learning Test	HVLT
Intraclass Correlation Coefficient	ICC
Loss of Consciousness	LOC
Mental Fatigue Scale	MFS
Magnetic resonance imaging	MRI
Mild traumatic brain injury	mTBI
Modified Taylor complex figure	MTCF
National Adult Reading Test	NART
Principal Axis Factoring	PAF
Principal component analysis	PCA
Post-concussive Syndrome	PCS
Post-traumatic amnesia	PTA
Rey-Osterrieth Complex Figure Test	RCF
Rivermead Post Concussion Symptoms Questionnaire	RPCSQ
Standard deviation	SD
Single Photon Emission Computed Tomography	SPECT
Thematic analysis	TA
Traumatic brain injury	TBI
Taylor Complex Figure	TCF
Trail Making Test - A/B	TMT A/B

Chapter 1 - Long-term complaints following mild traumatic brain injury – an introduction

1.1 Chapter premise

The following chapter introduces the topic of mild traumatic brain injury (mTBI) and the long-term complaints that can result from it. mTBI is an injury affecting approximately 1.4 million people annually in the UK. The consequences of mTBI can place significant pressure on public healthcare resources and social services. However, there has been a long-standing debate about the aetiology of long-term complaints following mTBI, bringing into question the authenticity and validity of these complaints. As a result, public perception of mTBI or concussion as an injury is that it is not long-lasting – it is 'just a concussion' (McKinlay et al, 2011; Merz et al, 2017). While it is recognised that most of those who experience an mTBI make a full recovery, those experiencing persisting symptoms can have a significant effect on their life. This chapter will first define a mTBI, addressing prevalence, causes and social outcomes. Secondly, it will discuss the symptoms of mTBI and consider the current understanding of the cognitive and pathological processes that underpin these symptoms. Finally, it will set out the rationale for the current research programme and address the research question of the thesis.

1.2 Acquired brain injury

An acquired brain injury (ABI) is damage to brain tissue that occurs during or following the birthing process (Elbaum, 2019). ABI can result from neurodegeneration, stroke, substance abuse, oxygen deprivation, infection, tumour, haemorrhage, traumatic event – though some argue that neurodegeneration is a separate category (Coetzer & Balchin, 2014). The cognitive, functional, and behavioural consequences of ABI vary depending on factors such as the location and nature of tissue damage and demographic, social and psychological factors (Evans, 2010, Ravindra & Hawryluk, 2020).

1.3 Traumatic brain injury

Traumatic brain injury: *"A physiological disruption to brain function due to a traumatic injury"* (American Congress of Rehabilitation Medicine, 1993)

Traumatic brain injury (TBI) is a sub-category of ABI where brain tissue is damaged because of an external force that causes disruption to the normal functioning of the brain. TBI may include alterations to consciousness, amnesia, neurological deficits or altered mental state at the time of the injury (Ravindra & Hawryluk, 2020). TBI is a blanket term that encompasses the mechanism, location, and severity of the injury – all of which play a role in the injury's development and potential outcomes.

Classification of the mechanisms of injury vary. Injury can result from an object entering the skull, and brain tissue, referred to as an open or penetrating wound. The classic case of Phineas Gage (Damasio et al, 1994), where a tamping iron pierced his skull and brain tissue, is an example of this type of injury. Blunt force or closed head TBI is an object hitting the head sufficiently to cause tissue damage to the brain, but the missile does not penetrate the skull. Brain damage resulting from the head being hit by an object would be a typical example. Injuries from these mechanisms often result in localised or focal lesions, where there is damage to or atrophy of the brain tissue within a specific area (Ravindra & Hawryluk, 2020).

Alternative mechanisms of TBI include the rapid movement of the neck and head, resulting in acceleration and deceleration forces being applied to the brain. These forces can either be lateral or rotational. The forces can shunt the brain into the hard skull bone, causing crushing and bruising to the brain tissue. Further, and more importantly, as will be discussed further on, the rapid movement can distort the brain's structure, leading to stretching and sheering of the neurons making up the brain structure (Chancellor et al, 2019; Romeau-Mejia et al, 2019). Similarly, blast injuries result from an external blast reverberating through the body, which may or may not cause the body to move. The reverberations cause widespread stretching, sheering and damage to the neurons, similar to acceleration-deceleration injuries. Such injuries are typically found in military service men and women. Acceleration-deceleration and blast injuries are more likely to result in diffuse injuries, with micro-lesions and metabolic disruption spread throughout the brain's structure rather than in a specific area.

Classification of injury is also considered regarding the location of lesions. Injury location implicates what functions are affected and to what extent. For example, right-hemisphere damage has been indicated to affect global-level processing, whereas left-hemisphere injury involves local-level processing (Akshoomoff, Fereleto, Doyle & Stiles, 2002). Further, considering the differences between localised versus global damage, focal injuries typically have specific deficits, whereas diffuse injuries are non-specific (e.g., Ariza et al., 2006). There may also be differences in the severity of focal versus diffuse injuries - focal injuries to the frontal regions have shown greater executive functioning impairment than diffuse injuries (Wallesch et al, 2001). TBI is also classified by severity; a commonly used system is the Mayo Classification System of TBI Severity (Malec et al., 2007). The system classifies injury based on length of loss of consciousness (LOC), length of post-traumatic amnesia (PTA), and Glasgow Coma Scale (GCS)¹ score (Table 1.1).

Table 1.1
The Mayo classification system of traumatic brain injury severity

Classification	Criteria		
	LOC	PTA	GCS
Severe	> 7 days	> 7 days	≤ 8
Moderate	30 minutes to 7 days	24 hours to 7 days	9 – 12
Mild	≤ 30 minutes	< 24 hours	13-15

1.4 Mild traumatic brain injury

Mild traumatic brain injury (mTBI) refers to a TBI that is at the less severe end of the injury scale. According to the Mayo Classification system (Table 1.1), individuals classified with an mTBI will report a LOC of fewer than 30 minutes, including where there was no LOC. PTA will have been experienced for less than 24 hours, including the reporting of no PTA. GCS score will be within the ranges 13-15,

¹ The GCS (Teasdale & Jennett, 1976) is a triage assessment that considers how alert and responsive an individual is during healthcare admission. Individuals are scored on aspects such as responsiveness to instructions, and scores range from 3 to 15, with a higher score indicating a greater level of consciousness. Whilst being a well-known and commonly used scale, evidence to support its predictive accuracy in terms of recovery and social outcomes is weak. Alternative models indicate that observation of LOC and PTA are more accurate - particularly the latter (Lezack et al, 2004). PTA is the inability to remember details of the injury, of themselves or form new memories following the injury.

indicating that the individual was close to or at full alertness at the assessment time. However, it is essential to note that GCS scores can vary between the point of injury and hospital admission. This variation may account for the limited reliability in GCS scores predicting recovery and social outcomes, as recorded assessment depends on the point at which GCS was recorded. mTBI may also be divided into complicated versus uncomplicated mTBI. A complicated mTBI is where there is evidence of skull fracture and, or observable lesions within the brain tissue (Veeramuthu et al, 2017)

1.5 Concussion

Concussion: “A complex pathophysiological process induced by biomechanical forces which typically results in the rapid onset of transient neurological dysfunction (McCrory et al, 2017)”

It is not possible to examine mTBI without talking about concussion. The term concussion has become controversial over recent years, with it being suggested that the term should no longer be used. Many argue that a concussion should be referred to as an mTBI (e.g., Lumba-Brown et al., 2018). The recent increased interest in sports based TBI has demonstrated the long-term implications of repeated sub-concussive events, highlighting the need to take 'concussions' seriously. (e.g., Rawlings et al, 2020). Some healthcare professionals posit that 'concussion' is preferable when speaking with patients as it suggests a more transient and resolvable condition than terms such as 'brain damage' or 'brain injury' (Sussman et al., 2018). However, those who prefer 'mTBI' argue that 'concussion' detracts from the potential long-term consequences and the need to allow the injury time to heal to avoid further, more severe injury. To an extent, the latter is correct. Research indicates that public perceptions of the term 'concussion' typically not do not consider it to be a brain injury. However, participants often acknowledged that a concussion could lead to long-term complications (McKinlay et al., 2011; Merz et al., 2017).

The similarities between the formal definitions of TBI and concussion supports the argument to refer to the latter as mTBI (see above and section 1.3). Further, UK healthcare guidelines from the National Institute for Health and Care Excellence (NICE) refer to a

concussion as a minor injury to the brain. Concussions are defined as *"The common result of a blow to the head or sudden deceleration usually causing an altered mental state, either temporary or prolonged. Physiological and/or anatomical disruption of connections between some nerve cells in the brain may occur."* (NICE, 2019). A similar position is taken by the US Centers for Disease Control (Lumba-Brown et al., 2018; Romeu-Meija et al, 2019). Therefore, for this thesis, concussion will be considered a mTBI. Thus, the term mTBI will be used to refer to the case of mTBI and concussion, except where the term concussion, or derivatives of, is used in formal clinical diagnoses and measurement tool titles.

1.6 Epidemiology and prognosis of mTBI

Globally, TBI is a significant health and social care issue. Whilst figures in the UK vary, estimates of attendance to emergency healthcare services for TBI are between 600,000 and 1.4 million people per year (Pandor et al, 2011; NICE, 2018). However, these figures are considered to be underestimates due to individuals not seeking healthcare support at the time of the injury. Of all TBI cases, 60-95% are mild (Roebuck-Spencer & Cernich, 2014; Brazinova et al, 2021). The most common cause of injury in Europe is road accidents, followed by falls, sports injuries, and assault, though evidence suggests falls will overtake road accidents in Western, affluent countries (Roebuck-Spencer & Cernich, 2014; Brazinova et al., 2021). Males form the majority of patients, whilst the most common age group are those aged 15 to 34 (Lefevre-Dognin et al., 2020).

The prognosis for mTBI is favourable, with most mTBI cases fully recovering within three months (Karr, Areshenkoff & Carcia-Barrera, 2014). However, a sub-set will continue to experience symptoms beyond these three months, sometimes for many years. Meta-analyses indicate that the frequency of persisting complaints ranges from 2% to 24% (Pertab, James & Bigler, 2009; Karr, Areshenkoff & Carcia-Barrera, 2014). Again, this figure is suggested to be an underestimation of how many experience symptoms for some time after their injury. Recovery indicators are typically based upon functional recovery, such as returning to work (e.g., Broshek, De Marcho & Freeman, 2015). Figures will not account for individuals who have returned to work despite still experiencing symptoms. The high percentage of elite sport-based incidents also challenges this figure. Research demonstrates underreporting of symptoms to return to the game as soon as possible (Ferdinand Pennock et al, 2020). Therefore, the prevalence of individuals living with post-concussive symptoms may be larger than official figures. In some cases, the persisting experience of symptoms may result in a post-concussive syndrome (PCS) diagnosis.

1.7 Pathological processes of mTBI

Evidence of brain cell death is typically found in moderate to severe TBI; however, this is rarely the case in mTBI (Romeau-Mejia, Giza & Goldman, 2019). Rather than evidence of lesions, research has indicated that mTBI typically results in neurological function disturbance (Chancellor, Franz, Minaeva & Goldstein, 2019). Neuropathological theories of mTBI are usually grounded in the metabolic cascade, triggered by axonal stretching, and shearing due to acceleration-deceleration forces.

Reticular, Centripetal, and Pontine cholinergic hypotheses all posit that damage to cells results from the ionic homeostatic changes triggered by the stretching and shearing of the axons (Chancellor et al., 2019; Romeau-Mejia et al., 2019). Excitatory neurotransmitters (such as glutamate) are activated by the injury and trigger potassium to leave cells, triggering depolarisation. To counter the depolarisation, an influx of sodium and calcium to the neuron occurs, which damages the cell and mitochondria (DeLellis, Kane & Katz, 2009; Romeau-Mejia, Giza & Goldman, 2019). The pontine cholinergic hypothesis posits that injury triggers the uncontrolled release of the neurotransmitter acetylcholine (McGuire, Ngwenya & McCullumsmith, 2017; Chancellor et al., 2019). ACh is involved in frontal cortex and hippocampus function and regulates sleep, attention, memory, cognitive processing, and executive function. The metabolic cascade is also known to cause neurofilaments and microtubule deformation, disrupting axonal message transporting (Romeu-Mejia et al., 2019). Synaptic plasticity is inhibited due to disruption to ligand-gated NMDA excitatory receptors and inhibitor GABA interneurons. Neuroinflammation is caused by the activation and upregulation of microglial cells and inflammatory cytokines. These move across an altered blood-brain barrier, resulting in further neuroinflammatory damage - though this process is not fully understood yet (Romeu-Mejia et al., 2019)

Several mechanical factors of the injury influence the downward impact of this metabolic cascade. Rotational acceleration has been associated with greater biomarker evidence of mTBI than linear acceleration (DeLellis, Kane & Katz, 2009). Explanations include the effects of rotation on the integrity between the grey and white brain matter. Variation in the density of the tissue results in different kinetic impacts, such that the adjacent tissues move differently from one another, exacerbating shearing and stretching (DeLellis, Kane & Katz, 2009). The magnitude of the force influences the biomechanical outcomes, which is captured within the centripetal hypothesis (Chancellor et al., 2019). Lower rotational force is argued to cause superficial damage, which resolves quickly. However, the greater force causes the kinetic energy to penetrate deeper into the issue,

causing more severe damage over a wider brain area. The spread of the force load across the tissue also plays a role; the same force magnitude spread across a broader impact area result in lower individual forces. (Chancellor et al., 2019). Impact velocity is also being examined to understand the influence of velocity versus force.

1.8 Post-concussion syndrome and the symptoms of mTBI

Post-concussion syndrome (PCS), also referred to as post-concussive syndrome or postconcussional syndrome, is the presence of concussive symptoms following mTBI (DeLellis, Kane & Katz, 2009). As discussed in section 1.6, it is accepted that most concussive symptoms resolve themselves within three months. However, PCS may be formally diagnosed if symptoms are experienced beyond this timeframe. In such circumstances, PCS may also be referred to as Persisting PCS to differentiate between symptoms experienced within the typical recovery period and those which become a chronic problem (DeLellis, Kane & Katz, 2009).² PCS is considered a syndrome as this indicates a collection of characteristic symptoms, which may be underpinned by differing neuropathological causes (Beaumont et al., 1999). Whilst there is variation in the terminology used to describe the symptoms of PCS, these typically fall into three categories: cognitive/fatigue, emotional/affective and somatic/physical symptoms (e.g., Potter et al., 2006; Hermann et al., 2009). One of the most well-known measures of PCS is the Rivermead Post Concussion Questionnaire (RPCSQ) (King et al., 1995). The RPCSQ lists 16 symptoms: fatigue, headaches; sleep disturbances; poor concentration, being irritable; taking longer to think; forgetfulness, feeling frustrated; dizziness, noise sensitivity; restlessness; feeling depressed; light sensitivity; nausea, blurred vision, and double vision.

PCS is often considered to be a heterogenous syndrome. This variation comes from several factors, including the position in the recovery journey, environmental and activity factors, and individual risk factors. Presentation of symptoms may be referred to as progressive, with some experienced primarily in the acute phases (e.g., nausea; blurred vision), whilst others present later or are more persisting (e.g., fatigue; poor memory) (Beaumont et al., 1999). Of all the symptoms, fatigue is considered the most common and persistent symptom following mTBI (Lannsjo et al., 2009; Roe et al., 2009; Moller et al., 2019). There is also evidence to suggest that some symptoms are more enduring and trait-based, while others are dynamic, state-like, and activity-dependent (Medvedev et al., 2019; Moller et al., 2019). These include fatigue, concentration, restlessness, irritability,

² For the purpose of this thesis, PCS will refer to the experience of symptoms beyond three-months post injury

headaches, and decision making (Medvedev et al, 2019; Moller et al, 2019). Noise sensitivity, nausea and sleep disturbances are considered more enduring and stable (Medvedev et al., 2019).

A number of biopsychosocial factors underpin the risk of experiencing persisting PCS: the severity of symptoms is the most consistent predictor of recovery (Iverson et al., 2017); poor mental health and co-morbidities such as ADHD increase the risk of PCS and delayed recovery. Sex-based differences also play a role; males are more likely to experience an mTBI. Conversely, evidence indicates females experience significantly more severe symptoms and take longer to recover (Gallagher et al, 2018; Bretzin et al, 2021). Explanations include the difference in neck circumference and strength³ and menstrual hormonal changes (Alsalaheen et al, 2019; Gallagher et al, 2019). Age at the time of injury is also an essential factor. Immature, unmyelinated axons are more vulnerable to damage and axonal swelling, thereby increasing the risk of persisting symptoms in children and adolescents (Iverson et al., 2017; Romeu-Meji et al., 2019). Older age is also a risk factor at the other end of the lifespan (King, 2014). Additionally, a previous history of mTBI also increases the risk of being diagnosed with PCS and the severity and duration of symptoms (Iverson et al., 2017). Finally, engagement in litigation and compensation-seeking cases are possible predictive factors of higher reporting of PCS symptoms (King & Kirkwilliam, 2011). Although these individuals likely have more severe and life-changing symptoms, they are seeking financial support.

One of the challenges of recognising PCS as a neurologically based condition is that many of the symptoms are expressed in other conditions. For example, in Multiple Sclerosis, fatigue and similar cognitive challenges are reported (Manjalay et al., 2019). Individuals with chronic pain have also scored equivalently on the RPCSQ as individuals with PCS (Smith-Seemiller et al., 2009). Further, a high frequency of PCS symptoms is reported within a healthy non-TBI group (Dean, O'Neill & Sterr, 2012). However, a reason for the heterogeneous nature of PCS and non-diagnosed persisting symptoms of mTBI is that there are sub-categories with different presentation and aetiologies rather than one umbrella syndrome (Lumba-Brown et al., 2019). Others argue that rather than sub-classifications, it is multi-dimensional, and each dimension should be examined and treated individually (Vos et al., 2019)

³ It is posited that the reduced neck girth and strength leads to greater acceleration-deceleration forces, which causes greater stretching and sheering, and metabolic disruption (Alsalaheen et al, 2019).

1.9 Outcomes of mTBI and PCS

Long-term symptoms of TBI are associated with unfavourable outcomes, although it is recognised that symptom severity predicts functional outcomes, with more severe symptoms associated with adverse outcomes (Ponsford, Draper & Schonberger, 2008; Mani et al., 2017). The effects of the mTBI may affect education, work, relationships, homelessness, and civil freedom. Whilst most individuals with persistent symptoms return to work, a systematic review and meta-analysis indicated that 12% do not return to work within 12 months (Bloom et al., 2018). Return to work is challenged by limited workplace accommodations, lack of support from healthcare providers and issues with work colleagues not understanding the nature of the injury (Graff et al., 2021). Paediatric and adolescent experience of TBI has been shown to affect academic attainment by as much as 60% compared to peers and 28% to siblings (Sariasalan et al., 2016). As much as 50% of relationships result in breakdown following TBI (Claude Blais & Boisvert, 2005). Explicitly focusing on mTBI, the long-term duration of symptoms is a risk factor for relationship dissatisfaction (Bannon et al., 2022). Research by Oddy et al. (2012) and Durand et al. (2017) suggests that TBI history is a risk factor for homelessness and incarceration.

History of TBI has also been found to be associated with poorer health outcomes in later life. Depression and other affect symptoms form part of PCS screening inventories (e.g., King et al., 1995; Lumba-Brown et al., 2019). Recent reports indicate as much as 44% increased reporting of anxiety and depression symptoms in individuals with mTBI compared to controls (Ashina et al., 2021). Unfortunately, experiencing mTBI has been indicated to increase the risk of death by suicide two-fold (Fralick et al., 2019). Experience of mTBI, particularly repeated injury, has also been shown to increase the risk of having a neurodegenerative condition later in life (Gardner & Yaffe, 2015), which has increased interest in the sports world recently (Newman & Dennis, 2021).

1.10 Neuroimaging of mTBI

One of the sub-classifications of mTBI is whether there is evidence of tissue damage, typically evidence through neuroimaging. Where evidence of damage exists, this may be referred to as complicated mTBI. Cases with no evidence of organic damage may be called uncomplicated mTBI. To capture tissue damage, a variety of options for neuroimaging now exist. Computed tomography (CT) captures static images of cross-sectional planes via x-ray (Wilde, Hunter & Bigler, 2014). Magnetic resonance imaging (MRI) measures the electromagnetic signals of hydrogen nuclei in the tissue cells. Hydrogen-based protons in

the tissue absorb electromagnetic waves and then reradiate these, which are captured in the imaging. Like CT, static snapshots of tissue intensity are provided to identify where there has been a tissue integrity loss. MRI is shown to have better sensitivity in detecting axonal injury than CT (Wilde, Hunter & Bigler, 2014). However, whilst MRI is considered the more sensitive tool, it has been questioned for its suitability in detecting tissue damage following mTBI. Umile et al. (2002) found that 75% of individuals with persistent PCS following mTBI had normal MRI, and CT scans at the time of injury. However, these participants later indicated cellular damage on positron emission tomography (PET) and single-photon emission computed tomography (SPECT) imaging. Techniques that capture functional processing, such as PET and SPECT, have demonstrated the sub-structural and metabolic changes characteristic of mTBI that are not necessarily captured by CT and MRI (Chancellor et al., 2019). Other such techniques include diffusion tensor imaging (DTI), functional MRI (fMRI), and magnetic resonance spectroscopy (MRS) (Wilde, Hunter & Bigler, 2014; Chancellor et al., 2019).

Despite the evidence that CT lacks sensitivity in capturing cortical damage following mTBI, current healthcare guidelines in both the UK and US are that CT is the imaging technique of choice (NICE, 2014; Wilde, Hunter & Bigler, 2014) – if undertaken at all. However, Umile's work highlights the challenges of neuroimaging in emergency care. As discussed previously, neurological disruption occurs due to a metabolic cascade, which, though relatively rapid, happens over time, sometimes over several days. Therefore, whilst other neuroimaging options are available that are more reliable in detecting neurological disruption, they are unsuitable for emergency care. This highlights the importance of follow-up care, including neuroimaging, following mTBI.

1.11 Cognitive deficits following mTBI

The debate as to whether there are long-term cognitive deficits following mTBI has been long-standing. A meta-analysis of studies examining cognitive deficits following mTBI argued that there were limited significant differences between an mTBI and a control group (Binder, Rohling & Larrabee, 1997). Small and clinically insignificant effect sizes were observed in cases where differences were found. These results were supported by later meta-analyses (Schretlen & Shapiro, 2003; Frenchman, Fox & Maybery, 2005; Belanger et al., 2005; Rohling et al., 2011). Where differences did exist, it was argued that these were the result of factors such as poor mental health, pre-existing low-level function, and malingering (Binder, Rohling & Larabee, 1997).

This argument was supported by a lack of evidence of organic damage in participants. However, these meta-analyses typically only examined uncomplicated mTBI cases; they did not dispute the evidence of cognitive change in complicated cases. However, as discussed in the previous section, this lack of evidence of organic damage may result from limitations in neuroimaging techniques rather than a lack of damage. This is illustrated clearly in the work of Veermuthu and colleagues (2016), who found that the uncomplicated mTBI group presented with persisting complaints rather than the mTBI with visible lesions. Whilst the complicated group may present with more severe neuropsychological symptoms in the acute phase than those with an uncomplicated mTBI; these are quicker to recover. The evidence from neuroimaging and the work of those such as Veermuthu highlight that lack of evidence on CT scans should not be taken as evidence of a lack of long-term consequences following mTBI.

The conclusion that mTBI does not result in long-term cognitive consequences has been further challenged, in that it is simply that the methodology and methods being used to investigate this are not sophisticated enough to capture this evidence. It is possible that mTBI is not considered to result in long-term cognitive consequences due to the methodology and methods used to investigate it. Proponents such as Pertab, James and Bigler (2009) and Iverson (2010) challenge the populations under investigation. These authors argue that the mTBI group in these studies treat the participants as homogenous, investigating individuals who have made a full recovery and those still experiencing symptoms collectively. Iverson (2010) evidenced with real-life and hypothetical data that differences are masked due to statistical mechanisms when sub-sets of individuals still experiencing symptoms and recovered participants are combined. One scoping review argued that 80% of individuals experience chronic cognitive deficits following an mTBI (McInness, Friesen, MacKenzie, Westwood & Boe, 2017). However, criticisms were raised regarding the methodology used to arrive at this conclusion, such that their estimation was an inflation of the true extent of cognitive deficit following mTBI (Iverson et al, 2019).

Further methodology critique concerns the design of studies and what data is being analysed. A significant body of studies evidenced a lack of deficits by comparing group differences cross-sectionally and on an absolute performance basis. However, examining group mean performance differences does not consider individual differences. Such an approach does not consider the starting point of individuals. Not considering of pre-morbid characteristics of the individual can risk reaching a false-positive or a false-negative result (Brooks, Holdnack & Iverson, 2011). For example, consider an individual who had a pre-morbid above-average performance. Their post-morbid performance in low-average and

average ranges could be regarded as unimpaired. However, this is relative to the population. That individual is experiencing reduced performance compared to their pre-morbid state. It is argued that cognitive change should be examined when a group differences approach is used (Iverson, 2010). This approach is used in sports-based literature, where pre-season baseline neuropsychological measures are taken (e.g., Collie et al., 2004; Moser et al., 2007). Although, in public healthcare, pre-morbid data is not typically available.

Where literature indicates a disruption to cognitive performance following mTBI, this is typically found in the executive functions; this is not surprising, given that neuropathological processes disrupt function more prominently in the frontal and temporal regions responsible for executive function. Commonly reported deficits are indicated in processing speed, attention, working memory, verbal and semantic fluency, and delayed recall (e.g., Frenchman, Fox & Maybery, 2005; Ponsford, Draper & Schonberger, 2008; Karr, Areshenkoff & Carcia-Barrera, 2014; Veeramuthu et al., 2016).

1.12 Sensitivity of neuropsychological measures – are we measuring the right thing?

A further challenge against the argument that mTBI does not result in long-term cognitive dysfunction concerns the neuropsychological tools that are used to measure cognitive function. It is recognised now that the behavioural and cognitive impact of mTBI is nuanced and dynamic, making them difficult to capture with classic neuropsychological tools. Therefore, tools that are more sensitive to the challenges of mTBI are required; this is captured in the following quote from Bigler and colleagues: *"neuropsychology's goal should be to refine its methods...not to continue arguing that permanent deficits do not occur from mTBI."* (Bigler et al., 2013).

One such study that demonstrates this is that of Geary et al. (2011). An examination of verbal memory indicated that individuals with mTBI had comparable verbal recall performance to healthy controls. But *how* they recalled the word list differed. Rather than using the semantic categories as a framework from which to recall individual items, as is typically the case with the control group, the mTBI group used either a serial or haphazard order. Whilst the absolute performance was comparable, the effort required to achieve this level of performance was greater. Evidence from Geary et al.'s work argues that we should be examining *how* an individual achieves a goal rather than only focusing on *what* they achieve. This theory correlates with evidence from neuropathology and neuroimaging that it is processing that has been disrupted rather than a complete loss of function. The study's

conclusions also support the theoretical fatigue model by Wylie & Flashman (2017), which argues that cognitive fatigue results from changes in effort and energy resourcing.

The Rey-Osterrieth Complex Figure Task (RCF) is a pen and paper-based task where individuals are asked to draw and recall a complex image. The RCF is a popular neuropsychological tool (Rabin, Barr & Burton, 2005; Rabin, Paolillio & Barr, 2016; Egeland et al, 2016) which indicates visuospatial function and visual recall. Researchers have used the task to observe an individual's drawing process to replicate the image. Typical drawing strategy reflects a global bias, with core elements drawn first (Rey-Osterrieth, 1941/44, as cited in Corwin & Bylsma, 1993). However, examination of drawing performance with clinical populations has demonstrated divergence from this typical model (e.g., Lang et al., 2016; Joliffe & Baron-Cohen, 1997). For example, individuals with a history of eating disorders and autism have been shown to adopt a featural and local bias, replicating the image with a piecemeal approach. This cognitive processing style has been argued to underpin behavioural functioning (e.g., Lopez et al., 2008).

As with the California Verbal Learning Task used by Geary et al. (2011), individuals with mTBI have been shown to have comparable visual recall performance to control groups (e.g., Zappala & Trexler, 1992). Yet, few studies examine drawing process within mTBI populations. Only one study was found that examines RCF strategy use by mTBI populations. Whilst the findings did not find differences between the mTBI and control group; the strategy was measured using a subjective, binary measurement of global or local bias (Zappala & Trexler, 1992). More sophisticated measures of drawing strategy are available, such as the Boston Qualitative Scoring System (Stern, et al, 1994) (BQSS), but which have not been used with an mTBI population. This highlights a gap in the literature about our understanding of performance on the RCF task following mTBI. Whilst visual recall may not be affected by mTBI, visual perception and problem-solving may be. Given the findings from Geary et al (2011), there is an indication that observing the drawing process an individual uses to replicate the RCF may support a diagnosis of PCS within neuropsychological assessment, and provide an explanation for persisting symptom experience following mTBI.

1.13 The research question

The research question for this thesis is: *'Does observing and measuring the drawing process used for the Rey-Osterrieth Complex Figure Task support the neuropsychological*

assessment for individuals with long-term symptoms following mTBI?' Based upon the findings of Geary et al., 2011, hypothesis one was that observing RCF drawing process constructs would support the diagnosis of PCS in individuals with persisting symptoms following mTBI. The second hypothesis was that drawing process scores would be associated with the symptoms of PCS and behavioural challenges commonly reported by this subset of individuals.

The thesis will seek to answer this question by addressing limitations of previous work that examines mTBI that have been discussed in this chapter. Three methodological questions underpinned the work conducted. These questions were: *Are we measuring the right thing? How do we measure it? How do we analyse it?* This work was achieved via a series of tasks that examined drawing performance in individuals without a history of mTBI and individuals with and without persisting symptoms following mTBI. The methodology and methods of these works will be discussed in the following chapters.

Chapter Two – Methodology & Methods

2.1 Chapter Premise

The following chapter lays out the methodology and methods used within the current thesis. The chapter discusses the cognitive neuropsychological mixed-methods approach that was used to address the research question. It provides details of participant inclusion and exclusion criteria, and why such criteria were put in place. Lastly, it provides details and an explanation of the neuropsychology assessment protocol that was used throughout the thesis.

2.2 Cognitive Neuropsychological Approach

The current research had an applied focus, which sought to develop and validate a clinical neuropsychological tool. The aim of the thesis was to present evidence that use of this tool supports clinical formulation for individuals experiencing long-term challenges following mTBI. As the thesis sought to understand behaviour following brain dysfunction, it made sense to adopt an investigatory approach that focuses on brain dysfunction. Whilst the cognitive neuroscience approach examines human behaviour of the healthy brain, cognitive neuropsychology focuses on the damaged brain and how this correlates with atypical function and behaviour (Evans, 2010). Therefore, the methods in this thesis were informed by the cognitive neuropsychological perspective that places individuals with brain dysfunction at its centre.

Cognitive neuropsychology as a research approach has been shown to address the knowledge and methodology gaps of other approaches. For example, neuroimaging can provide evidence of lesions underpinning behavioural dysfunction. However, as discussed in Chapter One, brain tissue damage following mTBI is not usually visible via CT (Umile et al, 2002; Wilde, Hunter & Bigler, 2014). In such cases, neuropsychology is argued to be a better indicator of pathological processes and localization of damage (Evans, 2010). Further, whilst neuroimaging may be able to provide information as to where the damage is, it provides little information about what problems the damage may be causing. As a practice, clinical neuropsychology identifies and triangulates the cognitive, emotional, and behavioural consequences of brain dysfunction (Evans, 2010). It seeks to both describe and understand the cause of the observable signs and symptoms of cognitive disruption. It takes account of the nature and extent of cognitive impairment, as well as the practical and

functional consequences of cognitive impairment – the impact on daily life, work, education, leisure, and social relationships.

Another advantage of the neuropsychological approach is that it takes into account individual differences (Evans, 2010). It recognises population variation, such that cognitive function is on a continuum of performance, rather than a functioning/impairment dichotomy. Individual performance is considered in terms of population norms, but more importantly on an individual relative basis. By this, normal performance within average ranges for one individual may be impaired performance for another. Demographics and premorbid status, such as IQ, educational performance, employment, and health provide a basis from which to assess relative change in function, rather than focusing only on absolute performance. Without considering the premorbid performance and the full cognitive profile, post-injury performance can be over or underestimated, particularly where no pre-morbid baseline data exists, as is usually the case.

Given the aims of this thesis, and the advantages that come from adopting a cognitive neuropsychological approach to investigate the outcomes of mTBI, the methods within this thesis were designed to reflect the principles of cognitive neuropsychology.

2.3 Research Design and Methods

The clinical neuropsychological assessment process acknowledges the patient's qualitative view of problems and administration of quantitative standardized tests as equally important evidence (Evans (2010). By its nature, neuropsychology reflects a mixed methods approach, considering both quantitative and qualitative evidence when investigating the nature of an individual's neurological condition. Therefore, a mixed methods research design was used within the current thesis. Through triangulation, a mixed methods design addresses the limitations of one method by use of the other, and vice versa, to give a more accurate picture of what is being investigated (Todd et al, 2004). Yardley and Bishop (2017) suggest that whilst quantitative data can be used to assess and measure change, qualitative data provides an in-depth understanding of the experience of those changes.

One of the challenges to our understanding of the long-term outcomes of mTBI is that there is a lack of evidence that triangulates the subjective reports from patients with cognitive and organic evidence. The current thesis approached this limitation to our understanding of mTBI by using quantitative neuropsychological tests alongside qualitative subjective reports from participants who have a history of mTBI. This was achieved by firstly

conducting a quantitative validation study of digital tools to record performance on the RCF task, which was conducted with healthy, non-TBI participants. Next, normative data of a novel scoring system for assessing the drawing process adopted when completing the RCF was collated. The normative data was produced from healthy, non-mTBI population sample. Whether the drawing constructs of the RCF supported a diagnosis of PCS was examined quantitatively with a population sample of both healthy, non-mTBI and mTBI participants. Finally, quantitative, and qualitative subjective reports from participants of the symptoms and challenges of their injury were examined via a mixed-methods study.

2.4 Participant recruitment

Participants in the non-TBI sample were recruited via social media advertisement, university SONA system, and post-graduate networking groups. Recruitment for the mTBI sample was done so via similar means, as well as via head injury charities and support groups. It was decided to recruit outside of a medical setting as the study was designed to examine performance of both individuals who have fully recovered from their mTBI, and those still experiencing symptoms. By recruiting only through a medical setting, this would restrict the available opportunity sample to those who were still reporting symptoms. All participants had to meet the inclusion/exclusion criteria (see 2.5). Eligibility was determined via an automated pre-screen, which included details of the study. Individuals who were not eligible to take part were advised by automated message that they were not eligible. Individuals who met the criteria and were happy to participate were invited to leave contact details. Upon leaving details, participants were contacted to arrange a time to take part in the study. Unfortunately, details of the number of non-TBI individuals who did not meet the criteria were lost. For the mTBI participants, 101 individuals completed the pre-screen. Of these, 17 met the criteria. Upon further discussion, it was decided that one participant did not meet the criteria of mTBI (see 2.5.3), and their participation in the study was closed.

2.5 Participant Inclusion & Exclusion criteria

Several inclusion and exclusion were included within the research design. These inclusion and exclusion criteria were applied to participants both within the healthy, non-TBI sample and the TBI sample.

2.5.1 Inclusion criteria

All participants were required to meet the following inclusion criteria:

Over 18 years of age

As the purpose of the research was to investigate the neuropsychological impact of mTBI on adults, only those over the age of 18 years were included within the current study. Research, such as that conducted by Waber and Holmes (1985), Akshoomoff et al (1995), and Anderson, Anderson and Garth (2001), indicates there is a developmental trajectory to the drawing process that is adopted by typically developing children. As children mature, they move from a haphazard, piecemeal approach, towards a strategic, hierarchical drawing process similar to a typically performing adult. As a result of this developmental trajectory and a need to reduce variables that may confound any performance patterns found, children were not included within the current study. No upper age limit was included within the research design. Whilst there is recognition that aging has an impact on neuropsychological performance (Boone et al, 1993, Cherrier et al, 1999, Gallagher & Burke, 2007), participants were age matched to accommodate for age related differences. Further, accelerated cognitive aging as a result of neurodegenerative conditions was accommodated for within the exclusion criteria.

English speaking

IQ has been found to be influential in a number of neuropsychological assessments, including the RCF (Boone et al, 1993; Gallagher & Burke, 2007). To include examination of this, estimated IQ was recorded via a standardised reading test that was administered in English. Further, verbal memory, verbal and semantic fluency and inhibitory control were all conducted via English language tasks. As such, participants were required to speak English as either a first or second language.

2.5.2 Exclusion criteria

Given the nature of several of the neuropsychological assessments that were used within protocol, it was essential that individuals were physically able to participate. To accommodate this, physical and sensory eligibility restrictions were put in place:

Vision impairment

Many of the neuropsychological assessments, including the RCF task, required individuals to be able to clearly view and read material either on printed material or on a screen. As such, individuals who had vision impairment that could not be corrected by spectacle or contact lenses were excluded from the research.

Colour blindness

As the task used to measure inhibitory control – the Stroop - requires perception of colour, individuals with colour blindness were excluded from the study.

Hearing impairment

All instructions for the protocol tasks were delivered verbally. Further, for some tasks the verbal stimuli could only be delivered once – information cannot be repeated if it has not been heard fully. To ensure that participants were not impeded in their performance due to hearing impairments, individuals with hearing impairments that could not be corrected with hearing aids were excluded from the current research.

Mobility difficulties

Several tasks, including the RCF, required participants to be able to hold a pen or pencil and draw without mobility restrictions. Individuals with mobility issues that affect their ability to write and draw, or who find holding a drawing instrument difficult were excluded. This was to ensure any differences observed within scores were from differences in cognitive function rather than mobility.

In addition to accounting for physical factors within the exclusion criteria, cognitive and neurological factors were also accounted for.

Diagnosis of self or immediate relative of Autistic Spectrum Condition

Understanding of the cognitive processing differences of Autism Spectrum Condition (ASC) shows that individuals with ASC approach the completion of the RCF task in a piecemeal style (Booth, 2006). Further, studies have shown that immediate relatives of individuals with a diagnosis of ASC, e.g., parents and siblings, are also found to complete

the RCF task with a piecemeal bias, regardless of ASC diagnosis (Happé et al, 2001). Therefore, individuals with a diagnosis of ASC, or with an immediate blood relative with a diagnosis, were excluded from either data collection point.

Eating disorder history of self, or immediate relative

Research indicates that individuals with a history of eating disorders perform atypically on qualitative measures of the RCF task (Lopez et al, 2008), such that a more piecemeal approach is adopted within the drawing process. Literature also shows that immediate relatives, who may or may not have a history of eating disorder, adopt a localised approach (Tenconi et al, 2010). Therefore, individuals with a history of eating disorder or who had an immediate blood relative with an eating disorder were excluded.

Diagnosis of neurological, developmental, or psychiatric condition

Cognitive differences are well documented across neurological, developmental, and psychiatric conditions. This includes deficits in executive functions such as organisation and planning, as indicated by the RCF. Dyslexia has been shown to be characterised by RCF drawing differences, such that there is greater holistic processing (Klicpera, 1983; Rajabpour Azizi, Akhavan Tafti & Mohsenpour, 2021). Attention Deficit (Hyperactivity) Disorder is associated with organization and planning deficits - though it is unclear the impact this has on performance on the RCF (Molitor et al, 2019; Syrjälä, 2020). Psychiatric conditions such as schizophrenia and obsessive-compulsive disorder have been linked to less accurate replications and taking longer to replicate the image (Heinrichs & Bury, 1991; Sullivan et al, 1992; Savage et al, 1999; & Vanborg et al, 2014). Though, it is argued that the anticholinergic medication prescribed for schizophrenia may be causing this, rather than the condition (Georgiou, Lamnisos & Giannakou, 2021). Neurodegenerative conditions with known cholinergic disruption, such as dementia, have been shown to have significant deficits in RCF replication (Cherrier et al, 1999; Martini et al, 2020). Due to the known effects of either the condition or pharmaceutical interventions to address symptoms, anyone with a diagnosis of a neurological, development or psychiatric condition was excluded from the study.

Diagnosis and Treatment for Cancer

Evidence demonstrates there can be long-term cognitive impairment following a cancer diagnosis (Hede, 2008, Moore, 2014). It is unclear whether the impairment is the

result of the cancer itself, or the variety of cancer treatments (Hede, 2008; Moore, 2014). Further, the length of time the impairment lasts for is variable across cognitive functions, though, higher order cognition impairments persist for as much as 20 years, such as executive function, planning, and metacognition typically persist (Hede, 2008; Moore, 2014). Therefore, individuals who have received a diagnosis of cancer were excluded from the study.

In addition to physical and cognitive factors that have been known to influence performance on neuropsychological assessment, two further exclusions were included to account for external influences on performance:

Neuropsychological assessment history

Research evidence indicates that neuropsychological instruments, including the RCF, can be affected by practice effects, such that performance improves through repeated administration (Loring, Martin & Meader, 1990; Mitrushbina & Satz, 1991; Levine et al, 2004; Calamia, Makon & Tranel, 2012; Yamashita, 2015;). Test-retest analysis typically indicates a retest period of 12 months to account for any practice effects (Basso, Bornstein & Laing, 1990). Therefore, any individuals who have undertaken a neuropsychological assessment within the previous 12 months were excluded from the study.

History of traumatic brain injury

As the aim of the thesis was to investigate long-term symptoms and challenges following mTBI, it was important that all participants within the healthy, non-TBI sample did not have a history of mTBI. Any non-mTBI individuals with a history of mTBI or concussion were excluded from the study.

Litigation proceedings for accident

Involvement within litigation proceedings has oft been posited as influencing the reporting of long-term post-concussion symptoms (Binder, Rohling & Larrabee, 1997; King & Kirkwilliam, 2011). To avoid doubt regarding participant malingering, individuals who were currently involved in legal proceedings regarding their accident were excluded from the study. This exclusion criterion was only necessary to apply to the recruitment of the TBI population sample.

2.5.3 Mild Traumatic Brain Injury Classification and Time Since Injury

For participants with mTBI, at least 3 months needed to have passed since injury occurred. This is in line with NICE health guidelines for referral to a neurotrauma team for persisting symptoms, and with existing literature that indicates the majority of mTBI patients make a full recovery within 3 months (NICE, 2019). Participants with mTBI were required to meet the criteria of a mTBI. The Mayo TBI Severity Classification model (Malec et al, 2007) was used, such that participants who experienced a TBI with LOC of more than 30 minutes and PTA of more than 24 hours were excluded from the study. As the population sample was recruited outside a clinical setting, healthcare records including GCS scores were not available. In addition, several participants reported not visiting healthcare services till some time had passed after their injury. Therefore, whilst GCS scores form part of the Mayo TBI classification system, GCS scores were not included as an inclusion/exclusion parameter within the current research.

2.6 Protocol

Throughout this thesis, the same protocol (Appendix One) was used for all participants. Standardised neuropsychological screening inventories that are quick to administer have been developed, such as the Addenbrooke's Cognitive Examination (Mathuranath et al, 2000) and the Montreal Cognitive Assessment (Nasreddine et al, 2005). However, the sensitivity of these scales is low, and not sufficient for the purpose of this thesis. Therefore, a more extensive protocol was required. The protocol within this thesis was modelled on a neuropsychological screen used within clinical practice by the project supervisory team. Used within an outpatient neurotrauma clinic at Addenbrooke's Cambridge University Hospital, a major trauma care centre within the UK, it has been found to provide a robust indication of an individual's cognitive and psychological profile (Edginton et al, in prep; Loveday et al, 2006, Edginton et al, 2010) following TBI. It provides coverage of the cognitive profile, assessing each of the core cognitions, and addressed both verbal and visual function (Table 2.1). To address the aims of the research question, additional items were added. This included inventories that measured subjective daily challenges and symptoms. To ensure consistency of effort and fatigue levels across participants, all task items were conducted in the same order by each participant.

Table 2.1*Neuropsychological screen protocol, including protocol section, item name and function measured*

Protocol section	Protocol item	Function measured
Cognitive	Hopkins Verbal Learning Task – Revised*	Verbal memory
	COWAT Verbal Fluency Task*	Verbal fluency
	COWAT Semantic Fluency*	Semantic fluency
	D-KEFS Category Switching *	Verbal semantic productivity; mental flexibility
	Trail Making Test*	Processing speed, mental flexibility, divided attention, perceptual tracking
	WAIS Letter-Number Sequencing*	Digit span, working memory
	Stroop Colour-Word task*	Selective attention, inhibitory control
	Rey-Osterrieth Complex Figure Task*	Visual memory, perceptual processing
Intelligence	National Adult Reading Task*	Estimated IQ
Psychological	Hospital Anxiety & Depression Scale*	Anxiety, depression
	Becks Depression Inventory – II*	Depression
Symptoms	Rivermead Post-concussion Symptom Questionnaire	Post-concussion symptoms
	Mental Fatigue Scale	Fatigue levels
Self-report daily challenges	Cognitive Failures Questionnaire	Behavioural and daily functioning challenges
	Dysexecutive Scale	Behavioural and daily functioning challenges
Patient History	Demographic, Health, and Lifestyle Questionnaire	Demographics, handedness, brain injury details**

Table 2.1. Research Protocol: protocol items by section, with function measured. * Protocol items included within the Edginton et al protocol. ** Sub-section of Patient History section only included within protocol with participants with a history of mTBI.

2.6.1 Estimated Intelligence

Estimated IQ is a standard measure within neuropsychological assessment as predictive statistical analysis indicates that for many clinical assessments, including the RCF, IQ is an influencing factor (Boone et al, 1993; Gallagher et al, 2007). Further, IQ is found to be a compensatory factor, such that those with a great IQ are demonstrated to better recover and compensate following mTBI (Oldenburgh et al, 2016; Fraser et al, 2019;). The National Adult Reading Test - 2 (NART) (Nelson & Wilson, 1991) is a verbal reading task that provides an estimate of pre-morbid intelligence. Intellect is considered to be resistant to neurological damage, including closed head injuries, normal and accelerated

cognitive aging (Strauss, Sherman & Spreen, 2006). Individuals are presented with a list of 50 non-phonetic English words. The greater the number of words pronounced correctly, the higher the estimate of premorbid IQ. The NART has high interrater reliability, and practice effects are considered to be of low clinical significance (Strauss, Sherman & Spreen, 2006).

2.6.2 Neuropsychological assessments

Details of each neuropsychological assessment in the protocol are given below, with the exception of the RCF, which will be discussed comprehensively in Chapter Three. Each assessment has published population normative scales, which are used widely within clinical neuropsychological practice.

Controlled Oral Word Association Test

The Controlled Oral Word Association Test (Benton, 1967) is a verbally administered assessment, whereby spontaneous production of words is evaluated. The test is formed of two sections. The first is known as word, letter or verbal fluency, or the FAS test. The individual is asked to say aloud as many words as they can think of beginning with a given letter within a 60 second window. Restrictions of eligible words include no proper nouns can be used, or words that have the same beginning e.g., run, runner, running. The individual is asked to do this with three different letters. The most commonly used trio of letters used is FAS, though other combinations may be used within clinical practice (Strauss, Sherman & Spreen, 2006). Total number of words across all three letters is recorded. The second section is known as category or semantic fluency. The individual is given a category and is asked to produce as many words as fit within that category within 60 seconds. The category used within the current thesis is 'Animals,' which is the most commonly used category (Strauss, Sherman & Spreen, 2006). Different breeds, gender or developmental stages are restricted e.g., duckling, molly, ewe, corgi. Total number of words produced is recorded.

The use of different fluency forms gives an indication of variation in lexicon retrieval function. It is suggested that verbal fluency relies on organisation and non-habitual searching strategies, whereas semantic fluency utilises hierarchical memory organisation such as category and sub-categories as searching tools (Strauss, Sherman & Spreen, 2006). Whilst mild effects of age are found within adults across both forms, education and IQ have been found to be significantly correlated with performance (Strauss, Sherman & Spreen, 2006). Test-retest reliability is considered to be high and interrater reliability is particularly good, with near-perfect correlations often reported.

Delis-Kaplan Executive Function System – Category switching

The Delis-Kaplan Executive Function System (Delis et al, 2001) is a compilation of nine tasks that assess executive function, which were derived from existing cognitive measures. The individual tasks can be administered as stand-alone assessments. Each task provides multiple scores that isolates individual aspects of cognitive executive function. The inclusion of category switching to the verbal fluency test was done so with the specific intention to make the test more sensitive to mild brain injury (Strauss, Sherman & Spreen, 2006). The individual is asked to produce as many words as possible within 60 seconds that fall within one of two categories, alternating between the categories. The categories used within the current research were Fruit and Furniture. The task is considered to have low test-retest validity and is vulnerable to short-term practice effects, therefore requiring the exclusion of participants who have completed recent neuropsychological assessments (Strauss, Sherman & Spreen, 2006). Studies within TBI populations have indicated that the test statistically differentiates between individuals with TBI and matched controls (Strong, Tiesma & Donders, 2010), however, it is not clear how sensitive the test is at differentiating between those classified in the mild ranges from matched controls (Anderson et al, 2017).

Hopkins Verbal Learning Task

The Hopkins Verbal Learning Task - Revised (Brandt & Benedict, 2001) is a 3-stage measure of verbal learning and memory, measuring learning, delayed recall, and recognition. In the learning phase, (HVLT-L) a list of 12 words within three semantic categories is read out, and the participant recalls aloud the list in any order. The list is read out twice more with the participant recalling the list after each time. Total learning is indicated by the number of items recalled across the three rounds. Delayed recall (HVLT-D) is the absolute recall rate of the list of words after 30 minutes. Discriminative recognition (HVLT-R) is examined immediately after delayed recall. The administrator reads aloud a second list of words which contain the original 12 words, plus six semantically related words and six semantically unrelated words. Discriminate recognition is measured by subtracting the number of incorrectly identified semantically related words from the number of correctly identified original words. The inclusion of the recognition phase isolates memory recall function from memory storage. The order within which words are recalled across the phases is also indicative of storage and recall processes, such as whether individuals use the semantic category framework to recall or not. Research has indicated that participants with mTBI perform equivalently to controls. However, they recall with less observation of the semantic categories (Geary et al, 2011). The HVLT is highly affected by age variation and to

an extent by education and IQ (Strauss, Sherman, & Spreen, 2006). As a result of practice effects, alternate versions of the list are available. In the current study, version one was used, with the semantic categories being precious gems, places of accommodation and animals.

Stroop Task

The Stroop task is a visual task which gives an indication of inhibitory control, selective attention, and cognitive flexibility. The version used within the current protocol follows the traditional colour-word interference format. Two lists of 112 items randomly ordered from the words Red, Green, Blue and Tan are presented to the participant. The Colour list is presented with ink colour that is congruent to the word printed. For the Colour-Word list, the ink colour and word are incongruent. The participant is first asked to read aloud the Colour list as quick as possible. They are then asked to complete the Colour-Word list, but to say aloud the colour of the ink, not the word. They are given two minutes to complete the task. Total number of correctly called colours is recorded. Age is a significant influencing factor on Stroop performance, though education and IQ are only considered to have a small effect (Strauss, Sherman & Spreen, 2006). The test is indicated to have good test-retest reliability; however, significant practice effects have been found. The colour-word effect has not found to be a discriminate measure for mTBI patients compared to controls (Strauss, Sherman & Spreen, 2006).

Trail Making Test

The Trail Making Test is a paper and pen-based task used to measure attention, processing speed and mental flexibility (Iowa Trail Making, n.d.). The individual is asked to connect 25 circled numbers in numerical order in the first phase (TMT-A) and 25 encircled numbers and letters in numerical/alphabetical order, alternating between the digit forms, in the second phase (TMT-B). Whilst both phases give an indication of perceptual tracking and processing speed, TMT-B also indicates mental flexibility and divided attention ability. Differences in clinical ranges across the two phases can be indicative of impairment in ability to switch attention between two tasks. Time taken to complete each stage is recorded in seconds. A practice stage is given in advance of each timed phase. Age, education, and IQ are all indicated to influence performance on the test (Straus, Sherman & Spreen, 2006). Test-retest reliability is considered to be adequate, with practice effects found in the short term, but not at 12-months. Interrater reliability is considered good. The test is considered to be sensitive to closed head injury severity, though results are varied within mild case

research (Straus, Sherman & Spreen, 2006). That said, research has indicated it to particularly useful in identifying cases of PCS from controls (Cicerone & Azulay, 2002).

Wechsler Adult Intelligence Scale - Letter-Number Sequencing

The Wechsler Adult Intelligence Scale (WAIS) Letter-Number Sequencing task (Schoenburg et al, 2007) examines digit span and working memory capacity. The administrator reads aloud to the participant a string of numbers and letters that are in a randomised order. The individual is asked to rearrange the characters so that all the numbers are given first in ascending order, followed by the letters in alphabetical order. The string of characters begins with two digits, with an extra character being added to the string after every third trial, till a total of nine characters. The test continues to such a point where the participant makes three errors in a row. The test is one of the few verbal subtests of the wider WAIS protocol that is susceptible to age differences. Education is also considered to be an influencing factor. The task is also considered to be sensitive to the severity of TBI (Strauss, Sherman & Spreen, 2006).

2.6.3 Mental Health Inventories

Standard neuropsychological assessment takes account of psychological variables. Depression and anxiety have been found to be associated with outcomes and recovery following TBI (Pineau, Marchand & Guay, 2015). Indications of mental health status were measured via self-report inventories, which were delivered to and completed by participants via Qualtrics. Copies of the measures can be found in the Appendix (3 and 4)

Hospital Anxiety and Depression Scale

The Hospital Anxiety and Depression Scale (HADS) (Zigmond & Snaith, 1983) is a 14-item inventory which provides a clinical indication of both anxiety and depression. The inventory asks participants to answer multiple choice questions based on experience over the previous fortnight. Each answer can score between 0 and 3 points. Seven items correspond to anxiety and seven items to depression, with a range of total score of 0 to 21 for each.

Becks Depression Inventory – second edition

The Becks Depression Inventory -II (BDI) (Beck, Steer & Brown, 1996) is a 19-item depression specific clinical scale. As with the HADS, the BDI-II asks participants to answer multiple choice questions based on the previous fortnight. Each answer can score between 0 and 3 points, with a range of total scores between 0 and 57. The BDI provides a more sensitive indication of what is troubling the individual. It also asks whether the individual has experienced any suicidal thoughts. As it is acknowledged that having the opportunity to verbalise suicidal thoughts can help the individual to work through these thoughts, contact details for mental health support organisations were built in-to the questionnaire.

2.6.4 Symptoms and behavioural challenges

In addition to the clinical measures of cognitive function and mental wellbeing, several other health and TBI focused inventories are included within the assessment battery. All Self-report inventories were completed via a Qualtrics survey by all participants in all population sample groups. Copies of each measure can be found in the Appendix (5-8)

Rivermead Post-Concussion Symptoms Questionnaire

Given the thesis seeks to provide evidence of clinical indicators of post-concussive symptoms, there is a need to measure and analyse the symptomology reported by participants. The Rivermead Post-Concussion Symptom Questionnaire (King et al, 1995) is a self-report 16-item inventory which assesses the subjective severity of post-concussion symptoms. Each symptom is rated on a Likert scale of 0 - 4, with zero representing not experienced at all, and four being a severe problem. The RPCSQ is widely used both within clinical practice and research and has been subject to many analyses regarding psychometric properties and usage (e.g., Potter et al, 2006; Dean et al, 2006; Barker-Collo et al, 2018; Barker-Collo et al, 2019; Lannsjo et al, 2021). The RPCSQ is administered within the current protocol as it provides both the number of and extent of symptoms experienced by individuals and has been used to define those individuals with mTBI who are considered to be asymptomatic, and those who are symptomatic. Individuals with mTBI were asked to complete the questionnaire relative to their pre-morbid state. Non-mTBI participants were asked to complete the questionnaire relative to subjective performance 12-months previously.

Mental Fatigue Scale

Fatigue is a commonly reported symptom following TBI, and of PCS (Lannsjo et al, 2009; Roe et al, 2009; Moller et al, 2019). Post-concussive fatigue is specific to cognitive fatigue, rather than physical fatigue or normal tiredness. The Mental Fatigue Scale (MFS) (Johansson & Ronnback, 2014) is a self-report multidimensional inventory. The scale was developed specifically for use within neurological examination practice, to measure symptoms of mental fatigue after neurological disorders or injuries. It contains 15 items that examine affective, cognitive, and sensory symptoms: sleep duration and variation of symptoms across the course of the day. Ratings addresses intensity, frequency, and duration, so as to develop a fuller description of the individuals experience of mental fatigue. Scoring is based on a Likert Scale between 0 and 3, with the ability to select half points (e.g., 0, 0.5, 1, 1.5 etc.). A cut-off score of clinical significance is suggested at 10.5. The MFS was chosen over other measure of fatigue such as the Fatigue Impact Scale (Fisk et al, 1994) as it specifically addresses mental and cognitive fatigue, rather than physical fatigue. Psychometric analysis of the MFS found sufficient internal consistency, with a Cronbach's Alpha of 0.944 (Johansson & Ronnback, 2014)

Cognitive Failures

Following TBI, individuals may report continuing cognitive failures, such as disruptions to memory, word finding or planning. The Cognitive Failures Questionnaire (CFQ) (Broadbent et al, 1982) is a 14-item multiple-choice scale, asking participants or patients to respond to the frequency with which they experience a list of 'cognitive failures,'. These include behavioural challenges such as entering a room and forgetting why they went in there. Each item is rated on a Likert scale between zero and four, with zero being no issue, and four being a significant challenge. Total scores can range from zero to 56. The CFQ has been indicated to have strong test-retest reliability, as well as good predictive validity (Bridger, Johnsen & Brasher, 2013). All participants, regardless of mTBI status, completed the CFQ.

Dysexecutive Scale

The Dysexecutive Scale (DEX) (Burgess, 1996) is similar to the CFQ, in that it measures subjective behavioural challenges associated with neurological conditions. The Dysexecutive scale looks specifically at behavioural challenges that are considered to be

underpinned by executive functions. The scale has 20 items, which are scored between zero and four, with a total score range of zero to 80. A higher score is indicative of more experience of dysexecutive challenges. The DEX has been indicated to be a sensitive measure of real-life outcomes following mTBI (Bar-Haim Erez et al, 2009),

2.7 Participant history

Reflective of the neuropsychological approach to patient assessment, demographic details of participants were recorded. Further, mTBI participants were asked to provide details of their injury, such as time since injury and case. All details were recorded at the start of the participant session, after the study briefing and consent being given, and before the cognitive assessment (Appendix Two).

2.7.1 Demographic details

The following demographic details of each participant were recorded during the history taking part of the participant session:

Age of participant

As discussed earlier in the chapter, age is indicated to be an influencing variable of RCF performance (Boone et al, 1993, Cherrier et al, 1999, Gallagher & Burke, 2007). It is also considered to be a predictive variable of epidemiology and prognosis of mTBI (King et al, 2014; Iverson et al, 2017; Romeu-Mejia et al, 2019; Lefevre-Dognin et al, 2020). All of the clinical neuropsychological tasks used within the current protocol are standardized accounting for age related differences. Therefore, age of participants was recorded.

Sex recorded at birth

As sex-based risk and recovery differences have been reported following mTBI (e.g., Sufrinko et al, 2017; Gallagher et al, 2018; Alsalaheen et al, 2019; Bretzin et al, 2021), it was considered important to understand the split of male to female within the current data set. Further, as examined in greater detail in Chapter 3 (section 3.9), there is evidence of sex-based differences in how an individual completes the RCF task.

Years of education/Highest level of education completed

There is evidence to suggest that the number of years in formal education (and therefore, indicative of highest level of education completed) is a predictive factor in how an individual performs on neuropsychological assessments, such as the RCF (Rosselli & Ardilla, 1991, Berry et al, 1994; Strauss, Sherman & Spreen, 2006; Tsaltali et al, 2020). Therefore, years of education and the highest level of education participants had completed was recorded.

Handedness

A study conducted by Rakhmanov and Dane (2020) within university athletes indicated that left-handed participants have superior recall performance on the RCF. Considering drawing processes adopted to complete the RCF task, there is some indication to suggest that handedness is indicative of hemispheric processing biases. Those who are left-handed are more likely to adopt a global, hierarchical approach, whereas those who are right-handed are more likely to adopt a featural approach (Weinstein, Kaplan, Casey & Hurwitz, 1990), though this was carried out only within women. Whilst evidence to support handedness differences is limited, handedness of participants was measured via the Edinburgh Handedness Inventory (EHI) across all participants, so as to account for this in analysis. The shortform EHI can be found in Appendix Two.

2.8 Traumatic Brain Injury History

Participants who had had an mTBI were asked additional questions relating to their injury. Details of how long had passed since the injury, how the injury occurred; and where the injury was to the head were recorded. Whilst participants had already been screened to ensure they met the criteria of a mTBI, they were asked for more specific details about loss of consciousness and post-traumatic amnesia. Details were recorded during the Patient/Participant History section of the protocol. Participants in the mTBI data collection were also given the option to give details of the effects of their injury in their own words. Participants were asked "If you had to describe to someone how your injury has affected your day-to-day life, how would you do this?" Responses were given in a free-format text box once all of the self-report inventories had been completed. The responses given in this section were analysed using Thematic Analysis, reported in Chapter Seven.

2.9 Variations in administration of protocol

Due to the aims and design of the studies conducted within this body of research, a number of administrative variations were made to how the protocol was delivered to the participants. In Chapter Four, the first study examined use of a digital pen versus a normal pen. Details of the variations to administration for this study are given within the methods section of the first study in Chapter Four. For the second study in Chapter Four, and the remainder of the studies within this thesis, data collection had to be adapted to account for social distancing guidelines as a result of the Covid-19 Pandemic. Details of these adaptations are given below:

2.9.1 Covid-19 Pandemic adaptations

In March 2020, the UK government instructed a national lockdown as a means to control the spread of a viral coronavirus known as Covid-19, which had been declared a pandemic by the World Health Organisation. Universities were instructed to close their campuses and move to online teaching. Only essential travel was to be made, and there was to be no non-essential face-to-face contact with anyone outside of your immediate household. Anyone who could work from home was instructed to do so. As a result, all face-to-face research was suspended. Restrictions on face-to-face research remained in place till late 2021. This had a significant impact on the methods used within the current thesis. Neuropsychology assessment is traditionally conducted in-person. As such, the design of the study protocol & administration methods to be used within this thesis reflected this. Participants were either to be seen within a testing lab on the University campus or within their home. When the lockdown came into effect, data collection with non-TBI participants was in progress. As it was no longer possible to continue doing this in-person, adaptations to the methods and protocol were required.

Despite its history as a face-to-face process, research investigations into the use of video-based teleneuropsychology have been undertaken prior to the pandemic (e.g., Cullum et al, 2014; Brearly et al, 2017; Marra et al., 2020; Harder et al, 2020;; Parks et al, 2021; Algret et al, 2021; Tailby et al, 2020; Chapman et al, 2021; Gnassounou et al, 2021). Whilst this data indicated feasibility of the use of videoconferencing in neuropsychological assessment, the literature examining validity and reliability was limited. The Covid-19 pandemic presented an opportunity to extend research in this area, which would also support the continuation of the current research as it was originally intended, albeit with some administration adaptations. As a result of the on-going social distancing rules,

participants within the mTBI data collection were also seen via videoconferencing. To accommodate the use of videoconferencing, and the physical separation of the test administrator from the participant, administration adaptations were made:

Test environment adaptations or instructions

At the start of each video-based study session, participants were asked to change the settings of the video-conferencing platform to 'Speaker View' so that they could see the researcher in the full screen. This meant that the experience of speaking via videoconferencing was closer to an in-person experience, rather than a split screen with themselves also on the screen. For participants seen in-person, the mental health and daily challenges inventories were all presented via the administrator's laptop. Under the video-conferencing conditions, participants were emailed links to the surveys and task and asked to complete as soon as possible after the video-session had finished.

Whilst a neuropsychological assessment in person is cognitively demanding, it is recognised that video conference adds an extra layer of demand. It is noted in research in practice that participants and patients are observed to tire quicker via video-based assessment (BPS, 2020). This was taken into consideration, and participants were advised at the start of each session that they could take a break if needed – which was reflective of instructions given to participants taking part in in-person sessions. At the mid-point of the protocol participants were asked if they needed to take a short break.

To be able to conduct the RCF and TMT tasks, it was necessary that the participant had a household or support bubble member present. In face-to-face neuropsychological assessment, whilst family members may be present in the patient history part of a session, it is usual that they wait outside of the appointment room whilst the assessments are being undertaken, so as to remove risk of interruption, embarrassment or assistance. During data collecting for this thesis, where possible household/support bubble members were asked to distance themselves from the participant and to not interrupt when not involved in these tasks.

Materials pack

To accommodate participation being conducted with the participant and administrator in separate locations, some materials were sent to participants in advance of the test

session. A materials pack was sent via Royal Mail to all participants who took part in the research via videoconferencing. Each pack contained:

- Three separate blank sheets of paper for the RCF task, marked with either a C, I or D for each phase.
- A copy of the TMT task trails
- A copy of the Stroop task lists
- Instructions for the supporting household/support bubble member
- A stamped addressed envelope to return all of the materials

Each of the RCF sheets, the TMT trails and Stroop lists were placed in individual, numbered envelopes, number one to five. Participants were told to not open these envelopes until instructed during the study session, so as to ensure no prior information regarding what they would be asked to do was given away. The household/support bubble member instruction sheet was placed inside an envelope marked "HMI." Participants were instructed to give the envelope to their household/support bubble member, who could open it before the meeting, but were not to look at it themselves. The household or support bubble member was also shown a short 5-minute video at the start of the study session to give further instructions as to what they would need to do. Funding to support the costs of mailing the materials packs to participants and the postage on the return envelopes was granted the British Psychological Society Cognitive Psychology Section and was received with thanks.

NART task adaption

When presenting a direct digital replication of the NART word list via screen, the font size of the target words became illegible. To address this, each word was added to a separate page of a PowerPoint document in the same font and size as appears on the physical stimulus sheet typically presented to participants seen in-person. Participants were instructed to read aloud each word as it appears or pass where they did not know the word. Errors were recorded by the test administrator as they would be in an in-person setting. Once participants correctly read a word, the administrator would progress the slides.

Trail Making Task

Copies of the four TMT sheets were sent to participants in advance of the study session, in a sealed, numbered envelope to be opened as instructed during the session.

The household/support bubble member was instructed in advance via the household/support bubble member instruction sheet and video to tell the participant when to start and to say aloud when they finished. They also observed the participant completing the task and would stop them and correct them if they went wrong, as the task administrator would within an in-person session. The researcher would time how long it took that participant to complete the tasks based on the stop/end announcements made by the household member. The completed sheets were sent back to the administrator via the return envelope.

RCF

For each phase, the participant was instructed to open the relevant number envelope and to place the blank sheet of paper in front of themselves in landscape orientation. The RCF image was added to a word document in the same dimensions as used in an in-person setting, which was then shared with the participant via the screensharing when the participant and household/support bubble member were ready to begin. As a smartphone was being used to record drawing process for participants in the normal pen condition in the first study (see section 4.2.6), the same method was used to record the drawing process when the protocol was conducted via videoconference. The household/support bubble member used a smartphone as directed to record the participant's drawing process. Recordings were using a secure file sharing platform after the session was finished. When the participant was happy, they had finished, they were instructed to fold the sheet in half when and place into the stamped-addressed envelope, so that it was not visible for later recall stages.

2.10 Participant incentives

Participants in the current study were not paid for their time to take part. In part, this was due to funding levels not being sufficient to accommodate payment that accounted for the estimated sample size needed to reach statistical power, and the time investment of each participant. Further, research into the payment of participants within clinical populations has indicated that participants are motivated to participate to supplement income (from employment or benefits), regardless of the risks involved (Govender, Naidoo & Naidoo, 2019), and particularly influences those at the lower end of the pay scale (Gyll, Spoth & Redmond, 2003). Such findings raise ethical and moral implications in paying participants to take part in research. However, studies have showed that survey participation & completion rates increase as incentive rate increases (Perez, Nie, Arden, Radhu & Ritvo, 2013), indicating that the greater a reward, the greater chance the study has of recruiting sufficient

participants to reach statistical power. Further, the nature of an incentive can influence the type of participants who volunteer, such as where a lottery-based reward is used versus a charitable donation (Hsieh & Kocielnik, 2016).

With such knowledge in mind, and the limit on funding available for the study, participant incentives were offered in the form of a lottery-based reward. Each data collection point had its own reward draw. For the non-TBI population, two prizes of £50 were awarded in the form of shopping vouchers. For the TBI population, funding for participant incentives was awarded by an external awarding body (PsyPAG, a member network of the British Psychological Society). Two prizes of £50 and four prizes of £25 were awarded. Prize winners could individually choose to receive their prize as a shopping voucher, or for a charitable donation to be made in their name to a brain injury charity. Participants were also asked if they wished to take part in the prize draw as part of the informed consent process.

2.11 Chapter Concluding remarks

The current thesis used a cognitive neuropsychology approach within a mixed methods design to address the research question so as to triangulate across cognitive measures, psychological and social factors and client reports, reflective of clinical neuropsychological formulation. Two data collection points were conducted, one with healthy, non-TBI participants, and the second with individuals who have experienced a mTBI. All participants completed a cognitive screen, a series of self-report inventories and supplied demographic details. The first set of studies examined the validity of new technology-based research methods, which could both inform future clinical practice, and validate the use of methods in later studies within the current research. The second study examined population data from a non-clinical group, so as to determine normative performance across existing methods and a new scoring method to measure drawing process used to complete the RCF. Neuropsychological assessment performance across a non-TBI control group and mTBI group were examined via regression analysis to examine which variables, including RCF constructs, best predicted post-concussive symptoms, and behavioural challenges. Finally, a mixed-methods study, using thematic analysis, examined the reported experiences of individuals who have previously had a mTBI, to understand themes in lived experiences of those with an mTBI, and how individuals experiencing long-term difficulties following their mTBI describe their experience. Whilst this chapter has addressed the methodology, methods and procedures used within this thesis, further details specific to each study are given within each chapter. Further, as the main focus of attention in this research, the RCF task has been addressed comprehensively in Chapter Three.

Chapter Three – The Rey-Osterrieth Complex Figure Task

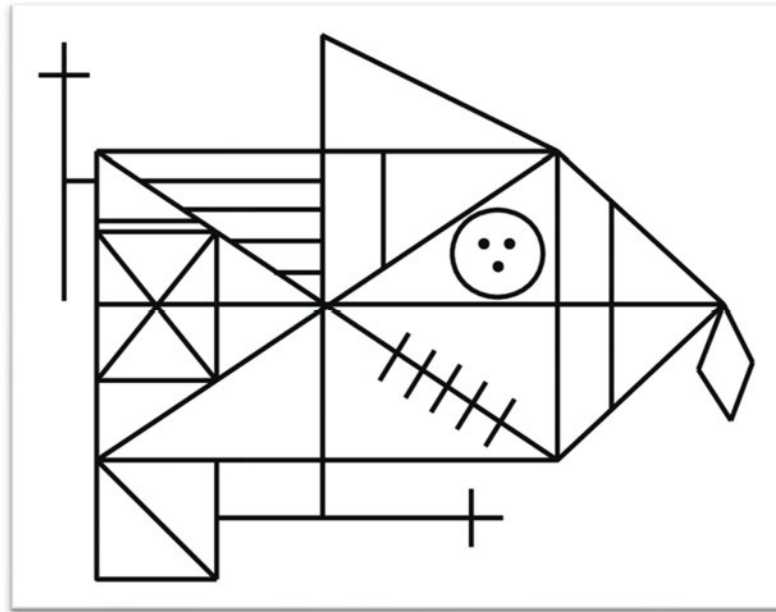
3.1 Chapter Premise

As the Rey-Osterrieth Complex Figure task (RCF) is central to this thesis, it is sole focus of this chapter. Firstly, a brief history and description of the RCF task is given, as it was designed by the original authors and the alternative options that have since been developed. The research that has informed how the RCF task has been administered in this thesis is presented, addressing factors such as the number of phases and time-points used, the use of alternate figures, and the influence of demographic factors. In the second half of the chapter, details of the method of measuring drawing process used in this thesis are given. This presents the challenge in choosing a suitable scoring system to use with a TBI population and explains how the four constructs of drawing process were developed.

3.2 The Rey-Osterrieth Complex Figure Task

The Rey-Osterrieth Complex Figure Task (RCF - figure 3.1) is a non-verbal drawing-based neuropsychological task (Rey, 1941; Osterrieth, 1944, as cited in Corwin & Bylsma, 1993). The RCF measures visual memory, visuospatial construction ability, and other executive functions such as planning, organisation and problem-solving (Strauss, Sherman & Spreen, 2006). Patients are asked to copy the complex figure as accurately as they can. The template and the patient's reproduction are removed, and they are asked to redraw the image from memory. The reproduction may then be scored for accuracy and the strategy used to complete the task. Due to the multiple indications of cognitive function it offers, the RCF is commonly used within neuropsychological assessment (Rabin, Barr and Burton, 2005; Rabin, Paolillio & Barr, 2016; Egeland et al, 2016). Given the prevalence of its use within neuropsychological practice already, and therefore the familiarity clinicians have with the task, developing the use of the RCF task would not add additional assessment time to an already somewhat lengthy process for the patient, or require significant learning of administering a new task for the clinician.

Figure 3.1
The Rey-Osterrieth Complex Figure



The RCF has also been argued to provide an indication of malingering (Meyer & Volbrecht, 1999). The individual must have significant knowledge of impaired patterns of performance to be able to complete the task (and the assessment as a whole) in such a way that their performance could be considered equivalent to that of an individual with impaired function (Meyer & Volbrecht, 1999). Even those with some understanding of neuropsychology struggled to convincingly mangle consistently throughout the task. As discussed earlier in this thesis, malingering and involvement in litigation following mTBI is a contentious issue. Given the findings from Meyer and Volbrecht indicate that even for mTBI cases the RCF has sufficient discriminative properties in identifying malingering cases, this further rationalises the inclusion of the RCF task within a neuropsychological screen or assessment given the amount of information that can be gained from a task that takes a relatively short time to complete.

Performance on the RCF task can be considered on two bases. Firstly, accuracy of reproduction can be assessed – *what* the individual draws. Secondly, how the individual goes about completing the task can be examined – *how* they draw. Rey highlighted the importance of this differentiation, positing that “*all the processes involved in making the copy*” need to be examined (Rey, 1941; cited in Corwin & Bylsma, 1993) as well as recording how accurate the ‘segments’ of the figure were reproduced. Rey recognised that

both aspects needed to be considered to determine the presence and nature of any underlying cognitive dysfunction.

3.3 The original system for measuring accuracy

To measure *what* is drawn, Rey identified 43 'segments' including four segments that consisted of multiple lines (the circle with 3 dots; the diamond; the four lines in the top left quadrant; and the 5 lines crossing the diagonal in the lower right quadrant) (Rey, 1941; cited in Corwin & Bysma, 1993). These four complex segments carried two potential points whilst all others carried one point. Where a segment was correctly replicated and positioned, points were awarded, with maximum overall score of 47. Rey provided a percentile-based scale for individuals over 15 years of age indicating what was considered poor recall performance.

Osterrieth considered accuracy assessment of the copy phase to be as important as the memory phase and argued that both should be scored in a systematic way (Osterrieth, 1944; cited in Corwin & Bysma, 1993). He also posited that accuracy should be measured based upon 18 'elements' rather than 43 'segments'. Osterrieth argued that due to the high level of distortion that could occur in replications it was difficult to assess which segment each pen stroke represented. Using eighteen elements simplified & sped up scoring and missing or distorted elements were easier to identify (Figure 3.2 & Table 3.1). Reproductions are scored on accuracy of placement and accuracy of reproduction of each of the 18 elements. Two points are awarded if the object is accurately placed and accurately replicated. If the object is accurately placed but is distorted in its reproduction, or vice versa, one point is awarded. If there is an indication that the object has been recalled, but is both distorted and placed incorrectly, half a point is awarded. If the object has been omitted, no points are awarded (see Table 3.2). A maximum of 36 points is awarded for each phase completed. Osterrieth's scoring system is reflected in the majority of standardised protocols and normative data scales that are utilised within neuropsychological assessment today⁴.

⁴ Many journal articles reference EM Taylor (1959). Taylor's text, as she appropriately references, is a translation of the Osterrieth system. The translation into English by Corwin & Bysma (1993) was not published till many years after EM Taylor's work was published. The referencing of EM Taylor rather than Osterrieth may be a reflection of what was available to authors at the time. This thesis will reference Osterrieth as the author of the accuracy scoring system, though it is recognised this is the same system as is referenced to EM Taylor in other articles.

Figure 3.2
Rey-Osterrieth Complex Figure with numbered elements (taken from Martin & Velay, 2012)

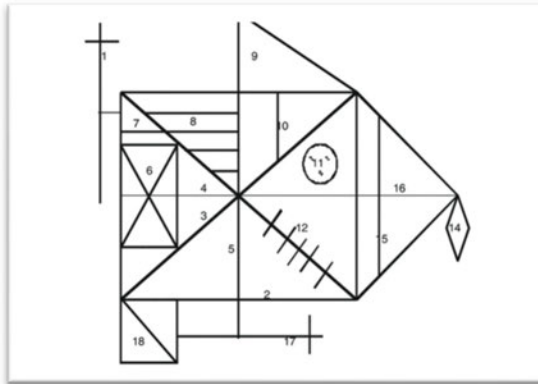


Table 3.1
Objects of the Rey-Osterrieth Complex Figure task, reproduced from Spreen & Strauss (2006)

Object	Details
1	Cross upper left corner, outside of rectangle
2	Large rectangle
3	Diagonal cross
4	Horizontal midline of 2
5	Vertical midline
6	Small rectangle, within 2 to the left [including internal cross]
7	Small segment above 6
8	Four parallel lines within 2, upper left
9	Triangle above 2 upper right
10	Small vertical line within 2, below 9
11	Circle with three dots within 2
12	Five parallel lines with 2 crossing 3, lower right
13	Sides of triangle attached to 2 on right
14	Diamond attached to 13
15	Vertical line within triangle 13 parallel to right vertical of 2
16	Horizontal line within 13, continuing 4 to right
17	Cross attached to low centre
18	Square attached to 2, lower left

Table 3.2
Scoring of accuracy for the Rey-Osterrieth Complex Figure Task.
Reproduced from Spreen & Strauss (2006)

Reproduction	Placement	Score
Correct	Placed properly	2
	Placed poorly	1
Distorted or incomplete but recognizable	Placed properly	1
	Placed poorly	½
Absent or not recognizable		0

3.4 Development of the Rey-Osterrieth system for observing drawing process

Rey also provided guidance on how to examine *how* people drew the figure – the drawing process an individual used to replicate the figure. Rey's emphasis was upon drawing what he referred to as the 'central armature' - the large rectangle, diagonals, and bisecting lines - as the first steps in the drawing process (Rey, 1941; cited in Corwin & Bylsma, 1993). This was called a centralised approach and departure from this was considered to be indicative of atypical function. Examples were provided of these inferior approaches; however, observations were on a descriptive basis rather than a quantitative measure. Little attention was given to the order the other details were drawn in. Osterrieth's second revision to Rey's work was to formulate an indexed system of measurement for

drawing process. Taking Rey's descriptive observations, an index of seven categories was presented. However, whilst this provided consistency in reporting the type of drawing process used and an indication of severity of departure from a centralised approach, assessment remained subjective.

3.5 Revisions to the Rey-Osterrieth scoring system

Given the limitations of the Rey-Osterrieth scoring system, it has been revisited by many times. A review of the literature identified 29 scoring systems for the RCF. Of these, 11 systems measured accuracy (Table 3.3) and 20 measured drawing process (Table 3.4). Two systems include a unique scoring system for both accuracy and drawing process.

Table 3.3
Accuracy based scoring systems of the Rey-Osterrieth Complex Figure Task.

Author	Year	Population
Andronikof-Sanglade, Fjellestad-Paulsen, Ricard-Malivoir & Evain-Brion	1997	Dwarfism, Children
Becker, Boller, Saxton & McGonigle-Gibson	1988	Alzheimer's Disease
Berry, Allen & Schmidt	1991	Geriatric
Duley, Wilkins, Hamby, Hopkins, Burwell & Barry	1993	HIV
Denman	1984	Epilepsy
Fastenau, Bennett & Denburg	1996	Adults
Loring, Martin & Meador	1990	ABI, Psychiatric
Meyers & Meyers	1992	Stroke
EM Taylor	1959	Children
Rey & Osterrieth*	1941/44	Obsessive compulsive disorder
Waber & Holmes*	1985	Children

* Methods also include a unique scoring methodology for drawing process.
Unless stated, is for use with adult populations.

Most systems for measuring accuracy are elaborations or extensions of the original scoring devised by Rey & Osterrieth and are upon accuracy of reproduction and placement of 18 elements. Some systems have fewer elements (Becker et al, 1988), and others increase the number of elements (Denman, 1984). Variations across the scoring systems are either to accommodate differences between clinical populations (Andronikof-Sanglade et al, 1997; Becker et al, 1988;), or to provide more precise measurements of accuracy (Berry et al, 1991; Duley et al, 1993; Fastenau et al, 1996). Others focus on the type and number of errors made, rather than what has been done correctly (Andronikof-Sanglade et al, 1997; Loring, Lee & Meador, 1990; Rapport et al, 1995). One system also includes a recognition phase (Meyers & Meyers, 1992), reflecting the various aspects of memory recall assessed within verbal recall tests, such as the HVLT.

For drawing process – often referred to as qualitative characteristics – several systems, such as the Gasparini et al (2008) or the BQSS (Stern et al, 1994), are extensions

to Osterrieth index system. These systems examine how strongly individuals adopt a configural approach to drawing. Other systems examine discrete aspects of the drawing process, such as order (e.g., Visser, 1972) or continuity (Waber & Holmes, 1985). Further, more recent systems are adaptations or extensions of earlier revisions (Figure 3.3). As with scoring systems for accuracy, there is a variation in the populations the systems have been developed for. Observations and baseline data from these studies need to be approached with caution as they may not be generalizable to an adult TBI population. There is also variation over whether the systems use categorical or continuous data.

Cultural validity of the RCF has also been examined. It is recognised that understanding or normative behaviour has been developed with English-speaking, educated American populations, which may not be appropriate for use globally. Work to produce culturally specific normative data typically uses the standardized system developed by Osterrieth (e.g., Rivera et al, 2015 (Latin America); Vogel et al, 2012 (Danish); Tremblay et al, 2015 (French-Canadian); Caffara et al, 2002 (Italy)).

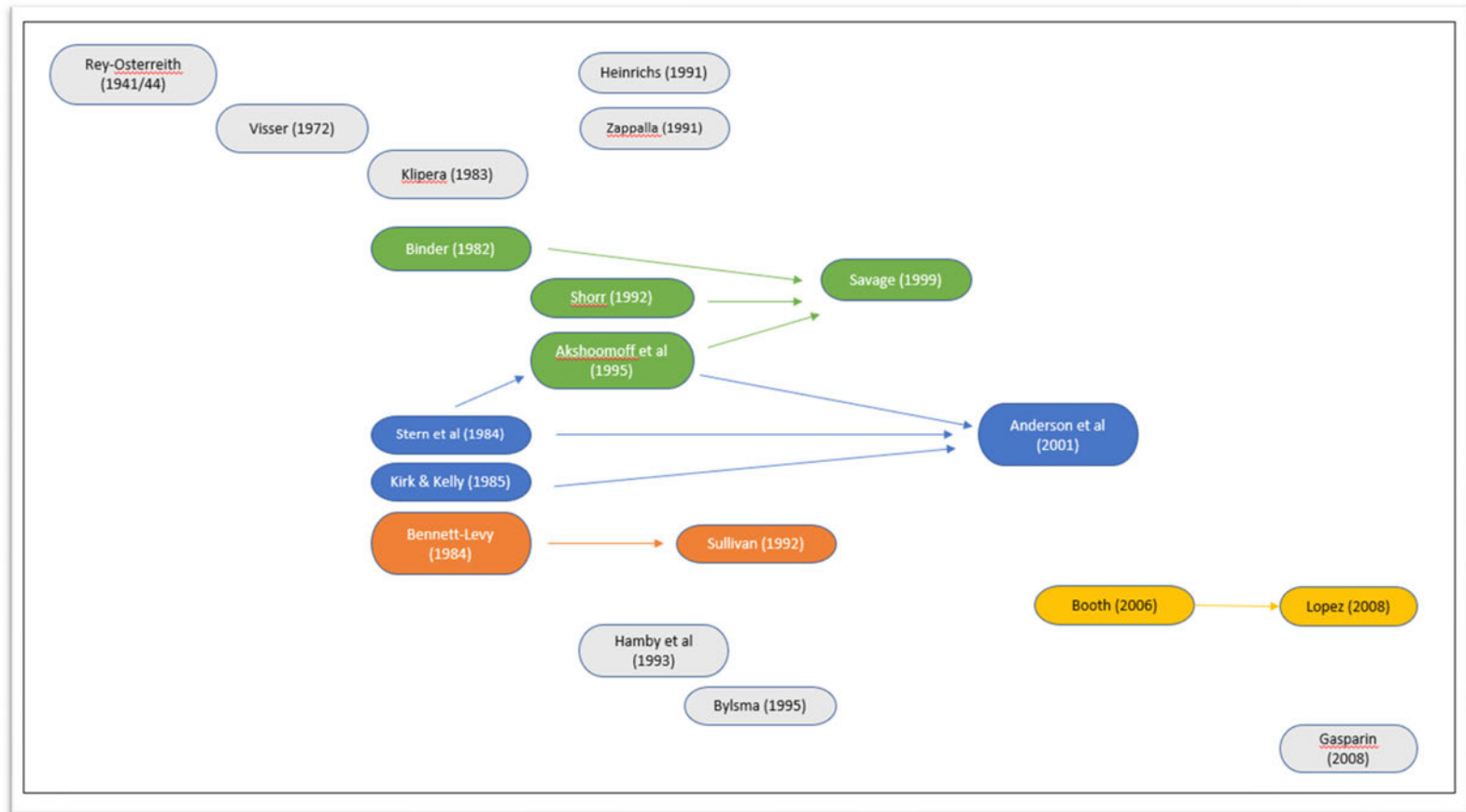
Table 3.4
Drawing process scoring systems of the Rey-Osterrieth Complex Figure Task

Author	Year	Population
Anderson, Anderson & Garth	2001	Children
Akshoomoff & Stiles	1995	Children
Bennet-Levy	1984	Adults
Binder	2008	Stroke
Booth	2006	Autism, Children
Bylsma, Carlson, Schretlen, Zonderman & Resnick (Q-system)	1997	Adults
Gasparini, Masciarelli, Vancore, Ottavani, Salati, Talarico, Lenzi & Bruno	2008	Dementia
Hamby, Wilkins & Barry	1993	HIV
Heinrichs & Bury	1991	Schizophrenia
Kirk & Kelly	1986	
Klipera	1983	Dyslexia, Children
Lopez, Tchanturia, Stahl, Booth, Holliday, Treasure	2008	Eating disorders
Rey & Osterrieth*	1941/44	Adults
Savage, Baer, Keuthen, Brown, Rauch & Jenike	1999	Obsessive Compulsive Disorder
Shorr, Delis & Massman	1992	Military Neuropsychiatric
Stern, Singer, Duke, Singer, Daughtery & Kaplan (BQSS)	1994	Various ABI
Sullivan, Mathalon, Nim Ha, Zipursky & Pfefferbaum	1992	Schizophrenia, Chronic Alcoholism
Visser	1972	ABI
Waber & Holmes*	1985	Children
Zappala & Trexler	1992	Mild head injury

* methods also include a unique scoring methodology for accuracy
n.b. unless stated, is for use with adult populations

Figure 3.3

Timeline and family tree of 19 scoring system methodologies of qualitative features of the drawing process for the Rey-Osterrieth Complex Figure Task.



Colours & arrows indicate 'family' connections.

3.6 Psychometric properties and selection of system

3.6.1 Inter-rater & intra-rater reliability of Osterrieth accuracy scoring

The RCF has been well examined for its psychometric properties. It is generally accepted that the copy and recall phases using the Osterrieth scoring system for accuracy has good to excellent interscorer reliability, with standard correlation coefficients typically above 0.8 (e.g. Carr & Lincoln, 1988; Loring, Martin & Meador, 1990; Liberman, Stewart, Seines & Gorden, 1994; Berry, Allen & Scmitt, 1994; Fastenau, Bennet & Denburg, 1996; Rapport, Charter, Dutra, Farchione & Kinglsey, 1997; Tupler, Welsh, Asare-Aboagye & Dawson, 2008). However, much of this research only uses Pearson or Spearman coefficients. Whilst these coefficients indicate a consistent linear relationship between raters, they do not indicate systematic difference or agreement. More rigorous analysis of interrater reliability employs intraclass correlation analyses (ICC). Liberman et al (1994) found good to excellent linear correlation, but also identified systematic differences by examining ICCs. Smith et al (2007) found low ICCs for copy & recall. However, Tupler et al (2008) indicated excellent ICCs between 0.93 and 0.96 and minimal rater bias.

Further, interrater reliability is not always consistent. Reliability has been found to be consistently lower for the copy phase, and the size of the individual differences are greater (Tupler et al, 2008; Liberman et al, 1994; Rapport et al, 1997; Berry et al, 1994). Explanations are that it is easy to determine the presence or absence of an element in the recall phases. However, accuracy judgement of placement and replication that is characteristic of the copy phase is more subjective. Inconsistency of reliability between elements has also been found, with the complex elements demonstrating more frequent and greater differences than simple elements (Tupler et al, 2008). Further, lack of consistency across the range of scores has been observed, with the lower scores demonstrating less reliability (Liberman et al, 1994).

Consistency of scoring by the same clinician is also an important consideration. Intra-rater reliability is not as well examined as inter-rater reliability. However, evidence indicates that there is good to excellent reliability, and is stronger than interrater reliability (Liberman et al, 1994; Tupler et al, 2008). Such findings indicate that whilst there may be judgement differences between clinicians, over time the clinicians are consistent with their scoring. To account for inter-rater bias, Liberman and colleagues (1994) recommend dual scoring of the RCF, and use of ICC analysis to determine the reliability of accuracy data, a recommendation also made by Charter et al (2000) and Smith et al (2007).

3.6.2 Inter-rater and intra-rater reliability of other accuracy scoring systems

The level of psychometric analysis of other accuracy scoring methodologies is more limited than for the original scoring system. What analysis has been conducted questions whether the adoption of a new scoring system is necessary (Fastenau et al, 1996). The Rey-Osterrieth accuracy system demonstrated better reliability and validity than the new system developed by Fastenau and colleagues. Further, the new system took longer to score. The Denman memory (1984) system has very precise criteria for scoring. However, whilst reliability was found to be stronger than the Rey-Osterrieth system, differences between the coefficients were not statistically or clinically significant (Rapport et al, 1997).

Loring et al (1990) argued if a system is too precise it may no longer be measuring non-verbal recall and visuoperception, but rather laziness or drawing ability. An important consideration when selecting which system to use is to determine the purpose for which it is being used. What it is you wish to measure will determine the complexity of the system that needs to be used. A more complex system that takes time to measure and is sensitive to small effect sizes may be necessary for within research. However, within clinical practice, a system that is quick to learn and apply may be preferable. Given the evidence supporting the reliability of the Rey-Osterrieth scoring system; the limited psychometric data for newer systems; and the question over what is actually being measured with more precise systems, the Rey-Osterrieth scoring system was used to measure copy accuracy and non-verbal recall in the current protocol. Further, all RCF reproductions were scored by two raters, and interrater reliability assessed using linear correlation coefficient analysis and ICC analysis.

3.6.3 Psychometric analyses of qualitative scoring systems

Psychometric data for qualitative scoring systems is also limited – if not more so. Anderson et al (2001) indicated good to excellent inter-rater reliability (0.85 – 0.92) and moderate to excellent intra-rater reliability (0.79 – 0.94). Hamby et al (1993) presented a kappa statistic of 87.6% for interrater reliability; and Sullivan et al (1992) indicated intra-rater reliability scores between 0.90 and 0.94. Deckersback et al (2000) compared their new system to that of Shorr et al (1992). They found moderate to high interrater agreement for organization of the individual elements, and excellent interrater reliability for the whole organization score. They also found equivalent interrater reliability for Shorr's organization score. Smith et al (2007) found a perfect ICC score for the Deckerbeck et al (2000) organization score. Not all systems, however, indicate consistently strong scores. Folbrecht et al (1999) presented inter-rater reliability scores for the Fragmentation and Planning

measures of the BQSS (Stern et al, 1994) between 0.59 and 0.90, and intra-rater scores between 0.76 and 0.83, except for one rater who had an intra-rater reliability score of 0.35 for fragmentation. The findings from this study are particularly interesting given the BQSS system is copyrighted and must be purchased for full details of their system yet has variable strength of psychometric properties.

One paper reported a comprehensive analysis across ten qualitative scoring systems (Troyer & Wishart, 1997). They found inter-rater reliability across the systems to be moderate to high, with coefficients between 0.62 and 0.98, all of which were statistically significant. Consistent with Folbrecht and colleagues' findings, Troyer and Wishart found the BQSS to have the lowest psychometric characteristics, and by some way compared to other systems. They argued a preference for the Q-system (Byslma, 1997, cited in Weider et al, 2016) or the Bennet-Levy (1984) scoring system, positing that the wide range of scores, measuring of multiple aspects and varying the weight across elements increased the reliability of the systems. However, as commented by Anderson et al (2001), it is not surprising that Troyer & Wishart concluded that no single system provided a sufficiently robust scoring system for all aspects that can be measured using the RCF, given the range of aims and focus of the measures. With this in mind, and that the only qualitative system specifically designed for use with a mTBI population found no statistically significant differences in performance when compared to a control group (Zappalla & Trexler, 1992), a more detailed review of each scoring system was required before determining how to score the drawing process within this thesis. This is presented in the second half of this chapter.

3.7 Copy and recall phases

3.7.1 Number of phases

The original protocol for administering the RCF stipulated carrying out a copy phase, then a recall phase after a 3-minute interlude (Rey, 1941; cited in Corwin & Bylsma, 1993). However, alternate protocols instruct that the early recall phase is carried out immediately, rather than after a 3-minute delay. A delayed recall phase has been included but, as with the early recall phase, the length of this delay varies between administration protocol, being anything from 15 minutes to beyond an hour (Strauss, Sherman & Spreen, 2006). Further, there is variation in whether both recall phases are used (Knight, Kaplan & Ireland, 1994). Given this variation, literature was examined to determine which phases to include, and at which time points.

Whilst the original Rey-Osterrieth protocol instructed to carry out the recall phase after a 3-minute break, other protocols such as those used by Loring, Martin & Meador (1990) use an immediate recall. Meyers and Meyers (1995), examined what, if any, material implications there are from varying the point at which the early recall phase is carried out. Their findings indicated that there was no difference on accuracy scores on a 30-minute delayed recall whether the recall was immediate (zero minutes) or 3-minutes. No other studies have been found that examine whether the early recall phase done immediately or after 3 minutes.

However, there is evidence to suggest that inclusion of an early recall phase influences delayed recall, with as much as a 6-point increase in scores (Loring, Martin & Meador, 1990). Conversely, Meyers and Meyers (1995) found no evidence of significant differences in delay recall scores. However, whilst their findings conflicted with Loring et al's, they highlighted that this difference may be due to variations between the scoring system used between the two studies. Yamashita (2015) addressed this using the traditional 36-point scoring that Loring et al used. Yamashita agreed with Loring et al that the inclusion of an immediate recall phase did result in a significant improvement in delayed recall scores – although the average difference found was smaller at 2.2 points.

3.7.2 Length of recall delay

With regards to the length of time of the delayed recall, Berry & Carpenter (1992) indicated that the length of delay for the delayed recall phase had no impact on accuracy scores whether carried out at 15, 30, 45 or 60 minutes. However, differences were found when the delay was over an hour. Whilst this is the only study that examines the impact of the time delay on delayed recall score for the RCF, Wood et al (1982; cited in Cermak, 2014) indicated that a similar examination had been carried out using the Taylor Complex Figure (TCF) – an alternate version to the original complex figure. Whilst only pilot data, no significant differences were found whether the delayed recall phase was carried out at 30 minutes or 60 minutes.

It is also important to consider practice effects over a longer period. Yamashita (2010) found that participants were able to recall elements of the figure one year after they initially completed the task. Importantly, he argued that whilst they may not explicitly remember the task, there was an indication of incidental learning. They recognised the procedure of being asked to draw something and then recalling it at a later stage. Therefore, they expected a recall element when repeat tested, even though they were not told this

would be the case. They may have employed encoding strategies whilst completing the copy task resulting in the greater recall. However, Mitrushina & Satz (1991) found no significant difference in copy scores when participants were tested annually - though there was within the recall phases. Levine et al (2004) found comparable results, such that whilst the copy phase was not found to suffer from long term practice effects, the recall phases did. Both studies further support the idea of incidental learning for participants who have may completed the task previously. In line with the evidence that practice effects become clinically insignificant after 12 months across, this was taken into account within the exclusion criteria of the study.

Given the evidence to support that use of two recall phases as an indication of non-verbal memory dysfunction and the variable evidence that an early recall phase influences delayed recall, the current protocol will include a copy, early recall, and delayed phase. To reflect the original Rey-Osterrieth protocol, the screen the protocol was modelled on (Edgington et al, in prep; Loveday et al, 2006; Edgington et al, 2010), and that no differences have been found whether early recall is conducted immediately or after three minutes, the early recall phase will be carried out after three-minutes. Finally, given that delayed recall scores were found to be equivalent as long as it is administered within 60 minutes, it will be carried out at the end of the protocol, approximately 30 minutes after the copy phase is carried out. Further, an exclusion criterion will be applied such that individuals who have completed a neuropsychological assessment within the last 12 months will not be eligible to take part.

3.8 Alternate figures and group design

Given that the empirical evidence discussed previously indicates that the RCF is sensitive to practice effects, the use of alternate complex figure versions was considered. Further, the use of alternate versions facilitates a within-participants design, thereby removing individual performance variance between comparative groups. Whilst there is not a Rey-Osterrieth authored alternative version, others have developed alternate versions such as the TCF (Taylor, 1969), the Modified TCF (Hubley, 1999), the Medical College of Georgia versions (MCG) (Meader et al, 1991), and Mack complex figure (Frazier et al, 2001). All are constructed with one central parallelogram and several smaller details in and around the central shape. The number of elements and scoring systems reflect the original RCF design, apart from the Mack which has 22 elements. However, whilst, these alternate versions are available, evidence indicates use of these alternate figures is not directly comparable to the RCF.

When measuring copy accuracy, evidence suggests the alternate figures could be used interchangeably as no significant differences were found between versions (Vingerhoets et al, 1998; Yamashita, 2006; Kuehn & Snow, 1992; Strauss & Spreen, 1990; Tombaugh & Hubley, 1991; Hubley & Jassal, 2007; Yamashita & Yasugi, 2008; Peirson & Jansen, 1997; Frazier et al, 2010; Meador et al, 1993). The comparability was found across healthy adults, older adults, healthy children, children with ADHD, neurological patients, and individuals with sleep disorders. However, some differences were found. Copy accuracy of the MTCF was not found to be comparable to the RCF, a difference that was borderline clinically meaningful (Hubley, 2010). Though these findings conflicted with previous studies looking at the MTCF (Hubley & Jassal, 2007; Hubley & Tremblay, 2002; Yamashita, 2006). A point to note is that most studies only examined significant differences, and not agreement. Only one study examined agreement (Frazier et al, 2010). However, generally, it is accepted that there is sufficient evidence to support the use of alternate figures when examining copy accuracy.

For recall, however, varying results have been found depending on the alternate figure being examined, and the population being considered. Recall scores for the TCF have repeatedly been found to be significantly higher than for the RCF (Vingerhoets, 1998; Yamashita, 2006; Spreen & Straus, 1990; Tombaugh & Hubley, 1991; Hamby et al, 1993; Pierson & Jansen, 1997). Explanations for this difference are that the TCF is constructed with less lines and more concrete shapes (Hubley & Tremblay, 2002), and that it is easier to verbalise its construction compared to the RCF, aiding its encoding and retrieval (Klipera, 1983; Vingerhoets et al, 2008). Kuehn & Snow (1992), however, indicated that whilst percentage change differed between the RCF and TCF, this disappeared when examining absolute scores. Further, evidence suggest differences across the lifespan. Sadeh et al (1996) found no significant differences between the two forms within healthy children, and children with ADHD. However, Gagnon et al (2003) found older adults were influenced by the order in which the alternate versions were presented. They suggested that the simpler construction of the TCF acts as a template for the RCF for older adults, enabling transference of learning to the more complex RCF. The reverse of this effect was not found nor did not appear in the younger adults. Investigations over a longer period indicated that rate of forgetting for the RCF and TCF was comparable, and as such could provide equivalent indicators of rate of forgetting (Tombaugh & Hubley, 1991).

Given the limited evidence to support the use of the TCF in direct comparison to the RCF for recall, the MTCF, MCG and Mack versions were created. Whilst overall the

literature supports that these versions are better substitutes or direct comparators for the RCF, the evidence is still conflicting. Overall, there is support for the use of the MTCF (Hubley & Tremblay, 2002; Hubley & Jassal, 2007; Hubley, 2010; Yamashita, 2006) and the Mack (Frazier et al, 2010). However, the MCG has been found to differ from the RCF (Meador et al, 1993; Yamashita & Ysugi, 2008). The advantage of the MCG, however, is that it has 4 versions, which were found to be comparable with each other.

However, even with an alternate version that could be used is direct comparison to the RCF, this only addresses direct memory recall. It does not address the implicit learning of the task more broadly, as was discussed by Yamashita (2010). Given the evidence leans towards arguing that the alternate figures are not suitable alternatives for the RCF – and particularly for drawing approach – the use of alternate figures is not appropriate within the first two studies that compare performance across administration methods. Further, given the previously discussed practice effects of the RCF, this directs the design of the first two studies towards a between participants' design. An interesting observation from across the literature is the variance between the use of within and between participant studies, with as many studies adopting either approach. As such, the use of the RCF only in a between participants design was considered the most appropriate approach for this thesis.

3.9 Demographic factors

3.9.1 Age

As with many neuropsychological measures, the performance on the RCF and alternate forms has been indicated to be affected by demographic characteristics of the individuals being assessed. Research has typically examined the influence of age, IQ, years of education and gender. How and to what extent these characteristics influence performance varies.

Age has been indicated to be a predictive factor for RCF accuracy performance (Boone, Lesser, Hill-Gutierrez, Berman, D'Elia, 1993). Developmental research indicates an improvement in both recall as children develop (Akshoomoff et al, 1995). At the other end of the life cycle, age strongly influences recall performance and drawing process adopted to complete the RCF task. Older adults demonstrate equivalent copy accuracy to younger adults but recall significantly less on recall phases (Gallagher & Burke, 2007; Gagnon et al, 2003). Correlation & regression analysis also indicates a strong relationship between recall accuracy and age (Boone et al, 1993; Vingerhoets et al 1998; Hubley, 2010; Hubley &

Jassal, 2007; Frazier et al, 2010). However, evidence is not conclusive, with some studies indicating no age-related differences in RCF accuracy (Janowsky & Thomas-Thrapp, 1993; Yamashita et al, 2015).

Research examining age effects on drawing performance is limited but differences across the life span have been indicated. Typically, younger children focus on localised detail, and struggle to process the hierarchical and gestalt characteristics of the figure (Akshoomoff et al, 1995). It has been suggested that this development trajectory is mirrored in older adulthood. Hartmann & Potter (1998) observed a trend towards a more disorganised approach in older adults, whilst Chiulli (1995) found that a greater number of older adults adopted a 'non-configural' approach versus a configural approach. However, Chiulli also observed that in general more adults adopted a non-configural approach than a configural one, regardless of age. Further, Janowsky & Thomson-Thrapp (1993) found no significant age-related differences in drawing of the rectangle, diagonal cross and horizontal and vertical bisectors; nor did Hartmann & Potter (1998) find significant age-related differences in the fragmentation score of the BQSS.

Hartmann & Potter's (1998) research indicates that age-related differences may be more nuanced than generalised lower performance; an observation that is also reflected when looking at age-related differences across the complex figure variants. For example, Gagnon et al (2003) found that whilst both younger and adults performed better on the TCF compared to the RCF, this improvement was greater in older adults, and was influenced by stimuli order. Culturally specific research suggests that age-related effects persist across diverse cultures and languages (e.g., Caffarra et al, 2002). Given the body of evidence that does exist, age of participants should be accounted for in any research using the RCF.

3.9.2 Intellect and education

An examination of IQ also indicates that it is an influencing factor on RCF performance. The majority of studies indicate that higher IQ correlates with higher RCF accuracy scores (e.g., Gallagher et al, 2007; Boone et al, 2001) – although these findings were not supported by Wilson & Batchelor (2015). As with age, it has also been suggested that IQ influence is not straightforward, and that it is not a linear relationship. Donnell et al (2011) argued that this positive linear relationship only stands within the average ranges of the normal distribution curve of performance. When IQ is at either tail-end, performance regresses back towards the population norm, having less influence on scores. Whilst not a direct comparison, but strongly connected, several studies consider length of time within

education – a factor that is often accounted for within population norms. Rosselli & Ardilla (1991) argue that education is a significant contributor as age. Berry et al (1994); Tsatali et al (2020) and Rosselli & Ardilla (1991) all found education to be associated with RCF performance; however, Wilson & Batchelor (2015) and Boone et al (2001) did not. A significant limitation to the evidence to support IQ and education effects is that all the above studies only examine copy accuracy and recall, and not drawing process.

3.9.3 Biological sex

Biological sex effects equally exert a complex influence on performance. Whilst many studies indicate that there are no effects of biological sex on accuracy scores of the participant (Luzzi et al, 2011; Tsaltali et al 2020; Boone et al, 2001; Berry et al, 1994; Frazier et al, 2010) others indicate that they do exist. However, where some studies indicate that men typically have higher copy and recall scores than women (Rosselli & Ardilla, 1991; Gallagher et al, 2007), others found that women had higher scores on other factors. For example, Peirson & Jansen (1997) found that women produced more organised figures on the copy phase. When considering alternate forms, further differences emerge. Men are typically found to be better on the RCF, and women on the TCF (Vingerhoets et al, 1998).

Given the trend towards evidence suggesting that age is a predictive factor for RCF performance, and the inconclusive evidence for IQ, years of education and biological sex of the individuals, demographic data will be collected from all participants. This data will be examined to determine via inferential and regression analysis to determine what if any of the factors influence RCF performance. Findings will underpin any recommendations for how these factors should be accounted for within clinical assessment and normative data development.

3.10 Development of how to score drawing performance for the RCF

The second half of this chapter moves its focus towards how to measure the drawing process an individual uses to complete the RCF task. A review of the literature indicated that no single system sufficiently measured discrete drawing constructs on a continuous basis. Therefore, a novel scoring system was developed. The remainder of this chapter discusses the findings from the literature review, and presents details of four discrete drawing constructs, and how they are measured.

Given that only one of the existing qualitative systems directly addresses mTBI (Zappalla & Trexler, 1992) and their findings indicate it is too blunt an instrument to discriminate between mTBI and non-TBI populations, another solution for this thesis was needed. With 20 possible options, it was important to determine which aspects of drawing process were to be measured, and which system – if any – provided the best solution. As such, an in-depth review was undertaken to examine what each system measured and how. As indicated in Table 3.4, there is variation in the populations the systems are designed for. Therefore, suggested indications of dysfunction given by the authors may not be transferable to an adult TBI population. Systems also differ in whether they provide wide ranging continuous data, which is better adapted to clinical normative scales, or a short range of either discrete or indexed categories, which may be more meaningful immediately during the assessment but typically lack sensitivity or ability to transform into normative data. Based on these factors alone, it is not possible to identify if one scoring system would be suitable for this thesis.

The language used to name, define, and describe these systems also serves to complicate matters. For example, the term ‘qualitative’ is used to describe scoring systems designed to provide detail other than accuracy of structure and placement objects. However, this term is used broadly, and not all systems that identify themselves as measuring ‘qualitative’ aspects evaluate drawing process. For example, as mentioned earlier in this chapter Loring, Lee & Meador (1988) developed a system that provided more context to the type of accuracy errors being made, rather than providing detail of the drawing process. An examination of the complexity of terminology and what each system purport to measure demonstrates this point. The journal articles for 19⁵ qualitative scoring systems were searched for the terminology used to describe what the system was measuring. Where words were derivatives of the same word, these were categorized under the same term e.g., organized, organization. Where a word is given in the article title, or specifically given as the name for what the system was measuring, these are indicated in dark blue. Where a word is used as the name of a second level unit of measurement, such as Fragmentation in the BQSS, this is indicated in light blue. All other words were used within the body of the article text. The results of this review are presented in Table 3.5.

⁵ It was not possible to obtain a copy of the Kirk & Kelly, Bylsma and Heinrichs & Bury scoring systems for review. Analysis of the Bylsma system was based upon description given within Weider, Indredavik, Lydersen & Hestad (2015). A review of the Stern et al system was from their 1994 paper, rather than the full scoring manual, which is only available upon purchase of the administration manual. Analysis of the Heinrichs & Bury system was made from the summary within their 1991 paper. In the absence of the Kirk & Kelly paper, it was only possible to include 19 systems within the system review.

As individuals with long-term symptoms following mTBI perform comparatively to health controls (Geary et al, 2011), and towards ceiling levels for copy accuracy, analysis of types of errors was not necessary. Several systems, such as Bennett-Levy (1984) considered symmetry of the figure as a qualitative aspect, however, as with error types, symmetry is a construct of accuracy of replication rather than specific to the drawing process being used. Terminology such as these were, therefore, not included within this review.

Fifty-two different words to describe aspects of drawing process were identified across the papers. Frequency of appearance ranged from appearing in a single article to appearing in 16 articles. The words 'Configural,' 'Continuation,' 'Organization,' 'Piece-meal,' 'Planning,' 'Process' and 'Strategy,' 'Fragmentation' and 'Detail' were used in at least half of the articles. The word or derivatives of the word 'Organization' was found in 16 of the 19 articles. This included seven articles using it within the title, and 3 further articles where it was used as the name of a unit of measure. However, it was found that terminology lacked clarity over what it measured. How terms were used did not always reflect the literal understanding of the word. For example, the word configuration means something that results from a particular arrangement of parts or components. Whilst 'configural approach' to the RCF was found to mean drawing in a particular order, a more accurate term to use might be 'holistic approach,' indicating a complete system rather than the analysis of dissection into parts. Further, the terminology was used to refer to different aspects of drawing. This was particularly the case for the most frequently used words, such as Organization or Continuation. For example, Booth (2006) and Waber & Holmes (1985) both refer to examining continuity. However, whilst Booth considers this to be the drawing all the lines of an element consecutively, Waber & Holmes look at points on the image where one line is bisected by another. Contrastingly, systems were identified to measure the same construct but used a different term to name it. Using Booth (2006) again as an example, where they examine whether elements are drawing in a piece-meal fashion in their Style Index, Stern et al (1994) refer to this as Fragmentation.

Table 3.5
Drawing process terminology review

System	Taxonomy																				
	Alignment	Approach	Armature	Arrangement	Chaining	Chunking	Clusters	Coding	Coherence	Conceptual	Configural	Contextual	Continuation	Combine	Consecutive	Construction	Detail	Featural	Fragmentation	Framework	Gestalt
Anderson et al	x									x	x		x						x	x	
Akshoomoff et al		x					x		x		x			x					x		x
Bennet-Levy		x											x			x					
Binder					x						x								x		
Booth		x							x		x		x			x	x	x	x		
Bylsma									x						x					x	
Gasparini et al									x		x							x	x		
Hamby et al											x		x				x			x	
Heinrichs et al						x						x	x					x			
Klipera	x	x		x				x			x		x				x				x
Lopez et al									x				x		x		x		x		x
Rey & Osterrieth		x	x										x			x	x				
Savage et al		x									x					x			x		
Shorr et al							x				x		x			x					
Stern et al		x					x				x						x		x		x
Sullivan et al						x							x			x			x		
Visser																			x		
Waber & Holmes	x										x		x				x				x
Zappala & Trexler							x														

Key:

	First-level descriptor
	Second level descriptor
	Appears in description

Continued overleaf.

Table 3.5 (cont.)
Drawing process terminology review

System	Taxonomy																			
	Intersection	Interruptions	Juncture	Local	Logic	Order	Organization	Parse	Part-oriented	Perseveration	Piece-meal	Planning	Problem-solving	Process	Progression	Qualitative	Segment	Sequence	Set shifting	Starting
Anderson et al				x	x		x			x	x		x	x	x		x	x	x	
Akshoomoff et al			x				x	x		x	x		x	x	x		x	x	x	
Bennet-Levy						x	x				x							x		
Binder							x				x	x					x	x		
Booth			x		x	x	x		x	x	x	x		x				x	x	
Bylsma					x	x	x				x			x		x			x	
Gasparini et al							x				x			x					x	
Hamby et al					x	x	x				x	x					x			
Heinrichs et al														x				x		
Klipera	x						x				x	x					x		x	
Lopez et al			x		x	x	x				x			x				x		
Rey & Osterrieth														x					x	
Savage et al							x				x							x	x	
Shorr et al			x				x				x			x						
Stern et al						x	x			x		x				x				
Sullivan et al							x				x							x		
Visser		x									x			x				x		
Waber & Holmes	x		x				x		x	x		x		x			x			
Zappala & Trexler							x			x						x			x	

Key:

	First-level descriptor
	Second level descriptor
x	Appears in description

A final challenge is that several of the systems provide a single score to measure several constructs. This consolidation of constructs was referred to by Troyer & Wishart (1997), where their recommendations indicated that stronger systems measured more than one aspect. The strongest example of this merging of constructs is where a system refers to Configural processing. There is a consensus across articles that a configural approach to drawing the RCF is that it is drawn with a global bias, with an emphasis upon core elements being drawn at the start, rather than a localised, piece-meal, or fragmented approach. Whilst systems contending to measure configural processing vary between a basic, binary categorical system as in the Zappala & Trexler (1992) system, to more complex, multi-level systems, common to all systems is the drawing of the base rectangle shape at the start of the drawing process. However, unpacking what is a configural approach indicates that it is not one single construct. Rather, it focuses on elements being drawn as wholes, and the order within which the elements are drawn in.

Given this complexity and ambiguity of terminology used to describe what each system is measuring, each of the 19 systems were examined in further detail, independent of the terms used in the title or construct/unit of measure names, to consider what it is that they measure and how. It was determined that four general drawing process constructs are measured across the systems, with several systems addressing more than one construct (Table 3.6). These four constructs were named *Wholeness*, *Order*, *Continuation* and *Strategy*.

Table 3.6
Drawing process systems organised by construct

Wholeness	Order	Continuation	Strategy
Rey-Osterrieth	Akshoomoff & Stiles	Bennett-Levy	Booth
Anderson	Anderson	Hamby	Anderson
Stern (BQSS)	Stern (BQSS)	Klipera	Stern (BQSS)
Binder	Booth	Shorr	Akshoomoff
Booth	Bylsma	Sullivan	Booth
Gasparini	Hamby	Waber & Holmes	Bylsma
Hamby	Heinrichs		Gasparini
Heinrich	Klipera		Hamby
Klipera	Lopez		Klipera
Lopez	Rey-Osterrieth		Lopez
Savage	Savage		Rey-Osterrieth
Sullivan	Sullivan		Visser
Zappala & Trexler	Visser		

3.11 Construct of Wholeness

The construct of *Wholeness* refers to whether each of the 18 elements are drawn as discrete objects, with each of the lines constructing its whole being drawn continuously or contiguously. Of the 19 systems examined, 14 examined the extent to which these elements were perceived and drawn as whole units of information. For example, systems typically consider the rectangle to be drawn as a whole when all four lines that make up the rectangle are drawn one after the other, without a break to draw part of another element. Going against this would include continuing the overall outline of the RCF to include the small square in the lower left-hand corner, and the smaller triangle above to top right hand quadrant.

Variations in the systems highlight several questions about how to measure wholeness. The first issue is how many of the systems do not measure Wholeness and Order as separate constructs (Anderson et al, 2001; Gasparini et al, 2008; Hamby et al, 1993; Heinrichs & Bury, 1991; Klipera, 1983; Rey-Osterrieth (1941/1944), Sullivan et al, 1992; Zappala & Trexler, 1992). Only seven of the 19 systems measure Wholeness as an individual construct. The second challenge is what to measure to indicate the extent of wholeness perception and individual observes. Most systems that measure Wholeness measure this as a binary concept of whether the element is drawn whole or not. Booth (2006) and Lopez (2008) addressed this by included gradient scoring within their Organizational Strategy. Shorr and colleagues (1992) examined how many of 'junctures' were drawn through, adjusted for how many elements were drawn, giving a maximum perceptual cluster count of 20. Waber & Holmes (1995) counted the corners of the larger rectangle as 'intersections' within their Style measure, to indicate the extent of fragmentation that occurred. Whilst Sullivan (1992) combined measuring Wholeness and Order, they included a measure of the extent to which wholeness and order were not adhered to within their strategy scoring, primarily focused on the level of wholeness.

The third challenge is to what extent across the RCF should wholeness be observed. Almost all systems only considered the core elements, and in particular the rectangle. Few systems included elements beyond the five core elements despite their similarity in complexity, such as the smaller rectangle with a cross inside. Further, there was variation in the number of core elements counted. Waber & Holmes (1995) included four core elements - the rectangle, the diagonal, the vertical and horizontal bisectors. Binder (1982) and Savage (1992) referred to all five core elements, though Binder grouped the rectangle and triangle structures together. Booth (2006), Lopez (2008), and Stern (1994) included the five core

elements, plus one additional element (Booth & Lopez, vertical in triangle; Stern, smaller rectangle). Shorr (1992) included the small rectangle and small square but considered the diagonals in the small rectangle as a separate element. Bylsma (cited in Wieder, 2016), included the most elements, leaving out the vertical and horizontal lines in the triangle, the single horizontal line in the top left quadrant, and the single vertical line in the top right quadrant.

Following a review of the literature, the following parameters were decided upon to measure the construct Wholeness:

- Wholeness is defined as measure of the extent to which the elements are drawn as whole units, separate from continuation or juncture points.
- All elements that can be broken in to more than one smaller section, including where the line is broken either by a juncture or a bisecting line, will be included within the analysis Therefore, 16 elements (Table 3.7) will be scored for Wholeness
- Wholeness score will be adjusted to account for how many of the 16 elements are drawn by the participant

Table 3.7
Elements included in Wholeness construct scoring

Element
Rectangle
Diagonal
Vertical
Horizontal
Triangle sides
Four lines
Small rectangle with diagonals
Small square
Small triangle
Circle with 3 dots
Five lines
Vertical in triangle
Horizontal in triangle
Diamond

3.12 Construct of Order

The construct of *Order* addresses the order with which an individual completes each element. Of the 19 systems included within the review, 12 were considered to measure Order. As discussed in the previous section, several systems combine observing order with

wholeness, including the original Rey-Osterrieth (1941/44) system. Stern et al (1994) provided an independent measure of wholeness within their Fragmentation score, however, their measure of Order interweaves both Wholeness and Strategy, giving stipulations such as the rectangle must be drawn first, with no fractioning. Five of the 12 systems were considered to give an independent measure of Order. The majority of the systems emphasise drawing of the core elements, and particularly the central rectangle, at the start of the drawing process. Zappalla & Trexler (1992) only considers whether the central rectangle is drawn at the start. Heinrichs & Bury (1991) proposed a similarly simple measurement, that a Configural approach was where core elements were drawn before any internal detail. Hamby et al (1993) considered any detail elements being drawn before the core elements to be a configural error, as did Klipera (1983). More complex scoring systems include Bylsma (as cited in Weider et al, 2015) addressed order by awarding points when certain lines of the figure are drawn at particular points. Several systems placed a greater importance on the central rectangle, such as that by Bylsma, who award 3 points to the rectangle, the diagonal cross and horizontal and vertical bisectors 2 points, and one to the remaining named elements.

As with the construct of wholeness, few systems consider the order within which the smaller detail elements are drawn in. However, some authors did. Booth (2006) and Lopez (2008) record the first 6 elements drawn – based on when drawing the object is initiated. Elements are categorised within four groups with varying weightings reflecting a configural hierarchy. Visser (1972) addressed not only the core elements, but detailed elements. A table of scoring is provided that gives specific instructions of how to award a point for specific elements being drawn at specific points before other elements.

Several systems looked at order on a basis of starting point and progression, such as Akshoomoff et al (1995) and Booth (2006). However, where Booth (2006) developed a continuous measure based upon the order of elements, Akshoomoff et al (1995) used four progression descriptions. Further, with the Akshoomoff category descriptions they combine looking at order with aspects such as fragmentation and how systematic their approach was. Savage et al (1999) also considered progression on a simple observation basis. It was to be noted that if an individual started with an internal detail rather than a core element, were the next two elements drawn core elements. This was to address the noted high percentage of participants who draw the vertical cross first before starting on the rectangle, a trend that was also observed by Hamby, Wilkins & Barry (1993), and Wilson & Batchelor (2015). Hamby and colleagues posited that this was an artefact of western writing, where written language is structured from left to right. Akshoomoff (1995) also noted that only 25% of

participants started with rectangle and took account of drawing the vertical cross as a first move within their scoring.

At which point the element is scored needs to be considered. Booth (2006) and Lopez (2008) are specific within their protocols that order is recorded based of the first line of the element that is drawn. Other systems are less specific, which brings with it an element of subjectivity as to when elements are drawn. By basing scoring on the point at which the element is first drawn, this assumes that the whole of the element has been perceived and is being completed. This is not the case when a more fragmented approach is taken. Further, the first line of an element may be drawn as a result of another element being drawn, rather than being explicitly decided to draw that element.

Considering previous literature, the following parameters were decided upon to measure the construct of Order:

- Order was examined based upon how many of the five core elements were drawn in the first five steps, and adjusted for how many were drawn overall
- All elements were examined to determine if the five core elements model of order is relevant
- Observation of order was based upon completion of the whole element

3.13 Construct of Continuation

The construct of *Continuation* refers to the continuous drawing of lines that are bisected by another line. For example, the top horizontal perimeter line of the large rectangle is bisected by the vertical central line within the rectangle. The concept of continuation is such that the line is perceived as a whole and is drawn continuously, and not stopping at the bisect. Continuation of lines is a term referred to in 11 of the 19 supporting articles and is a specific named construct within three of the systems. Klicpera (1983) measured points of continuity within the core elements of the rectangle and diagonal lines. Similarly, Waber & Holmes (1986) addressed continuity within their Style construct by recording whether sides of the rectangle and horizontal and vertical bisecting lines were drawn continuously. Whilst these systems only addressed some of the core elements, Bennett-Levy (1984) identified 17 points of 'good continuation' within the whole figure, plus one point of 'poor continuation' where typical performance is to break the drawing line

between the small triangle at the top of the figure and the large triangle to the right of the rectangle. Whilst there are substantially more bisecting points not included, these points were argued to be very rarely points of discontinuity. An advantage of the Bennett-Levy (1984) system is that it provides continuous data, which is better suited to developing clinical scales. As mentioned in the construct of Wholeness, 'junctures' and 'intersections' are also considered by some scoring systems (Waber & Holmes, 1986; Shorr et al, 1992).

Figure 3.4
Score points of Continuation construct

3.14 Construct of Strategy

The construct of Strategy refers to how the individual tackles replicating the RCF image, and whether they adopt a clear system within which to complete the RCF task. Booth (2006) acknowledged the importance of considering strategy, as they noted that drawing with a local, piecemeal bias does not necessarily correlate with a haphazard strategy – an individual can draw in a piecemeal style but in a very systematic way. Whilst 12 of the systems considered Strategy, none of the systems provided a systematic way of scoring for this construct. The extent to which the various systems considered Strategy also varied. Whilst being one of the earliest of the systems developed, strategy was most clearly addressed by Visser (1972), who noted that healthy adults typically reproduce the figure by observing the hierarchical structure, but also moving around the figure from left-to-right and top-to-bottom. Hamby et al (1993) also acknowledged the use of a systematic approach, that elements spatially near to one another were completed temporally near to each other. Moving through the figure left-to-right/right-to-left, anti/clockwise, and top-to-bottom/bottom-to-top were all considered acceptable strategies. However, these were recorded only as subjective observations rather than providing an objective means to score.

The lack of existing scoring systems to measure Strategy may be due to the fact that drawing strategy can have different characteristics, as Visser and Hamby acknowledge, depending on how you define a strategic approach. However, what is consistent between the existing scoring systems that discuss strategy to any degree is that individuals move between the elements in a systematic way, rather than haphazardly jumping around the image. To an extent, the construct of strategy could be considered an indication of efficiency in approaching the task – is a strategy adopted that has a clear plan to achieve the goal of completing the task in a quick, cost efficient, systematic way. Typically, a drawing process that has a strong strategy will have a strong temporal-spatial relationship. This is a consistent characteristic whether considering if someone draws systematically from one side to the next, anti/clockwise or from top to bottom, as all require the next object or line to be drawn to be next to the one that has just been drawn. Where this temporal-spatial relationship does not exist is when observing the hierarchical structure, such that the core elements are drawn first. It could be that there is a typical hierarchy within strategy, such that the global framework is drawn in first, then the local detail is drawn with a spatial-temporal relationship. To determine objectively if the strategy characteristics observed by previous authors do exist, it is necessary to develop a means to measure this. Given the consistency of the temporal-spatial relationship throughout the observations of strategy, this

relationship will be considered within a healthy population. This will be done by using factor analysis to examine order data, so as to identify consistent clusters within which elements are drawn together within the drawing process across a healthy adult population.

3.15 Concluding comments

The RCF task has a long and sustained history of use within clinical neuropsychology, supported with a significant body of literature demonstrating its psychometric properties. As a task that is a robust measure of visual memory, it is a viable candidate for extended use with a mTBI population to understand more about long-term cognitive implications. As with verbal memory tasks such as the California or Hopkins Verbal Learning tasks, it provides a means to look at not just what it reproduced, but how it is reproduced. Further, given its use within clinical assessment is significant already, its use as suggested within this thesis is unlikely to increase the effort and time requirement from patients – though further analysis will require more time investment from the administrator or clinician. However, whilst the Rey-Osterrieth system is identifiable as the preferred system for scoring accuracy and recall of the RCF – and will therefore be used within the current study - the multitude of ‘qualitative’ systems is such that it is not so easy to identify a single off-the-shelf solution for analysing drawing process.

It is important to acknowledge the length of time that has passed since each of the existing systems was developed. Even the most recent of systems developed by Booth (2006) and extended upon by Lopez (2008) are over 15 years old, with most existing systems being developed at least 20 years ago. This could be considered a strength, that these systems have been used extensively and have undergone several analyses for rigour and psychometric strength, as has been demonstrated within the chapter. However, these systems are based upon administration methods that pre-date more recent technological and digital advances. As acknowledged by Hamby et al (1991), one of the constraints to detailed analysis of the drawing process was a need to record the order of every pen stroke. Booth (2006) also acknowledged a potential need for a more detailed & easier to use recording process than of either the pen-switching or flowchart methodologies offered by video recording the drawing process of her participants. As recording technology has become more accessible and more advanced, this issue has potentially now been overcome, and will be examined within this thesis in Chapter Four. However, this does mean that the systems and normative scales developed decades ago with a less detailed record of drawing process may not be transferrable to using other recording methods – a

point highlighted by Troyer & Wishart (1997) in their own study. It suggests that these systems may no longer offer the maximum potential of analysis they previously did.

As has been discussed in this chapter, a scoring system that provides suitable, objective, and continuous data for four independently measured constructs is required in order to examine the drawing process of, firstly, healthy adults, and, secondly, individuals with long-term complaints following a mTBI. Data collected from a healthy adult population will form the base data from which performance of symptomatic and asymptomatic mTBI individuals will be considered against. Following the review of the existing literature examining the properties of the RCF task presented within this chapter, the RCF task will be carried out with the three time points of copy, 3-minute recall (immediate) and 30-minute recall (delayed). As the TCF task does not provide a suitable alternative to the RCF, where the same clinical population is being examined within different conditions, a between-groups design will be used within all participants drawing the RCF. To account for the known practice effects on the RCF, there will be an exclusion criterion such that anyone completing a neuropsychological assessment within the previous 12 months will not qualify to take part in any of the studies. Further, demographic details such as sex at birth, age, IQ, and education history will be taken and considered with data analysis.

Chapter Four: Use of digital technology in neuropsychological assessment

4.1 Chapter premise

One of the challenges with measuring drawing process is how to record what a person is drawing. Traditionally, RCF drawing process has been recorded via the use of coloured pens or the administrator mirroring what the individual is drawing on a blank template. However, both of these approaches come with limitations. As digital technology has become cheaper and more accessible, it offers a potential solution to accurately record how an individual draws. A number of digital recording solutions exist, which have been examined empirically. However, a digital pen that indicates the most promise requires further validation for use with the RCF task. The premise of this chapter is to evaluate the validity of using a digital pen to record drawing process of the RCF.

Secondly, the chapter evaluates the validity of conducting a neuropsychological screening protocol via a virtual meeting room. Whilst validation of the digital pen was always intended, the examination of video-based teleneuropsychology was a response to the Covid-19 pandemic. Teleneuropsychology provided a solution to working within social distancing guidelines brought in by the U.K. government. However, due to limited video-based teleneuropsychology validation research, further examination was warranted.

4.2 Use of a digital pen to record the drawing process used to complete the Rey-Osterrieth Complex Figure Task

4.2.1 Traditional methods of recording the drawing process

Once familiar, accuracy scoring of the RCF is quick enough that it can be done within the assessment session. It also lends itself to scoring away from the assessment room, as reproductions can be taken away. Scoring drawing process, however, is more complex. The scoring of multiple constructs and inter and intra-scorer reliability analyses all require a method where the drawing process can be revisited multiple times away from the assessment. Two methods have traditionally been used to address this.

The first method is the use of coloured pens. The individual is instructed to switch pen at set points, such as allotted time intervals, or a break in drawing pattern. The standard order of colours enables the drawing process to be followed. However, this process ‘chunks’

together drawing steps, with several individual pen strokes being in one colour. Detail and sensitivity is reduced, which more sophisticated scoring systems require. Additionally, the interruption to change pens may alert the individual to the monitoring of the drawing process or break their focus, or, potentially altering their drawing strategy (Ruffolo, Javorsky, Tremont, Westervelt & Stern, 2001). In contrast, it is argued that the use of the colours highlights drawing patterns, such as the temporal-spatial heuristic noted by Visser (1972) and Hamby et al (1993) as discussed in Chapter Three.

The second method is the flowchart or template method. As the individual draws, the test administrator records direction and order of pen strokes using arrows and numbers. However, this does not allow for large distortions from the template, such as those observed in individuals with dementia (Cherrier et al, 1999) (Figure 4.1). Whilst the method reduces participant interruption, keeping up with the participant is acknowledged to be challenging, and scoring can take longer compared to the coloured pens method (Ruffolo et al, 2001). A comparative study of the two methods indicated the pen-switching condition typically outperformed the flowchart condition (Ruffolo et al, 2001). However, overall differences were minimal and were clinically non-significant.

Figure 4.1
Example of significant distortion. Image credit: Cherrier et al, 1999



4.2.2 Digitisation of the administration of the RCF

Digital tools present a solution to the limitations of the analogue methods to record RCF drawing process. Digital platforms are already used within neuropsychological assessment, such as the Cambridge Neuropsychological Test Automated Battery

(Cambridge Cognition, 2019), Digitising the RCF has a number of potential benefits. It removes interruption risk and records in more detail. It allows the drawing process to be viewed repeatedly. It also offers the potential to automate scoring, reducing time commitments for scoring and clinician subjectivity. As technology has become cheaper and more accessible, several potential solutions now exist.

Sophisticated haptic-based systems (Figure 4.2) measure movement on a 3D basis, and can also record pauses, stutters, tremors, and pressure (Li et al, 2011; Kaber et al, 2014; Qin et al, 2015). However, the structure of the device with the stylus attached to an arm restricts natural movement (Li et al, 2011). Use of the system, whilst giving detailed information, does not feel like drawing with a normal pen and paper. Further, the cost of such systems is prohibitive to services and centres where resources are limited.

Figure 4.2
Example of a haptic stylus device



More economically viable options include graphic tablets that capture drawing process via a stylus on the tablet pad. What the individual draws on the pad is reflected to them on a screen. However, qualitative analysis indicated that the disconnect between motor movement and visual feedback was challenging for participants, even when a practice session took place (Riordan et al, 2013). Digital touchscreen tablets address this disconnect, with the digital drawing line appearing where the stylus is placed. Research has demonstrated the usability of touchscreen tablets with a digital stylus in neuropsychological research and clinical practice (Jurica et al, 2015). However, whilst there is a much closer motor input and visual output connect, the drawing position is slightly above the visual plane, with a slight temporal delay. It has been demonstrated that this disconnect between input and output – however slight – impacts both perceived and actual performance (Li, Clamann and Kaber, 2013). Attempts to overcome these disconnects placed a piece of paper over

the tablet and inserted a drawing cartridge into the stylus. This provided the input-output connect yet enabled digital recording of the drawing process (Smith & Hiller, 1996). Additional information such as time taken, hesitations, velocity, acceleration, and tremors could also be recorded. However, Smith and Hiller did not measure how participants found working with this set-up or examine how it compared to analogue methods.

This previous literature highlights the importance of ensuring user experience reflects traditional pen and paper, such that haptic, temporal, and spatial feedback is mirrored. This is a critical point to consider with developing digital equipment, where a key principle of interface development is that it should not require any user to do more 'work' through the use of the interface than is required to use an analogue version (Sonntag, 2018). This concept of transference of experience from traditional paper and pen to digital is critical. If differences in performance occur as a result of using digital means, then clinical standardized scales cannot be used to make a clinical assessment.

4.2.3 Empirical evidence for the use of digital pens in neuropsychological assessment

A neater solution to that suggested by Smith and Hillier is the use of a digital pen, such as the Neo Smartpen. Designed to look and feel like a standard pen, the pen writes with a ballpoint pen cartridge on normal paper printed with a digital matrix. A camera at the tip of the pen converts each pen stroke to a digital file. The file can be replayed via a smartphone or tablet app or downloaded to a computer. Costing around £100 for the pen, and the matrix free to download, this is an economically viable solution for both research and clinical settings.

Research examining the validity of using a digital pen within neuropsychological tasks, however, is limited. Sanada et al (2014) compared performance between the Wacom Inkling pen and the pen-switching method. Subjectively, participants found the Inkling pen as comfortable to use as the coloured felt-tip pens and was less distracting than switching pens. Objectively, differences between the methods were found - though direction of the differences was not consistent. Further, there were limitations in the data analysis methods used, as differences were assessed absolutely rather than statistically. Empirical research examining use of the Neo Smartpen has focused on the Clock face task (Prange & Sonntag, 2019) or Trail Making test (Prange et al, 2021), but not the RCF. Further the research assessed feasibility of automating scoring rather than validity of use compared to traditional pen and paper (Prange & Sonntag, 2019). Comfort and usability of the Neo Smartpen has been examined. Though, this was done by asking participants to hold the pen, but not to

draw with it (Sonntag et al, 2014). Whilst the currently available evidence provides support for the use of a digital pen, the evidence is extremely limited. As such, further validity analysis of the use of a digital pen whilst completing the RCF is needed.

4.2.4 Aims and Hypotheses

The aim of this study was to evaluate the validity of using the Neo Smartpen whilst completing the RCF task compared to accuracy performance using a standard pen and paper. It also evaluated perceived usability and comfort of the pen after it had been used by participants. It was hypothesised that RCF accuracy performance using the Neo Smartpen would be equivalent to performance using a standard pen in a healthy adult population. Further, it was hypothesised that users would find using the Neo Smartpen comparable to a normal pen and paper

4.2.5 Study design, group allocation and participants

A between-groups design was used due to the known rehearsal effects of the RCF (Loring et al, 1990; Levine et al, 2004, Yamashita, 2015)⁶. The two groups were a normal pen condition, and a digital pen condition. Participants were recruited through social media and research recruitment mailing lists. Snowballing recruitment techniques were also adopted. All participants were offered the opportunity to enter a prize draw to win one of two £50 Amazon vouchers. Twenty-seven participants were recruited, all of whom met the inclusion and exclusion criteria as detailed in Chapter Two. Group allocation was done so by alternating the conditions between participants to ensure equal allocation to each condition e.g., participant one used the digital pen; participant two used a normal pen, and so on. Thirteen participants were allocated to the normal pen condition, and 14 to the digital pen condition.

Twenty-two participants were female, and five were male. Whilst the total sample was uneven, there was even weighting between each condition. Left and right-handedness was even across the groups. No statistical differences in age, estimated IQ and years of education were found between the groups (Table 4.1).

⁶ See Chapter Three, section 3.8 for further discussion on the use of a between-group versus within-group design when examining RCF accuracy performance.

Table 4.1
Descriptive data of participants by condition: mean age; estimated IQ

Variable	Normal		Digital		<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Age	43.2	13.85	43.71	15.93	0.93
Estimated IQ	109	11.75	109	9.0	0.97
Years of educ.	15.00	2.55	15.07	2.81	0.61

4.2.6 Equipment and procedure

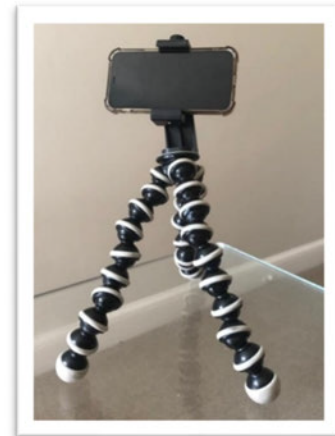
For the digital pen condition, a Neo Smartpen M1 and Neo Smartpaper was used (Figure 4.3). For the normal pen condition, a standard Bic ballpoint pen was used, with white A4 paper. To record the drawing process in the normal pen condition an iphone (generation 12) was attached to a Joby GorillaPod (Figure 4.4) and placed directly opposite the participant, focused only on the drawing page. In addition to the protocol and materials set out in Chapter 2, participants in the digital pen condition also completed a usability survey to understand user awareness and comfort (Appendix 9). The survey was administered via Qualtrics at the end of the protocol, to maintain consistency between the conditions.

All participants followed the protocol as set out in Chapter Two. This included a copy, immediate recall (3 minutes) and delayed recall (30 minutes) phase of the RCF task. All reproductions were scored for accuracy in accordance with EM Taylor's (1959) protocol.

Figure 4.3
The Neo Smartpen and NeoPaper



Figure 4.4
Camera and tripod set up



4.2.7 Results

The sample data was analysed via a 2x3 mixed methods ANOVA (between condition (2); within: phase (3) to examine if there were any statistical differences between the groups. Further, agreement between conditions was analysed via correlation coefficients, intra-class correlations and Bland-Altman plots.

4.2.7.1 ANOVA analysis

The data was reviewed to determine if the assumptions of ANOVA were met. Data was continuous and independent. Influential outliers were considered by examining the studentized residuals. One participant was found to have studentized residuals more than 3 standard deviations below the mean for the copy phase, and 2 standard deviations for the immediate recall phase. As being greater than 2 standard deviations below the mean is considered to be within the very poor/clinical boundaries, this participant was removed from the rest of the analysis. Normality of residuals was examined via visual inspection of Q-Q plots of the studentized residuals for each phase. All phases were considered to be close to normal distribution. Equivalence of variance was examined between the conditions. Whilst the recall phases were found to have equal variance, for the copy phase variance was not statistically equal ($p = 0.04$). Mauchly's test of Sphericity indicated that the assumption of sphericity was violated. As a result, a Greenhouse-Geisser adjusted was applied.

As expected, there was a main effect of drawing phase ($F = 172.42$, $p < 0.001$). Post-hoc analysis indicated that the Copy phase scores were significantly greater than both the immediate and delayed recall phases (both $p < 0.001$) (Table 4.2). No significant difference was found between the immediate and delayed recall phases ($p = 1.00$). No significant main effect of pen condition was found ($F = 1.161$, $p = 0.320$) nor was a significant interaction between variables found ($F = 0.01$, $p = 0.91$)

Table 4.2

Mean and standard deviations of RCF scores for each phase, split by condition (Normal versus Digital)

Phase	Normal		Digital		Total	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Copy	35.62	0.65	35.08	1.80	35.35	1.36
Immediate	22.65	6.18	22.62	2.08	22.64	5.03
Delayed	22.35	4.96	23.42	6.39	22.89	5.63

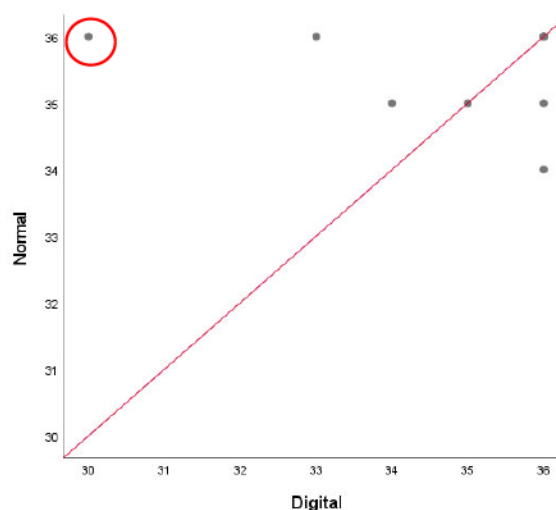
4.2.7.2 Correlation coefficient analysis

To enable correlation analysis to be conducted, participants in the two groups were matched. As evidenced in Chapter Three, age is a known factor for differences in accuracy performance on the RCF task, particularly on the recall phases (e.g., Vingerhoets et al, 1998; Boone et al, 1993; Gagnon et al, 2003; Gallagher & Burke et al, 2007; Hubley & Jassal, 2007; Frazier et al, 2010; & Hubley, 2010). Clinical scales for performance on the RCF are also age adjusted. With this in mind, participants in each group were ranked by age and then paired. Correlation of the data points was statistically analysed via a correlation coefficient. Given the existence of tied ranks in the copy phase – a large number of data points with the same scores - Kendall's tau correlation analysis was used. Analysis indicated a small relationship, but this was not statistically significant ($r(11) = 0.100$, $p = 0.710$). For the recall phases, Pearson's correlation analysis was used. For the immediate recall phase, a strong, statistically significant correlation was found between the two pen conditions ($r(11) = 0.620$, $p = 0.024$). For the delayed recall phase, a moderate correlation was found, however, this was not statistically significant ($r(11) = 0.358$, $p = 0.230$).

Visual inspection of a scatterplot (Figure 4.5) for copy scores indicated one data point fell some way from the rest of the data (see circled data point). The participant in the digital condition was the lowest scoring participant across the data set, which their matched pair scored the maximum available.

Figure 4.5

Scatterplot comparing copy accuracy scores of a normal versus digital pen. Line of perfect correlation indicated.



4.2.7.3 Intraclass Correlation Coefficient

Whilst Pearson's and Kendall's correlation coefficients demonstrate how closely the data points of the two conditions relate to one another, this only demonstrates if that relationship is linear and consistent. It does not reflect whether the scores of the two conditions agree with one another – that is, if an individual scores 32 on the delayed recall phase with the normal pen, do they score 32 with the digital pen? To address this, an intraclass correlation coefficient (ICC) was calculated for each phase. As the pen conditions were fixed, but the participants were randomly selected, a two-way mixed model was selected. An absolute agreement model type was selected as the purpose of the analysis was to consider agreement, and not how consistently the methods correlated, as this was already assessed. Due to the homogenous nature of scores within the copy phase and the sensitivity of ICC calculations to variation, only the recall phases were analysed. For the immediate recall phase the ICC indicated an acceptable level of agreement (0.780). For the delayed recall phase, even with consideration of the most liberal of guidelines of what is an acceptable ICC value, this fell just below the minimum acceptable level (0.527) (Fox-Fuller et al, 2021).

4.2.7.4 Bland-Altman plots of agreement

A Bland-Altman plot infers agreement between two measures if 95% of the data points fall between the 95% confidence interval limits (Bland & Altman, 1986). One data point for the copy phase (Figure 4.6 - overleaf) and for the delayed recall phase fell outside of the limits (Figure 4.7 - overleaf). Therefore, 92.31% of the data points were within the 95% confidence intervals. All data points for the immediate recall phase fall within the limits (Figure 4.8 - overleaf). A linear regression analysis of the Bland-Altman plot data informs whether any other potential bias exists. For the copy phase, the coefficient analysis indicates an unstandardized beta score of -1.660, which is statistically significant ($p = 0.002$), indicating the existence of systematic bias, where on average greater scores were found in the normal pen condition. For the immediate and delayed recall phases, the unstandardized beta scores are close to zero, and not statistically significant ($B = 0.024$, $p = 0.937$; $B = -0.370$, $p = 0.381$).

4.2.7.5 Usability and comfort

Participants in the digital pen condition completed the pen usability survey after using the pen (Table 4.3 - overleaf). All participants noticed the pen was not a traditional pen.

However, 85.71% reported using the digital pen felt no different to drawing with to a standard pen, and 100% saw no difference with the digital paper. 14.29% found using the digital pen somewhat harder to use than a normal pen, and less comfortable to hold. However, 28.57% found the pen more comfortable to use and 57.14% found no different to holding a standard pen. No participants considered their performance on the drawing tasks to be worse using the digital pen than if they used a normal pen.

Table 4.3

Usability, comfort and perception of performance using a digital pen (%)

	Much improved	Somewhat improved	No difference	Somewhat harder	Much harder
Ease of pen use	0	0	85.71	7.14	7.14
Comfort of pen	7.14	21.43	57.14	14.29	0
Use of paper	0	0	100	0	0
Perception of performance	0	7.14	92.68	0	0

Perceived performance and comfort were not found to correspond to objective performance. Two participants scoring the lowest scores found the pen as comfortable to use and did not feel their performance was affected by it. The participant who found the pen uncomfortable to hold and draw with performed within normal ranges across all three phases. However, whilst they found the pen uncomfortable, they did not feel it effected their performance.

Figure 4.6
Bland-Altman plot for Copy phase (digital -normal)

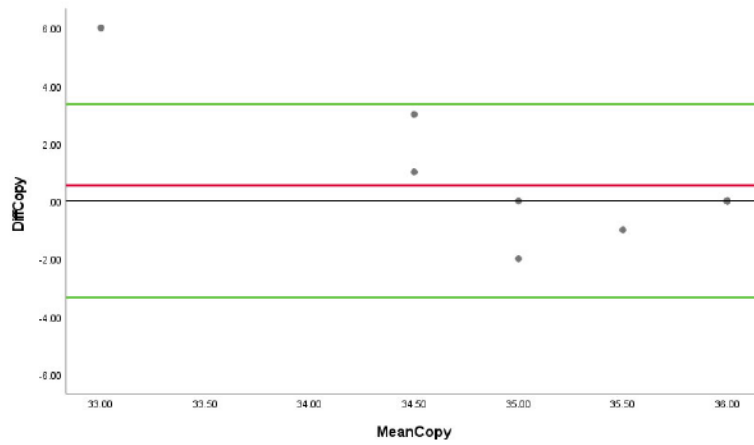


Figure 4.7
Bland-Altman plot for Immediate recall phase (digital – normal)

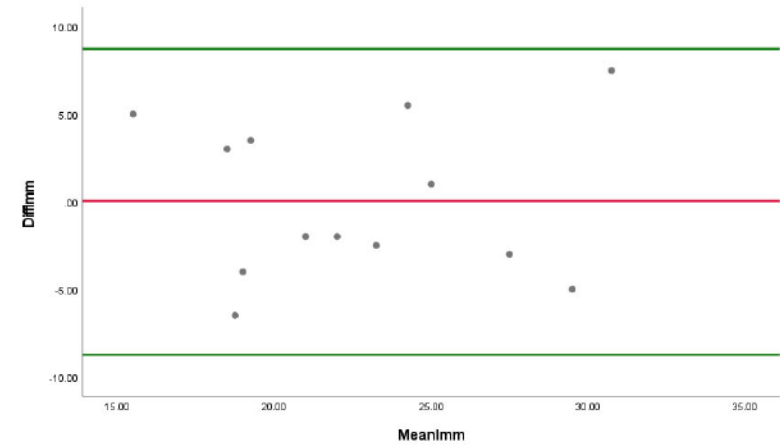
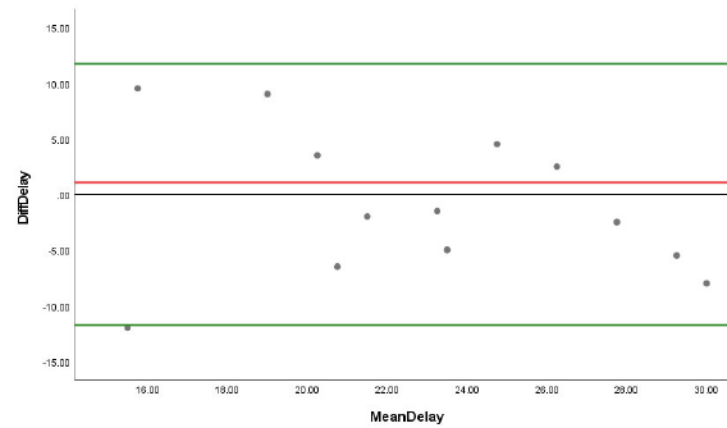


Figure 4.8
Bland-Altman plot for Delayed recall phase (digital – normal)



4.2.8 Study conclusions

The above study sought to consider the validity of using a digital pen to complete the RCF task compared to using a normal ballpoint pen or pencil. This was examined firstly by considering if there were significant differences, and then examining the extent of relationship and agreement between the conditions. The hypothesis that RCF accuracy performance using the Neo Smartpen would be equivalent to performance using a standard pen in a healthy adult population was accepted. No significant mean or variance differences in accuracy were found between the two pen types. The findings indicate that using a digital pen to complete the RCF task compares to using a normal pen within a healthy population. This reflects the findings from the study carried out by Sanada et al (2014) using the Inkling pen. However, correlation co-efficient analyses indicated varying levels of linear relationship. For the two recall phases, the relationship between the conditions was found to be moderate to strong. However, for the copy phase, this relationship was weak. Further, the relationship for the delayed recall was not statistically significant. As indicated in Figure 4.5, one matched pair was found to be an outlier, with the digital participant scoring some way below the rest of the group. Such a divergence of scores may account for the weak strength of correlation in the copy phase.

This study added to previous understanding of digital pen use in neuropsychological testing as it moved beyond inferential differences analysis to statistically examine agreement. A linear correlation co-efficient only examines consistency of relationship. It does not indicate the extent to which a pair of data points agree. An intraclass correlation co-efficient analysis provides the means to examine this. However, it is highly influenced by the range of scores, such that it has little interpretive value when the data set is highly homogenous - as with the copy phase scores in this sample. Therefore, only the recall phases were analysed. Whilst the immediate recall phase was found to have good to excellent reliability between the two pens, the delayed recall phase was found to be below accepted thresholds, even by the most liberal of guidelines (Fox-Fuller et al, 2021).

These findings correspond with the Pearson's correlation coefficient scores that were calculated. There are a number of possible explanations for such results, both of which are due to sample size. The Pearson's correlation did not reach a statistical power cut-off of 80%. Therefore, it is possible that the sample size was not large enough to identify a moderate correlation size that was statically significant. A priori analysis using G*Power (Faul et al, 2007; Faul et al, 2009) indicates that a sample size of 58 would be needed to do so. Further, whilst the range of scores within the recall phases were greater than for the

copy phase, this range is still limited in relation to the total possible range of scores, which is known to effect ICC analysis. It is suggested that the minimum sample size for an ICC to have interpretive strength is between 18 and 50 (Bujang & Baharum, 2017). Given that the current study has 13 matched pairs, the value of the ICC interpretations should be considered with caution.

To further analyse the level of agreement between the two pen conditions, data was plotted to a Bland-Altman plot. The plot infers an adequate level of agreement when 95% of the data points fall within the 95% confidence intervals (Bland & Altman, 1995). The immediate recall phase was indicated to have a good level of agreement, which aligned with the Pearson's correlation and ICC analysis. In the copy and delayed recall phase, the 95-95 criterium was not met – which again corresponded with the other analyses. However, this reflected a single data point falling outside of the interval limits. Due to the small sample size, a single data point had a relatively larger influence on the statistical outcome compared to the same scenario in a larger population sample. For the copy phase, when the outlier pairing was excluded and the data re-examined, no statistically significant bias was found.

The study also examined usability and comfort of the Neo Smartpen. The hypothesis that users would find using the Neo Smartpen comparable to a normal pen and paper was accepted. Whilst all participants in the digital pen condition were aware they had been given a non-standard pen; the majority of participants did not feel that it effected their performance. Even where individuals found the pen somewhat uncomfortable or felt it had impaired their performance, they still scored maximum points in the copy phase and in normal clinical ranges for the recall phases. Overall, these results compare to user feedback collected by Prange and colleagues (2019). However, this current data extends Prange's work, as these results were gained after participants used the pen to complete a task, rather than simply holding the pen.

Due to the between-participants design of this study, it was accepted that there would be some absolute differences between the groups as a result of individual differences. Such a design limits the how precisely agreement can be examined. However, as discussed in Chapter Three, a repeated measures design was discounted due to the known practice effects of the RCF. However, further validity analysis could be conducted using a counterbalanced repeated measures design with the RCF and other versions to support the findings in this study. Further, the study was conducted with healthy participants. It is recommended that validity analysis is conducted with clinical populations. This was an

intended part of this thesis, however, due to the Covid-19 social distancing restrictions, it was not possible to conduct this research.

Comparability of performance using a digital pen versus a standard ballpoint pen during the RCF task is critical to clinical use of the pen. Normative scales were developed using standard pen-and-paper administration; therefore, validity of the digital pen has to be examined to support the use of these scales when using a digital pen. The findings from this study do this. What makes a pen such as the Neo smartpen a viable solution is that it writes with a standard ballpoint cartridge on paper, mirroring the temporal, spatial and haptic feedback of analogue equivalents. It honours the principle of interface development that its use should not require additional effort from the user than its original format (Sonntag, 2018). Further, it is economically viable for use within public healthcare systems such as the NHS as it costs around £100 per pen – significantly less than more complex systems that are available. The digital pen enables the drawing process of the RCF to be recorded step by step enabling sensitive and detailed analysis, whilst not rendering the existing clinical norms redundant. Further, it has the potential to automate the scoring of a complex task such as the RCF, reducing the time and effort needed by clinicians to administer it. Overall, the findings from this study indicate that whilst RCF performance within a healthy population is not affected by using the digital pen. This provides sufficient evidence to support the use of the digital pen in later studies. It also supports grouping the data from both the digital and normal pen group as a single population in later analysis.

4.3 Video-based teleneuropsychology

4.3.1 History of teleneuropsychology

Teleneuropsychology is the practice of neuropsychology carried out remotely via communication technology, such as telephone or videoconference. Adopted since the 1980s, as the name indicates remote neuropsychological assessment was typically conducted via telephone. However, as accessibility to videoconferencing platforms has increased, this has presented new opportunities to conduct video-based teleneuropsychology.

There are several advantages to the use of teleneuropsychology. The first of these is increased healthcare accessibility for patients. Neuropsychologists are typically based in centrally located hubs, which can mean significant travel time and distances for patients who are more rurally located. This is a particular concern in countries where ‘medical deserts’

exist (Harder et al, 2020; Gnassounou et al, 2021; Tailby et al, 2021; and Chapman et al, 2021). It is, therefore, unsurprising that the majority of teleneuropsychology research comes from countries with very large land masses such as Australia and the USA. The use of teleneuropsychology helps to readdress the risk balance for patients who are immunocompromised – a concern that was been amplified during the Covid-19 health crisis – providing the opportunity to undertake necessary assessments, without further compromising their health. Teleneuropsychology could also be adopted to triage patients, to determine where a longer assessment is needed. This would focus limited resources where most needed and reducing waiting times for those individuals (Parks et al, 2021). Considering video-based teleneuropsychology specifically, it supports the use of non-verbal assessments and behavioural observations, increasing the amount of clinical information that can be gathered (Pulsifer et al, 2021).

However, there are a number of limitations. Whilst video-based teleneuropsychology can increase accessibility for some, it can be a barrier for others. Digital poverty is a global concern. Harder et al (2020) found that almost 25% of their US participants did not have access to the necessary equipment and connections. Particularly in countries where healthcare is not 'free at the point of access', this can further amplify the disparity of access to neuropsychological care. The same individuals are disproportionality effected both through limited access to healthcare services and through digital poverty for financial reasons. Infrastructure, even in developed, industrialised countries may also be a barrier. According to the Office of National Statistics, in 2019 7% of households in Great Britain were still without internet access (Office of National Statistics, 2019). As such, a complete move to online appointments is not feasible. Concerns have also been raised regarding ethics, privacy and confidentiality, limited control over testing environment, technological competence and comfort, and test security (Parks et al, 2021). Validity of use with standardized clinical scales that have been developed under in-person conditions has also been questioned. However, many of these concerns have been addressed either through the publishing of revised guidance, such as that by the British Psychological Society's Division of Neuropsychology (2020) or the Inter-Organizational Practice Committee (Bilder et al, 2020); or through empirical research.

5.3.2 Empirical evidence for the use of video-based teleneuropsychology

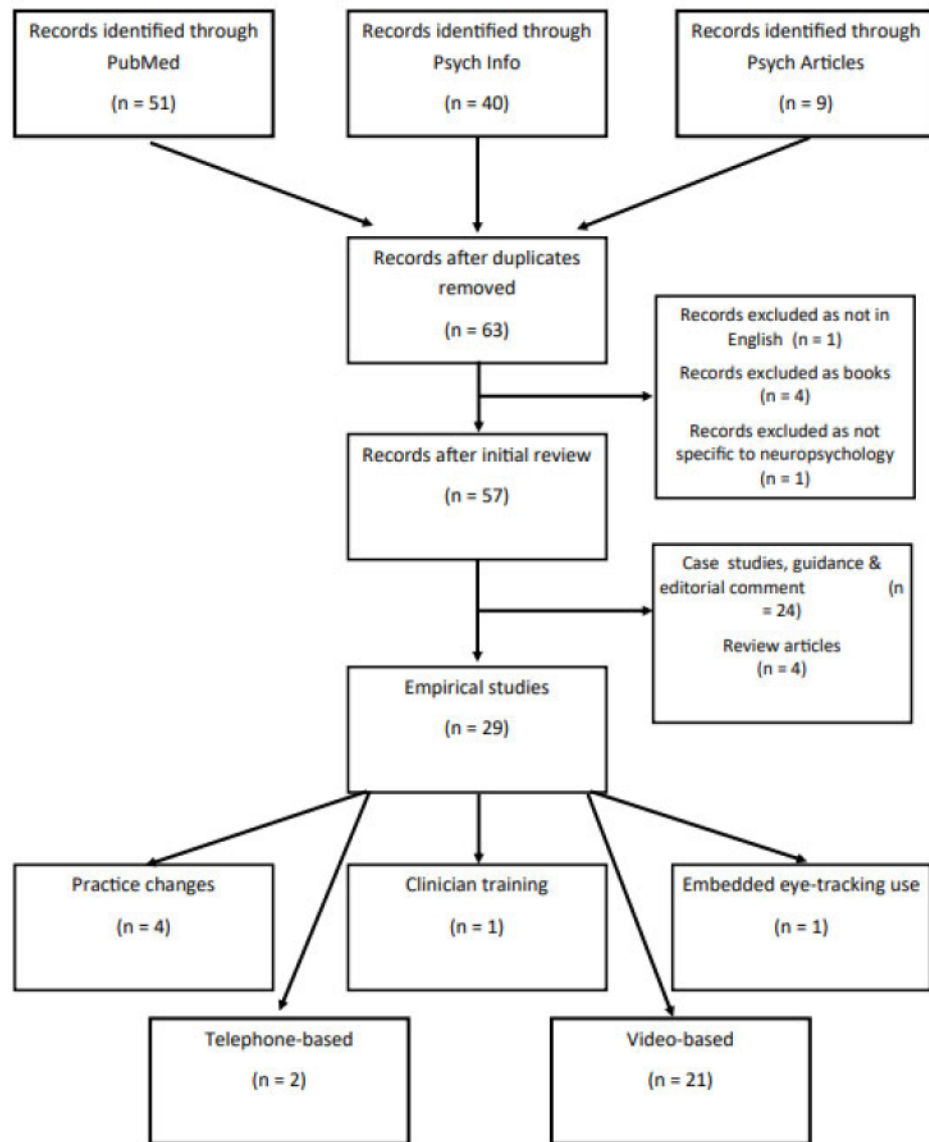
Whilst teleneuropsychology has being present within neuropsychological practice for several years, take up has been limited. Pre-pandemic, just 21% of neuropsychologists were using or intended to use video-based teleneuropsychology in their practice – with only

limited use for cognitive assessment (Hammers et al, 2020). This was often explained as a result of limited empirical evidence to support its use. Pre-2020, only one systematic review and meta-analysis addressing validity and reliability of video-based teleneuropsychology had been published. Brearly and colleagues (2017) concluded that videoconferencing was a viable option for undertaking a neuropsychological assessment. However, their recommendations across assessment types varied. Verbally mediated tasks were strongly supported, and visually supported tasks acceptable. However, the variability in motor-dependent tasks warranted further investigation. Non-timed tasks, tasks where instructions could be repeated, and tasks with repeated-rounds were more strongly supported than timed or single-round tasks. The breadth of tests examined was also limited due to the exclusion of studies that were not counterbalanced cross-over designs or where an administrative assistant was required.

Unsurprisingly, available literature into the use of video-based teleneuropsychology has increased since the Covid-19 pandemic. A search of the term 'teleneuropsychology' across three databases (Pub Med, Psych Info, Psych Articles; all fields) returned 57 peer-reviewed articles, of which 20 were published in 2020 and 29 in 2021, with four months of the year still to go (see Figure 4.9, overleaf). However, only 24 of the 49 articles published since 2020 were empirical studies. The remainder were practice recommendations, case studies and review articles. Twenty-one studies focused on the use of video-based assessment, though, this reduced to just eight that assessed the validity, reliability, or psychometrics of video-based teleneuropsychology. As such, the validity and reliability of video-based teleneuropsychology still remains a much under-researched resource. Whilst presentation of this search does not cover the full extent of the research that is available in this area given the variation in terminology that is used, it does illustrate how the knowledge landscape has evolved rapidly in this area of clinical neuropsychology.

Marra et al (2020) extended Brearly's review to capture more recent studies. However, in the context of the Covid-19 pandemic and the susceptibility of older adults to the virus, they focused their systematic review on studies addressing adults aged 65 years and above. Again, validity of video-based teleneuropsychology was indicated, though, some tests had more convincing evidence to support their use, mostly related to sample size. As with the Brearly study, only studies that were counterbalanced cross-over designs and that did not require administrative support were included.

Figure 4.9
Illustration of literature search of the term 'teleneuropsychology'



All but one of the tasks included within the current protocol (Letter-Number sequencing) had been examined within validity studies to some extent. Kirkwood et al (2000) concluded that the NART was more robust test for estimated IQ compared to the Quick Test IQ test. The HVLT-R had variable results. Chapman et al (2021) found that significantly less words were remembered in the HVLT-R learning trials under video conditions, albeit with a small effect size. They also found greater variance for HVLT-R total recall during the video condition. On the other hand, Cullum et al (2014) found scores to be slightly higher in the

video condition – a difference that was statistically significant. Parks et al (2021) did not find any significant differences in either the learning or total recall scores. Marra et al's (2020) meta-analysis indicated that validity for the learning phase was excellent and good-to-excellent for the total recall phase.

Verbal fluency tasks such as FAS letter fluency, Animals semantic fluency and Furniture/Fruits category-switching are found to be the most robust for use under videoconference conditions and are generally the most studied for validity – most likely due to their verbal-delivery format (e.g., Algret et al, 2021; Tailby et al, 2020; Chapman et al, 2021; Gnassounou et al, 2021; Parks et al, 2021). However, verbal fluency was found to be more robust than semantic fluency. Fox-Fuller et al (2021) posited this was a result of a single round for the Animals task compared to three rounds on the FAS task, one for each letter. This corresponded with Brearly et al's (2017) and Marra et al's (2020) conclusions that tasks with repeated rounds were more robust. Category switching is less examined than verbal and semantic fluency. However, evidence indicates there is validity of its use under video conditions (Harder et al, 2020). Other verbal based tasks such as the Stroop task, have also been examined, though only minimally. To modify for the videoconference environment, Chapman et al (2021) issued copies of physical test cards to the participant in separate envelopes, with instructions for use delivered verbal as standard. The colour and colour-word phases were found to be comparable between in-person and video conditions, though interference scores were found to be systematically greater in the video condition.

Tasks that measure non-verbal performance have been less examined, and what results exist are more varied. This includes tasks from the current protocol such as the RCF and Trail Making (TMT) test, all of which require physical materials to be provided to the participant in advance. For the RCF, literature mostly examined accuracy and not drawing process, due to the nature of needing to see what the individual was drawing. Chapman et al (2021) overall found consistency between the conditions. However, lower performers were found to be negatively affected by the video condition compared to better RCF performers. Gnassounou et al (2021) found no significant differences between the conditions. For the TMT task, some researchers chose to substitute the standard TMT task for the Oral TMT task in the video condition. For those that maintained the pen-and-pencil format, the results were variable. Whilst Chapman et al (2021) found no significant differences, Gnassounou et al (2021) found more errors in TMT-A in the video condition.

Whilst a key concern of practitioners is the comparability of using a video-setting and the validity of using the in-person based norms, it is useful to also examine the evidence

regarding the acceptability and feasibility of administering a video-based neuropsychological assessment. Patients represented a variety of clinical conditions, including neurodegeneration, neuro-oncology, psychiatric, neurodevelopmental, movement disorders, epilepsy, stroke, TBI, auto-immune conditions and infections. Tailby et al (2020) conducted a short screen of three tasks with adults who have epilepsy. They included anecdotal evidence in their report that several participants commented they were more comfortable and at ease being at home. Chapman et al (2021) indicated that their participants – stroke patients - were satisfied with either in-person or video-based, but there was an overall preference to in-person – though they were aware this came at the cost of travel and wait-list times.

However, practitioner subjective acceptability was found to be lower than the acceptability of patients. For example, Peterson et al (2020) found 100% of their patients responded to a survey that they were satisfied with their virtual assessment, compared to 90% of clinicians. Key concerns raised by clinicians were limited control over environment; privacy and confidentiality; technological competency; poor connection; test security; ethical and legal issues; lack of specific normative data; and practical implications of test administration (Rochette et al, 2021; Hammers et al, 2020; Eilam-Stokes et al, 2021). Several authors supported the use of video-based teleneuropsychology, but not as a complete replacement for in-person assessment. Eilam-Stock et al (2021) suggested that it should only be used as a proxy, and not full replacement. Parks et al (2021) suggested it be used for triaging, with full assessment taking place in person. Interestingly, research has indicated that the necessity of having to use videoconferencing during the pandemic has increased the subjective acceptability of its use by practitioners. One study indicated that whilst only 11% of responders had used videoconferencing pre-pandemic, 59% intended to continue to use it post-pandemic (Hammers et al, 2020).

Given the limited evidence that is available regarding video-based teleneuropsychology, the generalisability of the findings from those studies must be considered. An issue that is raised repeatedly within research published during the pandemic was location of the patient or participant during the assessment session. Particularly in earlier studies, assessment was operated on a clinic-to-clinic basis, to control the testing environment and consistency of internet connection (e.g., Gnassounou et al, 2021; Chapman et al, 2021). However, this still required travel to a suitable centre. Whilst non-essential travel was restricted in many countries and there were valid concerns regarding exposure to the Covid-19 virus in medical centres, particularly for those with additional health risks, this highlighted the need to consider clinic-to-home or home-to-home arrangements. Harder et al (2020) used a clinic-to-home setting with a paediatric population

who had various neuroimmune conditions, such as multiple sclerosis, encephalomyelitis and optic neuritis. They found no statistically significant differences between the face-to-face clinic setting and the home setting. However, only inferential statistical analysis of differences was examined, and not agreeability.

Algret et al (2021) addressed this by conducting a retrospective analysis of patient assessments, using invariance analysis with patients with dementia and a healthy control group. They concluded there was comparability assessment via videoconferencing. However, for individuals diagnosed with mild cognitive impairment the use of verbal fluency and digit span forward tasks was less well supported. A further finding from their study was that no difference was found between different devices – supporting the use of videoconferencing on mobile phones if a tablet or computer was not available. Tailby et al (2020) also found there to be no statistical differences between assessments undertaken at home or in the clinic.

A further key point to consider is the clinical population each of the existing studies address. A significant proportion of the studies address either older adults and/or dementia-based diagnosis, or paediatric populations. When considering TBI or concussion populations specifically, very limited literature is available. Of five articles that were identified that specifically addressed TBI or concussive populations, 3 of these focused solely on validating the use of communication assessments (Theodoros et al, 2003; Theodoros et al, 2008; Rietdijk et al, 2017). The remaining two articles provided guidance and considered feasibility qualitatively of delivering a video-based assessment (Ellis et al, 2019; Parlar et al, 2020). Whilst some empirical studies that examined validity and reliability of a mixed neurological population did include TBI patients, no studies to date have been found that consider video-based assessment for TBI or concussive populations exclusively (e.g., Parks et al, 2021).

Whilst there has been a significant shift in the knowledge landscape regarding the use of video-based neuropsychological assessment, the evidence that does exist is still very limited and heterogenous. Administration set up, clinical populations, comparable conditions, and tests used vary significantly. Further, the majority of research that does exist only considers inferential statistical differences, in the form of t-tests or ANOVA, with very few considering agreement statistical analysis, such as intra-class correlations or Bland-Altman plots. As such, there is still a need for validity studies to be carried out that examine the use of video-based teleneuropsychology.

4.3.3 Aims and hypotheses

Due to the Covid-19 pandemic and the social distancing measures mandated by the UK government, in-person data collection was suspended from March 2020. At this time, we were mid-way through the data collection with healthy, non-TBI participants. As it was anticipated (correctly) that in-person data collection would not be possible for several months, alternate approaches had to be considered. The use of videoconferencing offered a solution. As it was not necessary to travel for the study session, the potential participant pool was widened and enabled the sample to be more representative of the UK, rather than being a South-East micro-population. It also presented an opportunity to examine the validity and reliability of conducting the current screening protocol via video-based teleneuropsychology. The aim of this study was to examine the reliability and validity of undertaking the current neuropsychological screening protocol via videoconference, and in particular, the RCF. It was hypothesised that performance on the RCF and the remainder of the screening protocol would be comparable between the in-person and video-conference settings.

4.3.4 Conditions and allocations

Participants were allocated to either the in-person or to the video condition. In-person participants were those participants who took part in the digital pen study, reported in the first half of this chapter. For the video condition, participants participated via the online videoconferencing platform, Zoom. Zoom has been used in previous research and is an approved system under US HIPPA rules (Fox-Fuller et al, 2021). Participants were instructed to use either a stand-alone monitor, laptop or tablet, but not a mobile phone. This was in accordance with guidance set by the BPS Division of Neuropsychology (2020) for video-based teleneuropsychology. Participants were instructed to set the viewing option to Speaker view, so that it was as close to face-to-face interaction as possible. All participants who took part prior to April 2020 were in-person. From April 2020 to September 2020 participants took part via video-meeting.

4.3.5 Participants

Fifty-four participants were recruited to the study. This sample size was the result of matching the number of participants in the video condition with the number of participants who had already taken part in-person. Each condition had 27 participants. All met the

inclusion and exclusion criteria as detailed in Chapter Two. Recruitment was via the same methods as in the first study in this chapter.

Age of participants ranged between 22 and 70 years of age (mean 44.61 yrs, SD = 12.47). No statistically significant difference in age was found ($t(52) = -0.66, p = 0.51$). Forty-four participants were female and 10 male. Whilst this split was not equal across the whole sample, the split between the conditions was (22 female: 5 male). The age of leaving school was equally matched across the two groups. Estimated IQ, as measured by the NART, statistically differed between the two groups ($t(52) = 0.43, p = 0.01$). Participants in the videoconferencing condition on average had higher estimated IQ than those in the in-person (Table 4.4)

Table 4.4
Mean and standard deviation of age, estimated IQ and years of education by group

Variable	In-person		Videoconference		Total
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>p</i>
Age	43.48	14.70	45.74	9.93	0.51
Estimated IQ	108.25	3.12	115.56	9.01	0.01
Years of Education	15.04	2.64	15.67	2.17	0.34

4.3.6 Procedure and online adaptations

All participants were administered the same protocol as indicated in Chapter Two. For the in-person condition standard neuropsychological assessment administration was used. For the video condition, administration was preserved as far as possible to ensure consistency between conditions. Where adaptations were made, these were as per the adaptations reported in Chapter Two, section 2.8.1. This included posting materials to the participant in advance, and the participant asking a household or support bubble member to assist.

4.3.7 RCF Results

The data was analysed via a 2x3 mixed-methods ANOVA, with condition as the between groups variable (in-person and video), and the RCF phase as the repeated measures variable (copy, immediate recall, and delayed recall). Further agreement between

the conditions was analysed via correlation coefficients, intra-class correlations and Bland-Altman plots.

4.3.7.1 ANOVA analysis

The data was reviewed to determine if the assumptions of ANOVA were met. Data was continuous and independent. Influential outliers were considered by examining the studentized residuals. One participant was found to have studentized residuals more than 3 standard deviations below the mean for the copy phase (-5.61). Normality of residuals was examined via visual inspection of Q-Q plots of the studentized residuals for each phase. All phases were considered to be close to normal distribution. Equivalence of variance was examined between the conditions. All phases were found to have equal variance. Mauchly's test of Sphericity indicated that the assumption of sphericity was met. No action was taken due to the robustness of ANOVA, however, the influence of the outlier was considered during interpretation of data.

As expected, there was a main effect of drawing phase ($F(1, 52) = 340.29, p < 0.001, \eta^2 = 0.87$). Post-hoc analysis indicated that copy phase scores were significantly greater than both immediate and delayed recall phases (both $p < 0.001$) (Table 4.5). No significant difference was found between the two recall phases ($p = 1.00$). No significant main effect of pen condition was found ($F(1, 52) = 0.75, p = 0.39$), nor was a significant interaction between the variables found ($F(1, 21) = 1.83, p = 0.18$).

Table 4.5

Mean and standard deviations of RCF scores for each phase, split by condition (In-person versus videoconference)

Phase	In-person		Videoconference		Total	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Copy	34.93	2.56	32.26	1.16	35.09	1.97
Immediate	22.24	5.34	20.57	5.96	21.41	5.67
Delayed	22.48	5.90	20.91	5.80	21.69	5.85

4.3.7.2 Correlation coefficient analysis

As in the previous study, groups were matched for age. Correlation of the data points was statistically analysed via a correlation coefficient. Given the lack of normal

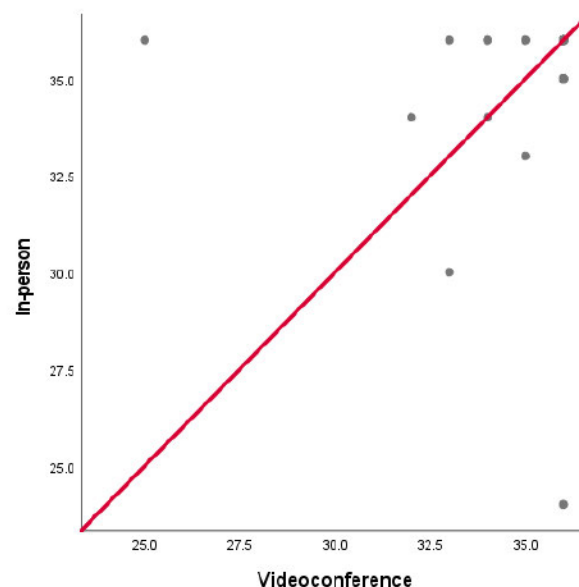
distribution of scores for the copy phase, plus the existence of tied ranks – several data points with the same scores - Kendall's tau correlation analysis was used for the copy phase. Kendall's correlation is also less prone to influence of outliers. Analysis indicated a small relationship, but this relationship was not statistically significant ($r(27) = 0.170, p = 0.353$). For the immediate and delayed recall phases, Pearson's correlation analysis was used. For both recall phases a small relationship was found between the scores for the environment conditions, however, neither relationship was found to be statistically significant (Immediate recall: $r(27) = 0.170, p = 0.396$; Delayed recall: $r(27) = 0.184, p = 0.359$). As with the previous study, a visual inspection of the scatter plot (Figure 4.10) illustrated the influence of the participant who was a significant outlier.

4.3.7.3 Intraclass correlation coefficient

Intraclass correlation coefficients (ICC) was examined for the two recall phases to determine the level of agreement between the two conditions, as in the previous study. Neither recall phase was found to reach an acceptable level (immediate recall = 0.286; delayed recall = 0.310). As a value of 0.6 is the most lenient minimum acceptable level of agreement between scores (Fox-Fuller et al, 2021), the data suggests there is limited agreement between the pairings.

Figure 4.10

Scatterplot comparing copy accuracy scores of in-person versus videoconference conditions. Line of perfect correlation indicated.



4.3.7.4 Bland-Altman plots of agreement

Two data points for the copy phase fell outside of the limits (Figure 4.12, overleaf) , reflecting 92.59% of the data points falling within the 95% confidence intervals. For the recall phases, four data points fell outside the upper limit for immediate recall (Figure 4.13, overleaf), and two outside each limit for delayed recall (Figure 4.14, overleaf) . Both reflected 85.19% of data points falling with the 95% confidence intervals. For the copy phase, the bias value was very close to zero, however, for the immediate and delayed recall phases the bias value indicated slightly superior average performance in the in-person condition. However, linear regression analysis indicated the proportional differences were not statistically significant (Copy: $B = 0.215$, $p = 0.591$; Immediate recall: $B = -359$, $p = 0.288$; Delayed recall: $B = -0.120$, $p = 0.721$).

4.3.8 Cognitive screen results

Data from the rest of the screening protocol was also examined for agreement between conditions. Analysis was carried out via independent samples t -tests, with neuropsychological test score as the dependent variable, and test condition (in-person versus videoconference) as the independent variable. Further, agreement between conditions was analysed via correlation coefficients, intra-class correlations and Bland-Altman plots.

4.3.8.1 Group differences analysis

The data was reviewed to determine if the assumptions of the independent samples t -test were met. Box and Whisker plots for each neuropsychological test were reviewed to identify significant outliers. Extreme outliers were identified in delayed verbal memory ($N=1$) (Figure 4.15, overleaf) ; TMT-B ($N=1$) (Figure 4.17, overleaf) ; and the Stroop task ($N=3$) (Figure 4.16, overleaf). No participants were identified as an outlier for more than one test. Normality of distribution for each measure was evaluated via visual inspection of Q-Q plots. All tests were considered to be close to normal distribution. Equal variance for each test was also examined, via Levene's Test for Quality of Variance. HVLT Recognition and TMT-A were found to violate this assumption, therefore a correction was applied. Significant differences were found between the conditions on the TMT-A task ($t(45.02) = -2.48$, $p = 0.02$). For all other tests, no significant differences were found (Table 4.6, overleaf).

Figure 4.12

Bland-Altman plot for Copy phase (in-person – videoconference)

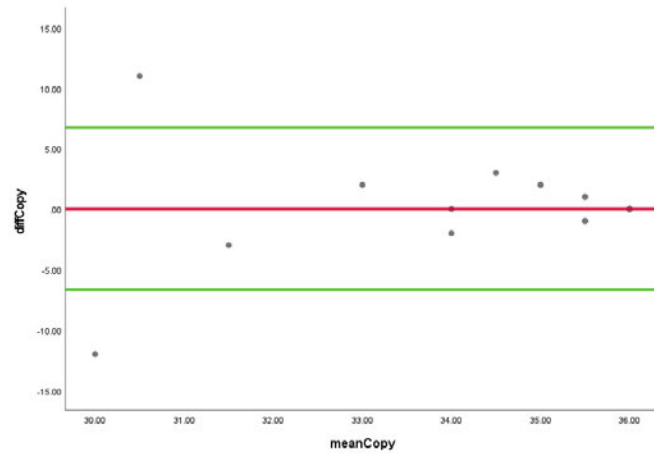


Figure 4.13

Bland-Altman plot for Immediate recall phase (in-person – videoconference)

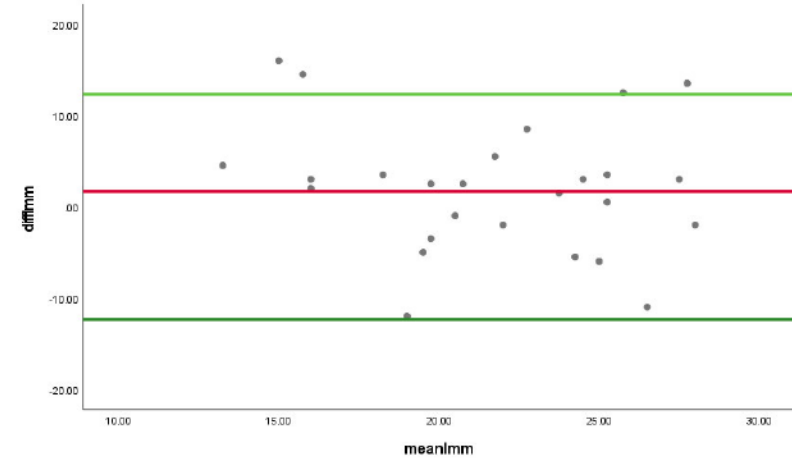


Figure 4.14

Bland-Altman plot for Delayed recall phase (in-person – videoconference)

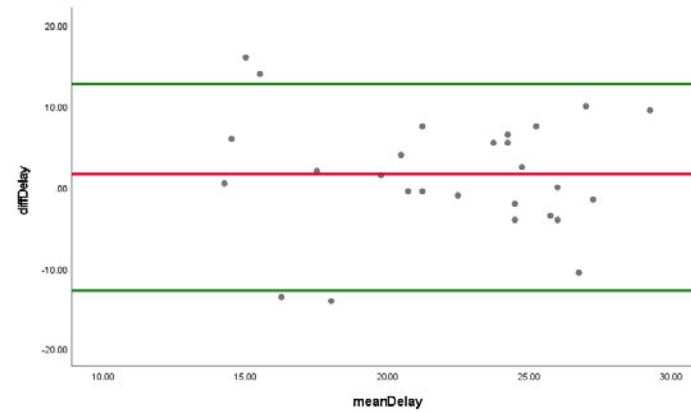


Table 4.6*Mean, standard deviation & significance for each test, by condition (in-person versus videoconference)*

Test	In-person		Videoconference		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
HVLT learning	29.37	3.22	29.74	3.50	-0.41	0.69
HVLT delayed recall	10.44	2.44	10.37	1.71	0.13	0.90
HVLT Recognition	1.11	0.97	10.93	1.69	0.49	0.62
FAS verbal fluency	44.15	11.87	19.15	14.28	-1.40	0.17
Animals semantic fluency	24.04	4.93	14.11	4.70	-0.06	0.96
Category switching	16.67	2.75	17.85	3.29	-1.44	0.16
Trail Making Test – A	24.15	5.86	29.22	8.88	-2.48	0.02
Trail Making Test B	55.56	15.83	60.96	23.39	-1.00	0.32
Letter-number sequencing	11.19	2.96	12.88	3.46	-1.90	0.06
Stroop	106.44	10.19	107.04	10.27	-0.21	0.83

Figure 4.15

Box and Whisker plot for Delayed HVLt recall

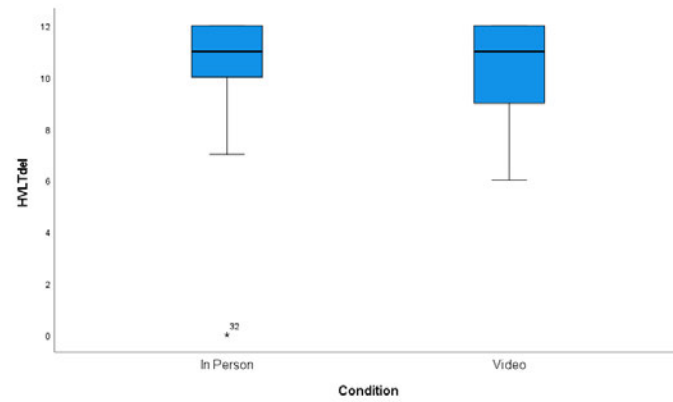


Figure 4.16

Box and Whisker plot for Stroop task

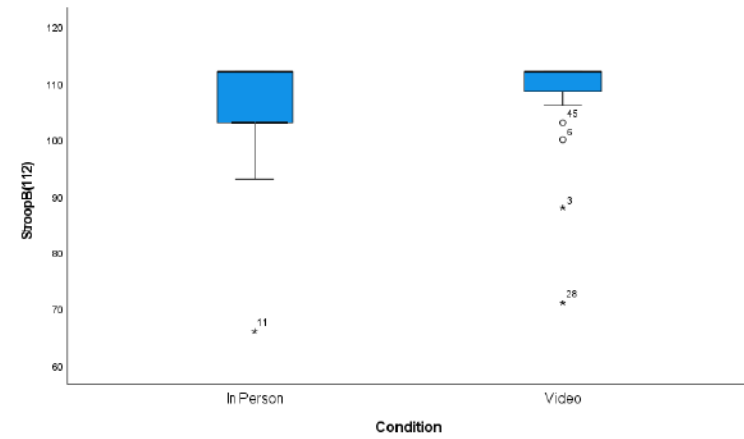
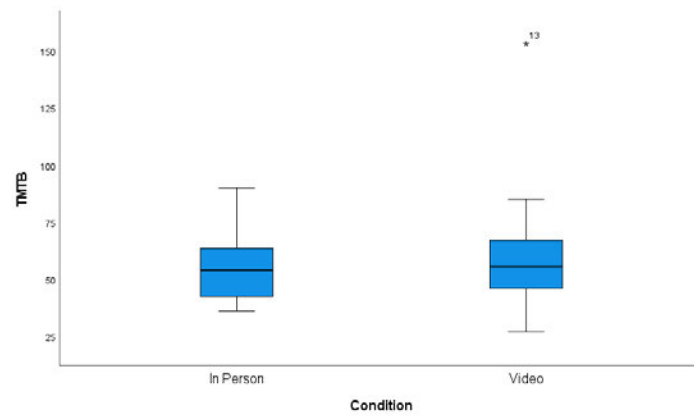


Figure 4.17

Box and Whisker plot for TMT-B



4.3.8.2 Correlation coefficient analysis

Correlations between the conditions for each test were also statistically analysed with correlation coefficients (Table 4.7). As the HVLT-R delayed recall, HVLT-R recognition and Stroop tasks had tied ranks, Kendall's tau correlation analysis was used for these tasks. For all other tasks, Pearson's correlation analysis was used.

Table 4.7
Correlation coefficients for neuropsychological tests

Test	Analysis	<i>r</i>	<i>p</i>
HVLT learning	Pearson's	-0.06	0.77
HVLT delayed recall	Kendall's tau	0.17	0.28
HVLT Recognition	Kendall's tau	0.22	0.21
FAS verbal fluency	Pearson's	0.10	0.62
Animals semantic fluency	Pearson's	0.26	0.20
Category switching	Pearson's	0.09	0.66
Trail Making Test – A	Pearson's	0.01	0.97
Trail Making Test B	Pearson's	-0.20	0.33
Letter-number sequencing	Pearson's	0.17	0.41
Stroop	Kendall's tau	0.19	0.26

HVLT-R delayed recall, HVLT-R recognition, FAS verbal fluency, Animals semantic fluency, TMT-B, Letter-number sequencing and the Stroop task were all found to have small, positive linear relationships, whilst also other tasks indicated close to absence of linear relationship. However, none of the correlation coefficients were found to be statistically significant.

4.3.8.3 Intraclass correlation analysis

Due to the lack of heterogeneity of scores for the HVLT-R delayed recall, HVLT-R recognition, Category switching, TMT-B and Stroop tasks, ICC calculations were not conducted on these tasks. ICCs were conducted for the remainder of the tasks (Table 4.8, overleaf). None of the analysed tasks were found to demonstrate a high level of agreement between conditions. However, as with the simple correlation analysis, none of the ICC coefficients were found to be statistically significant.

Table 4.8*Intraclass correlation coefficients between in-person and videoconference conditions*

Test	ICC	<i>p</i>
HVLT-R learning	-0.13	0.62
FAS verbal fluency	0.17	0.31
Category switching	0.15	0.48
TMT-A	0.01	0.48
Letter number sequencing	0.26	0.20

4.3.8.4 Bland-Altman plots of agreement

Data points for each test were plotted to a Bland-Altman plot (Figures 4.18 – 4.27, overleaf). For the Category switching, TMT-A and Letter-number sequencing tasks, all data points within the 95% confidence intervals. HVLT-R learning, HVLT-R delayed recall, HVLT-R recognition and FAS verbal fluency were found to have one data point each outside of the 95% limits. 96.30% of data points for each of these tasks were within the 95% intervals. The Animal semantic fluency, TMT-B and Stroop tasks each had two data point outside of the confidence intervals, reflecting 92.59% of data points within them. Both the HVLT-R recognition and TMT-A tasks were found to have significant proportional bias (Table 4.9). The regression analysis indicated that participants had lower HVLT-R recognition discrimination scores in the video condition than the in-person condition ($B = 787$, $p = 0.006$). For the TMT-A tasks, participants generally took longer to complete the task in the video condition, indicating slower cognitive processing speed ($B = -0.781$, $p = 0.042$).

Figure 4.18
Bland-Altman plot for HVLT learning

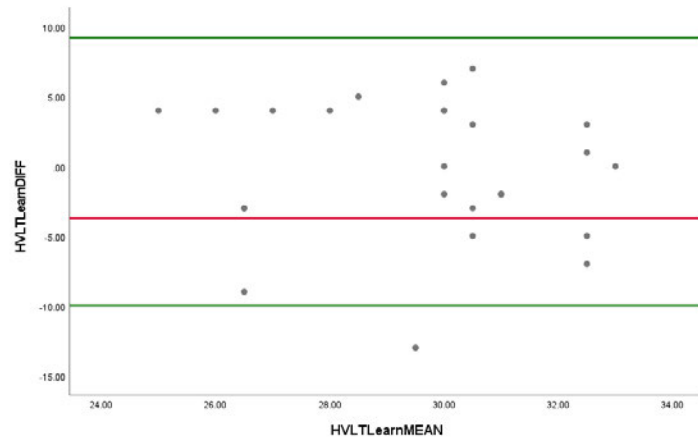


Figure 4.19
Bland-Altman plot for HVLT recognition

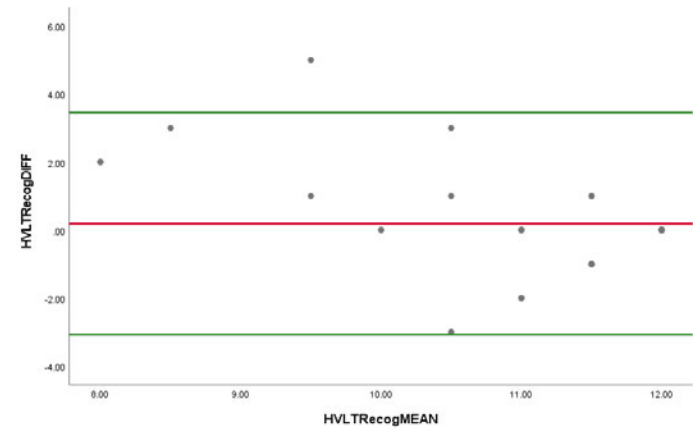


Figure 4.20
Bland-Altman plot for HVLT delayed recall

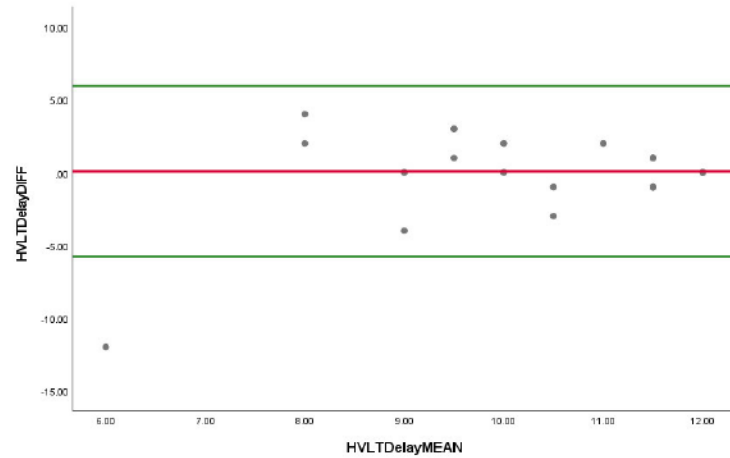


Figure 4.21
Bland-Altman plot for FAS verbal fluency

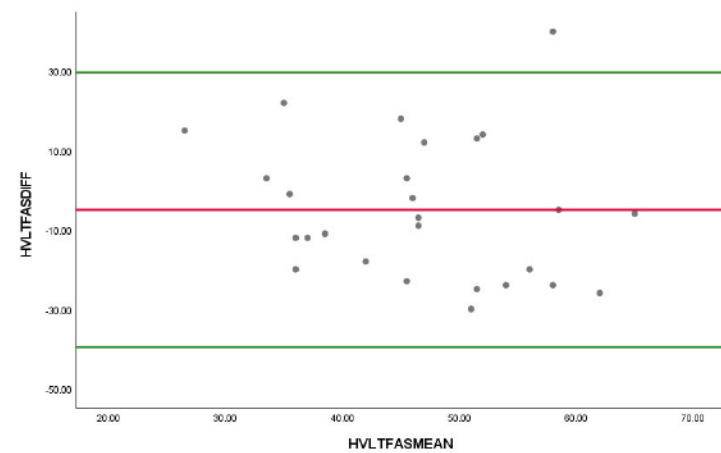


Figure 4.22

Bland-Altman plot for Animal semantic fluency

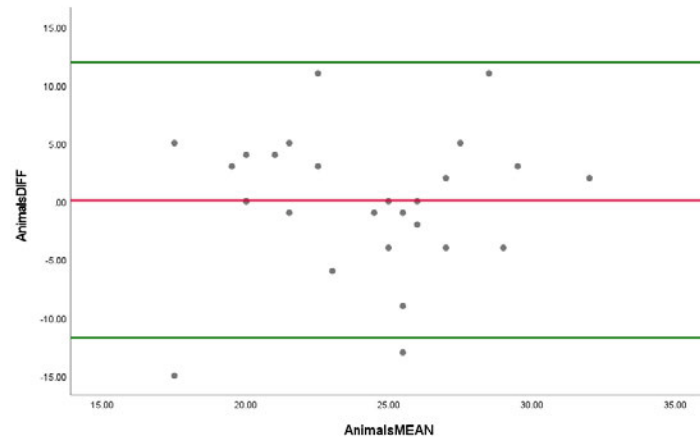


Figure 4.23

Bland-Altman plot for Letter-number sequencing

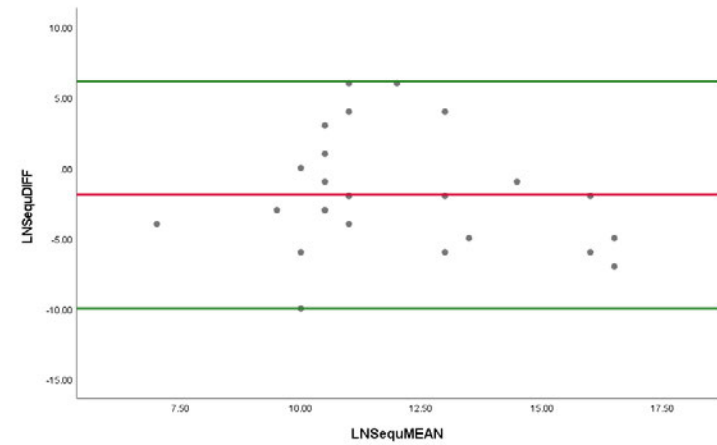


Figure 4.24

Bland-Altman plot for Category switching

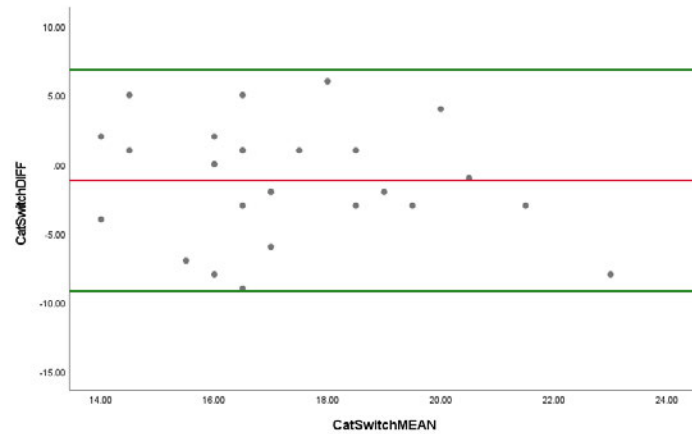


Figure 4.25

Bland-Altman plot for Stroop

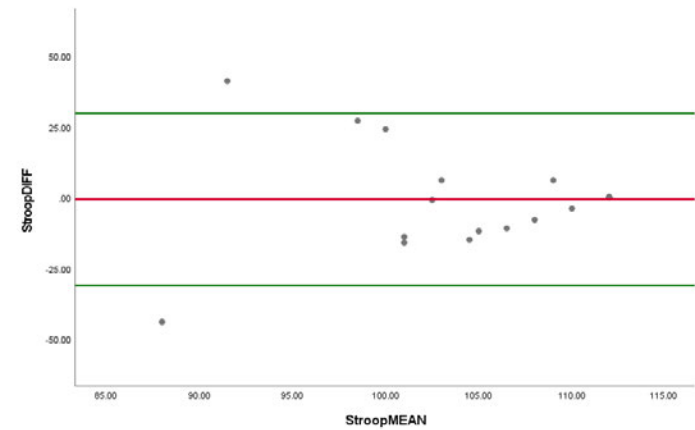


Figure 4.26
Bland-Altman plot for TMT-A

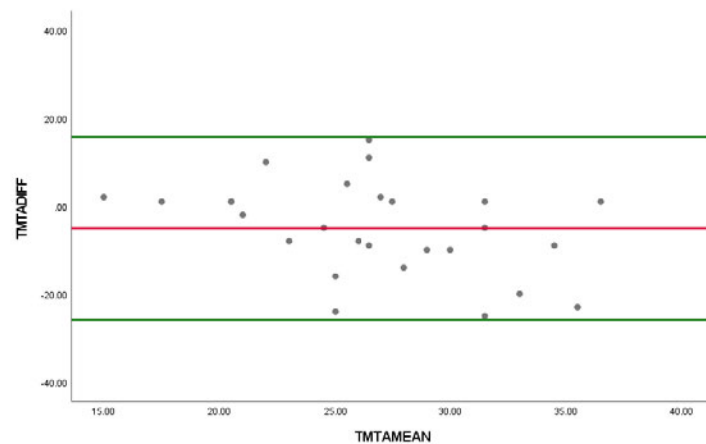


Figure 4.27
Bland-Altman plot for TMT-B

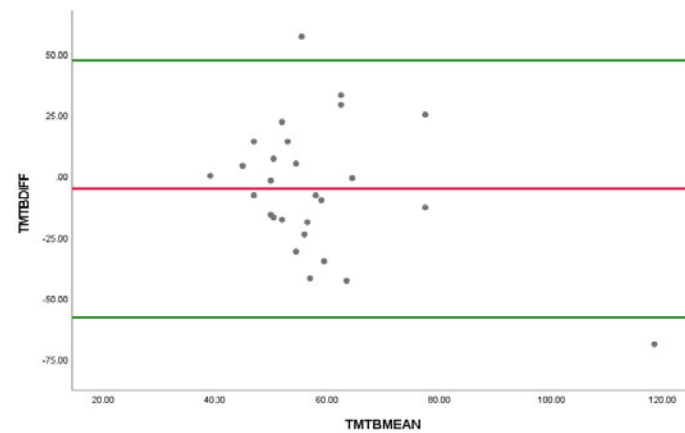


Table 4.9

Unstandardized beta and significance values to test proportional bias between in-person and videoconference conditions

Test	Unstandardized beta	<i>p</i>
HVLT-R learning	-0.17	0.69
HVLT-R delayed recall	0.69	0.08
HVLT-R recognition	0.79	0.01
FAS verbal fluency	-0.33	0.36
Animals semantic fluency	0.08	0.81
Category switching	-0.33	0.37
TMT-A	-0.78	0.04
TMT-B	-0.68	0.06
Letter number sequencing	-0.29	0.41
Stroop	0.01	0.98

4.3.9 Study conclusions

Video-teleneuropsychology is a relatively new area of neuropsychology, which has only been used and researched on a limited scale. Whilst feasibility of video-teleneuropsychology has been demonstrated, validity evidence is limited and varied. Given the limited body of evidence regarding video-neuropsychology, the aim of this study was to examine the validity of carrying out a cognitive triage screen via video-based teleneuropsychology.

No statistically significant mean differences were found between the two environmental conditions, except for TMT-A. Participants in the videoconferencing condition took significantly longer than those completing the task in-person. Correlation coefficient analyses were conducted to examine to what extent performance between the in-person and videoconferencing conditions agreed. The HVLT-R recall and recognition phases, animal semantic fluency, letter number sequencing, the Stroop task and the three phases of the RCF task were all found to have a small, positive relationship. However, the remainder of the tasks were only minimally correlated. As linear correlation coefficients only consider consistency of relationship, and do not consider the extent to which the data pairs agree, Intraclass correlation coefficient analyses were also conducted. Due to its sensitivity to homogenous data and data that does not have normal distribution, not all of the tasks were included in this, including the copy phase of the RCF. For those that were examined, none of the tasks met the minimum accepted level of reliability between the two conditions.

However, none of the coefficients were found to be statistically significant, therefore, it was not possible to rule out that these findings were purely due to chance, rather than there being lack of agreement. Bland-Altman plots were also examined. The majority of the tasks were found meet the 95-95 criterium, indicating an acceptable level of agreement. Animals semantic fluency and each of the RCF phases were also found to breach the rule. However, given the small sample size, a single data point accounted for almost 4% of the total sample. Regression analysis indicated that both the HVLT-R recognition and TMT-A tasks had significant systematic bias. No other tasks, including the RCF phases, were found to have significant bias.

The findings from the study do not provide a clear picture as to the validity of carrying out the screening protocol via video-based teleneuropsychology versus in-person assessment – reflecting the results of existing literature. The direction of mean differences varied, with some of the tasks having superior performance in the in-person environment, whilst videoconference had superior performance in others. Furthermore, whilst some tasks met acceptable statistical significance parameters, others did not. If only group differences were considered, as is the case with much of the existing literature, this would suggest that carrying out the current cognitive protocol via video-teleneuropsychology is equivalent to it being carried out in person. However, given the aim of this study was to determine if there was agreement between the conditions, this method of data analysis alone was not sufficient. Whilst correlation coefficient analyses did not provide strong support that the environmental conditions provided equivalent results, only two of the tasks were indicated to have systematic bias, with the in-person condition performing at a superior level.

Interestingly, the results did not seem to support previous findings that verbally delivered, multiple-round tasks were more robust than once-only or motor-based tasks. Whilst the motor-based TMT-A did indicate a significant difference between the conditions and significant proportional bias towards the in-person condition, TMT-B did not. An explanation for this may be due to the repeated format of the task, with several practice opportunities available. By completing TMT-A and both practice rounds, participants were better orientated to the TMT-B trial than TMT-A. Further, the TMT was one of the tasks that was adjusted to accommodate videoconferencing. When carried out in-person, the administrator both tells the individual when to start and records the time take to complete the task. In the video setting, the household member told the participant when to start and verbally confirmed when the participant finished, from which the administrator recorded the time based on this. This may have resulted in a disconnect between the actual point of

finishing and when the timer was stopped – however, this would have been likely to impact both stages of the task, and minimally.

Another interesting result was the evidence of systematic bias in the FAS, in favour of the video-condition. Due to the forced allocation of participants to the conditions, there was a significant difference in estimated IQ, with the videoconference group having a higher estimated IQ mean than the in-person group. This may account for the systematic difference between the groups on this task. However, agreement was found between the two conditions via the Bland-Altman plot. This agrees with Brearly et al's (2017) and Marra et al's (2020) conclusions that the FAS task is a robust assessment to use via video-based teleneuropsychology.

As important as it is to examine the statistical validity of administering these tasks via videoconferencing, it is also important to consider feasibility and practicalities. For the letter-number sequencing task, two of the participants in the video-condition did not complete the task due to bandwidth issues. Due to the nature of the task, the verbal-stimuli cannot be repeated. Therefore, the task is more vulnerable to connection issues. However, no significant difference between mean scores was found, a small Pearson's correlation coefficient was indicated, and all data points fell within the 95% confidence intervals on the Bland-Altman plot. These findings indicate this task adapts well to video-teleneuropsychology. However, this is dependent on there being a strong and consistent connection. The conclusions from this task support those of Brearly et al (2017) who posit that once-only tasks were not as robust in video-teleneuropsychology as tasks where instructions can be repeated.

As this was a retrospective study that was driven by necessity because of the Covid-19 pandemic, there are limitations to the design that must be addressed. It was not possible to randomly allocate participants to either condition, which may result in unknown bias. Possible unknown factors could be that those who were furloughed were more likely to take part as they had more available time than those who were continuing to work either from home or those who were considered 'key workers', such as grocery shop workers, delivery drivers, and those within the medical services. Further, as it was a requirement that the participant had an available H/SB member present to provide administrative assistance, this excluded individuals who lived alone and lacked a nearby support bubble. It was also not possible to operationalize the study to be a within-participants study. Whilst participants were matched for age, it is acknowledged that even when age and education history is taken into account there is still wide variation in individual abilities. However, as argued previously in

this thesis, practice effects are a known issue with neuropsychological assessments. Therefore, there are differing limitations to using a repeated measures design that may confound results.

Overall, the findings from this study provide some support for the use of videoconferencing within neuropsychological assessment. However, as has been suggested in previous work by Brearly et al (2017) and Marra et al (2021), there is variability in the strength of support across the tasks, indicating that some may be more suitable for use within this setting than others. As the participants within this study were healthy controls, cognitive assessment via videoconference with a clinical population should be reviewed cautiously as it is not possible to determine if these findings are generalisable to individuals who may have significant cognitive deficits. Further, the methodological limitations of the study – albeit unavoidable as a result of the world-wide health crisis taking place at the time of data collection – makes it difficult to determine if the limited agreement found here is true, or if the differences are being over-inflated as a result of individual differences that could not be accounted for. The real value of video-teleneuropsychology may not be that it fully replaces in-person assessment, but it may be good-enough to be used within the triage of patients, with a robust in-person assessment being undertaken later if deemed necessary. This has the potential to reduce waiting lists to assessment and avoid unnecessary travel for the patients.

4.4 Chapter Discussion

The use of digital technology within neuropsychological assessment is not a new consideration, however, it is one that is continually evolving. This presents researchers and clinicians with opportunities that may not have previously been accessible technologically or economically. These opportunities encompass how we collect data, where we collect data, and how we analyse it. They can also help to re-address accessibility to neuropsychological healthcare, and how generalisable research findings are across various populations.

Previous operational challenges, such as how data is recorded, could now be addressed through new digital tools. The use of a digital pen to record the drawing process someone uses when copying a complex image is an example of this. As our understanding of cognition and brain function has developed, we have moved away from focusing on modalities of function to a more nuanced understanding of network systems, including white matter function. With this evolution of understanding, the tools we use to illustrate and measure these processes need to be more sensitive. Drawing tasks such as the RCF have

the potential to provide rich and robust information about cognitive functioning. However, this requires detailed recording and scoring of the drawing process an individual adopts. Analogue methods such as the use of coloured pens or template flowcharts have provided a solution in the past, but it is acknowledged that these methods lack sensitivity and operational weaknesses. Digital recording of the drawing process provides superior sensitivity by recording individual pen strokes. It also has the potential to reduce subjectivity within scoring and reduce administration time, with research teams currently developing automated scoring platforms using the data gathered by a digital pen, and which is fed back into an app on a smartphone (e.g., Canham, Smith & Tyrell, 2000; Canham, Smith & Tyrell, 2005; Jung Hyun et al, 2018; and Prange and Sonntag, 2019).

A weakness of research in this area is that the demographic of the participants is often very narrow. The acronym WEIRD acknowledges the bias in scientific research towards a sample pool that is white, educated, industrialized, rich and democratic. Feminist psychology also highlights that much of our foundational understanding of psychology recruited participants from university students that at that time were, in the majority, males in their early 20s (e.g., Pownall and Stainton-Rogers, 2021). Whilst current research is moving away from this, there are still concerns about how generalisable research can be when a population demographic is relatively homogenous. Even in the UK there is a recognised social and education divide between the north and south of the country. Where research studies require direct interaction between a participant and researcher, participant recruitment has typically been opportunistic and thereby limited to those who are local and/or have the means to travel. Limited availability of research funding increases this as researchers are not able to cover travel expenses for participants. Digital technology such as videoconferencing can therefore reduce demographic bias within research studies, by opening up the potential for participants who may not have been able to participate for geographical reasons, not only on a national basis but also internationally. This has the consequence that the recommendations and guidelines that are given to clinical practitioners will also be less biased, and relevant to a wider range of patients. Further, where research examines rare conditions, this can also increase the access to individuals if geography is no longer a barrier to participation, increasing the statistical power, rigour and the strength of research studies.

Within clinical practice, the use of videoconferencing also has the capacity to increase access to neuropsychological healthcare and assessment. Particularly in rural locations, or in large land-mass countries where 'medical deserts' exist, videoconferencing provides an opportunity for individuals to access specialist healthcare. This is a particular

benefit for conditions that may be debilitating or rare and require specialist knowledge, or when there are other health factors, where visiting a hospital or healthcare facility may come with additional health risks.

However, whilst there are many benefits to the use of digital technology within neuropsychological practice, its use should still be adopted with caution. Whilst videoconferencing, for example, may increase accessibility to healthcare for some, for others it may be a barrier. Digital poverty is a concern even in Western countries such as the UK or US, but video assessment guidelines typically advise using a full screen monitor over a smartphone screen – although initial investigation has indicated that the use of a smartphone screen is equivalent to a full screen monitor (Algret et al, 2021). Additionally, the quality of the connection has an impact on the reliability of the assessment, particularly for tests that are once-only or non-repeatable such as letter-number sequencing, as was found within our own study. The high price tags of digital technology are not just prohibitive to the patients but can also be a barrier within healthcare practice. Within the NHS, current budgets are stretched, and the benefit of adopting expensive technology has to significantly outweigh the associated costs.

Further, whilst there has been a boom in the amount of research that has been undertaken to validate the use of digital technology in neuropsychological practice, this is still an under-researched area. Neither of the technologies examined within this chapter have a large body of evidence supporting their use, and what research has been undertaken is limited in its statistical power, the breadth of assessments and populations considered, and the appropriateness of the statistical analysis that has been used. Whilst the findings within this chapter go some way to extend our understanding of the use of digital technology within clinical practice, and in particular alongside the RCF task, this work is not without its own limitations and weaknesses. The sample sizes were intended to be larger, however, the Covid-19 pandemic prevented these targets from being met. Replication of the studies is required before considering adoption within clinical practice.

The participants within these studies had no history of neurological conditions. It is essential to determine if there is variance in suitability of digital methods across clinical populations. These differences may be condition-specific rather than general to all neurological populations, as was found in the Algret et al (2021) study. As such, validity and reliability analysis should be carried out specifically with the clinical populations for whom these tools are intended to be used with.

A limitation of the study is that it was not possible to account for individual differences in the data analysis. Whilst alternative complex figure models exist, such as the Taylor Complex Figure (TCF), studies comparing performance between the two tasks indicates that they are not wholly comparable, with the TCF task demonstrating on average higher recall scores than the RCF task (Tombaugh & Hubley, 1991). As such, adopting the use of different figures within a repeated measures design to avoid practice effects could introduce its own systematic bias. However, whilst accounting for individual differences is a relevant consideration, in the current study it is likely that by using a between-participants design it undervalued the level of agreement between the conditions, making the variance between conditions seem greater than it actually was.

In conclusion, the findings from this chapter provide support for the use of digital technology within neuropsychological practice. However, whilst the Covid-19 pandemic has accelerated interest in this area, this is a much under-researched area that requires more attention, and the current studies should be replicated.

Chapter Five: Rey Complex Figure Drawing process in healthy individuals

5.1 Chapter Premise

The method of measuring the drawing process an individual adopts to complete the RCF has typically emphasised the drawing of five core elements as whole components as a priority over the drawing of the more localised, featural elements. Any deviations from prioritising these elements in the drawing process have been considered to be an indication of differential, and in some cases deficient, cognitive processing. However, as demonstrated in Chapter One, research now indicates that healthy adults may not adhere to this bias as strongly as was previously considered. In addition, whilst several authors have indicated the need to consider the strategy adopted, such as a spatial-temporal pattern, as well as order and wholeness, little guidance has been given on how to record strategy in an objective and measurable way.

The premise of the following chapter is to examine completion of the RCF task in a healthy, adult population. The order in which the 18 elements of the RCF are drawn will be examined, to consider if historically used models of order do represent the drawing characteristics of a healthy, adult population, or to determine if there is an over-exaggeration of the priority of these elements. In addition to Order, Wholeness of drawing each of the elements will be considered, to determine to what extent healthy adults perceive and replicate the core elements as a unit of information. Continuity will be examined as an indication of the extent of fractionalisation that is typically observed. Finally, the chapter will propose an objective means by which to measure drawing Strategy, by examining the patterns in which a healthy adult population draws elements with the same spatial-temporal location.

5.2 Participants

Participants in the current study were those participants who took part in the studies in Chapter Four. Following validation of equivalent performance across the three environmental conditions (digital pen, normal pen, videoconference), all 54 participants were considered to be a single healthy adult population sample for the current study. Forty-four participants were registered as female at birth, and 10 as male. Ages ranged from 22 to 70 years of age, with an average age of 44.61 years (SD 12.48). Whilst there was a greater number of participants within the late 40s and early 50s range, overall, there was a good

spread across the age range, with the exception of those in their later 60s. However, this was not considered a problem, as previous research has indicated older adults to be their own population group, and clinical knowledge would benefit from focused research across the older adults age span. One participant was removed from the analysis due to the video recordings of the drawing process being corrupted. As such, drawing process data was analysed on 53 participants. The sample size of 53 was considered suitable for development of exploratory normative data based upon recommendations by Bridges & Holler (2007), which suggests a sample size between 50 and 75.

IQ of participants ranged from 87 (low average) to 128 (superior), with an average of 112 (high average). Overall, the IQ of the participants within the study were skewed towards a high level of IQ, with relatively few within the low average range. Due to the opportunistic recruitment method used within the study, this skew was unavoidable, though it should be considered when drawing conclusions from the overall study. This spread of IQ is somewhat reflective of the generally high levels of education in participants who volunteered for the study, with 13 individuals completing post-graduate training at either post-graduate diploma or master level, and 2 individuals completing doctoral education. Two-thirds of all participants had completed a bachelor's degree level of education. Handedness of participants was recorded as the hand used to write, and a score of strength of handedness, using the Edinburgh Handedness Inventory (EHI). A total of 45 participants reported being right-handed whilst nine reported being left-handed, notably none of the participants reported writing with both hands. Whilst the EHI scores were recorded in this study, the purpose was to examine a relationship between strength of handedness against RCF drawing scores, and therefore a mean score has not been reported here.

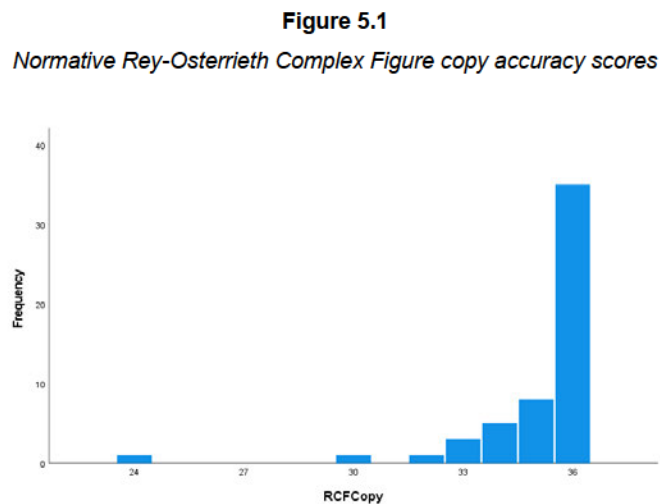
5.3 Protocol and scoring procedure

All participants followed the protocol as per Chapter Two and Appendix One, completing a short form neuropsychological screen which included a copy, a 3-minute and a 30-minute delay phase of the RCF task. Materials and equipment used were dependent upon which group they were part of within the studies in Chapter Four. All RCF replications were scored for accuracy in accordance with the Osterrieth (1944) scoring system and all replications were scored for wholeness, continuity, order, and strategy as per the methodology detailed in Chapter Three. Primary scoring was conducted by the author of

this thesis and the second-rater scoring was completed by two BSc (Hons) Psychology final year students, one for accuracy analyses and one for the drawing process.

5.4 Accuracy scoring population data

Each copy, immediate and delayed recall replication was scored for accuracy in accordance with the Osterrieth (1944) scoring system, observing accuracy of placement of each element and accuracy of replication. For the copy phase, scores ranged from 24 to 36, with 36 being the maximum score possible. Mean score for copy replications was 35.09 (SD 1.98). As illustrated in Figure 5.1, the distribution of scores was heavily skewed towards the maximum range of scores, which is in keeping with known ceiling effects within a healthy adult population.



An examination of the box and whisker plot for the scores indicated two data points were extreme outliers. Whilst the two participants were included within the remainder of the analysis in this chapter, they were not considered for participant matching within the analysis reported in Chapter Seven, given their copy accuracy fell outside typical ranges for the population.

For the immediate and delayed recall phases, accuracy scores fell across a wider range (Immediate: 7 to 35; Delayed: 7 to 34), with mean scores of 21.41 (SD = 5.67) and 21.69 (SD = 5.85) respectively. Distribution of scores (Figure 5.2) for immediate recall was representative of normal distribution, as confirmed by a non-significant one-sample

Kolmogorov-Smirnov test ($t = 0.08$, $p = 0.20$), however, there was deviation from normal distribution within delayed recall scores (Figure 5.3; $t = 0.14$, $p = 0.01$). Skewness towards ceiling scores was not unexpected, and is representative of typical population normative analysis (Caffarra et al, 2002)

Figure 5.2

*Normative Rey-Osterrieth Complex Figure
immediate recall accuracy scores*

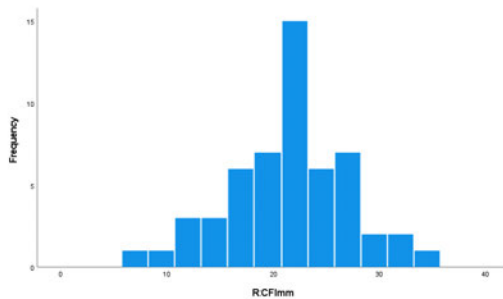
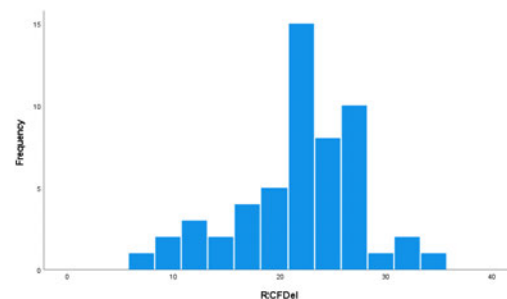


Figure 5.3

*Normative Rey-Osterrieth Complex Figure delayed
recall accuracy scores*



Accuracy scoring for all three stages was carried out by a second scorer to assess inter-rater reliability. Kendall's tau coefficient was used to determine the level of correlation between scorers for the copy phase due to the limited range of scores, and the high level of tied ranks. This was found to be positive, strong ($r^2 = 0.61$) and statistically significant ($p < 0.001$). Pearson's correlation coefficient for immediate and delayed recall scores between the two raters were found to be positive, very strong ($r^2 = 0.96$, $r^2 = 0.963$), and were both statistically significant ($p < 0.001$).

To examine the presence and extent of agreement between scorers, intraclass correlations were also calculated. As the two conditions (raters) were fixed and all participants were scored by both raters, a two-way random effects model was carried out. An absolute agreement model was selected as the purpose of the analysis was to consider agreement between the raters. All three conditions were found to have excellent reliability as there were all over 0.9 (copy ICC = 0.967; immediate ICC = 0.972; delayed ICC = 0.951) and were statistically significant ($p < 0.001$). The level of agreement was considered to be excellent, which agrees with previous literature examining the psychometric properties of RCF accuracy scoring as discussed in Chapter Three. Due to the high level of agreement

between the scores, later analyses using accuracy scores used the scoring carried out by the author of this thesis.

5.5 Wholeness Construct population data

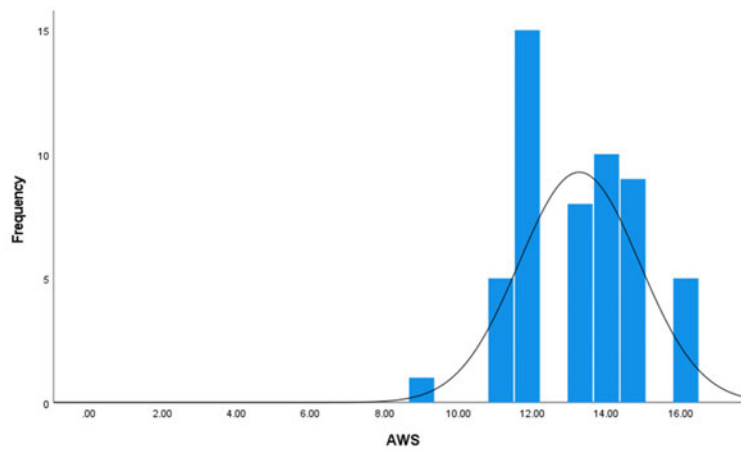
The construct of Wholeness is a novel construct that was derived from the literature review reported in Chapter Three. Wholeness is the extent to which individuals replicate the RCF by drawing its complex elements as whole units. The participants' copy phase drawing process was scored to consider whether objects were perceived and drawn as an individual, whole object. Objects included within the Wholeness construct were all those that had more than a single line. Sixteen of the objects were considered, with the vertical line in the top right-hand corner and the horizontal line above the small rectangle with diagonal not being included (See Figure 3.1). Raw wholeness scores were adjusted to account for how many elements were drawn, to give a more accurate representation of the extent of wholeness with which individuals drew the RCF, and to separate the construct of Wholeness from Accuracy. To calculate the Adjusted Wholeness Score (AWS) the following formula was applied:

$$AWS = x \frac{y}{z}$$

where x represents the raw wholeness scores, y represents the total number of elements in the template image (in this case, 16), and z represents the number of elements that were drawn by the individual. Note that for the copy phase, all but one participant drew all 16 of the elements included with the Wholeness Construct.

Current population AWS ranged from 9 to 16, with a mean score of 13.26 (SD 1.63). A one-sample Kolmogorov-Smirnov test indicated that distribution differed from normal distribution, with a slight negative skew (-0.06) (See Figure 5.4). The distribution of AWS scores indicated that a typical adult with no neurological diagnosis will draw the RCF with some elements *not* drawn as wholes, however, overall, the majority of elements are drawn as whole units. Examining the frequency data of the individual elements indicated that the elements most likely not to be drawn as whole elements were the core rectangle and the diagonal. Only 41.5% of participants drew the rectangle as a whole unit, and only 37.7% drew the diagonal as a whole unit.

Figure 5.4
Normative scores for Adjusted RCF Wholeness



The remaining core elements of the horizontal line, vertical line and sides of the triangle were completed as whole units by the majority of participants – 41 (77.1%), 50 (94.3%) and 43 (81.1%) – though, this is still indicative that a percentage of a healthy, adult population does not adhere as strongly to the global bias as existing scoring systems would suggest. The remaining, more localised elements were almost entirely drawn as complete units (Table 5.1).

The current data indicates that rather than healthy adults perceiving and drawing the five core elements as whole units as typical behaviour, there is a population spread similar to normal distribution, with a population peak of 3 core elements drawn as wholes and drawing of all none or all 5 elements as a whole unit being an extreme behaviour. This differs from most of the previous literature that considers typical behaviour to be skewed towards drawing of all 5 elements as whole. Rather, the current findings demonstrate that performance towards either end of the scale to be extraordinary, as was found by Wilson and Batchelor (2015).

The current population ASW scores were transformed into z-scores to give an indication of what scores would fall within clinically average ranges. A score between 12 and 15 is considered within average ranges (12 = low average, 15 = high average). An ASW score of 11 would be considered borderline, with a score of below 10 being very poor. An individual scoring an ASW of 16 would be within the superior ranges.

Table 5.1
Percentage of elements drawn as whole units

	Element	Whole (%)
Core elements	Rectangle	41.5
	Diagonal	37.7
	Vertical	94.3
	Horizontal	77.4
	Triangle sides	81.1
Local elements	4 lines	96.2
	Small rectangle	64.2
	5 lines	100
	Circle with 3 dots	100
	Vertical cross	100
	Small triangle	79.2
	Diamond	100
	Horizontal cross	88.7
	Small square	67.9
	Vertical in triangle	96.2
	Horizontal in triangle	96.2

A Pearson's correlation coefficient indicates only a small positive relationship between how many of the core elements and the remainder of the elements are drawn as wholes, which was not found to be statistically significant ($r^2 = 0.11$, $p = 0.43$). Inter-rater reliability analysis indicated a very strong correlation between the two raters ($r^2 = 0.90$, $p < 0.001$), and intraclass correlation analysis were found to have excellent agreement (ICC = 0.88, $p < 0.001$). The results indicates that we should not generalise the extent to which an individual adopts a piecemeal approach across the whole of their drawing process based upon whether the core elements are drawn as whole elements or not. Whilst an individual may fragment drawing the large rectangle, this does not necessarily mean that they will fragment the smaller rectangle. An atypical piecemeal approach would present with drawing the localised elements as smaller parts, rather than complex wholes. This provides support for a Wholeness score to include all 16 multi-line elements rather than basing a judgment of wholeness on just the core elements.

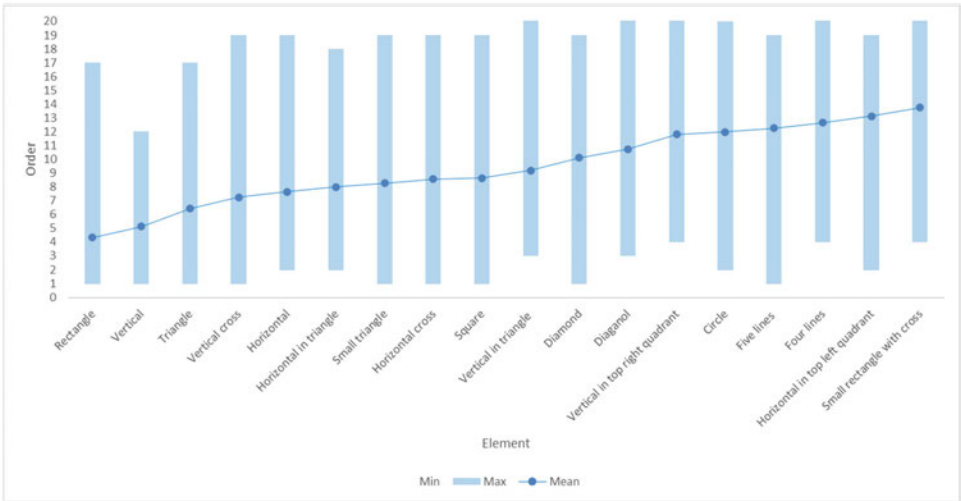
5.6 Order Construct population data

The Order Construct is an index of the order in which each complete element is drawn. It has typically been examined as a combined observation with Wholeness, rather than as a discrete construct. This analysis aimed to establish how much variation and consistency there is within this typical population. The proportion of the sample population who drew the rectangle as the first complete element was analysed. Of the 53 participants, 26.4% (14) completed drawing the large rectangle as the first element. Such findings reflect

the pattern observed by Wilson and Batchelor (2015), who found that only 53% of their participants commenced drawing of the RCF by drawing the rectangle first. In the current population sample, the vertical cross was completed as the first element by 52.8% (28) of the participants, and of these, 15 (28.3%) completed the rectangle as the second element. In total, 52.8% of participants (28) proceeded with drawing the RCF by drawing the rectangle as the first complete element, or immediately after drawing the vertical cross. The data indicated that for almost half of the adult population the rectangle is not drawn as the foundational step in reconstructing the RCF.

Consideration of the elements, the mean order position and the concentration spread of order position for each element was examined and is illustrated in Figure 5.5. All elements presented a wide range of order scores, indicating that there was no absolute order bias, and that there was substantial variation in drawing order across a healthy adult population. That said, there was an indication of order trend for some individual elements. The vertical line was typically completed in the first two thirds of drawing order, with none of the participants completing the element after order position 12. Smaller, localised elements such as the four lines, the vertical in the top right quadrant and the small rectangle with cross tended towards being completed after several other elements had been completed first. However, all elements had a wide range of scores, of at least 10 order locations.

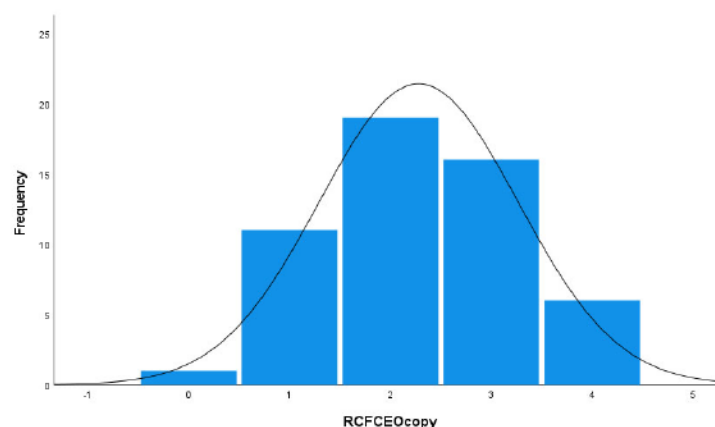
Figure 5.5
Mean and range of order position, in ranked order by mean, from earliest to latest



Of note, whilst the rectangle is not drawn by all participants at the start of the drawing process, the mean order value for the rectangle was closest to the start of the drawing process. Additionally, the vertical and horizontal bisectors and the large triangle mean order scores were also in the top 5 values. The vertical cross mean order score was ranked fourth across all elements, however, it also had one of the widest ranges of scores and 95% confidence intervals. Whilst the vertical cross may have been drawn at the beginning of the drawing process by some participants, it was often drawn by others as the last completed element. Of the five core elements, the fact that the diagonal was not drawn near the start stood out as being at odds with the emphasis that existing scoring systems place on it. Only eight participants drew the diagonal in the first five steps, with peak frequency observed in the third quarter of the drawing process.

Given how many of the scoring systems have historically focused on the order in which the five core elements are completed, this was examined in the current population sample. It was found that none of the participants completed all five of the core elements before moving on to the rest of the figure. The number of core elements drawn within the first five steps ranged from 0 to 4 elements being completed. The mean number of core elements drawn within the first five steps was 2.28 elements (SD 0.99). The distribution of how many of the core elements were drawn as the first five completed elements is indicated in Figure 5.6.

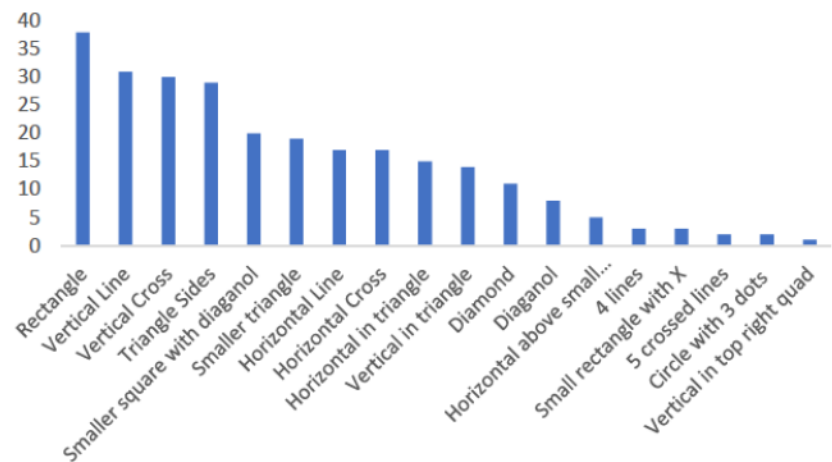
Figure 5.6
Normative scores for RCF Order (core elements only)



As observed with the data within the Wholeness Construct, the peak frequency is 2 core elements, rather than a skewness towards the population drawing the core elements as

the first five completed elements. It is as much an exception as for an individual to complete all five core elements at the start of the drawing as it is for them to start by completing none. All elements were examined to establish which of the core elements were typically not drawn in the first five moves, and to determine if there were other elements that were frequently drawn in the first five steps. Figure 5.7 illustrates the frequency data of how many times each element is drawn within the first 5 moves.

Figure 5.7
Frequency of elements completed within the first five completed elements



As with the mean order values of the individual elements, when looking at the frequency with which the core elements are drawn in the first 5 moves, the rectangle, vertical bisector and triangle sides are in the top five, as is the vertical cross. The diagonal appeared as the 12th element out of 18. Inter-rater reliability analyses were conducted by totalling the order positions of the five core elements across both scorers. Pearson's coefficient indicated a strong linear relationship between the two raters ($r^2 = 0.78$, $p < 0.001$), and a good intraclass correlation ($ICC = 0.78$, $p < 0.001$).

Whilst the data indicates that existing models overemphasise the extent to which the five core elements are drawn at the start of the drawing process, it is clear that the rectangle, vertical and horizontal bisectors and sides of the triangle are still foundational elements of the drawing process. However, these results suggest that it should not be considered a clinical deficiency for individuals to deviate from completing all of these elements at the start. Nevertheless, healthy individuals will typically draw at least some of these towards the start of the drawing process. However, the current evidence also indicates that drawing of the diagonals near the beginning of the drawing process is as likely to not occur as it is to occur.

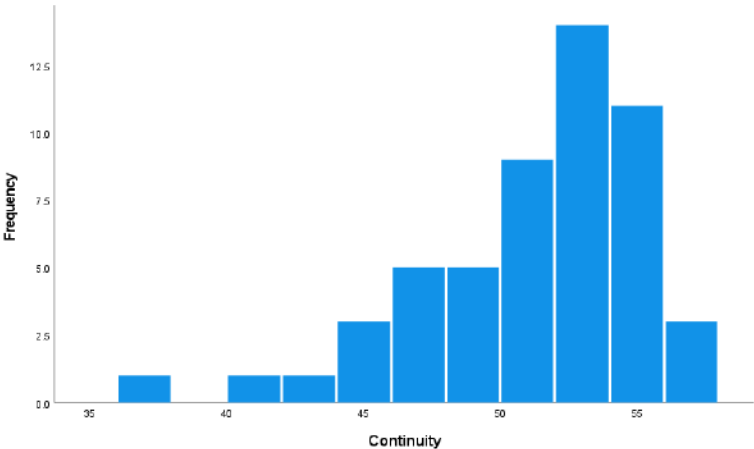
Therefore, not drawing the two bisecting diagonals within the first 5 steps is not necessarily indicative of deficient drawing process.

5.7 Continuity construct population data

Although not addressed by many of the scoring systems, some authors have considered whether participants perceive bisected lines within a larger element as continuous lines, or adopt a piecemeal approach, breaking up larger elements into smaller units. By observing all the points where there are bisecting lines, this gives a range of scores up to 56 data points, providing an indication of the extent of the fractionalisation an individual adopts.

In the current population sample, the smallest number of continuity points observed was 37 (indicating greater fractionalisation), and largest was the maximum of 56. The mean Continuity score was 50.68 (SD = 4.18). Continuity scores indicated a significant negative skew (-1.13, SE 0.33) (Figure 5.8), indicating that typical behaviour in healthy adults observes the majority of continuity points, though not necessarily all of them. This is not unexpected as some of the points where two lines bisect are at the border between elements, such as the bottom left-hand point where the rectangle meets the small square. However, the extent of deviation from drawing through all continuity points on average is greater than these border points, indicating that all healthy adults copy the RCF with some fractionalisation.

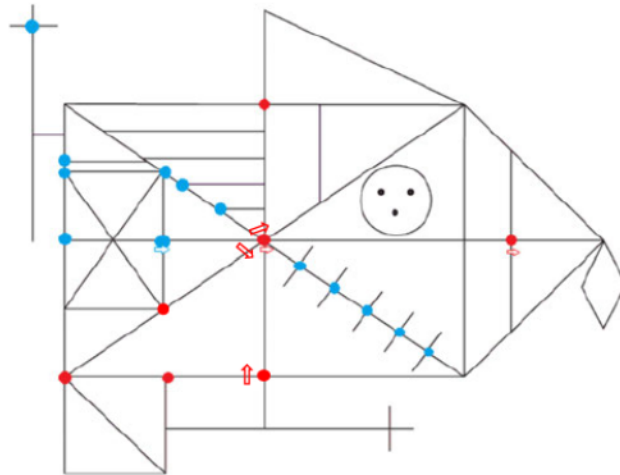
Figure 5.8
Normative scores for RCF Continuation



An examination of the individual points indicates that 19 points (Figure 5.9, blue dots) were drawn through by all of the participants. The four points where the four lines bisect with the central vertical were also drawn by all but one participant, who had significant distortion to the image and did not draw the central vertical. This would suggest individuals who do not draw through these 23 continuity points are performing atypically.

Figure 5.9

Points of continuation drawn through by all participants (blue) and by less than 50% of participants (red)



Nine further points (Fig 5.9, red dots) had a high frequency score in *not* being drawn through, with five of these not being drawn through by 50% or more of the participants. Of these, four are points that cross the border between elements, such as the point where the central rectangle and small square in the lower left corner meet. Of the remaining five points two are where the diagonal lines cross through the centre – as expected given the frequency with which the diagonal is not drawn as a whole element. Given the high percentage of the population sample not drawing through these points of continuity, this indicates that clinical judgements should not be based on individuals not drawing through these lines – rather fractionalisation at these points should be considered typical individual differences.

The continuity data was transformed into z-scores, to give an indication of clinical classifications. Continuity scores within the average range were found to be between 48 and 53, with score above 53 in the above average range, and scores between 45 and 48 low average. Scores between 42 and 45 fall within borderline ranges, and scores below 41 and below are considered very poor. Given these ranges, two participants within the current population fall within the very poor range.

Inter-rater reliability analysis indicated a very strong Pearson's coefficient ($r^2 = 0.92$, $p < 0.001$), and excellent agreement between two scorers ($ICC = 0.89$, $p < 0.001$). Such results indicate a suitable reliability between researchers and clinicians.

5.8 Strategy population data

5.8.1 Principal component analysis

As discussed in Chapter Three, whilst previous systems have provided a means to measure the first three constructs, the construct of Strategy has typically only been instructed as a subjective observation of whether the participant adopts a clear strategy such as moving left to right, top to bottom, or quadrant to quadrant around the image. As noted by authors such as Hamby et al (1993), participants who draw with a clear strategy typically adopt a temporal-spatial method, such that elements that are near each other are drawn at the same time. As no objective measure currently exists to measure whether participants adopt a temporal-spatial strategy, the current population was examined to consider if there were groups of elements that were commonly drawn together. To do this, order data of the individual elements was examined via an exploratory component analysis to determine if groups of elements were typically drawn together.

Principal component analysis (PCA) was used over principal axis factoring (PAF) as the advantages of PCA are that it is more robust when using small sample sizes and where there are collinearity problems within the data. Whilst PAF aims to identify latent factors, the aim of the current analysis was to look to determine if there are chunks of the RCF figure that are consistently used together, which is, to all intents, a reduction in dimensionality. Therefore, the data was analysed using PCA, with oblique rotation applied.

PCA analysis with an oblique rotation indicated that no correlations on the Component Correlation Matrix reached 0.3, therefore an orthogonal rotation was applied. As is considered normal practice in exploratory component analysis, various forms of rotation were applied, including multiple forms of orthogonal rotation give the absence of significant correlations found in the oblique rotation. Oblim, Promax, Varimax, Quartimax and Equamax rotations all indicated the existence of six factors (Table 5.2). Avoiding cross-loading of elements to multiple factors, and in consideration of previous theoretical positioning, an Equimax rotation suggested the best fit of rotation for the data. Some cross

loading was still evident. To address this, it was determined that four factors suggested a more suitable model.

Re-running the PCA analysis with an Equimax rotation restricted to four factors statistically supported the four-factor model (Table 5.3). The four factors were indicated to account for 67.34% of the total variation within the current population. Where the diagonal is positively cross-loaded to two factors, it was determined to associate this with the factor it had the largest loading to – factor four, the Core Structure (see where marked *). Whilst negative loadings within the factors are considered, given that the focus of the intended scoring model was whether elements were drawn together rather than those that are not drawn, any negative cross loading did not have to be associated to just one factor. The only other element that stood outside of what was expected in the factor split was that of the Vertical, which was loaded only to the second factor – internal right-hand side. Given the variation across the rotations which allocated this equally to the Core structure factor, and indications from order data that indicates this is often drawn within the first handful of moves along with other Core structure elements, it was determined that this would form a Core Structure element along with the Rectangle, Diagonal and Horizontal. Further, given the inclusion of the small triangle being drawn alongside the circle in 4 of the five rotations, it was determined to include the small triangle within the Internal Right Hand Side construct (Table 5.4).

Table 5.2
Principal components analysis of RCF Strategy, with rotation variations

Rotation	Factors					
	1	2	3	4	5	6
Oblim (structure) (Oblique)	Triangle Diamond Vert. in tri Horiz. in tri (Diagonal) (Four lines) (Small rect.) (Horiz. In TLQ)	Triangle Diamond Vert. in tri Horiz. in tri (Vertical cross) (Horiz. cross) (Square)	Rectangle Diagonal Horizontal (Small tri)	Five lines Circle Vertical in TRQ	Diagonal Vertical Vertical in TRQ	Circle Small tri Vertical in TRQ
Promax (structure) (Oblique)	Diagonal Four lines Small rectangle Vertical cross Horiz. in TLQ	Triangle Diagonal Horiz. in tri Vert. in tri (Vertical cross) (Square) (Horiz. cross)	Rectangle Diagonal Horizontal (Small tri)	Five lines Circle Vertical in TRQ (Small tri)	Diagonal Vertical Vertical in TRQ (Triangle) (Vertical cross) (Vert. in tri) (Horiz. cross)	Circle Small tri Vertical in TRQ
Varimax (Orthogonal)	Diagonal Four lines Small rectangle Horiz. in TLQ (Triangle) (Diagonal) (Vert. in tri) (Horiz. in tri)	Triangle Horiz. in tri Vert. in tri (Vertical cross) (Square) (Horiz. cross)	Rectangle Diagonal Horizontal (Small tri)	Five lines Circle Vert. in triangle	Diagonal Vertical (Vertical cross) (Vert. in diagonal)	Circle Vert. in TRQ (Diagonal)
Quartimax (Orthogonal)	Diagonal Four lines Small rectangle Vertical cross Horiz. in TLQ (Triangle) (Diamond) (Vert. in tri) (Horiz. in tri)	Triangle Circle Vert. in tri Horiz. in tri (Vertical cross) (Horiz. cross) (Square)	Rectangle Diagonal Horizontal (Small tri)	Five lines Circle	Diagonal Vertical (Vertical cross) (Vert. in tri)	Circle Vert. in tri (Diagonal)
Equimax (Orthogonal)	Four lines Small rectangle Horiz. in TLQ (Triangle) (Diamond) (Horiz. cross) (Vert. in triangle) (Horiz. in tri)	Triangle Diamond Vert. in tri Horiz. in tri (Vert. cross) (Horiz. cross)	Five lines Circle Vert. in TRQ	Rectangle Diagonal Horizontal (Small triangle)	Diagonal Vertical (Vertical cross) (Vert. in tri)	Circle Vert. in tri Small tri (Diagonal)

Table 5.3*Principal components analysis with equimax rotation, restricted to four factors*

Element	Factor			
	Internal LHS	Internal RHS	Core structure	Triangular structure
Four lines	0.839			
Small rectangle	0.925			
Horizontal TLQ	0.920			
Five lines		0.780		
Circle		0.872		
Vertical in TRQ		0.602		
Vertical		0.488		
Rectangle			0.798	
Diagonal	0.336*		0.529	
Horizontal			0.683	
Triangle	-0.558			0.669
Diamond	-0.693			0.348
Horizontal in triangle	-0.643			0.665
Vertical in triangle	-0.500			0.728
Vertical cross		-0.425		-0.588
Horizontal cross				-0.736
Square				-0.882
Small triangle			-0.723	

Table 5.4*RCF Strategy clusters with underlying components*

Strategy construct	Elements included
Core structure	Rectangle Diagonal Horizontal Vertical
Triangular structure	Triangle Horizontal in triangle Vertical in triangle Diamond
Internal Left-hand side	Four lines Rectangle with diagonal Horizontal line in top left quadrant
Internal right-hand side	Five lines Circle Vertical in top right quadrant Small triangle

5.8.2 Strategy construct population data

The four strategy 'clusters' were used to assess each individual participant as to the extent to which their drawing process observed these clusters of elements. Reflecting the method indicated in Geary et al's (2011) paper whereby they investigated the extent of use of semantic clusters in the CVLT, points were awarded each time a participant completed drawing of elements within the same cluster consecutively. For example, when the triangle and diamond were completed consecutively, one point was awarded. When all four elements of the 'Triangle cluster' were drawn consecutively, three points were awarded. A maximum total of 11 Strategy points could be awarded. Due to the fact that not all participants completed all elements, strategy points were adjusted to reflect any missing cluster elements. To calculate the Adjusted Strategy Cluster Score (ASCS), the following formula was applied:

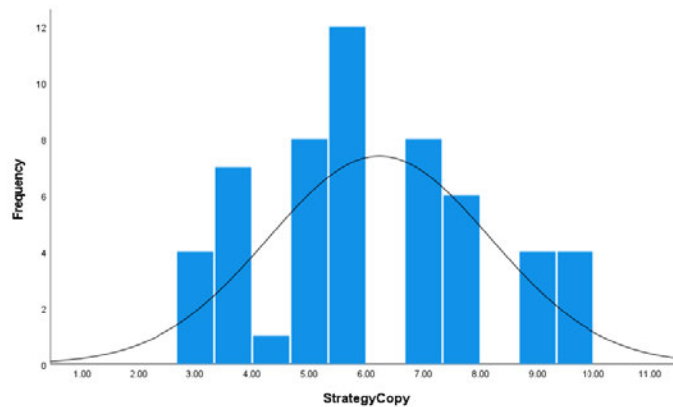
$$ASCS = a \frac{b}{b - c}$$

where a represents the raw Strategy Cluster Score, b represents the maximum total of points that could be awarded all elements are drawn (in this case $b = 11$), and c represents how many of the cluster elements have not been drawn.

The mean ASCS across the population was 6.23 (SD = 1.94), with a lowest score of 2.75 and a highest of 10. The distribution of scores (Figure 5.10) indicated that the majority of healthy adults adopted a temporal-spatial strategy of the four Strategy constructs, but only to an extent. The majority of adult non-TBI, healthy adults scored in the middle ranges. Strict adherence to a temporal-spatial strategy or a very limited adherence, in the context of the four strategy constructs, should be considered extraordinary or atypical performance.

Transforming the scores to standardized z-scores supports the use of clinical descriptors. Converting the ASCS population into z-scores indicated that those average Strategy performance falls between an ASCS score of 4 to 8 (low average = 4, above average = 8). An ASCS of 3 would be classified as borderline and a score of 0-2 as very poor range.

Figure 5.10
Normative scores for RCF Strategy



5.9 Time to completion

Whilst there is little evidence of normative scales of time to completion, and little evidence of use in clinical practice, some of the existing administrative guidelines state a time limit or observation of the time taken to complete the task as an indicator of dysfunctional cognitive processing. Typical time to completion of the copy phase is considered to be between 60 and 120s (Strauss, Sherman & Spreen, 2006), and 180 seconds as a cut-off. In the current study, time to completion data was collected to gather an indication of normative performance, and to determine if any of the constructs or self-report daily challenges inventories correlate with time taken to complete.

Time taken to complete the copy task ranged from 57 seconds to 3 minutes 53 seconds, with an average time of 118.57 seconds (SD 35.96). A one-sample Kolmogorov-Smirnov test indicated that the population spread was close to normality ($t = 0.10$, $p = 0.20$). The current population data indicates that individuals who take less than 1 minute and 8 seconds are performing within the superior ranges. At the other end of the scale, those taking between 2 mins and 49 seconds and 3 minutes 10 seconds are performing in the borderline ranges, with those beyond 3 mins 10s in the very poor ranges. These times reflect the times that are typically considered to be normal performance times to copy the RCF figure.

5.10 Handedness

Data was analysed to determine if there were any differences in drawing process on the RCF between those who were left or right-handed. One-way ANOVA analyses for each of the RCF measures indicated that there were no significant differences between overall scores based on whether an individual was left or right-handed (all $p > 0.05$).

5.11 Demographic factors

As discussed in Chapter Three, demographic characteristics have been examined to determine if any of these have an influence on completion of the RCF. Whilst the majority of these focus on accuracy, some do consider drawing approach. Indications are that age influences memory recall, however, the evidence for influence of age, IQ, years of education and biological sex on drawing process are less clear.

Correlation coefficient analysis was carried out to identify any significant relationships between the various RCF measurements and age, IQ and years of education. No significant linear relationships were found between any of these demographic categories and any of the RCF measures. Independent sample t-tests were carried out to determine if there were significant differences between the biological sexes for any of the RCF scores. Only Strategy score indicated a statistically significant difference ($t(15.39) = -2.22, p = 0.04$) between biological sexes, with males on average scoring a higher Strategy score than females. However, given the limited number of males within the sample population, these results should be considered with caution.

5.12 Correlation between RCF measures and other cognitive measures

A number of significant correlations were found between the RCF scores. Accuracy across all three phases was found to be moderately and positively correlated with Order score, based upon how many of the five core elements were drawn within the first five moves, with a stronger relationship within the two recall phases (Copy $r^2 = 0.27, p = 0.048$; Immediate $r^2 = 0.43, p = 0.001$; Delayed $r^2 = 0.43, p = 0.002$). Continuation scores were also found to be significantly and moderately correlated with the recall phases (immediate: $r^2 = 0.35, p = 0.01$; delayed: $r^2 = 0.34, p = 0.01$). Time taken to complete the copy phase was moderately correlated with copy accuracy, such that the longer an individual took to complete the task, the less accurate their reproduction was ($r^2 = -0.28, p = 0.05$).

Continuation scores were found to significantly correlate with Order ($r^2 = 0.43$, $p = 0.001$), however, not with Wholeness scores. Time taken to complete the copy phase was negatively correlated with Wholeness, measured by AWS, such that the more elements that were drawn as whole units, the quicker the completed the task. Strategy, as measured by the ASCS, was not found to be correlated with any of the other measures of the RCF.

Considering relationships between the RCF measures and other cognitive measures, RCF Copy accuracy was positively correlated with verbal fluency ($r^2 = 0.28$, $p = 0.04$). RCF Immediate recall accuracy was also found to be positively correlated with verbal fluency ($r^2 = 0.294$, $p = 0.03$) and with Category Switching ($r^2 = 0.31$, $p = 0.02$) and HVLT learning ($r^2 = 0.30$, $p = 0.03$). Strategy scores, as measured by ASCS, were found to be significantly correlated with HVLT recall ($r^2 = 0.28$, $p = 0.04$), verbal fluency ($r^2 = 0.284$, $p = 0.04$), letter-number sequence ($r^2 = 0.274$, $p = 0.05$) and Stroop scores ($r^2 = 0.28$, $p = 0.042$).

5.13 Chapter discussion

The premise of this chapter was to examine the performance of a healthy adult population on the RCF task. This included a detailed analysis of the four drawing process constructs identified in Chapter Three, so as to gain a better understanding of the typical drawing processes used to complete the RCF and to develop population normative data for further analyses.

Whilst the focus of this thesis is an examination of the drawing process, the accuracy with which the RCF is replicated is also important. Heavy distortion in the copy phase, and significant differences between the recall stages may be indicative of unknown diverse or deficient neurology, which we sought to control for within the participant pre-screen. Further, the underlying rationale for examining drawing process within an mTBI population is such that individuals are expected to have the same absolute recall of visuo-spatial items as healthy controls, albeit with a different means of recalling. Therefore, it was important to examine this data to understand if the current sample was representative of the non-clinical adult population, and to identify any individuals who would not be suitable matches for the control group in later studies. Accuracy scores in the current population sample across all three phases were found to be comparable to previous understanding of accuracy performance by healthy adults with no neurological diagnoses. Overall, participants performed at, or very near ceiling levels on the copy phase, and with a skew towards the ceiling level for the two recall phases. No significant difference was found between the two

recall phases, and scores between the two phases were found to strongly correlate, as expected within a healthy population. This supports the use of the current population sample as a normative baseline and a control group for later analysis.

Previous literature posits that typical drawing performance on the RCF is predicated upon drawing the five core elements, in particular the large rectangle, at the start of the drawing process, and as whole units. Examination of the two constructs of Wholeness and Order indicated that previous literature overemphasises the extent to which this occurs. Less than half of the sample population were found to draw the central rectangle or the diagonal as whole units, and only a quarter of the population drew the rectangle as the first complete element. Whilst these findings conflict with the majority of scoring systems that emphasise priority and coherent drawing of the core elements, they do support those of Wilson & Batchelor (2015) who indicated that only 53% of their participants commenced by coherently drawing the central rectangle. As was noted by Wilson and Batchelor (2015), Askhoomoff et al (1995), and this current study, a significant percentage of individuals began the drawing process by drawing the vertical cross first - argued to be an artifact of Western left to right reading patterns (Hamby, Wilkins & Barry, 1993). As all participants in the current study were English speaking, it was not possible to validate this theory. This is an important limitation of the current study, as it is not possible to determine the generalisability of the current normative data to other language speaking populations, and in particular those which follow a different written language structure, such as Arabic or East Asian.

Whilst the Order data provides a clear argument against models that prioritise drawing of the rectangle as the beginning point of the drawing process, there is evidence to suggest that some of the core elements do provide a basis from which the rest of the image is drawn – though not necessarily as the foundational step. The rectangle, vertical bisector and triangle sides trended towards being drawn near the start of the drawing process, with all three – alongside the vertical cross – being the elements most frequently drawn in the first five moves. However, the diagonal was only drawn within the first five moves by a small number of participants. This brings into question whether the inclusion of the diagonal's position within existing scoring methods is suitable.

As well as examining the constructs of Order and Wholeness which are the main focus of existing scoring methodology, the current study also examined Continuation. Whilst some scoring methods purport to measure fragmentation/fractionalisation, they are typically measuring whether the core elements are drawn as whole units, and often on an all or nothing basis. Continuation as scored here is a more accurate measurement of

fragmentation as it gives an indication of the extent to which an individual breaks up the image into smaller components. Whilst Wholeness is a measure from a top-down perspective, Continuation measures from a bottom-up perspective, which is theoretically more closely aligned with measuring whether someone is perceiving the image from a localised bias. This is supported by the lack of significant correlation between Continuation and Wholeness scores, indicating that they are measuring different variables.

The Continuation and Wholeness constructs also indicated that reducing drawing process to a frequency-based measure of the extent to which an individual draws the RCF image with strong adherence to Wholeness or Continuation only gives a partial picture. Where the deviation occurs is also important. It could be argued that existing scoring methodology does this by focusing on the core elements, however, the current data indicates that this focus may be misplaced. Whilst drawing of all of the core elements as coherent wholes may not be typical performance, the data indicated that the smaller, localised elements were more likely to be completed as a whole unit before moving on to the next. This suggests that rather than focusing only on whether the core elements are drawn as coherent, whole elements, the extent to which the localised elements are drawn as whole units should also be included. Whilst some fragmenting of the core elements is typical behaviour, fragmenting of the localised elements is not, and may therefore be a better indicator of neurological diversity or dysfunction. With regards to points of continuation, rather than attention being paid to all of the bisectors that form the central structure, it may be more pertinent to pay attention to those 23 points that are typically always drawn through by healthy adults to give an indication of atypical fractionalisation. Importantly, it is not just how much fractionalisation there is, but where it takes place.

Whilst the methods used in this study to measure Order, Wholeness and Continuation were developed from methods used by previous authors, the construct of Strategy used here was a novel development. Though strategy has been discussed by previous authors such as Visser (1972) and Hamby et al, (1993), no scoring systems are available to systematically measure strategy. Taking into consideration the observations that individuals typically adopt a temporal-spatial framework to replicate the RCF, and that non-clinical adults are more likely to use a semantic category framework to recall words in the CVLT (Geary et al, 2011), the current study sought to identify if there were 'clusters' of elements that were typically drawn consecutively. The component analysis identified four 'clusters.'

One of these clusters was formed by core elements, minus the triangle sides, indicating that there is merit in existing scoring systems that are based upon these elements being drawn together. However, the analysis here suggested that this core structure cluster is simply one of four clusters, which should all be considered within the drawing process. By ignoring the extent to which the other clusters are drawn consecutively would be akin to only measuring the animal semantic group when scoring the HVLT and ignoring whether the precious gems and places of habitation were recalled as a group. This provides further support to the argument that too much emphasis is placed upon whether the core elements are being drawn accurately and coherently within existing scoring systems. That said, similar to the other constructs, the current data indicates that Strategy has a distribution close to normal distribution, and that whilst some adherence to the four clusters occurs as typical behaviour, total adherence is indicative of extreme performance.

A number of important patterns were identified between the different RCF, and cognitive measures used within the current study. The Order construct was moderately correlated with accuracy on all three phases and Continuation was found to be moderately correlated with the recall phases. The more core elements drawn in the first five moves and the less fragmentation, the more accurate an individual replicated the RCF. Order and Continuation were also found to be moderately correlated with each other. Such findings were not surprising, as previous research (e.g., Shorr et al, 1992) has indicated that the more structured an individual drew the image, the more they recalled. Theoretical explanations are such that the core elements are used as a framework from which to hang the remainder of the image from, similar to the semantic framework within learning tasks such as the HVLT and CVLT. However, these constructs only provide a partial explanation for drawing recall, and therefore other factors must be involved.

Time taken to complete the task was negatively correlated with both Copy accuracy and Wholeness, such that the longer an individual took to draw the image, the less elements that were drawn coherently, and less accurately. This is not unexpected, as the most time efficient way to reproduce the image is to draw it with as few steps as possible. Time taken may also provide an important indication of the internalised problem solving and planning that is taking place. As correlations are not indicative of direction, it may be that time taken is representative of more effortful executive functioning, and that lower construct scores are a representation of this. Whilst observing the pattern of drawing is indicative of problem solving, planning and task management to an extent, use of 'metatalk' can provide useful information of the internalised processing that is taking place, and general complex problem solving and task management ability. Previous research that examines collaborative

discourse alongside visual problem solving has found that participants with amnesia use more words to discuss and take longer to complete the task than non-amnesic patients, argued to be an indication of increased cognitive effort to tackle the visual task (Duff et al, 2006 and 2008). The use of metatalk alongside RCF task completion provides an interesting avenue for future research to understand complex problem solving, planning and task management.

The Strategy construct, as measured by ASCS, did not correlate with any of the other RCF variables, however, it was found to have a small but significant correlation with HVLT delayed recall. This is a finding of interest, as the use of cues and frameworks, such as the semantic categories in the HVLT task, are argued to underpin more efficient memory consolidation and recall. However, whilst there is a correlation with HVLT delayed recall scores, there is no correlation with RCF recall scores. This might suggest that visual recall is not as well supported by 'chunking' information as verbal recall. This connects in part with literature regarding drawing of the RCF when compared with the Taylor version (TCF). It is suggested that the TCF is easier to verbalise which makes it easier to reproduce and recall than the RCF. Therefore, it is possible that the use of a chunking information is more strongly connected with verbal processing than visual processing. Alternatively, it may be indicative that the structure of strategy developed here does not provide a sufficient explanation for a framework individuals adopt to consolidate and recall the RCF task. Given that the model developed here indicated some of the clusters were highly correlated with one another, this is not an improbable explanation.

Clinically, these findings have implications upon any assessment being made solely based on the drawing process used to complete the RCF task. Importantly, judgement should be made based on clinical normative scales that account for the close to normal distribution of scores across each of the constructs. This includes Wholeness and Order, as assessments should not be made based on whether the core elements are drawn as a priority or coherently. Strong adherence to this should be considered as superior performance and not the norm, particularly with regards to the diagonal. Some deviation from drawing all five of the core elements within the first five moves may be the result of pre-existing, typical processing, and planning, rather than an indication of a deficit caused by a brain abnormality. However, the data indicates that typical drawing process does still move towards the perception of these core elements towards the start of the drawing process, but this is slightly delayed and requires additional processing time, rather than being an immediate observation.

That there were minimal significant correlations between the drawing process constructs indicated that each of these constructs are standalone processes. One of the strengths of this study is that it separates examination of Order and Wholeness from each other. The use of this approach has enabled it to be identified that Order and Wholeness of the core elements have different importance within drawing process in a non-clinical population. Whilst there is a trend towards the core elements – excluding the diagonal bisectors – being drawn near the start of the drawing process, it is less evident that the core elements are drawn as coherent, whole elements. Whilst core elements may be completed near the start of the drawing process, their replication may be fragmented. The lack of significant correlation between the two constructs also indicates that strong observation of Order is not predicated on observing Wholeness, and vice versa.

Reliability between clinicians is a key characteristic for any tool that is going to be used for diagnostic purposes. The inter-rater reliability analysis across the drawing process constructs supported the robustness of the scoring system between different raters. This supports the use of the four constructs as a model within clinical practice, and that equivalent scores would be recorded by different clinicians. Whilst this reliability analysis is restricted to inter-rater reliability, for the purposes of this exploratory study, this is sufficient to support further investigation and development of formalised normative sample tables, upon which more extensive reliability and validity analysis could be carried out.

As important as understanding reliability of the scoring of the constructs is determine whether demographic factors have an influence upon performance. Previous research has indicated that age influences RCF recall performance, which has subsequently been taken into account with the widely accepted recall norms. Influence of demographic characteristics upon drawing process are more variable, as addressed in Chapter 3. The current study found no significant correlations between age, estimated IQ or years of education with any of the RCF measures, including recall. It is important to recognise the skewness of estimated IQ scores in the current population, with a high incidence of those who would be classified in the above average and superior ranges. More extensive normative data scales would need to include a wider breadth of estimated IQ levels and years of education, particularly if the intention is to use these with a TBI population where there is a bias towards individuals with lower IQ ranges (Lefevre-Dognin et al, 2020) experiencing TBI more frequently. There was a good spread across the age range between 18 to 60 years of age, however, future normative development would need to examine older adults more closely, as there is some evidence to indicate differences in drawing strategy in older adults, both as a result of normal cognitive aging, and accelerated neurodegeneration. In this study, no differences between

biological sex at birth were found across the measures, with the exception of Strategy, where males were found to score with a higher Strategy score than females. Given the limited number of males within the current population sample, this should be reassessed with a more even biological sex-based split.

In summary, the current study has illustrated the presence of four different constructs which describe the drawing process individuals use to replicate the RCF image. It has also highlighted the need for more robust and inclusive normative data tables to be developed across the constructs to account for the close to normal distribution of population scores. The study indicates that drawing performance of the constructs reflects normal distribution, rather than strict adherence to observing wholeness, order, continuation, and strategy.

Chapter Six – Cognitive, psychological & demographic predictors of post-concussive symptoms

6.1 Chapter Premise

This chapter moves the focus of this thesis to individuals who have experienced an mTBI at least three months earlier. The debate in the literature that underlies this thesis is whether or not there are long-term cognitive deficits following a mTBI. As evidenced by literature previously presented, unquestionably there are some individuals who experience significant challenges in their day-to-day lives beyond the typical three-month recovery window. However, the aetiology of these symptoms and challenges remains unclear, due to limited evidence of cognitive deficits following mTBI. Those who argue cognitive dysfunction underpins these complaints posit that the methodology and methods used to measure and analyse cognitive performance are flawed. How we understand and measure cognitive dysfunction following mTBI was addressed in the previous chapters. This chapter will address the statistical limitations of previous literature.

The limitations of using group differences methodology to evidence cognitive change following mTBI will be explored. Regression analysis methodology was used to examine which measures of cognitive function, psychological status and demographic factors are the best predictors of post-concussive symptoms and behavioural challenges. The chapter also considers if observations of RCF drawing process can support the diagnosis of post-concussive syndrome. The chapter closes by considering whether or not there is evidence to suggest there are cognitive deficits that underly the daily challenges that some individuals continue to report some months – and even years – after their mTBI.

6.2 Introduction

Whether there is objective clinical evidence of the cognitive dysfunction that is subjectively reported by individuals following mTBI is a long-standing debate within clinical neuropsychology. One of the challenges is the limited power of studies due to small sample sizes. This can – and has been – addressed through the use of meta-analyses. A meta-analysis by Binder, Rohling and Larrabee (1997) indicated that whilst some statistical group differences existed, these were clinically nonsignificant. They acknowledged there was evidence of an association between mTBI and cognitive dysfunction, however, they argued

that other factors contributed to this such as litigation involvement and pre-existing low-level functioning. Later, the team corroborated their conclusions, re-analysing the original report (Rohling et al, 2011). Their results were supported by Schretlen & Shapiro (2003). Schretlen and Shapiro's analysis indicated no differences between individuals with mTBI and their control group one-month after injury. Frenchman, Fox, and Mayberry (2005) also found no significant cognitive differences between mTBI and control groups in the post-acute phase, nor did Belanger, Curtiss, Demery, Lebowitz & Vanderploeg (2005).

These conclusions have been challenged by others. Pertab, James and Bigler (2009) argued that previous studies treated all mTBI cases as a homogenous group, regardless of recovery status. They posited that the sub-group of individuals still reporting symptoms should be analysed separately from those who recovered. Their re-analysis of Binder and colleagues' 1997 work indicated that when the symptomatic sub-group were considered separately, cognitive dysfunction was more apparent. That reduced performance of a small, symptomatic sub-group is masked by the recovered majority was supported by Iverson (2010). Iverson presented real and hypothetical evidence to illustrate that statistical mechanisms reduce the visibility of the sub-group cognitive deficits when grouped together with a recovered majority. Whilst Belanger and colleagues (2005) found limited evidence of cognitive deficits, they challenged the suggestion that involvement in responsibility litigation was a cause for symptom exaggeration. Participants legally seeking compensation were found to present with neuropsychological deficits. However, the direction of this relationship is unclear – maybe those individuals are seeking support because their symptoms and complaints are more severe?

However, determining who is symptomatic and who is not is problematic. There is a lack of clarity over a cut-off point for symptom screens such as the RPCSQ. Some suggest a score of 16 as the indicator between someone who is symptomatic and who is not (Thompson et al, 2016). For others, a more conservative cut-off of thirty-five is posited (De Guisse et al, 2016). Such methodological issues illustrate the complexities in reaching an agreed and validated conclusion regarding cognitive impact.

Definitions of who does and does not have PCS aside, the issue of majority-minority masking can be addressed by analysing the symptomatic group separately. However, this does not address the limitations of using a group differences design within neuropsychological studies. As Binder and colleagues (1997) highlighted, individual baseline performance should be considered. Though, whilst they purported this as an argument for long-term cognitive deficits not existing, others argue that not accounting for

baseline performance can lead to false-negatives as well as false-positives. For an individual with pre-existing high functioning, presenting within clinically average performance can be an 'invisible' reduction in function (Brooks, Holdnack & Iverson, 2011). Demographic matching can reduce between group effects of co-variants such as IQ, education, and age. However, this does not address within group differences when considering consolidated absolute group performance.

Individual differences can be accounted for by considering individual change, rather than averaged cross-sectional performance. This approach is commonly used within sports-related concussion research and practice (e.g., Collie et al, 2004; Moser et al, 2007). Pre-season neuropsychological screening gives a baseline from which to judge recovery status. However, in public healthcare settings this is typically not available. An estimate of pre-morbid functioning has to be made, accounting for variables such as age, IQ and education, as well as considering patient history.

This subjective judgement, however, is challenged by another phenomenon of neuropsychological assessment - the probability of a healthy adult recording a low score on tasks. Typical normative scales are based upon the 'bell curve,' or univariate Gaussian distribution. Whilst most healthy individuals perform within close range to the mean, there will be some who score very high or very low. The use of percentile rankings to make clinical judgements is suitable for each test in isolation. However, when multiple assessments are delivered – as is standard practice – statistical error is introduced, and the probability of recording low scores increases (Brooks, Holdnack & Iverson, 2011; Brooks, Iverson & Holdnack, 2013). This increases the probability of a false-positive conclusion, by overestimating the levels of cognitive deficits. Principles of multivariate base rates indicate both patient and protocol factors play a role (Brooks, Holdnack & Iverson, 2011). This argues for a move away from examining group-based differences; and to consider individual performance in the context of the specific set of neuropsychological tasks being used.

Recent research into cognitive outcomes of mTBI has moved away from group statistics, and towards the use of regression modelling (McCrea et al, 2005). Regression analysis allows for the inclusion of demographic data such as IQ, years of education and age within the modelling, so as to capture individual differences in pre-morbid functioning. It also supports identification of the variables that are indicators of post-concussive syndrome and behavioural challenges following mTBI.

Draper & Ponsford (2008) used logistic regression analysis to examine the association between cognitive performance and injury severity. Their findings indicated cognitive dysfunction in individuals with mTBI several years after injury. In later TBI work, they found that cognitive dysfunction was associated with poorer functional performance (Ponsford, Draper & Schonberger, 2008). Processing speed, working memory and executive function most strongly predicted functional outcome. Barker-Collo et al (2015) also identified that at 12-months post injury, processing speed and executive function were strong predictors of post-concussive symptoms and real-life challenges, measured by the RPCSQ and CFS. These findings were supported by a systematic review which indicated that cognitive function mediates return to work rates after mTBI (Mani, Cater & Hudlikar, 2017). Within veteran populations, regression analysis has been used to identify the predictors of functional disability of those with mTBI and PTSD (Stika et al, 2021). Regression analysis has also indicated processing speed correlates with subjective fatigue reports— the most commonly reported persisting symptom of mTBI (Johansson, Berglund & Ronnback, 2009).

However, there are limitations in what we can understand from previous literature about cognitive dysfunction following mTBI in a general adult population. Understanding of recovery from mTBI has been vastly informed by sport-related research. However, time since injury is typically short-term focused, considering days and weeks rather than months or years. Research is conducted with an emphasis on returning to play. Further, research has evidenced underreporting of symptoms, so as to return to play quicker (e.g., Meier et al, 2015). Additionally, a significant body of sports-based research is conducted with paediatric and adolescent populations, where the developmental trajectory of the brain and particularly the pre-frontal cortex has to be factored in. Research within military personnel and veterans has provided valuable insights, however, injury mechanics are often blast-related, which are not typically experienced by a general population. Further, additional psychiatric factors such as PTSD can confound results (Johansson, Berglund & Ronnback, 2009).

Research has been conducted with a general adult population - by Ponsford and colleagues, for example. However, work examines TBI across the severity spectrum, rather than focusing on mTBI, or only includes moderate to severe injuries only. As Ponsford and colleagues found, severity influences both cognitive function and functional outcome, and influencing factors vary across injury severity (Draper & Ponsford, 2008; Ponsford, Draper & Schonberger, 2008). Regression research conducted with a purely mTBI population has typically considered the influence of psychological factors, such as depression and anxiety, on cognitive function (e.g., Pineau, Marchand & Guay, 2015), or it seeks to determine acute-

phase risk factors of recovery (e.g., Caplain, Blancho, Marqu & Aghakhani, 2017). Therefore, there is a need to consider which cognitive, psychological, and demographic factors predict chronic symptoms and behavioural challenges of mTBI only, in a general, adult population.

The current study sought to examine which demographic, cognitive and psychological variables were the best predictors of long-term post-concussive symptoms and behavioural challenges following mTBI. Variables were included as potential predictors of PCS symptoms, as measured by the RPCSQ; of behavioural challenges, measured by the Cognitive Failures Scale (CFS) and Dysexecutive Questionnaire (DEX); and of cognitive fatigue, as measured by the Mental Fatigue Scale (MFS).

6.3 Study design

The study design was a multiple linear regression analysis, aimed to identify the most influential variables that predict four outcome measures of mTBI. It also sought to determine if constructs of RCF drawing process predicted mTBI outcomes. The four outcome measures were the RPCSQ, the DEX, the CFS and the MFS (Appendices 5-8). Whilst studies exist which consider the variable profile underpinning each of these measures, no studies examine them all consecutively, or considers cognitive, psychological, and demographic variables together. By examining each of them, it can be considered if different variables underpin different complaints. This aligns with psychometric studies that indicate PCS to be an umbrella term for different sub-types or dimensions (Keshavan, 1981; Levin et al; 1987; Martua, Lumba-Brown & Ghajar, 2018; Vos et al, 2019).

The RPCSQ was selected as a measure of severity of post-concussive symptoms. Whilst other studies measure severity based upon PTA, LOC and/or GCS scores, these were not suitable for this study. As recruitment was carried out via the general public rather than through healthcare services, GCS scores were not available. Further, not all participants sought emergency care, where GCS would be recorded. With regards to PTA and LOC, the majority of participants did not experience either. For those who did, this was less than 5 minutes. Therefore, there was little variability upon which to analyse severity. Additionally, evidence has indicated that clinical severity classification via these acute stage measures does not necessarily correlate to severity of symptoms in the post-acute and chronic phases (Barr, 2014; Hawryluk & Manley, 2015). Further, the measures of PTA, LOC and GCS fluctuate at the scene, in the ambulance and in the hospital; as well as inter-rater variation, impacting the reliability of this data (Bledsoe et al, 2015).

The DEX and CFS are indices of behavioural problems in everyday life. The CFS addresses subjective attention and memory-based challenges. The DEX considers subjective reports of executive function-based challenges. Executive function, memory and processing speed have been identified previously in the literature as the cognitive functions most sensitive to mTBI (Frenchman, Fox & Maybery, 2005; Karr, Areshenkoff & Carcia-Barrera, 2014;). Both the DEX and CFS have been used in research to understand outcomes of mTBI. The DEX has been evidenced to be a sensitive measure for evaluating impact of mTBI on real-life outcomes of everyday functioning (Bar-Haim Erez, Rothschild, Katz, Tuchner & Hartman-Maeir 2009). Regression analysis has indicated reduced processing speed and executive function predict scores on the CFS 12-months post injury (Barker-Collo et al, 2015).

The MFS is an index of mental and cognitive fatigue. Fatigue is reported as the most common and most persisting symptom of mTBI and PCS (Lannsjö, Geijerstam, Johansson, Bring & Borg, 2009; Roe, Sveen, Alvsaker & Autz-Holter, 2009; Moller et al, 2019). Whilst the aetiology of mental fatigue following mTBI remains unclear, research has indicated that processing speed correlates with fatigue in mild and moderate TBI (Johansson, 2021). Whilst the RPCSQ includes fatigue as an item, the debilitating nature of cognitive fatigue and our limited understanding of it warrants examination in closer detail.

Predictive variables included within the modelling analysis were cognitive, psychological, and demographic (Table 6.1). As indicated throughout this thesis, cognitive function has been indicated to underpin PCS symptoms and behavioural challenges. Cognitive measures were those that formed the protocol screen as detailed in Chapter Two (2.5). This included scoring of the RCF constructs as examined earlier in this thesis (chapters three and five). Psychological factors included measures of anxiety and depression due to the known influence of both factors on cognitive function. Anxiety was measured via the HADS, and depression via the HADS and BDI-II (Appendices 3 and 4). Demographic variables were also included, as research demonstrates age, IQ, education, and registered sex influence cognitive performance and mTBI recovery (Ponsford, Draper & Schonberger, 2008; Fraser, Downing, Biernacki, McKensie, & Ponsford, 2019).

Table 6.1
Predicting variables included within regression analysis

Variable category	Predictive variable
Cognitive (n = 18)	HVLT learning
	HVLT delayed recall
	HVLT recognition
	HVLT recall strategy
	RCF copy
	RCF immediate recall
	RCF delayed recall
	RCF wholeness (copy)
	RCF order (copy)
	RCF continuation (copy)
	RCF strategy (copy)
	RCF time (copy)
	FAS
	Animals
	Category switching
	TMT A
	TMT B
	Stroop
	Letter-number sequencing
Psychological	HADS Anxiety
	HADS Depression
	BDI – II
Demographic	Age
	Years of education
	IQ
	Birth registered sex

The study recruited individuals who had experience of mTBI at least 3-months earlier. Definition of mTBI is addressed in Chapter Two (2.4.3). Matched non-TBI control participants were included within the analysis, in line with methodology used by Draper and Ponsford (2008). Inclusion of healthy controls was adopted as it is argued that patterns of impairment can be more clearly attributable to injury rather than other factors when controls are included. It also allows for group difference comparisons on outcome measures, to validate increased symptoms and challenges in the mTBI population.

The research question for the study was ‘which cognitive, psychological and demographic factors predict symptoms and real-life challenges following mTBI?’ The first hypothesis was that participants in the mTBI group would report greater post-concussive symptoms and real-life challenges compared to the control group. The second hypothesis was that executive function, processing speed, working memory, mental health status, age and years of education would be the strongest predictors of PCS symptoms and real-life complaints. The final hypothesis was that different variables would predict each of the outcome measures.

6.4 Description of participants

Sixteen mTBI participants were recruited to the study. All mTBI participants met the inclusion criteria of PTA and LOC described in Chapter Two (2.4.3). At least 3 months had passed since their injury. A further sixteen non-TBI controls were included from the group of participants in the studies in Chapter Four who took part via videoconferencing, so as to mirror the conditions between the TBI and control participants. Control participants were matched on sex, age, and IQ as closely as possible to the mTBI participants. All thirty-two participants met the exclusion and inclusion criteria set out in Chapter Two (2.4)

The intended sample size for the study was 100-120 individuals with a history of mTBI. Such a sample size would reflect Green's (1991) rule of thumb that when testing importance of predictors, the sample size should be at least 104 plus the number of predictor variables. Due to the on-going social-distancing restrictions, and the suspension of many community and non-emergency health services, recruitment was significantly impacted, and the intended sample size was not achieved. However, it is commonly recognised that there are varying opinions on the minimum sample size for meaningful regression analysis. A recommendation by Jenkins and Quintana-Ascencio (2020) suggests a minimum sample size of twenty-five. Haenlein & Kaplan (2010) stipulate that the number of participants should at least be more than the number of variables. Both of these guidelines have been met with a sample of 32.

The age range of participants ranged between 20 to 69 years, with a mean age of 43.59 years. Average years of education was 16.94 years. Average estimated IQ was 114.81. Independent samples t-tests indicated no statistically significant differences in age, years of education or estimated IQ between the mTBI and control groups (age: $t(30) = -0.53, p = 0.60$; education: $t(30) = 0.13, p = 0.90$; IQ: $t(30) = -0.82, p = 0.42$; equal variance assumed for all). Twenty-two participants were female, and ten were male. Whilst there was an overall imbalance between males and females, the split was the same in the two groups.

6.5 Brain imaging report

Whilst participants did not undergo neuroimaging as part of the protocol, most participants in the mTBI group anecdotally offered information on whether brain imaging was done at the time of their incident. Of the sixteen mTBI participants, only five reported having

CT scans taken at the time of their injury; all of which were reported to be unremarkable. The lack of evidence of neurological damage via CT is consistent with prior literature.

6.6 Procedures

All participants completed the neuropsychological protocol detailed in Chapter Two. Given the ongoing social distancing restrictions, all participants took part via video-teleconferencing. Adaptions were made to the protocol administration as reported in section 2.8. All statistical analysis was carried out via SPSS.

Due to a corrupted video for one of the mTBI participants, it was not possible to record drawing process scores for that individual. As a result of bandwidth issues, it was not possible to record an accurate score for one control participant on the Stroop task, and one control and one mTBI participant on the letter-number sequencing task. One participant in the mTBI group misunderstood the list to be recalled in the HVLT delayed recall phase. Missing scores were substituted with the mean score across all other participants. The addressing of missing data points through the use of the total sample mean is standard practice in regression analysis (Kang, 2013)

6.7 Outcome measure descriptives and inferentials

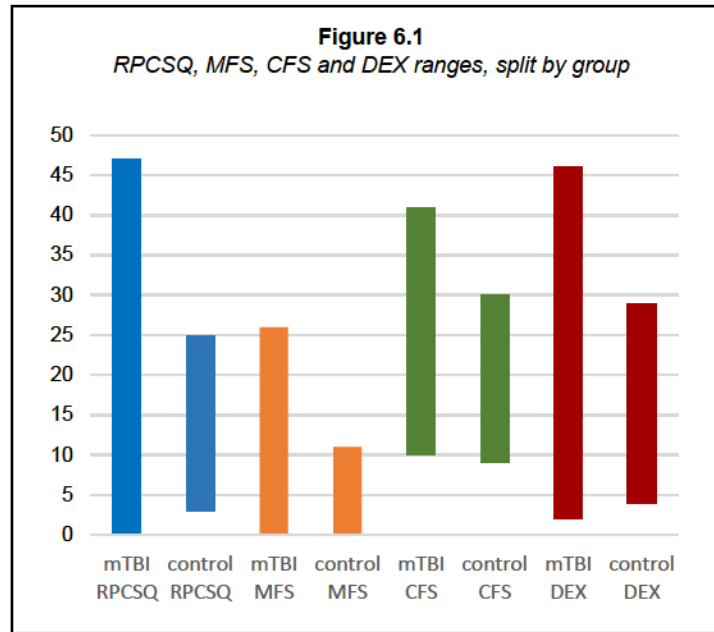
Independent sample t-tests were conducted to examine if there were differences in reported outcomes between the mTBI and control group. Results indicated statistically significant differences in RPCSQ and MFS scores (RPCSQ: $t(19.89) = 2.343, p = 0.03$; MFS: $t(20.40) = 2.31, p = 0.03$). Participants in the mTBI group reported more post-concussive symptoms and greater mental fatigue compared to the control group (Table 6.2). Whilst the mTBI group reported higher scores than the control group on both the CFS and DEX this did not reach statistical significance (CFS: $t(30) = 1.83, p = 0.08$; DEX: $t(22.82) = 1.01, p = 0.33$).

Variance of scores between groups for the RPCSQ, MFS and DEX were significantly different, as indicated by the Levene's test, with the mTBI group indicating a wider range of scores than the control group (Figure 6.1). Across all four outcome measures, the mTBI group presented an extended tail of higher scores.

Table 6.2
Group mean and standard deviations for the RPCSQ, MFS, CFS and DEX

Outcome	mTBI		Control	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
RPCSQ	21.13	15.11	11.56	6.19
MFS	10.69	8.83	5.13	3.81
CFS	22.00	8.09	17.56	5.32
DEX	17.38	12.53	13.81	6.65

M – mean. *SD* – standard deviation.



6.8 Cognitive predictor variable descriptives and inferentials

Basic descriptive statistics were calculated for each of the potential cognitive predictor variables (Table 6.3). RCF Wholeness and Category Switching were found to have statistically significant differences between the mTBI and control groups. For RCF Wholeness (copy), individuals with mTBI on average drew more elements ($m = 14.75$, $SD = 1.39$) as wholes than the control group ($m = 13.38$, $SD = 1.82$). In the category switching task, the mTBI group produced less words ($m = 15.31$, $SD = 2.63$) than the control group ($m = 17.94$, $SD = 3.64$). The Cohen's d for RCF Wholeness ($d = 0.85$) and Category Switching ($d = 0.84$) indicated large effect sizes. No other cognitive measures were found to have significant mean differences. However, HVLT Learning and RCF time scores were found to have statistically different variance between the groups. The mTBI group had a wider range of RCF time scores than the control group. The control group had a wider range of HVLT learning scores than the mTBI group.

Table 6.3
Mean, standard deviation, and independent t-test analysis for cognitive predictor variables, split by group

Variable	mTBI		Control		Total		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
RCF Copy	34.94	1.12	35.50	0.89	35.22	1.04	-1.57	0.13
RCF Immediate	23.00	4.99	21.25	6.78	22.13	5.92	0.83	0.41
RCF Delayed	23.13	4.86	21.41	6.94	22.27	5.96	0.81	0.42
RCF Wholeness (copy)	14.75	1.39	13.38	1.82	14.06	1.74	2.40	0.02*
RCF Order (copy)	2.94	1.24	2.63	1.20	2.78	1.21	0.72	0.48
RCF Continuity (copy)	50.94	3.59	51.12	4.67	51.03	4.10	-0.13	0.90
RCF Strategy (copy)	6.96	2.57	6.52	2.22	6.74	2.38	0.52	0.61
RCF Time (copy) (s)	132.00	59.32	129.56	31.77	130.78	46.83	0.15	0.89
HVLT Learn	30.19	3.08	30.63	4.79	30.41	3.97	-0.31	0.76
HVLT Recall	11.04	1.39	10.38	1.71	10.71	1.57	1.22	0.23
HVLT Recognition	11.38	0.81	11.38	1.09	11.38	0.94	0.00	1.00
HVLT Strategy (recall)	4.08	2.33	3.08	2.34	3.58	2.35	1.21	0.24
FAS	45.31	14.78	49.63	12.80	47.47	13.78	-0.88	0.39
Animals	25.56	4.29	24.38	4.95	24.97	4.60	0.73	0.47
Category Switching	15.31	2.63	17.94	3.64	16.63	3.40	-2.34	0.03*
TMT A (s)	24.94	7.82	30.13	9.82	27.53	9.12	-1.65	0.11
TMA B (s)	53.06	20.21	57.88	28.18	55.47	24.25	-0.56	0.58
Stroop	98.75	25.07	108.00	8.72	103.38	19.05	-1.39	0.17
Letter Number Sequence	13.19	3.21	13.38	3.34	16.63	3.40	-0.16	0.87

M – mean. *SD* – standard deviation. *t* – t-test statistic. *p* – significance value

* statistically significant at $p < 0.05$

6.9 Psychological predictor variable descriptives and inferentials

Basic descriptives and inferentials were also calculated for the psychological predictor variables, as measured by the HADS and BDI-II (Table 6.4). No statistically significant differences were found between the mTBI and control groups on any of the measures of anxiety or depression.

Table 6.4
Mean, standard deviation, and independent t-test analysis for psychological predictor variables, split by group

Variable	mTBI		Control		Total		<i>t</i>	<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
HADS Anxiety	5.75	3.82	6.56	3.20	6.16	3.49	-0.52	0.52
HADS Depression	4.06	3.235	4.25	3.49	4.16	3.31	-0.16	0.88
BDI - II	10.56	10.90	9.38	7.89	9.97	9.38	0.14	0.73

M – mean. *SD* – standard deviation. *t* – t-test statistic. *p* – significance value

6.10 Variable correlation coefficient analysis

Pearson's correlation coefficients were calculated across all outcome and variable measures. All predictor variables significantly correlated with at least one outcome variable,

except for Animals. All outcome measures significantly correlated with at least one predictor variable.

6.10.1 Predictor correlations with outcome measures

The four outcome measures were found to significantly correlated with one another (Table 6.5). All relationships were positive and had large strength. The greater a participant scored on one outcome measure, the more likely they would score highly on the other measures.

Table 6.5
Pearson's correlation coefficients and significance between outcome measures

Variable	RPCSQ		CFS		DEX	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
CFS	0.62	0.000				
DEX	0.59	0.000	0.68	0.000		
MFS	0.80	0.000	0.78	0.000	0.73	0.000

r – Pearson's correlation coefficient. *p* – significance value

Psychological predictor variables were also found to significantly correlate with the four outcome measures, with the exception of the CFS and HADS anxiety measures (Table 6.6). All correlations were positive, and moderate to strong in strength. The greater an individual scored for either anxiety or depression, the greater they would score on each of the outcome measures, and vice versa. For the BDI-II, all correlations were strong, whilst for HADS depression they were moderate, indicating the BDI-II is a more sensitive measure of the depression with regards to mTBI outcomes.

Table 6.6
Pearson's correlation coefficients and significance between outcome measures and psychological predictor variables

Variable	HADS anxiety		HADS depression		BDI - II	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
RPCSQ	0.44	0.01	0.62	0.000	0.69	0.000
CFS			0.42	0.02	0.71	0.000
DEX	0.48	0.01	0.38	0.03	0.65	0.000
MFS	0.43	0.01	0.52	0.002	0.80	0.000

r – Pearson's correlation coefficient. *p* – significance value

Whilst anxiety and depression were consistently found to be strongly related to mTBI outcome measures, only a handful of cognitive measures were found to do so (Table 6.7). Cognitive measures found to have a significant correlation coefficient with any of the outcome measures were RCF Copy (accuracy), Stroop, Category Switching and the FAS

verbal fluency tasks. All were found to be negative, indicating that the higher an individual scored on the outcome measures, the poorer their cognitive performance would be on those tasks.

The RPCSQ was found to have a moderate correlation with both the Stroop and Category Switching Tasks, as did the MFS. The CFS was indicated to have a moderate correlation with the RCF copy task. The DEX was found to moderately correlate with the FAS verbal fluency task.

Table 6.7
Pearson's correlation coefficients and significance between outcome measures and cognitive predictor variables

Variable	Stroop		Cat. Switch.		RCF Copy		FAS	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
RPCSQ	-0.54	0.001	-0.37	0.04				
CFS					-0.39	0.03		
DEX							-0.38	0.03
MFS	-0.50	0.003	-0.44	0.01				

r – Pearson's correlation coefficient. *p* – significance value

Demographic predictor variables were also found to reflect some significant correlation coefficients with mTBI outcome measures. Age was negatively correlated with the DEX ($r^2 = -0.42$, $p = 0.02$), indicating that the older an individual gets, the lower their DEX score. Estimated IQ was found to negatively correlate with RPCSQ scores ($r^2 = 0.39$, $p = 0.03$). Individuals with a higher IQ were indicated to report less RPCSQ symptoms.

Sex-based differences in outcome measures were examined via independent sample t-tests. No significant differences were found for the RPCSQ, DEX or MFS (all $p > 0.05$). However, a significant difference was found between men and women for the CFS ($t(29.97) = 2.13$, $p = 0.04$). Men were found to report fewer cognitive failures ($m = 16.80$, $SD = 3.62$) than women ($m = 21.14$, $SD = 7.92$). The variance in scores between men and women was found to significantly differ, with women presenting with a wide score range scoring than men.

6.10.2 Correlation coefficients between predictor variables

Correlation coefficients between the predictor variables were examined. Age was found to have a significant negative correlation with years of education ($r^2 = -0.41$, $p = 0.02$). Older participants typically reported fewer years in education. This is likely an artefact of societal changes in Further and Higher Education, with greater numbers completing

undergraduate degrees more recently (The Commons Library, 2022). No other demographic variables were found to significantly correlate with each other. An independent samples t-test indicated a significant difference in IQ between men and women ($t(30) = -2.76, p = 0.01$), with a higher average estimated IQ in men ($m = 119.90, SD = 18.37$) than women ($m = 112.50, SD = 6.38$). No sex-based differences were found in either chronological age or years of education.

Only estimated IQ was found to significantly correlate with any psychological variables. Both the HADS depression and BDI-II were negatively and moderately correlated with IQ (HADS: $r^2 = -0.58, p = 0.001$; BDI-II: $r^2 = -0.45, p = 0.01$). As estimated IQ increased, reports of depression decreased.

All four demographic variables were found to have some significant correlations with cognitive variables (Table 6.8). Age was only found to correlate with RCF measures. There were positive correlations of Age with RCF copy and continuation scores. Older participants were found to copy the RCF with greater accuracy, and with greater continuation observed. Age was negatively correlated with time taken to complete the copy phase, such that older participants took longer to complete the task.

Table 6.8
Pearson's correlation coefficients and significance between demographic and cognitive variables

Variable	Age		Education		IQ	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
RCF Copy	0.40	0.02				
RCF Cont.	0.61	0.04	-0.35	0.05		
RCF Time	-0.38	0.03			-0.42	0.02
HVLT Recall					0.37	0.04
HVLT Recog.			0.41	0.02		
FAS					0.63	0.000
Cat. Switch.					0.45	0.000

r – Pearson's correlation coefficient. *p* – significance value

Education was negatively and moderately correlated with RCF Continuation. The more years of education, the fewer continuation points were observed. Years of education were also positively and moderately correlated with HVLT recognition scores – the more years in education, the more words correctly recognised. IQ positively correlated with HVLT recall, FAS verbal fluency and Category switching. The relationship was moderate for HVLT recall and Category switching, and strong for FAS verbal fluency. IQ was also negatively

and moderately correlated with RCF Time; those with higher estimated IQ were quicker to copy the RCF.

A number of sex-based differences were found in cognitive variables (Table 6.9). Men were found to score higher than women on HVLT recall, HVLT strategy and TMT B. Sex-based differences were also found on both depression indices. Women, on average, reported more symptoms of depression than men.

Table 6.9
Mean, standard deviation, and independent t-test analysis for predictor variables, split by sex

Variable	Male		Female		Total	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>t</i>	<i>p</i>
HVLT recall	1.77	0.49	10.23	1.66	-4.00	0.000
HVLT strategy	4.78	2.05	3.03	2.31	2.22	0.04
TMT B	42.20	12.37	61.50	26.08	2.2	0.04
HADS depression	2.30	2.63	5.00	3.30	2.28	0.03
BDI-II	4.60	3.17	12.41	10.28	3.24	0.003

M – mean. *SD* – standard deviation. *t* – t-test statistic. *p* – significance value

Psychological variables were found to have strong positive correlations with one another, as expected (Table 6.10). Only two significant correlations were found between psychological and cognitive variables. Both Depression measures were negatively and moderately correlated with FAS verbal fluency. Individuals reporting higher levels of depression produced fewer words in the FAS task.

Table 6.10
Pearson's correlation coefficients and significant between psychological and cognitive variables

Variable	HADS depression		BDI-II		FAS	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
HADS anxiety	0.52	0.002	0.59	0.000		
HADS depression					-0.42	0.02
BDI-II	0.79	0.000			-0.40	0.03

r – Pearson's correlation coefficient. *p* – significance value

Finally, correlation coefficients between the cognitive measures were considered. A number of non-RCF tasks were found to significantly correlate with one another. HVLT learning had a positive, moderate correlation with letter-number sequencing ($r^2 = 3.78$, $p = 0.03$). Individuals scoring highly on HVLT learning were found to score highly on letter-number sequencing. HVLT recall and HVLT strategy were positively and strongly correlated ($r^2 = 0.50$, $p = 0.004$). Individuals adopting a strong semantic strategy recalled more words from the HVLT word list. A strong, positive correlation was found between FAS verbal

fluency and Category Switching ($r^2 = 0.65$, $p = 0.000$). Individuals producing more words on the FAS task typically produced more words in the Category Switching task. The two trail-making phases were also found to be strongly and positively correlated with each other ($r^2 = 0.54$, $p = 0.002$). Individuals who took longer to complete the numbers only phase (TMT A) also took longer to complete the letter-number phase (TMT B).

Some non-RCF tasks were found to have significant correlations with RCF measures. HVLT recall was moderately and positively correlated with RCF immediate recall ($r^2 = 0.36$, $p = 0.04$). However, no significant correlation was found between HVLT recall and RCF delayed recall. TMT A was found to negatively correlate with both RCF recall phases (Immediate: $r^2 = -0.41$, $p = 0.02$; Delayed: $r^2 = -0.38$, $p = 0.03$). Individuals taking longer to complete TMT A typically recalled less on the RCF recall trials. Lastly, RCF continuation was positively and moderately correlated with the letter-number sequencing task ($r^2 = 0.38$, $p = 0.03$). Individuals who copied the RCF image with less fractioning were found to re-sequence larger strings of digits.

A number of significant correlation coefficients were also found between the RCF measures (Table 6.11). RCF copy accuracy was not found to be correlated to any other RCF measures. As expected, the two recall phases correlated positively with each other and strongly, with an almost perfect correlation. Both recall phases were also strongly and positively correlated with Wholeness scores. Individuals who recalled the RCF with more elements drawn as whole units were found to have stronger RCF recall after 3 and 30 minutes. RCF delayed recall also had a strong, positive correlation with RCF order. Those who observed a greater number of core elements as the first five completed elements typically had stronger delayed recall. RCF Wholeness was found to be positively correlated to RCF continuity, as was RCF order, though the relationship was stronger with Wholeness than Order. Order was moderately and positively correlated with RCF strategy – those who observed a strong cluster strategy were found to complete more of the core elements in the first five moves. Finally, RCF time to completion of the copy phase was negatively correlated with RCF continuity. Individuals who adopted a more piecemeal approach typically took longer to complete the task.

Table 6.11
Pearson's correlation coefficients and significance between RCF measures

Variable	RCF imm.		RCF delay		RCF whole		RCF order		RCF cont.	
	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>	<i>r</i>	<i>p</i>
RCF delay	0.96	0.000								
RCF whole.	0.54	0.002	0.58	0.001						
RCF order			0.53	0.002	0.59	0.000				
RCF cont.	0.53	0.002			0.68	0.000	0.40	0.02		
RCF strategy							0.49	0.01		
RCF time									-0.47	0.01

r – Pearson's correlation coefficient. *p* – significance value

6.11 Predictor variable screening

Unnecessary predictor variables were screened for and removed from the regression models. Where there is a large number of variables, removing unnecessary variables reduces the risk of multicollinearity and weakening of the predictive model. The correlation coefficients between the predictor variables and outcome measures were examined. Where a coefficient was less than 0.1 – the lower cut-off for a small strength correlation according to Cohen (1988) – the variable was not included in model for that outcome measure.

6.12 Regression analysis assumptions

Each of the assumptions of multiple linear regression analysis were reviewed for each outcome measure before results of the regression were interpreted. Assumptions were reviewed on each model after minimally correlated variables were removed. Where assumptions were violated, remedial action was taken as appropriate; regression models re-run; and assumptions re-checked.

6.12.1 Variable characteristics

The first assumption of linear regression is that the dependent variable, or outcome, is continuous data. Each of the four outcomes measures met this requirement. The second assumption is that there are more than two independent, or predictor variables. The current modelling had thirty-two potential predictors and therefore met this assumption.

6.12.2 Independence of observations

Independence of observations is the assumption that performance of one participant is not influenced by another. Given the design of the study, theoretically there was

independence of observations. This was statistically supported with Durbin-Watson tests, where a score above 2 is indicative of independence.

6.12.3 Linearity

Collective linearity between the dependent and independent variables was examined by plotting studentized residuals and unstandardized predictive values against one another. Linearity was found to be observed for all outcome measures. Individual linearity between outcome measures and each of the predictor variables was also examined. All predictor variables were found to have linearity with the outcome measures.

6.12.4 Homoscedasticity

Visual examination of the plot of studentized residuals versus unstandardized predictive values indicated homoscedasticity for each outcome model.

6.12.5 Multicollinearity

Two correlation coefficients were found to be greater than 0.7. As two measures of depression were included which strongly correlated, one was dropped. BDI-II was indicated to be more sensitive to the outcome measures, therefore HADS depression was removed from the regression. The second strong correlation was between RCF immediate and delayed recall. The strength and significance of the correlation between the two recall stages remained consistent when the groups were considered separately. This correlates with our understanding that where recall between the two phases is not consistent, this is an indicator of a neurological or pathological concern, such as dementia (e.g., Cherrier et al, 1999). Further, there is no indication of differences in recall in mTBI (Geary et al, 2011). Therefore, one of the RCF recall variables was removed from the regression. Examination of Tolerance/VIF values indicated a larger value for RCF immediate, therefore this was removed from the analysis. Upon re-running of the regression for each outcome with these variables removed, no issues of multicollinearity were found.

Other regression analyses such as Partial Least Squares which are more robust with regards to multicollinearity were considered. However, partial least squares regression reduces dimensionality of variables, whereas the goal of this analysis was to consider the impact of the individual predictor variables on the different outcomes. Stepwise regression has been argued to manage multicollinearity through its iterative process of removing weak

variables. However, stepwise regression has been argued to overfit models and theoretically sound variables are not included within the model, due to the nature of the sample used in the regression. Therefore, multiple linear regression was selected as the best method of analysis for the purpose of this study.

6.12.6 Outliers

Cook's distances were examined for each outcome measure. One case was found to be an influential outlier for both the RPCSQ and MFS, who had a history of mTBI. As the individual was within the mTBI group where some variation from normal performance is expected, all cases were included within the regression.

6.12.7 Normality of residuals

Examination of standardized residuals and studentized residuals via P-P and Q-Q plots, respectively, indicated approximate normal distribution of the residuals.

6.13 Multiple linear regression analysis

Multiple linear regression analysis was conducted to examine if there were differences in regression models for each of the outcome measures. Variables, as indicated in Table 6.1, were input into each model after variables with very limited correlation or multicollinearity were removed. Variables were entered into the model using the "Enter" method.

6.13.1 Rivermead post-concussion symptoms questionnaire

Seventeen predictor variables were included in the model for the RPCSQ (Table 6.12.) All four demographic variables, and anxiety and depression were included, along with eleven of the nineteen cognitive variables. The model was indicated to be statistically significant ($F(17,14) = 2.90, p = 0.03$). R^2 for the overall model was 77.9%, with an adjusted R^2 of 51.1%, indicating a large effect size. The model indicated that 51.1% of the variance in the RPCSQ was explained by variance in the seventeen predictors included.

The analysis indicated that the Stroop task was the most influential predictor of RPCSQ scores ($\beta = -0.42$). The relationship indicated that individuals with a higher RPCSQ typically

scored lower on the Stroop task. The Stroop task was also found to be a statistically significant predictor ($t = -2.67$, $p = 0.02$). No other variables were found to be significant predictors.

Table 6.12
Multiple linear regression model with RPCSQ as the outcome measure, in order from most influential to least influential variable

Variable	B	β	t	p
Constant	112.56		1.45	0.17
Stroop	-0.27	-0.42	-2.67	0.02
HVLT recall	2.51	0.32	1.59	0.14
Sex	-7.65	-0.29	-1.39	0.19
<i>BDI-II</i>	0.36	0.27	1.10	0.29
RCF order	-2.06	-0.20	-1.04	0.32
<i>HADS Anxiety</i>	0.71	0.20	0.92	0.37
RCF copy	-2.33	-0.20	-1.08	0.30
RCF delay	0.40	0.19	0.98	0.34
Age	-0.20	-0.19	-0.94	0.37
HVLT strategy	-0.69	-0.13	-0.75	0.46
Cat. Switching	-0.46	-0.13	-0.66	0.52
FAS	-0.10	-0.11	-0.44	0.67
<i>TMT B</i>	0.06	0.11	0.63	0.54
RCF strategy	0.49	0.09	0.56	0.58
L-N	-0.28	-0.07	-0.43	0.67
IQ	0.09	0.06	0.27	0.79
Education	-0.20	-0.04	-0.24	0.82

Variables in italics indicate measures where a higher score is indicative of poorer performance

Two RCF measures were found to be in the top five most influential variables⁷ – RCF order and copy accuracy (both $\beta = -0.20$). The model indicated that individuals with less accurate copy replications and who completed fewer core elements near the start report higher RPCSQ scores. Sex was also indicated to be one of the most influential predictors, with males predicted to report less symptoms than women ($\beta = -0.29$, males to females). Psychological variables were also very influential predictors for RPCSQ scores, with higher scores of depression ($\beta = 0.27$) and anxiety ($\beta = 0.20$) predicting higher reporting of post-concussive symptoms. HVLT delayed recall was the second most influential variable ($\beta = 0.32$). The model indicated that individuals with greater delayed recall ($\beta = 0.32$) would report more RPCSQ symptoms. Of the demographic variables, whilst all four were included within the model, IQ and education were found to have the least influence on RPCSQ scores of all the variables.

⁷ Three variables shared a ranked score, with a β of +/- 0.20. All three variables were counted as in the top 'five' strongest predictors.

6.13.2 Mental Fatigue Scale

Eighteen predictor variables were included within the model for the MFS (Table 6.13). As with the model for RPCSQ, all four demographic variables and the two psychological variables were included. Twelve cognitive variables were included in the model. R^2 for the model was 95.6%, whilst the adjusted R^2 was 74.9%. The model indicated that variance in the variables explained 74.9% of the variance in self-report fatigue scores. The model was indicated to be statistically significant ($F(18,13) = 6.13, p = 0.001$).

Table 6.13
Multiple linear regression model with MFS as the outcome measure, in order from most influential to least influential variable

Variable	B	β	t	p
Constant	26.91		0.84	0.42
<i>BDI-II</i>	0.43	0.56	3.15	0.01
Stroop	-0.11	-0.29	-2.53	0.02
Cat. Switch	-0.59	-0.28	-1.93	0.08
Sex	-4.02	-0.26	-1.64	0.12
RCF order	-1.54	-0.26	-1.84	0.09
IQ	0.20	0.21	1.34	0.20
HVLT recall	0.99	0.21	1.63	0.13
RCF strategy	0.58	0.19	1.56	0.14
Education	-0.48	-0.17	-1.28	0.22
RCF delay	0.20	0.17	0.98	0.35
RCT cont.	-0.24	-0.14	-0.79	0.45
<i>TMT B</i>	0.04	0.13	1.06	0.31
Age	-0.07	-0.10	-0.77	0.45
L-N sequence	-0.14	-0.06	-0.45	0.66
<i>HADS Anxiety</i>	0.12	0.06	0.38	0.71
RCF Copy	-0.38	-0.05	-0.40	0.69
Animals	0.06	0.04	0.35	0.74
FAS	0.00	-0.005	-0.03	0.98

Variables in italics indicate measures where a higher score is indicative of poorer performance

Depression, as measured by the BDI-II, was indicated to be the most influential predictor of MFS scores ($\beta = 0.56$), and was found to be a statistically significant predictor ($t = 3.15, p = 0.01$). The Stroop task was also statistically significant predictor ($t = -2.53, p = 0.02$), and was the second most influential predictor ($\beta = -0.29$). Individuals scoring more highly for depression symptoms were predicted to report more fatigue symptoms, as were those who scored lower on the Stroop task. No other variables were found to be statistically significant individual predictors.

One RCF variable was in the top five most influential variables⁸. As with the RPCSQ model, Individuals who drew more of the core elements as the first five completed elements were predicted to report less mental fatigue ($\beta = -0.26$). The Category Switching task was also one of the top five most influential variables of MFS. Individuals producing more words in the task were predicted to score lower on the MFS ($\beta = -0.28$). Also mirroring the RPCSQ model, women were predicted to report more symptoms of mental fatigue ($\beta = -0.26$). In addition, IQ was found to be in the top five most influential predictors, with individuals with higher IQ predicted to report higher mental fatigue symptoms ($\beta = 0.21$). The least influential variables for predicting mental fatigue were FAS verbal fluency ($\beta = -0.005$), Animals semantic fluency ($\beta = 0.04$) and RCF copy ($\beta = -0.05$), though the direction of influence remained the same as with the RPCSQ.

6.13.3 Dysexecutive Questionnaire

For the DEX, fifteen variables were included within the model (Table 6.14). Both psychological variables were included, however, only two demographic variables were indicated to have a correlation with DEX scores – Age and estimated IQ. The model was found to be statistically significant ($F(15,16) = 2.92, p = 0.02$). Variance in the variables were found to explain 48.2% of the variance in DEX scores ($R^2 = 73.3$, adjusted $R^2 = 48.2\%$). Whilst the overall model was statistically significant, none of the predictor variables were individually significant.

Table 6.14
Multiple linear regression model with DEX as the outcome measure, in order from most influential to least influential variable

Variable	B	β	<i>t</i>	<i>p</i>
Constant	-41.04		-0.59	0.56
<i>BDI-II</i>	0.47	0.44	1.91	0.07
IQ	0.47	0.36	1.77	0.10
RCF cont	0.74	0.30	1.36	0.19
FAS	-0.21	-0.28	-1.08	0.29
Age	-0.24	-0.28	-1.41	0.18
<i>RCF time</i>	0.06	0.27	-1.17	0.26
L-N Sequence	-0.84	-0.27	-1.33	0.20
HVLT learn	0.67	0.26	1.45	0.17
RCF order	-1.62	-0.20	-1.19	0.25
<i>HADS anxiety</i>	0.47	0.16	0.76	0.46
<i>TMT B</i>	0.06	0.14	0.71	0.49
<i>TMT A</i>	-0.14	-0.13	-0.66	0.52
RCF copy	-1.02	-0.11	-0.53	0.60
Cat. Switch	0.18	0.06	0.28	0.78
Stroop	-0.01	-0.02	-0.10	0.92

Variables in italics indicate measures where a higher score is indicative of poorer performance

⁸ Two variables shared a rank score, with a β of -0.26. Both variables were counted as in the top 'five' strongest predictors.

As with the model for mental fatigue, the BDI-II was indicated to be the most influential variable ($\beta = 0.44$). Great reporting of depression symptoms predicts higher DEX scores following mTBI. Both IQ and Age were indicated to be top five most influential⁹ predictors of DEX scores ($\beta = 0.36$; $\beta = -0.28$, respectively). Older participants were predicted to report lower DEX scores, as were individuals with higher IQ. RCF variables were also indicated to be two of the most influential predictor variables. Greater observance of RCF continuation points – a less piecemeal approach – predicted higher DEX scores, ($\beta = 0.30$), as did taking longer to copy the figure (RCF time: $\beta = 0.27$). One other cognitive measure was also found in the most predictor variables. FAS verbal fluency ($\beta = -0.28$) was indicated to predict DEX scores such that fewer words produced on the FAS task predicted higher DEX scores. In the reverse of the RPCSQ and MFS, the Stroop task was found to be the least influential of the included variables ($\beta = -0.02$), with Category switching as the second least influential ($\beta = 0.06$).

6.13.4 Cognitive Failures Questionnaire

The final model examined predictors of the CFS. Ten predictor variables were included in the model (Table 6.15). Of the demographic variables, IQ and Sex were included. Both psychological variables were included, as were nine cognitive variables. The R^2 for the model was 77.7%, and the adjusted R^2 , 61.6%. The model was found to be statistically significant ($F(13,18) = 4.825$, $p = 0.001$), and indicated that variance in the variables explained 61.6% of the variance of CFS scores, indicating a strong effect size.

Table 6.15
Multiple linear regression model with CFS as the outcome measure, in order from most influential to least influential variable

Variable	B	β	<i>t</i>	<i>p</i>
Constant	68.82		1.73	0.10
<i>BDI-II</i>	0.66	0.87	4.76	0.000
<i>RCF time</i>	-0.06	-0.36	-2.21	0.04
<i>HADS anxiety</i>	-0.69	-0.34	-2.19	0.04
Stroop	-0.08	-0.21	-1.46	0.16
RCF copy	-1.37	-0.20	-1.52	0.15
Sex	2.95	0.20	1.28	0.22
HVLT recog.	1.24	0.16	1.33	0.20
RCF order	-0.63	-0.11	-0.84	0.42
IQ	-0.09	-0.09	-0.50	0.62
Cat. switch	0.17	0.08	0.44	0.67
<i>TMT B</i>	0.02	0.08	0.62	0.54
HVLT learn	0.10	0.05	0.40	0.70
FAS	-0.02	-0.04	-0.17	0.87

Variables in italics indicate measures where a higher score is indicative of poorer performance

⁹ Four variables shared two ranked scores, with a β of -0.28 and +/-0.27. All four were counted as top 'five' predictors.

Again, the BDI-II was found to be the most influential predictor ($\beta = 0.87$). Higher reporting of depressive symptoms predicted higher CFS scores. BDI-II was indicated to be a statistically significant predictor ($t = 4.76, p < 0.001$), as were RCF time ($t = -2.21, p = 0.04$) and HADS anxiety ($t = -2.19, p = 0.04$). RCF time and RCF copy were both indicated to be two of the most influential predictors in the model ($\beta = -0.36, \beta = -0.20$). Those drawing more accurate copies of the image were predicted to report lower CFS scores. An unexpected result was that individuals taking less time to copy the image reported higher CFS scores. In terms of demographic variables, men were predicted to report higher CFS scores compared to women. FAS verbal fluency ($\beta = -0.04$) and HVLT learning ($\beta = 0.05$) were indicated to be the least influential variables on CFS scores.

6.14 Chapter discussion

The aim of the current study was to examine which cognitive, psychological, and demographic variables best predict experience of long-term symptoms and behavioural challenges following mTBI. In particular, it considered whether measures of RCF drawing process predicted these outcome measures. This was operationalised by conducting a predictive multiple linear regression analysis for four mTBI outcome measures. The four outcome measures examined post-concussive symptoms, mental fatigue, and behavioural challenges, as measured by the RPCSQ, MFS, DEX and CFS. A separate regression analysis was conducted for each outcome measure. The analyses were conducted with data from individuals who had experienced mTBI at least three months previously, and demographically matched controls. All models were found to be statistically significant, and with large effect sizes, indicating that the underlying variables were strong indicators of post-concussive symptoms and behavioural challenges.

Three hypotheses were set. The first stated that participants in the mTBI group would report greater post-concussive symptoms and behavioural challenges compared to the control group. Independent samples t-tests were conducted for each outcome measure. The mTBI group were found to report significantly more post-concussive symptoms overall, and mental fatigue symptoms specifically. However, whilst the mTBI group on average reported more behavioural challenges, as measured by the DEX and CFS, these differences were not statistically significant. This is not a surprising finding. The RPCSQ was specifically designed to capture the experience of PCS symptoms following mTBI, and fatigue is commonly reported to be the most prevalent and persisting of these symptoms (Lannsjo, Geijerstam, Johansson, Bring & Borg, 2009; Roe, Sveen, Alvsaker & Autz-Holter, 2009; Moller et al, 2019). Therefore, observing a significant difference between the two

groups on the RPCSQ and MFS was expected. However, CFS and DEX are general measures of behavioural challenges in everyday life and may not be as sensitive to mTBI differences. The results support those of Zargar et al (2015) who also did not find a significant difference in CFS scores between an mTBI and control group.

Whilst there were no statistical differences between the means for the DEX and CFS, the range of scores on the DEX in the mTBI group was found to be significantly wider than the control group. These results support Bar-Haim Erez et al's (2009), work who supported the DEX as a sensitive measure of evaluating real-life impact of mTBI. However, this was not the case for the CFS, where the variance was statistically equal. The findings from the current study support that the hypothesis - that individuals with mTBI report greater experience of post-concussive symptoms and real-life challenges than individuals with no mTBI history - should be accepted. However, it is also indicated that the CFS is a less sensitive measure of problematic outcomes following mTBI, compared to the RPCSQ, MFS and DEX.

The second and third hypotheses considered which of the variables were the most sensitive predictors of the outcome measures. The second hypothesis proposed that executive function, processing speed, working memory, mental health status, age and years of education would be the strongest predictors of PCS symptoms and real-life complaints. The third hypothesis stated that different variables would predict each of the outcome measures. Both hypotheses were examined via multiple linear regression.

Differences were found in the models across the four outcome measures. This took the form, firstly, of what was included in the model, and secondly, which of the variables were the most influential. As a large number of predicting variables in a regression model can increase the risk of multicollinearity, and cause a model to be less specific, variables were screened for inclusion in each model. This was based upon their individual correlation coefficient with the outcome measures. Variation was found in what was included in the model. Firstly, there was a distinction between what was included in the RPCSQ and MFS models versus what was included in the DEX and CFS. The RPCSQ and MFS included recall tasks, but not learning. The DEX and CFS reflected the opposite, including learning but not recall. The RPCSQ and MFS included all four demographic variables, whilst the DEX and CFS dropped years of education. DEX included age, but not sex, where CFS was the reverse. Finally, the RPCSQ and MFS included task strategy measures, whereas neither the DEX nor CFS did. The correlation coefficients indicate a difference between

post-concussive symptoms and behavioural challenges in the variables that significantly relate to these self-report symptom inventories.

These differences were examined further via the models that resulted from the regression analysis for each outcome measure. Similar to the correlation coefficients, there were differences in the influence demographic factors had on the outcome measures. The number of years of education is taken into account on several standardised neuropsychological assessment standardised scores. Therefore, it was included as a possible predicting variable of post-concussive symptoms and behavioural challenges. However, the regression analysis indicated that the influence of education has a different impact depending on which outcome measure is being considered. As discussed in the previous paragraph, education was minimally correlated with behavioural challenges, and not included in the models for the DEX and CFS. Further, whilst education was included in the modelling for both RPCSQ and MFS, for the RPCSQ, it was found to be the least influential predictor in the model. This differed from the MFS, where education was a moderate predictor. Years of education are indicated to be a factor of post-concussive symptoms, however, the results suggest that this is more specific to the symptom of fatigue, than general PCS symptoms and behavioural challenges.

These results were reflected to an extent when considering IQ. Unlike years of education, IQ was included in the model for all four outcome measures. However, as before, the influence of IQ had different magnitudes. For MFS and DEX, IQ was one of the most influential predictors. For CFS, it was one of the least influential variables, and the second least influential for RPCSQ scores. As with years of education, the regression indicated IQ to be one of the most influential predictors for fatigue, but not for PCS symptoms generally. This is an interesting finding as it indicates firstly that IQ plays a significant role in mental fatigue symptoms, but not other PCS symptoms. It also suggests that experience of mental fatigue may be more prominent in individuals with higher IQ. This does not necessarily mean, that individuals with lower IQ do not have mental fatigue challenges. But the change may be perceived more greatly in those with higher IQ. This has important implications on how we take into account base rates when understanding the long-term symptoms of PCS, both clinically and within research. Lack of reporting of mental fatigue symptoms does not necessarily mean a full recovery. Reported levels of fatigue have to be considered with consideration of years of education and IQ.

In addition to the variations that were found in which cognitive tasks were included in the models, there were differences in which tasks were most influential for each outcome

measure. The RPCSQ and MFS were similar in which cognitive variables were included in the model, and which variables were the most influential. For both, the Stroop task was a very influential and significant predictor, being the most influential for the RPCSQ, and second for the MFS. HVLT recall was also an important predictive variable, however, this was more so for RPCSQ scores than MFS scores. For the MFS, Category switching was the third most influential variable – thought it was only a moderate variable for RPCSQ scores. For the CFS, the Stroop task was also one of the most influential variables. However, for the DEX it was one of the least influential, along with category switching. For the DEX, the FAS verbal fluency and letter-number sequencing tasks were the most influential. This indicates a clear pattern that mental fatigue, and PCS symptoms more generally, are underpinned by different cognitive factors than other functional challenges that are captured by the CFS and DEX.

This pattern is further supported when looking which RCF variables were included within each of the regression models. Again, the RPCSQ and MFS reflected a similar structure – with RCF order scores being highly predictive for both. RCF copy accuracy scores were also found to be an important predictor of RPCSQ scores. For both the DEX and CFS, RCF time was a highly influential variable. RCF continuation scores were also important with regards to the DEX. These scores further support that there is variation in differences in underlying cognitive function for the different outcome measures. It also indicates the importance of including a measure of RCF drawing process, and one which separates order from completeness of the objects.

The positive correlation between RPCSQ and MFS scores with HVLT delayed recall was unexpected. Though effect sizes are usually very small, typically impaired memory is correlated with higher reports of PCS symptoms. However, the model suggested that better verbal recall performance predicts increased reporting of PCS and fatigue symptoms. Though, it is important to recognise that there was skewness in recall performance, with the majority of individuals performing near ceiling levels, and with low variance of scores overall. The limited variation and near ceiling performance across both groups correlates with previous work by Geary et al (2011), who evidenced that there is little difference in absolute delayed recall scores between individuals with mTBI and controls.

Whilst the direction of influence of HVLT recall appears unusual, how it interacts with *how* the words are recalled provides an important theoretical explanation. Whilst HVLT strategy was only a moderate influencer of RPCSQ scores, it was indicated that less use of the semantic category framework during recall predicts higher RPCSQ symptoms.

Therefore, if an individual has strong recall, but a weak HVLT strategy score, they are having to put more effort into recall, which would explain higher levels of RPCSQ and MFS. Again, this supports the findings of Geary et al (2011), who posited that individuals with mTBI were using more effortful strategies to achieve the same level of performance.

Conversely, the RPCSQ and MFS models indicated that individuals who adopted a drawing strategy that adhered strongly to the temporal spatial RCF clusters reported more symptoms. Additionally, completing more of the five core elements at the start of the drawing process predicted higher RPCSQ scores. Visual perceptual theory has historically argued that typical perception occurs with a global bias (Navon, 1997; Navon, 2003)). Considering the normative data that was reviewed in Chapter Five, it is indicated that normal observance of RCF strategy, order and wholeness are on a spectrum, with most individuals performing in the middle ranges. Typical function is some adherence to these heuristics, but not rigidly. Whilst there is some observation to the global and temporal-spatial heuristics, typically individuals are not tied to these, and are able to move between planes, such as shifting between perception from a top-down basis, and from a bottom-up position. Rigid cognitive flex – the reduced ability to flex between perceptual planes - may explain the unexpected direction of the model for RCF strategy. Differences that have been observed in a neurodiverse population drawing the RCF can help illustrate this.

Cognitive flexibility and set shifting are implicated in autism, along with a bias towards local processing (e.g., Happe, 1999; Sawaya, McGobigle-Chalmers & Kusel, 2019)). The local bias and reduced flexibility to shift away from this bias provides an explanation for the observation that neurodiverse individuals are more likely to be observed replicating the RCF in a piecemeal fashion. This provides a mirrored theoretical position to explain what we are observing in this current study. The reduced cognitive flexibility in individuals with long-term symptoms following mTBI suggests they are more rigidly stuck to global features, and struggle to tune in and out of detail. In a reverse of the well-known saying, they cannot see the trees for the wood (Navon, 1977). Individuals are unable to override a natural global bias and are therefore adhering to the global and temporal-spatial framework of the figure more rigidly. This also correlates with the higher scores of RCF order. These results are interesting, particularly when considered alongside the direction of influence of the category switching task, which was one of the most influential variables of MFS. Lower scores on the category switching scores predicted higher post-concussive symptoms. Category switching requires an individual to move between concurrently running tasks and is an indicator of cognitive flexibility.

The Stroop task was found to be a strong indicator of PCS symptoms and CFS scores. The task is an indicator of inhibitory control. This was not an unexpected finding, as inhibitory control is often impacted in milder forms of TBI, as a result of the mechanisms that cause them. Often, they are characterized by a whiplash type injury, and results in bruising to the frontal regions of the brain and neurotransmitter dysregulation. Inhibitory control and cognitive flexibility are considered to be distinct executive functions that are cholinergically mediated (Memisevic & Biscevic, 2018) therefore, it is conceivable that individuals with PCS and mental fatigue challenges are affected by both.

There is considerable evidence that argues against the use of group differences to evidence long-term challenges following mTBI. Statistical mechanisms are such that a non-symptomatic majority masks the deficits of a small, impaired subgroup. Further, individual baseline differences and change need to be examined, rather than post-morbid performance only. However, this data is typically not available. The theoretical arguments from Pertab and colleagues (2009), Iverson (2010) and Brooks et al (2011) provided the rationale for conducting the current study with regression analysis. However, interestingly the study did find some significant differences in cognitive function between the two groups. For both RCF Wholeness and Category switching, the mTBI performed statistically worse than the control group. This is unusual, as the mTBI group was treated as a homogenous group. Both those who would be considered recovered by cut-off points (e.g., De Guisse et al, 2016; Thompson et al, 2016) and those considered symptomatic were grouped together.

RCF Wholeness had a statistically significant difference between the groups. Individuals with mTBI were found to replicate the RCF image with more observation of the wholeness of elements than the control group. The effect size was also found to be large. However, it had very weak correlation with the outcome measures, and was also indicated to have multicollinearity with other variables so was removed from the modelling. Group difference analysis indicated movement of 1.5 standard deviations from the mean puts an individual in the borderline category, and 2 SDs into the very poor classification. Table 6.3 shows that one standard deviation in the current population sees a change in Wholeness score of 1.74. As indicated in Chapter Five, the mean normative score was 13.2. A reduction of 1.74 would result in a score of 11.5. Using the normative scores produced in Chapter Five, this would indicate a borderline score. This suggests that RCF Wholeness is a clinically sensitive measure for cognitive dysregulation following mTBI. However, this would require further investigation using more robust regression and minimal important change analysis.

Category switching group differences indicated that individuals in the mTBI group produced significantly less words than the control group. The task was also correlated with the outcome measures and included in the model for all four. Independently, it was significantly correlated with RPCSQ and MFS scores and was one of the most influential predictors of MFS scores in the regression analysis. The evidence of a statistically significant difference between the groups supports the theoretical argument that has been proposed that elevated levels of RPCSQ, and mental fatigue in particular, are underpinned by reduced cognitive flexibility.

Whilst there were a number of differences between the models, there were also some consistencies. All outcome measures had a mix of psychological, demographic, and cognitive variables indicated as the most influential. This supports the importance of taking account of the patient's history and their current psychological status when conducting a neuropsychological assessment. The most striking consistency was that depressive symptoms, as measured by the BDI-II, were the most influential variable for three of the models, and the second most influential for the RPCSQ. Poor mental health was posited by Binder and colleagues (1997) as the underlying cause of elevated levels of long-term functional symptoms following mTBI. The results of the regression models would support this theory.

As previous research and the current study indicates, mental health influences the extent of post-concussive symptoms and behavioural challenges an individual experiences following mTBI. However, a closer look at the individual correlation coefficients indicated that mental health status was significantly correlated with the outcome measures, but not with cognitive measures. This suggests that that poor mental health status has a greater influence on subjective perception of function than it does on objective cognitive function. These results conflict with the findings of Schonberger et al (2011), who found that depression and anxiety were not predictive of longer-term functional status. Their results indicated, however, that functional status was predictive of mental health status. The results also disagree with recognised understanding that depression and anxiety are understood to impair cognitive function (e.g., Iterwijk et al, 2021). The current study may suggest that psychological factors are a mediating influence. Rather than being the direct cause of symptoms and challenges, they exacerbate cognitive difficulties.

Addressing the second hypothesis first, measures of executive function, processing speed, working memory, mental health status, age and years of education were all included in at least one of the regression analyses. Other than years of education, each of these

variables was indicated to be a very influential predictor for at least one outcome measure. Therefore, with the exception of years of education, the hypothesis was accepted. However, in addition to these variables, IQ was also found to be an important variable with regards to post-concussive challenges. With regards to the third hypothesis, differences were found in the structure and importance of the variables that predicted each of the outcome measures. Overall, all of the outcome measures are sensitive to measures of Executive Function. However, this was more nuanced, with distinct aspects of executive function correlating with different outcome measures. RPCSQ and MFS is indicated to be most sensitive to inhibitory control and delayed recall and cognitive flexibility, whereas the DEX appears more sensitive to variation in working memory and effortful strategic processing. Further, whilst there is much similarity between the RPCSQ and the MFS, IQ plays a stronger role in MFS than it does the broader constellation of RPCSQ symptoms. Therefore, the third hypothesis was also accepted.

There were some limitations to the study. Variables for each model were selected on the basis of simple correlation coefficient relationships between the individual predictor variables and outcome measures. Whilst this removes potentially unnecessary variables from the regression, it may also remove variables where the effect is mediated by another variable. As correlation coefficients are only between the independent and dependent variable, it is not possible to consider interaction effects with other variables. Though the above models are statistically indicated to provide a strong predictive model of how the variables predict the outcome measure, further statistical examination would strengthen the validity of the findings.

In summary, the current study indicates that distinct aspects of the challenges that are experienced following mTBI are underpinned by different variables. This supports theories that suggest that PCS is an umbrella term for different sub-types of the condition, and which should be understood individually. However, whilst specific variables differ, on a general level, all outcome measures should be considered with regards to cognitive, psychological, and demographic variables. This highlights the importance of gathering all of this information as part of a neuropsychological assessment and mirrors the biopsychosocial theoretical model that underpins clinical psychology. The findings from this study provide evidence that observing drawing process constructs of the RCF tasks supports the diagnosis of PCS. Further, they suggest that cognitive flexibility and reduced inhibitory control underpins the symptom of fatigue, both of which are mediated by cholinergic functioning, which is known to be disrupted in mTBI. Finally, the study indicates the importance of understanding the mental health status of the mTBI patient.

Chapter Seven – Symptoms following mild traumatic brain injury

7.1 Chapter Premise

Throughout this thesis, limitations of the methodology and methods that have been used to examine persisting symptoms following mTBI have been addressed. This final chapter continues along this vein, to consider if how symptoms and challenges are being measured captures the experience of mTBI fully. As has been described previously, there has been an ongoing debate as to whether or not there are cognitive deficits following mTBI. However, it has to be questioned if how we understand the symptoms and challenges of mTBI are correct. Not fully capturing the challenges of persisting symptoms following mTBI may explain why there is difficulty in understanding resulting cognitive disturbance. What it is that is trying to be understood dictates how it can be understood. Therefore, unless we have clarity over what we are investigating, we cannot fully examine that phenomenon. If we are not measuring or understanding the symptoms or challenges fully, how can we correctly investigate the underlying aetiology?

This chapter will address the limitations of item-based questionnaires such as the RPCSQ and what they inform us about the experience of mTBI. Whilst there is important clinical and research value to quantitative inventories, this chapter addresses some of the limitations in what we can know from these methods and tools. It will highlight the importance of the qualitative patient report in the clinical neuropsychological assessment. It will also illustrate the nuances of the challenges following mTBI and suggest where future research of mTBI symptoms should focus.

7.2 Introduction

One of the challenges that has been highlighted with examining concussion and mTBI is the lack of agreement about how to define them. As discussed in Chapter One, there is variation in how to classify and define an mTBI, which has resulted in an extensive catalogue of inventory-based tools to measure and diagnose post-concussive symptoms. The chapter stated that for the purpose of this thesis, concussion was considered to be a mTBI. However, this understanding is not universal, with different perceptions in both medical and public groups of whether concussion is a brain injury (e.g., McKinlay, Bishop & McLellan, 2011; Merz, Van Patten & Lace, 2017). With the lack of agreement and clarity

about what mTBI – and concussion - is, this presents the challenge about how to understand what the outcomes are, which impacts how to investigate and the explain causes of these.

The Rivermead Post-Concussion Symptoms Questionnaire (RPCSQ) is a well-recognised and psychometrically tested measure of post-concussive symptoms (e.g., King et al, 1995; Eysers et al, 2005; Lannsjö et al, 2011; Balalla et al, 2020) – and is why it was selected as a measure of severity of symptoms and indicator of recovery status in the previous chapter. It can be used both as a diagnostic screen for post-concussive syndrome (PCS), and as a progress measure during recovery and rehabilitation. However, as indicated in Section 6.3, PCS has been argued to be an umbrella term for different types or dimensions of PCS (Keshavan, 1981; Levin et al, 1987; Eysers et al, 2006; Potter et al, 2006; Hermann et al, 2009; Maruta, Lumba-Brown & Ghajar, 2018; Barker-Collo et al, 2019; Vos et al, 2019). Supporting this theory, the findings in the previous chapter indicated that different symptoms and challenges are underpinned by different cognitive, psychological and demographic variables. Empirical studies have also indicated the dynamic nature of symptoms, such that not all are experienced all of the time (Medvedev et al, 2018; Moller et al, 2019). Further, it is recognized that there is a significant overlap of symptoms with other conditions, such as chronic pain (Snell et al, 2016) and also a substantial number of healthy non-neurological individuals reporting PCS symptoms (Dean, O'Neill & Sterr, 2012). This illustrates the complex landscape of understanding mTBI and PCS.

Self-report inventories can provide a quick-to-administer and useful snapshot of the challenges an individual is experiencing following their mTBI. However, this snapshot can be static in its nature, and not capture the nuances and dynamics that have been argued to be characteristic of long-term symptoms following mTBI and PCS. The limitations of itemised surveys to measure post-concussive symptoms were highlighted by Medvedev et al (2018), who argued that the RPCSQ is good at capturing enduring symptoms, but not those which are dynamic. It has been argued that diagnostic frameworks do not capture all challenges (Sveen, Ostensjö, Laxe & Soberg, 2012). Further, it has also been recognised that patients do not focus on just these classic symptoms when talking about their experience, but also the impact on real-life functioning (van Ierssel et al, 2021), which may be more important to the individual than the classic clinical symptoms (Banja & Johnston, 1994). Where functional measures are considered, it is argued that they are very narrow in what is covered (Levack et al, 2010). Therefore, there is a limit to what can be understood from the use of self-report inventories alone, in both research and clinical assessment.

The value of qualitative methods to clinical psychological understanding has become increasingly recognised (Harper, 2017). The use of qualitative research methods enables examination beyond the limitations of quantitatively analysed questionnaires, which may only provide a one-dimensional, static image, and one that may not capture the full experience of the individual (Priest, 2020). Itemised questionnaires are useful for providing real-time measurement, but do not necessarily provide understanding or multi-dimensional detail. Mixed methods, in particular, are argued to increase the validity of findings from differing methods due to the offsetting of biases and limitations within the individual methods (Jonsen & Jehn, 2009). The process of triangulation across data analysers and analysis methods forms a validation process, giving support to the findings from each of the methods. In addition, it provides a richness to statistical and static data. Triangulation between and across methods can help to ensure that we are fully capturing 'concepts' regarding a phenomenon. In the case of the current study, the use of mixed methods can support capturing 'concepts' more fully when we consider how a patient with long-term symptomology explains the real-life symptoms and challenges they face on a day-to-day basis.

Qualitative methods have been used to further our understanding of TBI. However, this research typically focuses on a single symptom or problem such as sleep (e.g. Mumbower et al, 2019); mental health (e.g., Beadle, Ownsworth, Fleming and Shum, 2016; Glinborg, Thomson & Hansen, 2017) or relationships (Power, 2016). Else it addresses experience and perception of rehabilitation and interventions (e.g., Graff, Christiansen, Paulsen & Egerod, 2017), or functional outcomes such as returning to work (e.g., Mansfield et al, 2015; Graff, Deleu, Christiansen & Rytter, 2020). Additionally, much research considers TBI generally, and does not focus on the specific and distinctive challenges of mTBI. Or it is very focused on a specific population of mTBI, such as sports players, military persona and children, for whom different biopsychosocial factors are at play than with a general adult population (e.g. Crowe et al, 2016; LaValle et al, 2019; Sicard et al, 2020). Even fewer studies use a mixed-methods approach - which is surprising given the synergies the methodology has with the biopsychosocial underpinning of clinical neuropsychology and the clinical formulation process. Therefore, very limited mixed-methods research exists which considers the general experience and real-life challenges of mTBI within a general, adult population.

The current study addresses a gap with the current literature, that considers how subjective symptom inventories correspond to the free-format description from individuals

about what their symptoms and daily challenges are. It sought to identify similarities and differences in the reports given via different methods.

7.3 Participants

Participants were the individuals in Chapter Six who reported a history of mTBI at least 3 months prior to taking part. One participant did not respond to the RPCSQ, therefore, quantitative analysis was conducted on the response from fifteen participants. However, the participant did include a free-format narrative, which was included within the qualitative analysis.

7.4 Procedure

All participants had completed the RPCSQ (Appendix 8) for the previous chapter study. Additionally, they were asked to provide a free-format narrative answering the question *“If you had to describe to someone how your injury has affected your day-to-day life, how would you do this?”* All responses were given online via Qualtrics. Qualitative analysis was carried out via traditional pen and paper methods, and NVivo. Quantitative analysis was conducted via SPSS.

A convergent mixed methods approach was used (Fetters, Curry & Creswell, 2013). Data was collected and analysed for each method concurrently. Integration was then carried out by merging data once individual analysis had been conducted via a narrative approach. A convergent approach was adopted as the focus of the study was to consider responses in parallel, not for one method to inform the other. It enabled simultaneous examination of quantitative and qualitative data regarding the symptoms and challenges of mTBI, to identify similarities and differences in participants self-reports. A side-by-side contiguous narrative approach was used to merge and report the findings. This method was chosen as it maintained the data type characteristics of the two sets of data, rather than transforming one into the same form as the other. This preserved the nature with which such data would be delivered in a clinical assessment, ensuring the ecological validity of themes identified in the data.

Qualitative data was analysed through Thematic Analysis (TA) (Braun & Clarke, 2006 and 2022). Other forms of qualitative analysis were discounted in part due to their deductive nature. For example, content analysis was not used as the focus of the study was

exploratory, and to look more generally across the participants' narrative. By investigating participant reports in an explorative way rather focusing on pre-determined symptoms being reported, this enabled the researchers to capture how a patient more generally describes the challenges that they have and identify aspects of their narrative that may not be covered by the self-report inventory.

The current study sought to identify patterns and themes within the reports of individuals who have experienced an mTBI and how they explain the impact to their life following the event. The use of TA provides flexibility in examining qualitative data to meet the aims of the study, but within a defined theoretical framework of investigation. A realist experiential epistemological position was taken, that the 'truth' of participants' experience, how they understand and describe the challenges they face following their injury is expressed through the language they use to describe the challenges. Not only do the concepts discussed exist in the mind of the participants, but also within the objective and social world (Jonson & Jehn, 2009). A semantic approach was adopted, as the aim of the study was to observe patterns in what participants were saying, rather than trying to interpret or infer meaning from what they were saying. The analysis was inductive, due to the limited nature of research specifically in this area, and the desire to focus on what the participants were saying that may not fit within pre-existing standardized symptom scales. Therefore, no existing codebook or predetermined concepts were used to analyse the text, other than a consideration for the research question being asked.

The six phases of thematic analysis as laid out by Braun & Clarke (2006 & 2022) were followed: familiarisation; coding; generating initial themes; developing and reviewing themes; refining; defining and naming themes; and writing up. Collaborative coding and theme development was undertaken, with the author of this thesis and another PhD researcher independently coding and developing initial themes. Once initial themes were developed, the two researchers compared coding and initial themes to reach the structure and definitions of the final themes. Whilst collaborative coding within TA is not a necessity, it is argued to enhance the quality of the coding process and the reliability of the themes developed (Vaismoradi et al, 2016).

Quantitative data was examined via frequency analysis, namely the chi-square test. Symptoms were analysed individually for frequency to each of the possible responses to the RPCSQ items – not at all; no more than before; mild change; moderate change; severe change. Analysis via frequency data enabled the examination of which symptoms were experienced the most and to what extent at least 3-months after an mTBI.

7.4.1 Reflexive positioning

The first coder - myself - was a 41-year old female, who had experience of working with individuals who have experienced an ABI, including stroke, dementia and TBI. As the developer of this programme of works and author of this thesis, I was embedded within the literature of symptoms and lived experience of individuals with persisting symptoms following mTBI. Further, I have previously been diagnosed with a concussion, from which I consider myself fully recovered. Whilst all attempts were made to consider the narratives of the participants with no preconceived ideas of what the analysis would show, it was important to recognise the existence of unconscious bias from this existing knowledge and experience. The second coder was a 30-year-old male, who was a qualitative researcher in Psychology, but not within TBI or neuropsychological assessment, and had very limited knowledge of persisting symptoms following mTBI. This offered a different lens through which they analysed the data and added to the reliability of the coding and themes.

7.5 Thematic Analysis

The thematic analysis identified five themes (Table 7.1). The themes were *Reconciliation and validation of recovery status*; *Tangible symptoms*; *Energy and effort*; *Living with mTBI within a social world*; and *Psychology of living with mTBI*. The first of these themes also included two sub-themes.

Participants' quotes are identified by anonymous participant number. Indication of extent of post-concussive symptoms and recovery for the participant being quoted is given by inclusion of RPCSQ scores. As indicated in Chapter Six, a score above 16 is considered indicative of post-concussion syndrome (Thompson et al, 2016); a score above 35 is considered moderate to severe challenges (De Guisse et al, 2016).

Table 7.1
Initial, final and sub-themes of thematic analysis

Initial themes		Final Themes	Sub-themes
Recovered	Recovery trajectory	Reconciliation and validation of recovery status	Recovery trajectory
Has a diagnosis	Length of time since injury		Support during recovery
Impact on life	Concerns for recovery		
Lack of understanding	Formal support		
Concerns there may be residual effects	Medical support		
Attributing symptoms to other factors	Non-formal support		
Solutions for complaints	Lack of support		
Cognitive impact		Tangible symptoms	
Sleep difficulties			
Physical injury			
Headaches			
Energy levels		Energy and effort	
Planning needed to do anything			
Brain fog			
Symptoms are limiting			
Need to recover energy levels			
Effort			
Detached from social world		Living with mTBI in a social world	
Social life challenges			
Noisy social situations			
Psychological impact		Psychology of living with mTBI	
Identity lost			
Traumatic incident			
Concerns regarding others' perceptions			

7.5.1 Reconciliation and validation of recovery status

This first theme encapsulates the individuals' perception, understanding and reconciliation of their recovery status. In particular, it addresses whether they consider themselves to have made a recovery from their mTBI. There was a clear distinction between those participants who identified themselves to be fully recovered from their mTBI from those who were still experiencing symptoms beyond the traditional 3-month recovery period. Individuals who considered themselves to have made a full recovery were often very specific in this regard, stating that they felt they had made a full recovery:

"I feel I have fully recovered from my injury" (Participant 7; RPCSQ - 15)

The conviction of this in some participants was sufficiently strong that they reinforced this point in several places in their narrative:

"I don't feel that my injury has had an effect on my life. Obviously initially I was incapacitated and had to go through a recovery process physically, but from day one my brain injury has not impeded my life" (Participant 4; RPCSQ - 6)

On the other hand, there was a second distinct group of individuals who reported experiencing problems from their mTBI as a very current situation, which had a significant impact upon their life. For some individuals, they were able to be very specific about the lack of full recovery, as a result of receiving a formal diagnosis of PCS:

"I've been diagnosed as suffering from Post-Concussion Syndrome" (Participant 16; RPCSQ 47)

But, for others the impact was discussed more generally in terms of the breadth and extent the injury is having upon their lives:

"The head injury feels like it has affected everything, my personal and family life and relationships, my working role, my health and friendships" (Participant 15; RPCSQ - 46)

"The effect of my accident has had a far greater impact that I could have imagined" (Participant 12; RPCSQ -32)

Individuals typically reconciled and validated their recovery status, regardless of what this status was. For individuals considering themselves to be recovered, they may justify any atypical behaviour or potential symptoms by attributing them to other factors, such as normal aging or pre-existing mental health concerns.

“For a time I was depressed and lacked motivation, however there were other contributing factors to this. The head injury certainly didn’t help in that respect” (Participant 8; RPCSQ - 23)

“I was 67 at the time of the accident, and am soon to be 70. I think a lot of my issues now come from aging rather than the injury, things like forgetting why you went in a room, or tiring more easily are I’m sure, age related.” (Participant 4; RPCSQ - 9)

Mirroring this, individuals still experiencing symptoms sought explanations from pre-existing health and life factors as a contributing factor to their lack of recovery:

“Maybe it’s because of my family history of migraines, or maybe it’s because I’d suffered whiplash injuries in the years preceding that left me susceptible. (Participant 16; RPCSQ - 47)

Whilst the narrative individuals used to explain their understanding of their recovery was mostly very clear, even in those individuals who stated themselves to be fully recovered, there were sometimes concerns that there were lingering effects:

“Occasionally I get very severe headaches and I will sometimes wonder if the injury caused this” (Participant 15; RPCSQ - 15)

Part of how individuals validated their recovery status forms the sub-theme of **Recovery trajectory**. Time since injury was used both as a means to explain recovery, but also as a characteristic that indicated the impact of the injury on their life.

For those individuals who considered themselves to be fully recovered, time since injury was often used as an explanation for their recovery:

“My injury was four years ago and now I feel fine” (Participant 5; RPCSQ - 10),

whereas for those who had not recovered, it was used as evidence to illustrate the impact the injury had on their life:

“Either way, it’s now been a long time since my actual injury.” (Participant 16; RPCSQ - 47)

However, for some participants, the passing of time presented the challenge that it was not so clear if there were any long-term effects, particularly if the incident had taken place during childhood

“As my injury happened when I was a young child, it is difficult to know if/how it has impacted me as an adult” (Participant 1; did not respond to RPCSQ)

Regardless of recovery status, participants indicated that there was a trajectory to recovery, and that whilst some of the participants were still experiencing symptoms and challenges, things had been worse:

“now I feel fine, however if I had answered this in the 2-3 months after my injury I would have answered completely differently” (Participant 5; RPCSQ - 10)

“For a little while I thought I might have to give up my career but 9 months on I am making some progress.” (Participant 12; RPCSQ - 32)

However, this acknowledgement of some recovery was still caveated with an understanding that things are still not back to full functioning, and also that recovery is non-linear, which brought with it uncertainty:

“I know I have got a lot better in the last couple of years, but I still feel like I’m stuck in a never-ending cycle and I have no idea when it will stop” (Participant 15; RPCSQ - 46)

“I should add that it was all much better March – end of May, then it got a lot worse” (Participant 10; RPCSQ - 28)

In addition to reconciling and validating the status of recovery, a sub-theme of **Support during recovery** was also identified, which was included in the narrative as either a

facilitator or as a limiter to day-to-day management of symptoms. Within this included awareness and understanding of the long-term effects of concussion:

“A disability which most people (including some GPs!) have never heard of – I certainly hadn’t before I found myself living with it.” (Participant 16; RPCSQ - 47)

and, for some, the lack of support lead to feelings of isolation:

“It feels quite [like] I’ve just been left to deal with it all by myself which is quite isolating” (Participant 12; RPCSQ – 32)

For those who had received an official diagnosis, this was acknowledged to be a good, but rare thing:

“I also know that I am very lucky to have had a diagnosis” (Participant 15; RPCSQ - 46)

Support came in the form of professional support such as pharmaceutical interventions to manage symptoms, or through community support groups:

“...some great support from my Community Head Injury team. Without them I’m not quite sure where I’d be.” (Participant 15; RPCSQ - 46),

as well as non-professional support from those close to them:

“I have had to lean on my partner more than normal, but he’s been great” (Participant 10; RPCSQ - 28)

In addition to support, individuals managed their experience of symptoms by finding practical solutions:

“I’m so lucky that I work part time now and could use that time to recover” (Participant 10; RPCSQ - 28)

“For quite some time I had to carry a Dictaphone with me” (Participant 12; RPCSQ - 32).

The theme of *Reconciliation and validation of recovery status* indicates that regardless of recovery status, individuals negotiate and evidence the status of their recovery. In most cases, individuals are clear in their narrative about whether they have recovered or not, particularly those who consider themselves to be fully recovered. For those experiencing challenges following their injury, this is often stated less directly, however, as the following themes will indicate, whilst a specific statement of not being recovered may not be given, narrative of the impact on real-life can be very detailed.

7.5.2 Tangible Symptoms

This theme reflects the describing of challenges that are typically tangible, recognisable and observable. They took the form of symptoms that could be clearly named and understood by others. This included discussion of what would be considered cognitive symptoms which most typically were either memory or concentration difficulties:

“I am unable to concentrate or process information” (Participant 14; RPCSQ - 45)

“Easily distracted” (Participant 9; RPCSQ - 24)

“I couldn’t remember things (from the simplest of things a family member would tell me...to having met someone in a meeting or anything that had been said in a meeting/project details etc). I couldn’t remember words the words for things.” (Participant 12; RPCSQ - 32)

However, discussion of cognitive symptoms only formed part of how participants discussed the impact of their mTBI. Some participants reported experiencing somatic/physical symptoms, which included difficulties with physical co-ordination, balance, or headaches. For others, this also effected their sleep.

“Sometimes it’s just struggling to stand because your balance is so out of whack that you’re falling backwards or forwards for no good reason. Or your spa[t]ial awareness and co-ordination isn’t working properly – no cooking with boiling water or using a sharp knife in case you have an accident (I painfully learn that quite early on)” (Participant 16; RPCSQ - 47)

“struggle to have a “normal” sleep schedule and will almost become nocturnal when not monitoring it closely. Sometimes (rarely) get headaches when before I never experienced a headache” (Participant 9; RPCSQ - 24)

Whilst these symptoms were acknowledged to be from the injury to the brain, participants did not necessarily see their physical injuries as a separate characteristic of their recovery or daily challenges, despite the focus of the discussion being on psychological and cognitive factors, and included discussion of them within their narrative:

“From a physical perspective, I had a couple of months where I was in a lot of pain and couldn’t do much at all. I had a broken wrist and a badly damaged knee. I’m still having to complete daily physiotherapy for my injuries which is very inconvenient” (Participant 12; RPCSQ - 32)

The theme of *Tangible Symptoms* was mostly identified within the participants who were still reporting symptoms – which is not a surprise given the nature of the theme is symptoms. However, some participants who were considered to be recovered did report experiencing somatic symptoms, though it was noted that these no longer occurred. The theme indicates that participants do include narrative about symptoms that can be clearly named and understood and may be very physical in nature such as orientation and spatial awareness difficulties. They also do not separate their physical injuries from their brain injuries in their recovery narrative, despite the focus of conversation being on TBI. However, whilst some of the report from the participants included these more tangible symptoms, a significant amount of their report focused on challenges which are not as observable to others. The next three themes address these.

7.5.3 Effort and Energy

The theme of *Effort and Energy* is arguably the most discussed across the participants yet indicates the challenges that individuals with long-term symptoms following mTBI have with being able to describe their symptoms, and for researchers and clinicians to be able to capture what the underlying pathological and cognitive complication is. Whereas the previous theme captures symptoms that are clearly named and understood by both specialists and within the public domain, the descriptions given by participants with regards to the consequences of their mTBI on effort and energy levels is less tangible and observable to those outside. It was described by several participants as having a feeling of a “foggy brain” (Participant 13; RPCSQ – 43) or feeling like they were “in a fog” (Participant 9;

RPCSQ - 24), and the experience of being somewhat removed from the environment around them:

“...like I’m “5 metres away from everything” (Participant 9; RPCSQ - 24)

...feeling detached – as if you’re not really connected to what you are doing, almost as if there’s a force field between you and the world.” (Participant 16; RPCSQ - 47)

This brain fog and feeling of detachment created daily implications on their ability to function and work. It was commonly described by participants that more effort was now required to complete tasks that were previously done with ease:

“I ‘think’ for a living, devising business strategy and solving complex organisation problems. I need to see complex patterns and facilitate workshops with clients. All of this has become much harder for me to do” (Participant 12; RPCSQ - 32)

The same participant again expressed this additional effort required further on in their narrative:

“One of the biggest impacts is the immense extra effort I now need when trying to work on my PhD studies” (Participant 12; RPCSQ - 32)

Alongside more effort being required to complete task, it was also commonly reported that individuals found themselves with limited capacity and energy, which was at a reduced level from their pre-morbid state

“I only have a certain amount of energy to “spend” on activities. This has meant that it has all got used up on my part-time work as a teacher, and I’ve had less/no time to work on my design work and building up that business. I’ve had to realise I can’t do as much as “normal” (Participant 10; RPCSQ - 28)

This combination of reduced energy and increased effort required meant that whilst they could function at the same or similar level that they did pre-injury, but their capacity to sustain this level of working was reduced:

"It reduced my sort of "base line" intelligence by a small amount, but when I do work hard, especially for extended periods of time it drops a great deal." (Participant 14; RPCSQ - 45)

The impact on effort and energy levels from the mTBI was also reflected in how well the participants could recover from mental exertion. It was recognised that not only is the exertion significant, but it also takes longer to recover from mental exertion - outside of the limits of what would be considered a normal recovery time:

"I still become very tired after doing anything that requires me to think or use cognitive capability" (Participant 12; RPCSQ - 32)

"...have to recharge become I carry on." (Participant 10; RPCSQ - 28)

"It's managing to (on a VERY good day) join your friends at a restaurant for a few hours because it's the only opportunity to see people in the 'real world' and having to allow for the fact that you'll consequently be wiped out with neuro-fatigue and migraines for the next 3+ day, unable to do anything" (Participant 16; RPCSQ - 47)

Incidentally, the participants from whom these last two extracts were taken both spontaneously reported experiencing physical symptoms following their participation within the study, in the form of migraines (with the onset for one participant being during the 90-minute session) and gastrointestinal complaints. This anecdotal evidence illustrates the extent to which the additional effort to complete complex tasks and remain focused for a substantial amount of time takes a toll upon wellbeing, both mentally and physically. Whilst only two participants reported very physical effects of cognitive effort, as participants were not asked about this, it may be that other participants also experienced very physical effects.

The consequences of the additional effort required, and reduced energy levels meant that participants needed to plan their lives more thoroughly than before, in such a way that it felt that their lives were limited and that their symptoms were dictating what they could do:

"Not being able to go out of the house without suitable planning and always being severely limited by what activities you can participate in" (Participant 16; RPCSQ - 47)

“I hate how I feel like the PCS dictates everything I do. I don’t like having to plan what I do around what I think might happen if I try to do what I normally would.” (Participant 15; RPCSQ - 46)

This need to plan ahead and loss of spontaneity and autonomy was further illustrated by the difficulties experienced with problem solving and task completion where a detailed plan has not been set in advance:

“Unable to do most tasks if I haven’t created a detailed list of things to do for the day the night before.” (Participant 9; RPCSQ - 24)

As the theme of *Effort and Energy* indicates, one of the most significant consequences of persisting mTBI and the most impactful on an individual’s life is the effect of cognitive fatigue. What it also reflects is one of the challenges that researchers and clinicians have with evidencing and understanding the basis of this fatigue, that it is not a static symptom. Whilst classic cognitive symptoms such as memory can be measured by administering a static recall task at any point in the day or neuropsychological assessment, cognitive fatigue only presents after a period of mental exertion.

7.5.4 Living with mTBI within a social world

The next theme encompasses the experience of individuals with persisting mTBI as they try to continue with their lives. Whether at home, in education or in work, this takes place within a social world. The experience of living with an mTBI in a social world took different forms, both as a symptom of their mTBI, but also in navigating living in society with their mTBI, both of which brought about different challenges.

Unlike physical injuries, one of the challenges that participants with mTBI face is that externally – at least once any physical injuries have recovered – they look well. Because of this, they experience challenges in people not understanding that they are struggling or find things difficult:

“Because the issues I face are ‘invisible’, no one really knows when I’m struggling and finding things challenging.” (Participant 12; RPCSQ - 32)

As a result of the invisible nature of their injury, and the intangibility of their difficulties, this created challenges with explaining to others the difficulties that they were facing and how this effected their behaviour:

“Sometimes it’s hard to explain this to people so that they don’t think that I’m being lazy or trying to get out of something” (Participant 10; RPCSQ - 28)

There were also concerns in communicating the challenges they were facing to others, that it may inconvenience them or that they were talking about it too much:

“...you don’t want to bore people with your problems or inconvenience them by asking for help.” (Participant 16; RPCSQ - 47)

As well as having to navigate the invisibility of the injury and symptoms with others, the injury itself also played a role in social challenges as a cognitive symptom of the injury. The effort required to manage complex situations was not isolated to just mental complexity as addressed in the previous theme, but also to managing social complexities, particularly with managing the ‘noise’ of social situations such that they became overwhelmed with the sensory overload from the situation:

“Often overwhelmed or feel very awkward in social situations with more than 2 or 3 people” (ZY55)

“It’s feeling completely overwhelmed from too much happening around you...” (Participant 9; RPCSQ - 24)

“If I am trying to work in the office whilst others are about, I now have to wear headphones. The noise makes me feel really distracted and almost disorientated” (Participant 12; RPCSQ - 32)

The theme of *Living with mTBI in a social world* illustrates the complexity of the challenges faced by individuals with persisting symptoms following their mTBI. It highlights the potential for oversimplification and lack of sensitivity of using post-concussive symptom/challenges inventories, and that they may not provide meaningful information about the nature of the social challenges the individual is facing. Whilst participants may respond that they have difficulties in social settings or working with others, this does not indicate if this is a cognitive symptom problem, or if it is a psycho-education/understanding

problem. This dichotomy could also be viewed as an internal/external division when considering how to best support the individual – does the intervention need to be focused on the individual or on the external situation?

7.5.5 Psychology of living with mTBI

The final theme addresses the experience of the participants' injury in terms of the psychological impact it was playing. Whilst the previous two themes were evident only in those individuals still experiencing symptoms, narrative that considered psychological and mental health impact was also found in those who considered themselves fully recovered. This theme took three forms, which included the trauma of the circumstances of the injury, the impact of the mTBI on cognitive affect, and the psychological impact of the persisting challenges following the accident.

As with the previous theme, cognitive affect was perceived by the participant, and in some cases by their family members, to be disrupted as a result of the injury to the brain:

“Feeling unable to have strong emotional reactions for things/people (dislike and affection)” (Participant 9; RPCSQ - 24)

As with the above participant, this took the form of reducing affect, whilst for others it took the form of heightened emotional response, such as increasing levels of irritability:

“I was irritable” (Participant 5; RPCSQ - 10)

Cognitive affect differences were considered by some to be a very current situation, as with the first participant quoted here, whilst for others, such as the second participant, this was a previous problem which had resolved itself during the recovery progress.

Whilst psychological impact for some was cognitive in nature, for others it was as a result of the trauma of the injury, particularly where the event was the form of a physical attack:

“Psychologically, after my injury I was very angry and wanted answers to why the attack took place” (Participant 2; RPCSQ - 5)

“The shock, hit from behind has made me more jumpy” (Participant 13; RPCSQ - 43)

Participants who considered themselves to be recovered, as was the author of the first quote here, were more likely to talk about the psychological impact of their injury, rather than as a cognitive impairment. However, participants who were still experiencing symptoms, like the second participant quoted, may talk about both.

The third form of psychological impact was where the persistence of the symptoms had affected their mental health. This theme was located within the narrative of those who were still reporting significant symptoms. For some, the impact of injury affected the individuals' sense of identity and that they were no longer the same person, as highlighted by the following extracts:

"...which makes me sad. I just want to be normal again" (Participant 14; RPCSQ - 45)

"I no longer feel like myself, as if I have become a different person" (Participant 15; RPCSQ - 46)

The difficulties individuals were facing also lead to feelings of embarrassment and shame:

"I often feel embarrassed in front of clients because I can't think as clearly or as quickly as I need to and used to" (Participant 12; RPCSQ - 32)

As a result of the injuries, several participants commented on the anxiety and stress that was caused as a consequence of the impact it had on their ability to work and earn an income:

"I also experienced a high level of stress within the first few months because I am self-employed and was unable to work. I lost out on several well-paid contracts. Doubting my ability to continue in my career and with my business, I became anxious that I wouldn't be able to regain my income levels. I am the main 'breadwinner' for the family so this was very concerning." (Participant 12; RPCSQ - 32)

The theme of *Psychology of living with mTBI* is similar to the previous theme that considers living in the social world, in that there can be different underpinnings to the psychological and mental health impact an individual is experiencing. It reflects the findings from previous literature that a loss of sense of self can come from the impact the mTBI has on their productivity, contribution and success (Levack et al, 2010).

It highlights that psychological impact may be the result of the challenges an individual faces in living with the symptoms of their mTBI, rather than the cause as some theorists argue. This also highlights the multi-factor and multidirectional nature of psychological wellbeing following mTBI.

7.6 Quantitative post-concussive symptoms

The frequency and severity of subjective experience of post-concussive symptoms were considered (Table 7.2). An increase in symptoms (response mild to severe) was indicated for all sixteen symptoms by at least some participants with mTBI. Concentration, longer to think and forgetfulness were the most commonly reported symptoms that had worsened since the injury had taken place. However, fatigue was the symptom most commonly reported as being a severe change. Nausea, blurred vision and double vision and depression were the symptoms reported the least.

Table 7.2

Percentage of experience and severity of post-concussive symptoms in individuals with mTBI at least three-months post injury (N= 15)

Symptom	Not at all	No more	Mild	Moderate	Severe	Increase
Headaches	20.0	20.0	20.0	20.0	20.0	60.0
Dizziness	33.3	6.7	33.3	26.7	0.0	60.0
Nausea	66.7	0.0	20.0	6.7	6.7	33.4
Noise sensitivity	33.3	13.3	6.7	20.0	26.7	53.4
Sleep	20.0	26.7	33.3	20.0	0.0	53.3
Fatigue	13.3	26.7	13.3	6.7	40.0	60.0
Irritability	33.3	26.7	13.3	20.0	6.7	40.0
Depression	33.3	40.0	6.7	20.0	0.0	26.7
Frustration	20.0	33.3	20.0	26.7	0.0	46.7
Forgetfulness	13.3	20.0	33.3	20.0	13.3	66.6
Concentration	6.7	26.7	26.7	13.3	26.7	73.4
Longer to think	13.3	20.0	26.7	20.0	20.00	66.7
Blurred vision	60.0	13.3	13.3	13.3	0.0	26.6
Light sensitivity	33.3	6.7	13.3	33.3	13.3	59.9
Double vision	80.0	6.7	13.3	0.0	0.0	13.3
Restlessness	33.3	26.7	13.3	13.3	13.3	39.9

Taking the Thompson et al (2016) RPCSQ cut-off of 16 as the indicator of PCS, chi-square tests were conducted for each of symptom to determine if there were statistically

significant proportion differences between the symptoms reported between 'recovered' and 'symptomatic' individuals. Assumption of chi-square analysis were considered. The data for each variable (recovery status; experience of symptom) were categorical and independent from one-another. To account for the small sample size, experience of symptoms was reduced down to two categories – no experience (not at all and no more) and experiencing (mild, moderate and severe). Due to not having a sufficient sample size to meet the expected count > 5 assumption, Fisher's Exact test results were taken.

Nausea, Noise sensitivity, Irritability, Frustration, Forgetfulness, Concentration, Longer to think, Light Sensitivity and Restlessness all had statistically significant association between recovery status and experience of symptom. For Nausea, symptoms were more frequently reported in the Symptomatic group. Associations for all variables were found to be at least moderate in size, with all of the significant associations being strongly associated. Concentration and Noise sensitivity were found to have the strongest association between recovery status and experience of symptom (both $V = 0.87$, $p < 0.001$).

Table 7.3
Fisher's Exact test, examining proportional differences of recovery status by experience for individual RPCSQ symptoms

Symptom	Sig. (p)	Cramer's V
Headaches	0.14	0.44
Dizziness	0.14	0.44
Nausea	0.04	0.58*
Noise sensitivity	0.001	0.87*
Sleep	0.32	0.33
Fatigue	0.14	0.44
Irritability	0.03	0.67*
Depression	0.10	0.49
Frustration	0.01	0.76*
Forgetfulness	0.09	0.58*
Concentration	0.002	0.87*
Longer to think	0.01	0.74*
Blurred vision	0.103	0.49
Light sensitivity	0.01	0.72*
Double vision	0.49	0.32
Restlessness	0.01	0.76*

* indicates approximate significance of Cramer's V to be $p < 0.005$

7.7 Chapter discussion

The current study was a mixed-methods examination of the report participants give regarding the challenges they may or may not face as a result of their mTBI. Participants responded to the recognised and psychometrically validated RPCSQ, as well as providing a free-format narrative of the impact and experience of their mTBI.

The thematic analysis identified five different themes in how individuals spoke about the experience of the mTBI. The first theme, *Reconciliation and validation of recovery status* talked broadly about whether participants perceived themselves to have made a full recovery or not. Overall, participants were very clear in their narrative as to their recovery status. Their perception of their recovery status was reflected in the responses given to the RPCSQ, with those individuals independently reporting to be fully recovered scoring near or below the validated cut-off score of 16 (Thompson et al, 2016). Correspondingly, those still reporting symptoms and challenges scored above the cut-off point. Whilst this could be considered as a confirmatory bias, it does support that the RPCSQ with a cut-off of 16 is a reliable, quick-to-administer screen of the extent of post-concussive challenges.

However, whilst perception of recovery initially appeared to be very binary, the theme indicated that participants internally negotiated, validated and reconciled their recovery status. For some this took the form of providing evidence either to rationalise why they were recovered, or why they may still be experiencing symptoms. Age, length of time since injury, and pre-morbid mental and physical health were all used as justifications, whether recovered or symptomatic. For others, even when they considered themselves to be recovered, they still harboured concerns that there was persisting damage. As part of the reconciliation of their recovery status, participants talked about their recovery as a journey or trajectory, regardless of where they were on that path. Support during recovery was also an important part of this journey. The nature of this theme reflected themes identified by Snell et al (2016). Their study also highlighted the importance of social scaffolding and validation from others during recovery. Whilst these factors may not improve the symptoms, it eased acceptance of and living with the injury.

The remaining four themes identified in the thematic analysis were more specific and were aligned with the different symptoms and challenges recognised as persistent post-concussive symptoms. As expected, much of this narrative came from individuals who were still experiencing symptoms. The theme of *Psychology of living with mTBI* is of particular interest to this thesis. Psychological factors have been posited by several authors such as

Binder and colleagues (1997) as an underlying aetiology of functional symptoms, rather than being the result of organic tissue damage. Interestingly, whilst psychology and mental health was identified as a theme in participants' narrative, Depression was one of the RPCSQ items indicated to have the least change since injury. The proportional difference between those scoring low and high on the RPCSQ was minimal, and not statistically significant. Irritability, Frustration and Restlessness, however, were found to have significant proportional differences, with no low RPCSQ scorers reporting any change. However, where higher scoring participants reported change, this was typically mild to moderate.

Convergence between the two data sets was found, with both indicating the importance of considering mental health factors after mTBI. However, the theme indicated a complex landscape as to the nature of these factors. They may be the result of the organic injury and are therefore a cognitive challenge. Association between TBI and emotional lability and inhibition is well documented (e.g., Archiniegas & Wotzel, 2014). Else, they may be the consequence of trauma and/or living with the challenges of a brain injury. It is acknowledged that poor mental health can exacerbate symptoms and cognitive function after mTBI (Pineau, Marchand & Guay, 2015), and therefore it is important to address this when working with a patient reporting mental health challenges. However, these may not necessarily be depressive in nature, and instead be trauma-based, frustration of living with the injury, or a cognitive dysfunction.

An interesting rhetoric within the narrative was observed that participants felt other people understood symptoms such as memory difficulties better than some of the other more invisible and less tangible symptoms. The theme of *Tangible symptoms* encapsulated these symptoms that are easy to name and explain, and included memory problems, headaches, and sleep difficulties. Patients also considered their physical injury recovery as an integrated part of their overall experience of mTBI. The Forgetfulness item on the RPCSQ was reported by the majority of participants to have changed to some extent since injury. A greater number of high RPCSQ scorers reported memory difficulties than low scorers, and to a greater extent. However, the proportional difference was not significant. This suggests that memory problems reported on the RPCSQ are not isolated to mTBI factors. This was captured in the previous theme, where a participant attributed memory challenges to age rather than their injury. This has important implications when considering population base rates during clinical formulation. Other 'tangible' RPCSQ symptoms of sleep and headaches were also reported to have changed by the majority of participants, indicating convergence between the two data sets. Further convergence was also found in that blurred and double

vision and nausea were the symptoms reported to change the least amount on the RPCSQ and were not discussed by participants.

The narrative regarding headaches highlighted an important difference in how interactions between symptoms and dimensions are understood within previous literature. The theme *Living with mTBI in a social world* captured functional aspects of post-concussive symptoms within a complex, social environment. Sensory overload, particularly noise sensitivity was discussed by several participants as something they needed to manage in their daily lives. Noise and light sensitivity in the RPCSQ were both indicated to have significant proportional differences between severity groups. Noise sensitivity was also the symptom second most frequently reported as being a severe problem following mTBI. Several factor and construct analyses of the post-concussive symptoms indicate these two symptoms to be associated with headaches and migraines (Keshavan et al, 1981; Maruta et al, 2018). However, in the current study they were considered to be a social problem. This aligns more closely with the factor structures produced by Eysers et al (2005) and Hermann et al (2009). This has interesting implications on our understanding of the dimensions of PCS. Whilst data-driven analysis may suggest environmentally effected symptoms are related to experience of headaches, and migraines, the current analysis indicates that for the individual they are more of a social challenge.

The invisibility and difficulties in others understanding some of the symptoms being experienced was captured within the theme of *Energy and Effort*. It was discussed in detail by several participants and appeared to have the greatest impact on daily functioning and quality of life. Not only did it impact their ability to carry out tasks, but it raised concerns that others would perceive them to be lazy. The RPCSQ items of Concentration and Longer to think were the two symptoms most reported to have changed following mTBI. Further, Fatigue had the highest frequency of being reported as severe. Concentration and Longer to think were also indicated to significant proportional differences, and with very large effect sizes – Concentration being the largest effect size found. The results from the two data sets indicated agreement that energy and effort are significant persistent challenges following mTBI.

However, as with the themes already discussed, participant narrative illustrated the complexity of cognitive fatigue. Analysis from the theme raises a number of questions regarding our understanding of cognitive fatigue following mTBI. Is it that there are less energy reserves to start with; that tasks are requiring more effort, so energy levels are being drained quicker; or that accessing the resources to undertake tasks is less efficient? Less

efficient resource pathways may provide an explanation that recovery takes longer, though it could be that more needs to be recharged because more has been used. Or a combination of both. From this analysis it is not possible to answer these questions, though, it does provide a clearer framework from which to consider researching the nature of cognitive fatigue, which has previously been difficult to capture and explain. Further, how participants talk about cognitive fatigue following mTBI supports the 'Fatigue as an indication of effort' model discussed by Wylie and Flashman (2017). Wylie and Flashman argued that fatigue as a result of effort was the most likely explanation of cognitive fatigue following mTBI.

Whilst there was a convergence of findings across the two data sets when talking about energy and effort there were also some disconnects. Firstly, the RPCSQ does not capture the feeling of fog or detachment that several participants spoke about. The use of the word 'detachment' by several participants was interesting, as it was used to explain a barrier between themselves and the context they were in. However, other research talks about detachment and disconnect as a feature of the psychological impact of mTBI, and how individuals felt disconnected from their pre-injury selves and/or from others (Levack, Kayes & Fadyl, 2010).

Secondly, the RPCSQ does not capture the dynamic nature of cognitive effort and fatigue or that it is task and environmentally dependent. This reflects the conclusions of Medvedev et al (2018), who argued that the RPCSQ is good at capturing trait-like symptoms, but not dynamic symptoms, such as chronic fatigue. The theme highlights the need to consider how symptoms are captured in self-report inventories such as the RPCSQ, but also how to measure it systematically. Tasks that assess absolute performance at a static point will not indicate the impact of these types of injuries where cognitive resourcing is damaged. Rather tasks need to be sufficiently complex and sustained so as to put a strain on functioning to get to the point at which reduced functioning and fatigue can be observed.

The final two themes highlight an interesting dichotomy about the need to understand the basis for these society-based challenges – a similar pattern found when talking about the psychology of mTBI. Whilst it was recognised that some of the social challenges faced were from the injury, it was also acknowledged that the invisibility of their injury and lack of understanding from others created problems for them. The issues highlighted in this study of the invisibility and lack of tangibility of their difficulties and the impact this has on their living within a social world correspond to the social disconnect discussed by Levack et al (2010), where looking uninjured and 'normal' raised questions about the validity of the

individual's disability. The findings here indicate the need to address both internal and external challenges when working with mTBI patients.

A limitation of the current study is that the qualitative data was collected via a qualitative survey rather than interviews. This presents two issues to consider. Firstly, the patient perspective is typically taken in a face-to-face interview between the patient and the clinician, which allows for follow-up questioning and the potential for more detail. However, it may be that the participants in the current study felt a freedom to be more honest about the challenges they were facing, as there was an element of distancing, and lack of scrutiny and judgement by completing the task via an online platform. Secondly, recommended minimum participant sample numbers for qualitative survey studies as part of a large project is 50 participants (Terry, Hayfield, Clarke & Braun, 2017), which the current study did not meet. As previously mentioned in this thesis, a larger sample size was intended, but the challenges of the environment that the study was being conducted in meant that this was not achieved. However, the richness of the reports, the consistencies across the different methods and the agreement with previous knowledge supports the validity of the findings in this study.

The examination of free-format responses within the theme of recovery status perception highlighted a methodological consideration when carrying out research that examines the long-term effects of TBI. Whilst overall patients were able to report concisely and coherently about their recovery status, some participants found it more challenging than others to report if there were differences between pre- and post-morbid status. These participants reported experiencing their mTBI many years previously, often as a child, and were therefore less able to identify if anything about their current behaviour or experience was a result of their mTBI. Whilst there is sufficient evidence that research should examine symptoms and cognitive function beyond the standard 12-month time frame typically observed within mTBI and concussion research (the point beyond which symptoms are clinically considered chronic), there is the risk that if injuries have occurred too long ago, it is not clear what is the result of the injury and what is the result of normal life span development, particularly if the injury occurred whilst the individual was a child.

The current study employed an under-used approach to understand the symptoms and the experiences of individuals who had an mTBI and how they reported these challenges. The use of the mixed-methods approach has greater synergies with the clinical neuropsychological process, where a verbal report of the patient's challenges is taken as well as using recognised self-report symptom and functional challenges scales, and cognitive assessments. The narrative individuals used to describe their challenges

corresponds with self-report inventories – as would be expected given the time and money that has been invested in to producing these – however, there is a greater richness in qualitative reports which can provide more insightful information about the challenges the individual is experiencing. It has also illustrated the complex nature of cognitive fatigue and has pulled apart some of the questions that need to be answered before this symptom can be fully understood and supported. The concurrent analysis of the results from the RPCSQ and the analysis of participants' narratives evidence a reliability and validity in the RPCSQ to indicate PCS may be present. However, the RPCSQ has limited use in understanding the challenges an individual with persisting post-concussive symptoms is experiencing. This highlights the importance of collecting subjective reports in patients' own words within healthcare pathways, both as further evidence to support diagnosis, but also to determine where support is needed.

Chapter Eight – Discussion

8.1 Chapter Premise

The premise of the following chapter is to consider the body of works that has been carried out as part of this thesis, and how the findings answer the research question posed in the thesis introduction. The chapter will summarise the key findings from each chapter and consider the clinical and theoretical implications that these findings have on our understanding of the long-term consequences of mTBI. It will address the limitations of the work, and how these limitations effect what we have been able to learn about the long-term and real-life impact of mTBI. Finally, it will close by making recommendations for future research and policy and to consider the wider impact of the knowledge that has been developed here.

8.2 The research question – a reminder

“There is nothing in the field of neuropsychology that is more divisive than the topic of mTBI.” (Barr, 2014)

“...neuropsychology’s goal should be to refine its methods...not continue arguing that permanent deficits do not occur in mTBI.” (Bigler et al, 2013)

The research question for the current thesis asked, ‘Does observing and measuring the drawing process used for the Rey-Osterrieth Complex Figure Task support the neuropsychological assessment for individuals with long-term symptoms following mTBI?’ The rationale for examining this question was underpinned by a long-standing debate regarding the aetiology of the long-term symptoms that are reported by a sub-group of patients who have experienced a mTBI or concussion. For one section of the clinical psychology community, these long-term symptoms – which may be diagnosed as Post-Concussive Syndrome (PCS) – were considered to be functional in nature. The evidence of cognitive deficits was limited, and neuroimaging was typically found to be unremarkable, with no sign of damage to brain tissue (Binder et al, 1997). This lack of neuroimaging evidence provided support to argument that where symptoms were reported, they were the result of pre-existing lower functioning, psychological distress and/or malingering, rather than the

result of the mTBI or concussion. Where cognitive deficits were found, they were considered to be negligible in size, and of no clinical consequence.

For another section of clinical practitioners and researchers, the lack of evidence of tissue damage or cognitive deficit was considered to be a consequence of limitations in methods rather than the nature of the symptoms. They argued that studies typically grouped all cases of mTBI together as a homogenous group, rather than considering those who had recovered separately from those still reporting symptoms. By examining cognitive performance collectively, evidence of deficits in the small sub-group of individuals with post-concussive symptoms was lost as a result of statistical mechanisms. Further, the importance of considering base rate performance and individual change was argued (e.g., Brookes, Holdnack & Iverson, 2011; Brooks, Iverson & Holdnack, 2013; Collie et al, 2004; Moser et al, 2007). Post-injury performance in average population ranges for a high functioning individual may be a reduction of performance for that individual, but crucially this detail is lost when examining absolute performance, rather than individual change. In recent mTBI research, the use of group differences methodology has given way to more sophisticated statistical methodology, such as regression analyses which provide a more nuanced understanding of the complexities.

To further support their case, advances in neuroimaging technology, such as the use of Diffuse Tensor Imaging, have allowed us to visualise functional neuron activity, rather than only providing a static image of brain structure and neuronal atrophy. This has enabled us to examine the sub-structural changes at an axonal level as well as disruption to neurotransmitter functioning and white matter connections that have come to be recognised as the neuro-anatomical consequences of milder TBI (DeLellis et al, 2009; Wilde et al, 2014; Romeu-Mejia, 2019). These advancements in imaging and biomarker techniques bring into question whether individuals who were considered to have no organic injury in older literature were misdiagnosed due to imaging limitations.

It has also been argued that there are limitations to the neuropsychological assessments that have been used to measure cognitive function (Bigler et al, 2013). Geary and colleagues (2011) evidenced that whilst there was comparable absolute recall performance between individuals with mTBI and a control group, the method used to recall the word lists of the California Verbal Learning Test (CVLT) differed. It was not *what* they recalled, but *how* they recalled it that was important. A further challenge to identifying the underlying cause of post-concussive symptoms is their heterogenous and dynamic nature. Importantly not all of the symptoms are experienced, and not all of the time. Psychometric

analysis of post-concussive scales such as the Rivermead Post-Concussive Symptoms Questionnaire (RPCSQ) have suggested there are different sub-types of PCS, whilst others have suggested there to be different dimensions that are experienced differently depending on the point of recovery a person is at, and the type of activities they are engaging in (Keshavan, 1981; Levin et al, 1987; Maruta, Lumba-Brown & Ghajar, 2018; Vos et al, 2019). The lack of evidence of a cognitive explanation for the symptoms may be due to neuropsychological assessments only capturing a static point of performance or they are not as complex, sensitive, or sustained enough to bring on the behavioural consequences of their injury.

The synergies between our understanding that it is anatomical disruption rather than atrophy, and cognitive change rather than loss of function highlights limitations of previous research examining cognitive outcomes of mTBI. There needs to be a move away from thinking about cognitive deficits of mTBI, and, instead, a focus on cognitive disruption. Whilst this nuance of language used may seem unnecessary, the shift of positioning directs what tools we should use in research to examine the cognitive impact of mTBI and the aetiology of long-term post-concussive symptoms. As per the quote at the beginning of this section by Bigler and colleagues, we need to refine our tools so that they measure the problem presented, rather than argue that they do not exist, simply because we are not able to measure them with what we currently have available.

The focus of this thesis was borne from the theoretical positioning of Geary and colleagues (2011) that individuals with mTBI can perform to the same absolute levels as controls, however, the cognitive processes that are required to achieve these goals are more effortful. This reflects the real-life reports of individuals who have returned to work following mTBI, who can still carry out their job almost as they did before, but it comes at the cost of cognitive fatigue – the most commonly reported and persistent post-concussive symptom (Lannsjo et al, 2009; Roe et al, 2009; Moller et al, 2019). It was hypothesised that this use of more effortful strategies to complete tasks may provide an explanation for the increased experience of symptoms in this sub-group of individuals. The RCF was adopted as the tool by which to investigate this as it is a sufficiently complex well-used neuropsychological tool which is known to capture cognitive functions such as problem solving, planning and perceptual processing. Further, the drawing process used by individuals to complete the task has been the focus of numerous empirical studies to provide an explanation for behavioural and psychological challenges (e.g., Sullivan et al, 1992; Savage et al, 1999; Cherrier et al, 1999; Booth, 2006; Lopez, 2008).

To answer the research question posed within this thesis, the multiple limitations and methodological issues associated with previous work and knowledge were addressed. The main methodological questions were: are we measuring the right thing? How do we measure it? and, how do we analyse it? A mixed-methods approach was used to explore these questions. This methodology was chosen in part due to the synergies it has with the neuropsychological assessment, where both cognitive measures and the patient's own description of their challenges are taken into consideration. Secondly, it was chosen as it provides the means to address where there is convergence and divergence between findings which may be lost if only one method is used, but further our understanding of persisting post-concussive symptoms.

The first half of the thesis addressed methodological and procedural challenges in examining the drawing process used to complete the RCF task. A review of the literature was conducted to identify currently available scoring systems that measure the drawing process – also known as qualitative characteristics – of the RCF task. As a result of this review, a novel scoring framework was proposed. The next set of studies explored how technology can be used to increase the functionality of the RCF task within clinical neuropsychology, which is economically viable to acquire, and easy to use by both the clinician and patient whilst being consistent with analogue methods already in use.

The second half of the thesis explored the nature of the drawing process used by both the general population, and those with a history of mTBI. In addition, it sought to consider how the drawing process used by an individual related to experience of post-concussive symptoms and real-life functional challenges. Firstly, normative data of RCF drawing process from a non-mTBI, healthy adult control population was examined to understand typical performance, and what would be considered the extremes of population differences. Next, limitations in previous literature that examined the cognitive function of individuals with mTBI were addressed using regression analyses to understand which cognitive, psychological, and demographic variables were the most sensitive indicators of post-concussive symptoms and behavioural challenges. This included examining if the drawing process adopted by an individual with mTBI to replicate the RCF gave insight into the experience of PCS. Finally, a mixed-methods approach was used to gain better insight into the challenges experienced by individuals who were in the chronic stages of post-concussive symptoms.

8.3 How to measure drawing process used for the Rey-Osterrieth Complex Figure

One of the first criticisms, raised by previous researchers, is that the cognitive assessments being used to measure cognitive function are not sensitive enough to capture the nuanced and subtle changes following mTBI. The first aim of this thesis was to provide evidence of an assessment that would overcome this limitation and provide a more comprehensive and varied set of measures. Observing how an individual goes about drawing the RCF was posited as an integral qualitative method to support this aim. Whilst there is a rationale and precedent for examining the drawing process used by an individual with mTBI (e.g., Zapalla & Trexler, 1992; Geary et al, 2011), it raised the question of how to go about this in an objective, practical and robust way. With verbal learning tasks such as the CVLT and HVL, a clear semantic category framework is visible which is easy to objectively record and analyse both quantitatively and qualitatively. The RCF task was originally designed to measure the drawing process an individual uses to replicate its structure (Rey, 1941, cited in Corwin & Bylsma, 1993). However, whilst a robust system for scoring accuracy has stood the test of time and psychometric evaluation, how to measure the nature of the drawing process an individual uses to replicate the RCF has been less clear.

A review of the literature in Chapter Three identified 20 independent drawing process scoring systems. Studies that have examined the psychometric properties of multiple scoring processes concluded that no single system provided a sufficiently robust means for measuring drawing all aspects of the RCF drawing process (Troyer & Wishart, 1997; Anderson, Anderson & Garth, 2001). Therefore, each scoring system was reviewed to determine which system - if any - would be suitable for use in the current thesis. Two review questions were asked - what is measured? and, how is it measured? The exploration of what was measured first, revealed a distinct lack of clarity over taxonomy, with over 50 different words identified across the systems to describe what they were measuring. Whilst there was some consistency with words being used more commonly than others, for example 'Organization' - there were differences in what drawing process constructs the words were referring to. Different systems that were measuring the drawing process used assorted terms for the same constructs, and vice versa. A further problem identified that some systems had developed two different constructs whilst another measured these as a single construct. This was evident in many systems, including the original format, where order and wholeness were measured in the same construct (e.g., Anderson, Anderson & Garth, 2001; Gasparini et al, 2008; Heinrichs & Bury, 1991; Rey, 1941 & Osterrieth, 1944, as cited in Corwin & Bylsma, 2007; Stern et al, 1994; Sullivan et al, 1992; Zappala & Trexler,

1992). The literature concluded there were four discrete drawing constructs of the RCF: Wholeness, Order, Continuation and Strategy.

Detailed examination of how each of the constructs were measured revealed further variation, and the most prominent observation was the level of objective measurement within the systems. Whilst some provided clear instructions on how to score the task (e.g., Bennett-Levy, 1984), others were more subjective (Hamby et al, 1993). Further, not all systems provided continuous data, using subjective categories instead (e.g., Gasparini et al, 2008). Whilst categorical and Likert scale systems can be beneficial for some measurement methods, for the purpose of clinical measurement, wide-ranged continuous data provides a more sensitive measure. Finally, several systems just measured a small handful of core elements, ignoring the localised elements (e.g., Shorr, Delis & Massman, 1992).

None of the existing systems that were identified provided an objective means of measuring all of the constructs individually on a continuous basis, that also considered all elements in the figure. Therefore, a novel method for measuring drawing process was developed based upon the four constructs identified in the literature review. The novel method was used in later analyses to firstly understand normative population performance in greater detail, and to then measure drawing process in an mTBI population.

8.4 Increasing functionality of the Rey-Osterrieth Complex Figure through the use of digital recording devices

Another operational challenge with observing the drawing process used to complete the RCF task was how to record patients' and participants' drawing process with enough detail to record each step. Scoring for multiple constructs requires the ability to return to the drawing process numerous times and this is even more important if dual scoring practices are being adopted - as is often the case in research. Traditionally, analogue methods have been used, such as using coloured pens, or the test administrator mirrors the steps on a blank template with numbers and arrows. However, both methods have limitations either with the level of detail they can capture, or the real-time practicality of their use. More importantly, they can bring the aim of the task to the attention of the participant or patient, thus confounding the information that can be gained from the task (Ruffolo et al, 2001).

Advances in technology have meant that digital recording technology is more accessible to both research and healthcare services. These new technologies provide a potential solution to recording the drawing process an individual uses to complete the RCF

task. A number of solutions exist, such as digital touchscreen tablets or sophisticated haptic virtual stylus systems. However, a digital pen that writes with a standard ball point cartridge on paper has been posited as a tool that requires no additional effort from the participant and reflects using a normal pen as closely as possible. The pen also provides the potential for digital scoring, and more detailed analysis of drawing factors, such as 'thinking' or 'hesitation' time. Research has previously been conducted using a pen such as the Neo Smartpen on drawing-based tasks with non-TBI clinical groups (Sanada et al, 2014; Sonntag, 2018; Prange and Sonntag, 2019). However, this research did not assess whether task performance was objectively equivalent to a normal pen and paper. Further, the statistical methods used to analyse the data focused on differences, rather than agreement. Additionally, whilst subjective usability analysis had been conducted, this was only on the basis of the pen being held, and not after being used. The focus of the first study reported in Chapter Three was to consider firstly whether use of a digital pen to complete the RCF task was equivalent to performance with a standard pen and pencil, using more suitable statistical methods for assessing agreement. It also assessed usability by the participants in a more rigorous way.

Whilst usability and ease of use from the point of view of the scorer was not assessed, unprompted anecdotal feedback from the scorers was provided. It was commented that scoring from the digital pen felt slower than with the video recordings. With the digital playback, it often had to be paused and replayed to capture where the next line was drawn. The image had to be scanned fully to determine where the next line was being drawn, whereas with the video-recordings the limb movement provided a clear cue as to where the next line was going to be, reducing the area of the image that needed to be scanned to identify what line was to be drawn next. Whilst digital recordings can be played back an infinite number of times, the additional time and effort required of the clinician or researcher may discourage use of the pen and the increased functionality of the RCF. However, this feedback was only anecdotal and collated from two individuals, highlighting the need to conduct usability research with the scorers, as well as with the participants and patients. It also supports the case for the work that is being conducted by researchers including Sonntag and Prange to automate scoring of the RCF task, however this technological advance is still some way from realisation.

Within this study the hypothesis that accuracy using the Neo Smartpen would be comparable to performance using a standard pen was accepted. There were no statistically significant differences found between the two conditions. More importantly, agreement analysis in the form of interclass correlations and Bland-Altman plots indicated strong

agreement between the two methods, with no evidence of systematic bias. Overall, subjective feedback from participants indicated high usability and comfort ratings. Interestingly performance did not correspond with comfort ratings and low scoring outliers felt comfortable using the pen, with the one participant who did not find the pen comfortable performing in similar ranges to the other participants. These results confirm previous research by Sanada et al (2014) and Prange & Sonntag (2019) and further extends their work as more robust statistical analysis was used to consider agreement. This addressed issues raised by Altman & Bland (1995) that the absence of evidence of differences is not evidence of agreement. It also more rigorously examined usability and comfort of a digital pen by asking the participants to evaluate their use of the pen. Further, it considered subjective comfort and usability in relation to objective drawing performance. However, there were some limitations to the study as it only examined accuracy, and not drawing process. To validate the use of a digital pen to record RCF drawing process, further agreement analysis should be conducted including the drawing process constructs. Unfortunately, the collection of data was stopped early in this study due to the Covid-19 pandemic, so only a small sample was used. Additional analysis with a second, larger sample size would increase the validity of its use.

In conclusion, recommendations have been made to conduct further validation analyses. Additionally, it is recommended that the pen is used for all pen and paper-based tasks in a multi-task protocol so as not to raise awareness to the unusual nature of the pen. However, overall, the study demonstrated that the use of a digital pen is a statistically reliable tool to use to record participants completing the RCF task. Further, it supported the combining of both sets of data into a single population sample, to be used in later analyses.

8.5 Adapting to Covid-19 – extending literature on video-based teleneuropsychology

The use of teleneuropsychology in Chapter Three emerged as a direct result of the Covid-19 pandemic which had a significant effect on research and healthcare practice globally between 2020 and 2022. Social distancing guidelines in the UK meant that face-to-face research and non-emergency healthcare was suspended. As it became apparent that the pandemic was not going to be resolved in a few weeks, or even months, solutions to contain exposure to the virus but to continue with healthcare and research were needed. The use of videoconferencing platforms became a much-used tool for people to keep in contact socially. It also provided a solution to the challenge of conducting a neuropsychological assessment when you cannot be in the same room as the participant or patient.

Teleneuropsychology has been in practice since the 1980s, however, this has predominantly been conducted via telephone, hence its name. This restricts the tasks that can be used, and also limits the behavioural observations that can be made. Empirical research into the use of videoconferencing or teleneuropsychology was being undertaken prior to the pandemic (e.g., Harder et al, 2020); mostly in countries where there are 'healthcare deserts', such as the US, France, and Australia. Whilst this research was generally supportive of using videoconferencing in neuropsychological assessment, it was limited. The pandemic provided an opportunity to add to this body of research, and to find a solution to be able to continue with the planned studies for this thesis.

With some adaptations made to how the protocol was operationalised, as indicated in Chapter Two, a similar reliability analysis to that which had been conducted for the digital pen was conducted. As well as validating the use of the RCF task via videoconferencing, all cognitive tasks in the screening protocol were considered. The results supported the use of videoconferencing in neuropsychological assessment to be a valid means of delivery, including the RCF task. However, there were some caveats to this as the trail making tasks were found to have weak agreement between the in-person and video conditions. These findings were in agreement with the work of some researchers (Gnassounou et al (2021), but not with others (Chapman et al, 2021). TMT-B appeared to be more robust than TMT-A, which may have been a result of having already completed the first trial. This aligns with previous work that suggests that multi-round assessments are more robust than single-round assessments (Brearly et al, 2017; Marra et al, 2020; Fox-Fuller et al, 2021). These findings indicated that caution needed to be taken when reaching conclusions in later analyses where the TMT tasks are used via videoconferencing methods.

There were some limitations to the study. As with the previous study, the sample size was limited. This was particularly observed when accounting for the impact that a single data point had on the data analysis outcome. For example, in the Bland-Altman analysis one outlying data point accounted for 3.7% of the total population. Therefore, a single outlier has a significant impact of whether the 95% interval threshold is met. With a larger number of participants, the impact of a single outlier is reduced, and more likely to meet the threshold. A further limitation of the study is that a between groups design was used rather than within groups, and therefore, individual differences may have impacted the extent to which agreement was observed. Whilst participants were matched for age, this does not remove all individual differences. Due to the pandemic, there was no choice but to use a between groups design, therefore this was unavoidable. However, there is

precedence for using either in previous research. Further, repeated measures research comes with its own challenges, such as the impact of practice effects, which was considered in Chapter Three in relation to the RCF.

The current study was carried out with a non-mTBI population, rather than a mTBI population. Whilst the current study and previous literature provides support in using videoconferencing to deliver the cognitive screen protocol, it does need to be held in mind that there may be a different outcome when conducted with individuals with mTBI. In fact, this was something that was raised by one of the mTBI participants in the later data collection, who commented that having to look at the screen for 90 minutes was challenging for them and was bringing on a migraine. It is recommended that similar research is conducted with a mTBI population to ensure the validity of videoconferencing with this clinical group.

A review of the literature indicated that the Letter-Number sequencing task has not been included in previous video-based teleneuropsychology validity studies. Whilst the current study indicated equivalence between the in-person and video groups, this task was found to be the most challenging to conduct online, particularly when bandwidth was limited. Bandwidth was found to be a challenge when delivering assessments that have a single-round only, or where the test stimuli cannot be repeated, such as in the letter-number sequencing task. Bandwidth also plays a role in being able to hear the participants response, which impacts reliable results recording. Harder et al., (2020) also commented on the challenges of bandwidth in their work that looked extensively at conducting a neuropsychological assessment via videoconferencing. Whilst overall the data in the study did not evidence any statistical limitations with using the letter-number sequencing task via videoconferencing, there is a need for further research to be conducted to ensure reliability of the current findings.

The findings from the study meant two things for the wider project. Firstly, it meant all of the data gathered from the non-TBI population could be considered as one sample to be employed in carrying normative data analysis later in the thesis. It also provided support for the ongoing use of videoconferencing when data was being collected with individuals with a mTBI history. The findings support previous work that video based neuropsychological assessment is a viable tool, though some tasks are more adaptable than others. However, the study also added to the existed literature. The majority of research that had been conducted used mean differences analysis, with very few conducting agreement analysis. The use of intraclass correlation analysis and Bland-Altman plots provides more robust

examination of the extent to which the use of videoconferencing is suitable. Further the study included tasks that have not previously been considered in validity analysis, such as the Letter-Number sequencing task.

8.6 Binary versus spectrum: why the rectangle is not a good indicator of cognitive deficit

Following the development of a novel set of drawing process constructs by which to measure RCF drawing process, it was necessary to understand typical population performance for each. Each construct was analysed to determine what normative population performance was, but also to consider whether demographic variables played a role. Findings from the study challenged scoring methods that argue that 'normal' performance is to draw core elements as whole units, and to draw them as the first few steps of the process (e.g., Anderson, Anderson & Garth, 2001; Bennet-Levy, 1984; Binder, 1982; Booth, 2006; Bylsma et al, 1997, as cited in Weider, Indredavik, Lydersen & Hestad, 2015; Gasparini et al, 2008; Hamby, Wilkins & Barry, 1993; Heinrichs & Bury, 1991; Klipera, 1983; Lopez et al, 2008; Rey-Osterrieth, 1941/44, as cited in Corwin & Bylsma, 2007; Savage et al, 1999; Stern et al, 1994; Sullivan et al, 1992; Zappala & Trexler, 1992). However, the findings from the current study contradict this position.

The RCF order construct data indicated that typical performance is not to draw all five core elements at the start. Rather, two or three is standard, with the Rectangle, Vertical line and Triangle sides being the core elements that are most likely drawn near the start. This does not mean, however, that they are the first three elements completed. Some systems, such as the Q-system (Bylsma, 2008, cited in Weider et al, 2015) place particular emphasis on the rectangle being drawn first. However, only 26% of participants in the current study completed the rectangle first. These results support the work of Wilson & Batchelor (2015), who argued that most adults do not start their drawing process by completing the rectangle. Examining RCF Wholeness also presented an important challenge to previous understanding as to whether or not the diagonal should be viewed as a core, foundational element. The diagonal was only drawn wholly by a third of the participants. More typically it was completed as separate arms, and as an element in each quadrant. This contradicts the majority of scoring systems which place emphasis upon drawing the diagonal as a whole unit, near the start of the drawing process.

Existing models of scoring typically take a binary view on assessing performance – either you do something, or you don't, thus not drawing the rectangle and other core

elements wholly at the start of the process was seen as indication of cognitive deficit or dysregulation. However, the findings in this study indicated that population drawing performance on the RCF should not be considered as binary phenomena. Rather each construct is on a spectrum, being distributed with the majority performing towards a central point, confirming that population performance is not skewed towards ceiling levels, as only a few participants perform at ceiling scores.

Whilst the Order, Wholeness and Continuation constructs were adaptations of previous systems, the construct of Strategy was only considered on a subjective, judgement basis in previous literature (Visser, 1972; Hamby Wilkins & Barry, 1993). Observations were that typically individuals moved strategically around the image, observing a top-to-bottom, left-to-right, quadrant-by-quadrant and/or a temporal-spatial pattern. To capture this in a systematic, measurable way, order data was examined via principal components analysis to determine if the 18 elements could be reduced to a smaller number of 'clusters' of elements which are typically observed to be drawn together. Four clusters were identified, which highlighted the qualitative strategy observations that had been made by others previously. Normative data analysis indicated performance to be on a spectrum, with typical behaviour observing RCF Strategy partially, albeit not rigidly.

One of the strengths of this study was the examination of Order as separate from Wholeness. Wholeness and Order were not found to be significantly correlated with each other, indicating that they should be measured as separate constructs. Further, the study included the construct of Continuation. Whilst it may seem that Wholeness and Continuation are measurements of the same underlying construct, the normative data indicated that they were not significantly correlated with one another. Therefore, they can be considered to be separate variables. It could be argued that measuring Wholeness considers perception from a top-down basis, whilst continuation is from a bottom-up basis. Continuation is a better indicator of the level of fractionation an individual is using, whereas Wholeness considers the level of holistic processing that is taking place.

An unusual result was that no significant correlations were found between scores on the RCF drawing constructs and age, IQ, and years of education. However, a significant difference was found between biological sex, with men on average scoring higher on RCF Strategy than women. This finding regarding sex-based differences corresponds with previous literature, such as that by Vingerhoets et al, 1998, where men have stronger strategy scores than women. As there were a limited number of men in the current study,

this needs to be explored further to validate the findings and to consider why this may be the case.

Limitations to the study should be considered, particularly as the size of the population sample was small. Whilst there is theoretical evidence to support the sample size in the current study, others would argue that for a more robust principal components analysis, a larger sample size is needed. Further, the study data was collected within three different 'environments,' in-person with a standard pen; in-person with a digital pen; and via videoconferencing. Whilst the previous studies in this thesis validated the collating of data from these conditions, the fact of whether there was an impact of this variable on performance was not considered. It is recommended that further population data is gathered, and that administration variations are also examined to ensure reliability across delivery modes.

The findings from the study provided a rationale to observe drawing performance on a spectrum basis in later analyses. It evidenced that the four constructs can be systematically, objectively, and continuously measured, indicating that they can be adapted for use within clinical neuropsychological assessment. The findings supported previous literature that argues that clinical judgements should not be made on single observations, such as whether the rectangle is drawn as a whole unit and at the start. It extended currently available scoring systems for RCF drawing process by providing normative data of a novel approach to systematically measure strategy, which has only been considered as an objective observation before. Further, it separates out the constructs that have typically been assessed when grouped together.

8.7 Examining drawing performance of the Rey-Osterrieth Complex Figure following mTBI

The previous chapters addressed limitations in the tools that were being used to measure cognitive performance following mTBI. This next chapter moved the thesis forward by addressing limitations in the data analysis methodology and methods that have been used in previous research. It also transitioned into examining the performance of individuals with mTBI, rather than a control population.

In this study, regression analysis was conducted to consider which cognitive, psychological, and demographic variables were predictive of post-concussive symptoms and behavioural challenges. It also examined whether observing the drawing process used to

complete the RCF task supported the diagnosis of PCS. It has been argued that in neuropsychological assessment we should not be looking at group differences, but individual differences. The use of regression analyses enables investigation of how change in predictive variables influences outcomes measures, irrespective of starting point.

The results supported the findings of other studies that indicated that executive function, processing speed, working memory, mental health status and age would be the strongest influencers of RPCSQ scores. It also identified that IQ is an important variable, but not education, a factor which had formed an integral part of one of the hypotheses of the study. With regards to the final hypothesis of the study, the results of the study indicated that the different variables influence the different outcomes measures. Interestingly, IQ was found to have a stronger influence on fatigue levels than general RPCSQ symptoms. The study also indicated that whilst executive function change underpins the symptoms and behavioural challenges of PCS, different functions influence different outcomes. One of the most important findings revealed that fatigue is more sensitive to change in inhibitory control and cognitive flexibility, however, functional challenges are more sensitive to variation in working memory. Such findings indicate that different aspects of PCS are underpinned by different aetiologies and support the theory that PCS is not a singular condition, rather it is an umbrella term for sub-types or dimensions of post-concussive symptoms.

Directly addressing the research question of this thesis, a key aim was to determine if the constructs of the RCF drawing process would support the diagnosis of PCS. The RCF constructs were found to be predictive of all four outcome measures, with each measure being of different importance. RCF Order was found to be an important variable of PCS symptoms and cognitive fatigue. However, RCF Continuation and Time to Completion were better indicators of functional and behavioural challenges. RCF Wholeness, however, was not included in any of the models as it was indicated to have very minimal association with any of the outcome measures, as well as multicollinearity. This illustrates the importance of measuring order and wholeness as separate constructs. However, RCF Wholeness was found to be a significant difference between the mTBI and control groups. Given this finding, and the understanding that individual correlation coefficients do not necessarily indicate strength of influence within regression modelling, it is recommended that the influence of this construct is examined further using regression methods that can manage multicollinearity.

Regression analyses have been used in previous studies to examine predictor variables of post-concussive symptoms and behavioural challenges following mTBI. However, the current study adds to this work by examining measures of symptoms and

functional challenges alongside one another. By doing so, it has produced the understanding that different challenges following mTBI are underpinned by different variables, including different cognitive functions. The analysis also evidences that examining the drawing process that an individual uses can support the diagnosis of PCS in individuals who are experiencing long-term symptoms following mTBI.

8.8 The patient report – understanding the nature of long-term post-concussive symptoms

The final study in the thesis sought to examine how the patients' subjective reports can help determine whether they have PCS. Previous literature has raised questions regarding the sensitivity and specificity of RPCSQ. Criticisms of PCS and tools to measure PCS focus on the fact that similar symptoms are reported in other conditions, such as multiple sclerosis (Manjalay et al, 2019), and more recently, long covid (Benjamin et al, 2022). Further, the general non-TBI population report experience of symptoms to an extent (Dean et al, 2012). This has fuelled the debate about whether PCS is organically underpinned, or if there are other contributing factors. In addition, given the number of post-concussive symptom inventories that are available, there are also questions about whether tools such as the RPCSQ capture all aspects of PCS, and thus measuring the right problems. One of the limitations of tick-box symptom screens are that they are typically one-dimensional in what they can tell us. A tick box item asking someone if they experience fatigue or not may tell us that the person experiences fatigue, for example, but it provides little information about the manifestation of that fatigue, or how it differs over time and activity. Qualitative data can provide more detail than numerical indications of the complex and sometimes invisible challenges individuals are facing following mTBI.

The last study in this thesis used a mixed methods approach to examine how participants responded to the RPCSQ, as well as analysing the qualitative data revealed in their responses to the challenges they faced. By doing so, it was possible to identify convergence in symptom reporting across the two methods, as well as divergence. The analysis indicated that there is convergence between scores on the RPCSQ, and the narrative individuals use to discuss the challenges they face following mTBI. Participant narratives aligned with the quantitative analysis from the RPCSQ, highlighting that cognitive fatigue and sensory overload are significant challenge for participants. The side-by-side analysis also supported that some symptoms, such as nausea, dizziness, and blurred vision, are not significant symptoms at the chronic stages of mTBI. Such results support the RPCSQ as a valid measure of severity of symptoms.

However, the analysis also indicated several areas where there was divergence, as well as lack of clarity about the experience of PCS symptoms. Psychometric modelling of PCS symptoms often considers sensory overload to be associated with headaches and migraines. However, the thematic analysis indicated that these sensory difficulties are more related to social challenges rather than being migraine related. Many individuals found it difficult to process complex sensory situations, such as multiple streams of conversation, suggesting that problems were more cognitive and effort-based rather than being associated with headaches. This correlated with commentary from participants about the need to rest and recharge for an extended period after socialising. Furthermore, it was clear that the RPCSQ did not capture the variability of symptoms or the sensation of 'brain fog' or detachment which several participants spoke about.

In addition to indicating where there were differences between qualitative narrative and quantitative measures of PCS via the RPCSQ, the analysis illustrated the complexity of PCS symptoms that are not necessarily captured in questionnaires. The theme *Energy and Effort* demonstrated that fatigue following mTBI is not one-dimensional, but a multi-faceted group of specific factors. Participants spoke about effort and energy as two interconnected but separate challenges. The manner in which fatigue was captured by the participants' accounts reflects the theoretical model of PCS fatigue proposed by Wylie and Flashman (2017), in that PCS fatigue is a result of increased effort. The analyses from the current study highlighted the need to understand more fully what the nature of the of cognitive fatigue is. Three aspects of effort and energy were identified – starting energy levels, the effort required to complete tasks, and finally, the extended recovery time to recharge energy levels. The findings provide direction for undertaking further research to understand post-concussive fatigue; to understand if it is one of these specific aspects, or if it is a combination of all three.

The complexity of cognitive fatigue was also reflected in the theme *Psychology of living with mTBI*. As discussed previously, mental health is of particular theoretical interest to this thesis. Therefore, the fact that such a strong theme relating to mental health following mTBI was identified in the narrative of the participants is significant. However, the thematic analysis within this study indicated an interesting pattern, that participants experienced frustration and irritability more than depression. The analysis also indicated that the psychological experience of mTBI is multi-dimensional and multi-directional. Whilst the findings from the study do not rule out the fact that mental health status plays a role in the symptoms of RPCSQ, it does illustrate that psychological factors may be an indicator of

organic damage, trauma, or chronic disability. Evidence of poor mental health should not be a reason to disregard the existence of long-term consequences of mTBI, rather, understanding of the cause and management of mental health difficulties should form part of rehabilitation formulations.

The findings from this mixed-methods study supported previous research regarding the validity of the RPCSQ in measuring post-concussive symptoms. In addition, it has supported research that argues that the RPCSQ is useful for screening symptoms at a static point, however, it is not suitable for capturing the dynamic nature of PCS (Medvedev et al, 2018; Moller et al, 2019). The current study furthered this research by adopting a mixed methods approach, analysing qualitative data as well as quantitative data to illustrate the dynamic and environmental nature of PCS, and in particular cognitive fatigue following mTBI. The convergence of both types of data validated the findings from quantitative data. The study also indicated the value of obtaining a patient's description and an account of their challenges in their own words to support a diagnosis of PCS as patterns in how a patient talks about the problems, they are facing can be indicative of PCS.

8.9 Answering the research question

The research question for this thesis asked whether or not examining the drawing process used to replicate the RCF provides supporting evidence for a diagnosis of PCS. As the results from the regression analyses in Chapter Seven indicate, RCF drawing process constructs provide a sensitive indicator of long-term post-concussive symptoms and real-life behavioural challenges. However, this is an oversimplification of the implications that result from this research.

Predictors of the MFS were examined due to previous evidence which indicates that fatigue is one of the most common and persisting symptoms of mTBI (Lannsjo et al, 2009; Roe et al, 2009; Moller et al, 2019). The significance of cognitive fatigue as a symptom following mTBI was supported by the findings from the final study in this thesis. Both the quantitative and qualitative analysis illustrated that fatigue is commonly experienced as a long-term or chronic symptom of mTBI. However - and more importantly for the participants in this research – it was one of the most impactful symptoms on their life. However, there is still very limited understanding about the nature and causes of this fatigue.

The results from the regression analyses indicated that reduced cognitive flexibility and inhibitory control underpin cognitive fatigue following mTBI and the narrative data from

the mixed-methods study provided examples of this in real-life situations. Whilst sensory overload has typically been associated with migraines and headaches, participants discussed this as a challenge of complex situations, particularly where there are several streams of information running concurrently. For example, the inability to ignore irrelevant sounds, and the reduced ability to focus in on a specific conversation provides a strong theoretical explanation for the challenges individuals with persisting mTBI symptoms face. The additional effort required to process all of the stimuli, and the additional effort to enforce inhibitory and flexibility functions can be argued to result in increased experience of fatigue. This theoretical explanation is supported by the model of 'Fatigue as an indicator of effort' posited by Wylie and Flashman (2017) as an explanation of fatigue following mTBI. This is a significant implication, as it provides support of a cognitive explanation for PCS; one which can also be explained by neurotransmitter dysregulation associated with cholinergic disruption (Sahakian et al, 1988). It also paves the way for investigation to further understand the causes of post-mTBI fatigue and to consider treatments and rehabilitation to address.

One of the challenges with arguing that there is an organic underpinning of long-term mTBI symptoms has been the evidence of increased depression and anxiety in these individuals. The results from this thesis do not dispute the role that mental health has on the subjective reporting of symptoms. The regression analyses indicated that change in depression was one of the strongest influencers of subjective reporting of general RPCSQ symptoms, fatigue, and behavioural challenges, indicated by the DEX and CFS. Participants' narratives also indicated the significance of mental health and the psychological impact of their injury and recovery. However, correlation analyses indicated that there was limited evidence of an association between mental health and cognitive function. A further interesting observation was that overall, the mTBI participants did not have clinically significant levels of depression or anxiety, yet these variables were still strong influencers of subjective outcome measures. Low mood or anxiety only need to be moderate to influence higher levels of symptom reporting. The findings throughout the second half of this thesis suggest that whilst mental health status is a factor of subjective post-concussive symptoms, it does not reduce the importance of cognitive dysregulation as a consequence of mTBI. Further, mental health status may be a mediator of symptoms, but it is not necessarily the cause. However, to determine this, further regression analysis is needed.

Throughout the thesis, three methodological questions underpinned the work that was carried out to investigate the thesis research question. Firstly, it was asked are we measuring the right thing? The findings from work conducted by Geary et al., (2011)

indicated that it is *how* individuals carry out tasks that should be examined. This tenet formed part of the theoretical framework that underpinned the current focus of this research. Rather than measuring for loss of function, a task was needed that indicated differences in function. The findings from this thesis support that assessments that examine functional processing rather than attainment are better indicators of post-concussive symptoms. The fact that recent imaging technology has concluded that mTBI is associated with disruption to axonal and neurotransmitter messaging, rather than atrophy or damage, that subsequently gives rise to the characteristic symptoms of mTBI further supports this theory of cognitive disruption rather than neurological loss.

The second methodological question asked how do we measure this disruption? Examining drawing process of the RCF provided a vehicle by which to observe this. However, answering this question more comprehensively fell into two parts. Firstly, how to score drawing process needed to be determined. Whilst there are a multitude of options already available by which to record drawing process, a review of this literature indicated that none of these options were found to be suitable. No system measured all four constructs that were identified in the review as independent constructs, or on a continuous basis. Therefore, a novel scoring approach to measuring drawing process was developed and validated for use in the other? studies. It is important to address why a different tool was used if research indicates that the CVLT already shows changes that are consistent with the neurotransmitter disruption. The findings from this study demonstrated that verbal and visual tasks can be affected differently, and therefore should both be included within a neuropsychological assessment. This reflects existing neuropsychological practice (Evans, 2010). However, it is also important to consider the differences between the two tasks. Verbal recall tasks such as the CVLT and HVLTL have a single level of semantic organisation, whereas the RCF has a more complex and hierarchical structure. Therefore, the two tasks are not wholly comparable. The advantages of the more complex structure are that it has provided an explanation of why there are processing differences. The population analysis also indicated a normal distribution pattern to performance, which is more conducive to the development of standardised clinical scales.

The second part of how to measure cognitive disruption was focused on the practicalities of how to record drawing process. As discussed, one of the factors that can limit the functionality of the RCF task is the need to record and revisit the drawing process an individual adopts. The use of a digital pen such as the Neo Smartpen provides an economically viable solution for recording drawing process without the need for cumbersome recording equipment such as the use of tripod and recording device. The validation of the

digital pen in this thesis indicated that individuals use it as they do a normal pen and paper, yet it provides the opportunity for a more detailed examination of how individuals complete drawing-based tasks. One of the benefits of using a digital pen is that the digital data can be downloaded and used to produce sophisticated data analysis. Such work is already being conducted with other clinical populations and with tools that are not as reflective of a standard pen (Smith & Hiller, 1996; Li et al, 2011; Riordan et al, 2013; Jurica et al, 2015; Kaber et al, 2014; Qin et al, 2015). This presents opportunities to move the pen beyond simply being a tool that records that requires analogue scoring. Automated scoring would reduce the effort load and subjectivity of scoring. Additionally, observations could be made that cannot be easily measured and analysed without the enhanced function of a digital platform, such as examining pause time.

The final methodological question asked how do we analyse performance? This question was in response to the concerns raised by authors of previous papers about treating individuals with a history of mTBI as a homogenous group within analysis. Further, it addressed the limitations of using a group difference design to provide evidence of cognitive deficits. The work in this thesis addressed this through the use of different types of analysis, such as regression analyses and a mixed-methods approach. The use of these approaches has identified which cognitive functions and assessments are the best indicators of cognitive dysfunction. It has also provided detailed information about the nuances of the symptom's individuals experience, which will inform future research and clinical work.

The research studies presented here addressed a number of the questions that have been raised by other researchers as to how understand, measure, and analyse the aetiology of persisting symptoms following mTBI. It also furthered our understanding of the outcomes of mTBI by focusing on those who would be considered to have chronic symptoms. Whilst there is a wealth of literature that considers the acute and post-acute symptoms of mTBI – typically within the three-month recovery window there is limited research that examines the long-term effects, particularly beyond 12 months. The average time since injury in the current study was 9 years.

8.10 Limitations

Whilst limitations specific to the individual studies have been discussed earlier in this chapter, there are additional limitations to the thesis, which are more general. As has been described throughout this thesis, some of the methods that were used were not what were initially intended. This had an impact on what could be investigated. That the majority of the

work was conducted via video-based teleneuropsychology means that the results that have been concluded here should be considered conservatively and, therefore, warrant further investigation to validate them in an in-person environment. Adaptions had to be made to how some of the tasks were administered, which may have impacted the validity and reliability of the scores. It was not always possible to see if participants were completing the tasks correctly, and therefore not possible to intervene, if necessary, which may have led to poorer performance than would be the case if conducted in-person. However, despite these constraints the research does provide piloted methodology protocols that can be employed within future research and clinical practice.

As opportunistic sampling was employed in the current study, this presented a challenge with ensuring a diverse sample that was representative of the wider population. Diversity of the population sample is even more of a problem when external challenges restrict recruitment, and results in small sample sizes. The participants who volunteered to take part were generally highly educated and typically had high average estimated IQ levels. As a result, the findings should be viewed with caution when considering individuals with lower levels of IQ and education. That said, one of arguments of Binder and colleagues was that where cognitive differences were found, these were due to reduced pre-morbid functioning and low intellect. That the current study was skewed towards those with higher levels of intellect argues against this, by providing evidence of long-terms symptoms and deficits in this demographic.

Another demographic skew was with regards to the uneven balance between males and females. However, previous research into the effects of mTBI and concussion have predominantly been with male participants, most likely an artifact of research being conducted within high impact sport and military personnel which were traditionally male dominated. Therefore, there is only limited empirical evidence to understand the impact of mTBI and concussion on women. Recent research has indicated that female sports athletes are almost twice as likely to experience concussion than men (Bretzin et al, 2021), and take longer to recover (Gallagher et al, 2018). Therefore, the imbalance in the biological sex of the participants in this study positively adds to our understanding of the effects of mTBI on women. However, it is necessary to replicate the research with a more balanced biological sex split, particularly as it was found to be an important predictor in the regression analysis in Chapter Six. A final demographic characteristic that also has to be considered is that all participants were English speaking either as a first language, or bilingual. Any individuals who were not British by birth were from Western Europe. Therefore, there is very narrow

cultural variance. How generalisable the findings here are beyond this demographic are not clear, and therefore, it is recommended that further research is conducted more globally.

Whilst recruitment is commonly a problem in research, particularly when it requires a significant time contribution from the individual, it is important to consider if this was exacerbated by the study design. It was questioned by the research team during the study if recruitment was being challenged as a result of the extensive inclusion and exclusion criteria that was set. Significantly more people completed the inclusion-exclusion pre-screen than were eligible to take part. However, each exclusion variables were included due to evidence that suggested results could be confounded by co-morbid conditions or demographics. Further, it cannot be ignored what was occurring globally at the time the research was being conducted, the pandemic and lockdowns were once-in-a-lifetime events, which had a significant impact on people's lives. For some, this unfortunately meant having to deal with loss of livelihood or loved ones. For others, it may have been coping with illness, isolation and loneliness and potentially also experiencing long-covid symptoms, which have been indicated to have a neurological effect, including cognitive fatigue (Ceban et al, 2022). Given some of the similarities in symptoms experienced with PCS and long-covid, the work conducted in this thesis may provide some insight into long-covid. Therefore, whilst it is recognised that the restrictions on participation did contribute to low recruitment, the rationale for these were sound and warranted. For future work, managing these challenges needs to be considered in the planning stages.

A further limitation of the work is that conclusions are based upon observation of behaviour. Whilst this is common practice in psychological research, it does identify where future research could be conducted to provide further support for the theoretical conclusions that have been made, that executive function underpins post-concussive symptoms. Eye tracking provides a neuroscientific approach to observing how an individual processes the complexity of the RCF. Patterns of saccades and eye-movement around the image may provide valuable information about the perceptual processing that is taking place in breaking down the image. Saccade performance has been investigated in relation to cognitive function (Moller et al, 2019) and may provide further support for cognitive disruption following mTBI. 'Think aloud' protocol allows the analysis of the internal problem solving that is taking place and would enable a qualitative approach to analysing how individuals attend to and perceive the complex nature of the RCF although it is recognised that this only captures explicit problem solving, and not implicitly, automatic processes that occur at a sub-conscious level.

This research originally intended to investigate potential bio-psycho explanations for atypical performance on the RCF tasks with a more detailed focus on cholinergic disruption and subsequent modulation for cognitive enhancement. The neurotransmitter Acetylcholine is known to be disrupted following mTBI, with cholinergic functioning being reduced (e.g., Sahakian et al., Mitra et al, 2020) leading to attentional, effortful cognitive processing and inhibitory difficulties. Nicotine, a cholinergic agonist, has been suggested as a treatment for conditions which are associated with cholinergic decline or disruption or neuroinflammation, such as TBI, Alzheimer's Disease and Schizophrenia (Heishman et al, 2010; Mitra et al, 2020). Nicotine has been found to enhance attentional and effortful processes and improve cognitive flexibility and inhibition and reduce intrusions (Edginton and Rusted, 2003); improve performance in visual searching tasks, with increased search speed and better focus (Rycroft, Rusted & Hutton, 2005); and increase semantic strategy processing in Parkinson's Disease (Holmes, Copland, Silburn & Chenery, 2011). Acetylcholine is also implicated in other neurological processes and conditions, such as cognitive aging, and is indicated to underlie decline in working memory and executive functions (Younger-Bernier, Kamil, Tramlay & Davidson, 2012). Whilst there are studies that examine the relationship between the cholinergic system and RCF recall, it does not appear that any focus has been given to drawing process. However, as recruitment was on an opportunistic basis, and no stratification for nicotine use was included, only a handful of the volunteer participants had a history of smoking. Whilst the low number of participants who smoke or have a history of smoking was good for their overall health, this was not so good for being able to examine nicotine and cholinergic effects in this thesis! A very recent systematic review by Rao, McConnel & Litofsky (2022) has indicated that more research is needed to understand the effects of nicotine-based cholinergic treatments for improving symptoms and cognitive function in TBI, which this thesis also recommends.

8.11 Impact

Since starting the work towards this thesis, there has been an explosion in the attention that is being given towards understanding the mechanics and outcomes of mild traumatic brain injury and concussion. This has mostly been geared by an increased interest in sports-related concussion. During the course of this PhD, a legal case has been filed on behalf of professional soccer and rugby football players for occupational injury, as a result of long-term concussive damage. As more sports become professional and commercial enterprises, the exposure to potential concussive and sub-concussive events by players is argued to increase (Newman & Dennis, 2021). Athletes compete more frequently,

train more frequently, and are stronger. Technology-advancement in sporting equipment has also increased the risk of concussion. For example, the nature of the modern soccer ball retains its velocity in wet weather, thereby increasing the opportunity for heading, which is being investigated as a cause of CTE in professional soccer players. The work produced in this thesis supports the existence of long-term consequences from concussion and sub-concussion events. It is important to acknowledge, however, that this is not an argument for banning these sports. Rather it is a call to reduce the risk of potential mTBI through better safety protocols.

It is also becoming increasingly acknowledged that TBI is a risk factor for homelessness, substance abuse/addiction, legal incarceration and, sadly, suicide (Ponsford et al, 2008; Mani et al, 2017; Oddy et al, 2012; Durand et al, 2017; Fralick et al, 2019). Whilst it is accepted that there is a small sub-set of individuals who experience chronic post-concussive symptoms, the debate regarding the aetiology of these injuries challenges the validity of PCS. This may underpin some of the challenges people with persisting symptoms have when trying to access rehabilitation services (Kander, 2020). There is also a disconnect between how the general public understands concussion, where it is not considered to be a form of brain injury, and therefore, not considered to be as serious. The findings here provide support that symptoms of mTBI and concussion have an organic basis. It argues for increased understanding within healthcare and the general population of the potential seriousness of mTBI, and that it is not 'just a concussion.' The conclusions in this study are somewhat at odds with the current healthcare provision for TBI, such as the recent closure of the Oliver Zangwill Centre in Cambridge.

8.12 Concluding summary

The long-standing debate regarding the existence of cognitive deficits following mTBI has been predicated on limitations in the methodology and methods that have been used to examine the phenomenon. The work in this thesis moves this debate forward by addressing several of these limitations. By doing so, the thesis provides evidence that supports the argument that mTBI can result in organic damage, which underpins the complaints patients can report for many months and years after their injury, and in particular, the experience of cognitive fatigue. Several recommendations are drawn from this research, some of which relate to future research and have already been given throughout this chapter. Further, it is recommended that including observations of the drawing process a patient uses to complete the RCF task can be indicative of reduced inhibitory control and cognitive flexibility, which may support a diagnosis of PCS. Whilst including detailed examination of the drawing

process used to replicate the RCF does increase administration time for the clinician, it is no more demanding on patients. This is an important consideration given that neuropsychological assessment is already typically quite long. However, the increase in administration time on the part of the clinician can be reduced through further investigation about how to automate scoring of drawing-based tasks. With regards to the use of the digital pen in research and clinical practice, it is recommended that the pen is used for all pen and paper-based tasks, so as not to draw attention to its nature.

The findings from this thesis have significant implications on the understanding of symptoms following mTBI. These have the potential to improve access to the correct support and rehabilitation for individuals experiencing long-term and debilitating symptoms. Further, they can increase public understanding of mTBI and concussion, and pave the way for improvements being made in activities that are known to increase the risk of these types of injuries.

To conclude, the following recommendations are given:

- Observation of the drawing process used to complete the RCF task is included within neuropsychological evaluation of individuals reporting long-term cognitive fatigue, including individuals who have PCS, as well as other conditions characterised by fatigue, such as long-covid and multiple sclerosis. Drawing process is assessed via the use of the novel scoring system that has been developed in this thesis, such that the constructs of Wholeness, Continuation, Order and Strategy are measured.
- Validated population normative scales are developed of the four RCF drawing constructs for clinical use.
- Attention is given in clinical assessment to mental health status of individuals with persisting symptoms following mTBI.
- Further analysis is conducted to assess the validity of a digital pen to record drawing process. This includes usability and feasibility analysis of analogue scoring of the digital playback, development of automated scoring, and validity analysis with clinical populations.
- Research is conducted that investigates the dimensions of cognitive fatigue, so as to understand the cognitive and biological aetiology underpinning each dimension. This includes further examination of the role of inhibitory control and cognitive flexibility play within cognitive fatigue.

Appendix One – Assessment protocol

Assessment	Form of delivery	Time
Demographic details	Questions verbally delivered by administrator	3-5 minutes
National Adult Reading Test	Participant to read aloud from list of words	3-5 minutes
Hopkins Verbal Learning Task – Learning phase	Verbal: researcher reads list of words to participant to recall. Participant recalls. 3 rounds	5-10 minutes
Rey Complex Figure Task – copy phase	Drawing: copy the RCF image	2-5 minutes
FAS verbal fluency Task	Verbal: participant to give as many words as can beginning with a given letter in 60 seconds. 3 rounds	5 minutes
Rey Complex Figure Task – immediate recall phase	Drawing: recall Rey image	2-5 minutes
Semantic verbal fluency Task	Verbal: participant to give as many animals as possible in 60 seconds	2 minutes
Category switching fluency task	Verbal: participants to alternate between fruit and furniture, giving as many as possible in 60 seconds	2 minutes
Trail Making Task – stages A and B	Pen & Paper: participant to draw continuous line on paper to join dots as quick as possible. Timed. A = numbers 1-25 B = alternating numbers and letters	5 minutes
Stroop Test – Word and Colour stages	Verbal: Participant to read aloud from printed list Word = read out word; Colour = read colour of ink. Timed – 2 minutes maximum	5 minutes
Letter-Number Sequencing	Verbal: Researcher reads out sequence of numbers and letters. Participant to rearrange in order of numbers then letters	5 minutes
Hopkins Verbal Learning Task – Delayed Recall Phase	Verbal: participant to recall words from earlier	2-3 minutes
Hopkins Verbal Learning Task – Recognition phase	Verbal: researcher reads aloud list of words which includes words from earlier phase. Participant to indicate yes or no if were in list	2-3 minutes
Rey Complex Figure Task – Delayed recall phase	Drawing: recall Rey image as per allocation, (see Table II.)	2-5 minutes
Hospital Anxiety & Depression Scale	Written: Participant to complete questionnaires on Qualtrics	2-3 minutes
Becks Depression Inventory – II	Written: Participant to complete questionnaires on Qualtrics	2-3 minutes

Assessment	Form of delivery	Time
Cognitive Failures Questionnaire	Written: Participant to complete questionnaires on Qualtrics	2-3 minutes
DEX	Written: Participant to complete questionnaires on Qualtrics	2-3 minutes
Mental Fatigue Scale	Written: Participant to complete questionnaires on Qualtrics	2-3 minutes
Rivermead Post-Concussion Symptoms Questionnaire	Written: Participant to complete questionnaires on Qualtrics	2-3 minutes
E-Pen usability survey	Written: Participant to complete questionnaire via Qualtrics (if applicable)	1 minute
Free-format report of symptoms	Written: Participant to complete questionnaire via Qualtrics (if applicable)	1-3 minutes

Appendix Two – Demographic Questionnaire

General information

The following data is used so that we can ensure we have a wide variety of ages and education, and balanced gender mix.

Current Age (in years)	
Date of Birth	
Gender at birth	
Highest level of education e.g. GCSE, A-Level, BTEC, HND, Degree, Masters	
Age left school	

Handedness

Please place a tick or 'X' in the box that most applies to which hand you would use for each of the below actions

	Always Left	Usually Left	Both Equally	Usually Right	Always Right
Writing					
Throwing					
Brushing teeth					
Using a spoon					

1. Do you regular take any prescribed or non-prescribed medication? If yes, please provide details.

--

2. How long ago did your injury take place? *

--

3. What was the cause of your injury? (only general details are required) *

--

4. Roughly, what area of your head/brain were injured? *

5. Did you lose consciousness, and if so, how long for? *

6. Did you experience memory loss, and if so, how long for*

7. Have you noticed any changes to your mood or personality since your injury? If so, in what way? *

** Indicates questions that were only delivered to individuals with a history of mTBI*

Appendix Three – Hospital Anxiety & Depression Scale

Tick the box beside the reply that is closest to how you have been feeling in the past week.
Don't take too long over your replies: you immediate is best.

I feel tense or 'wound up' (A)

3	Most of the time
2	A lot of the time
1	From time to time, occasionally
0	Not at all

I still enjoy the things I used to enjoy (D)

0	Definitely as much
1	Not quite as much
2	Only a little
3	Hardly at all

I get a sort of frightened feeling as if something awful is about to happen (A)

3	Very definitely and quite badly
2	Yes, but not too badly
1	A little, but it doesn't worry me
0	Not at all

I can laugh and see the funny side of things (D)

0	As much as I always could
1	Not quite as much now
2	Definitely not so much now
3	Not at all

Worrying thoughts go through my mind (A)

3	A great deal of the time
2	A lot of the time
1	From time to time, but not as often
0	Only occasionally

I feel cheerful (D)

3	Not at all
2	Not often
1	Sometimes
0	Most of the time

I can sit at ease and feel relaxed (A)

0	Definitely
1	Usually
2	Not Often
3	Not at all

I feel as if I am slowed down (D)

3	Nearly all of the time
2	Very often
1	Sometimes
0	Not at all

I get a sort of frightened feeling like 'butterflies' in the stomach (A)

0	Not at all
1	Occasionally
2	Quite Often
3	Very Often

I have lost interest in my appearance (D)

3	Definitely
2	I don't take as much care as I should
1	I may not take quite as much care
0	I take just as much care as ever

I feel restless as I have to be on the move (A)

3	Very much indeed
2	Quite a lot
1	Not very much
0	Not at all

I look forward with enjoyment to things (D)

0	As much as I ever did
1	Rather less than I used to
2	Definitely less than I used to
3	Hardly at all

I get sudden feelings of panic (A)

3	Very often indeed
2	Quite often
1	Not very often
0	Not at all

I can enjoy a good book or radio or tv programme (D)

0	Often
1	Sometime
2	Not often
3	Very seldom

TOTAL SCORE DEPRESSION (D) _____

TOTAL SCORE ANXIETY (A) _____

Appendix Four – Becks Depression Inventory – II

This questionnaire consists of 21 groups of statements. Please read each group of statements carefully. And then pick out the one statement in each group that best describes the way you have been feeling during the past two weeks, including today. If several statements seem to apply equally well, tick the highest number from that group. Be sure that you do not choose more than one statement for any group.

Sadness

0	I do not feel sad
1	I feel sad much of the time
2	I am sad all of the time
3	I am so sad or unhappy that I can't stand it

Pessimism

0	I am not discouraged about my future
1	I feel more discouraged about my future than I used to
2	I do not expect things to work out for me
3	I feel my future is hopeless and will only get worse

Past Failure

0	I do not feel like a failure
1	I have failed more than I should have
2	As I look back, I see a lot of failures
3	I feel I am a total failure as a person

Loss of Pleasure

0	I get as much pleasure as I ever did from the things I enjoy
1	I don't enjoy things as much as I used to
2	I get very little pleasure from the things I used to enjoy
3	I can't get any pleasure from the things I used to enjoy

Guilty Feelings

0	I don't feel particularly guilty
1	I feel guilty over many things I have done or should have done
2	I feel quite guilty most of the time
3	I feel guilty all of the time

Punishment Feelings

0	I don't feel I am being punished
1	I feel I may be punished
2	I expect to be punished
3	I feel I am being punished

Self-dislike

0	I feel the same about myself as ever
1	I have lost confidence in myself
2	I am disappointed in myself
3	I dislike myself

Self-Criticalness

0	I don't criticise or blame myself more than usual
1	I am more critical of myself than I used to be
2	I criticize myself for all of my faults
3	I blame myself for everything bad that has happened

Suicidal Thoughts or Wishes

0	I don't have any thoughts of killing myself
1	I have thoughts of killing myself, but I would not carry them out
2	I would like to kill myself
3	I would kill myself if I had the chance

Crying

0	I don't cry anymore than I used to
1	I cry more than I used to
2	I cry over every little thing
3	I feel like crying, but I can't

Agitation

0	I am no more restless or wound up than usual
1	I feel more restless or wound up than usual
2	I am so restless or agitated that that I have to keep
3	I am so restless or agitated that I have to keep moving or doing something

Loss of interest

0	I have not lost interests in other people or activities
1	I am less interested in other people or things than before
2	I have lost most of my interest in other people or things
3	It's hard to get interested in anything.

Indecisiveness

0	I make decision about as well as ever
1	I find it more difficult to make decisions than usual
2	I have much greater difficulty in making decisions that I used to
3	I have trouble making any decisions

Worthlessness

0	I do not feel I am worthless
1	I don't consider myself as worthwhile and useful as I used to
2	I feel more worthless as compared to others
3	I feel utterly worthless

Loss of Energy

0	I have as much energy as ever
1	I have less energy than I used to have
2	I don't have enough energy to do very much
3	I don't have enough energy to do anything

Changes in Sleeping Pattern

0	I have not experienced any change in my sleeping
1a	I sleep somewhat more than usual
1b	I sleep somewhat less than usual
2a	I sleep a lot more than usual
2b	I sleep a lot less than usual
3a	I sleep most of the day
3b	I wake up 1-2 hours early and can't get back to sleep

Irritability

0	I am not more irritable than usual
1	I am more irritable than usual
2	I am much more irritable than usual
3	I am irritable all the time

Changes in appetite

0	I have not experienced any change in my appetite
1a	My appetite is somewhat less than usual
1b	My appetite is somewhat greater than usual
2a	My appetite is much less than before
2b	My appetite is much greater than before
3a	I have no appetite at all
3b	I crave food all of the time

Concentration difficulty

0	I can concentrate as well as ever
1	I can't concentrate as well as usual
2	It's hard to keep my mind on anything for very long
3	I find I can't concentrate on anything

Tiredness or Fatigue

0	I am no more tired or fatigued than usual
1	I get more tired or fatigued more easily than usual
2	I am too tired or fatigued to do a lot of the things I used to do
3	I am too tired or fatigued to do most of the things I used to do

Loss of interest in sex

0	I have not noticed any recent change in my interest in sex
1	I am less interested in sex than I used to be
2	I am much less interested in sex now
3	I have lost interest in sex completely.

Appendix Five – Mental Fatigue Scale

This questionnaire will help us to find out you perceive your health. We are interested in your present condition, that is, how you have felt in the past month. Each question below is followed by four statements that describe: No (0), Slight (1), Fairly serious (2), and Serious problems (3). We would like you to select the figure before the statement that best describes your problems. Should your problem fall between two statements, there are also figures that indicate this.

Fatigue.

Have you felt fatigued during the past month? It does not matter if the fatigue is physical (muscular) or mental. If you recently experienced something unusual (for example an accident or short illness) you should try to disregard it when assessing your fatigue.

0	I do not feel fatigued at all
1	I feel fatigued several times every day but I feel more alert after a rest
2	I feel fatigued for most of the day and taking a rest has little or no effect
3	I feel fatigued all the time and taking a rest makes no difference

Lack of initiative

Do you find it difficult to start things? Do you experience resistance or a lack of initiative when you have to start something, no matter whether it is a new task or part of your everyday activities?

0	I have no difficulty starting new things
1	I find it difficult starting new things than I used to. I'd rather do it some other time.
2	I takes a great effort to start things. This applies to everyday activities such as getting out of bed, washing myself and eating.
3	I can't do the simplest of everyday tasks (eating, getting dressed). I need help with everything

Mental Fatigue

Does your brain become fatigued quickly when you have to think hard? Do you become mentally fatigued from things such as reading, watching TV, or taking part in a conversation with several people? DO you have to take breaks or change to another activity?

0	I can manage in the same way as usual. My ability for sustained mental effort is not reduced.
1	I become fatigued quickly but am still able to make the same mental effort as before
2	I become fatigued quickly and have to take a break or do something else more often than before.
3	I become fatigued so quickly that I can do nothing or have to abandon everything after a short period (~5mins).

Mental Recovery

If you have to take a break, how long do you need to recover after you have worked 'until you drop' or are no longer able to concentrate on what you are doing.

0	I need to rest for less than an hour before continuing whatever I am doing
1	I need to rest for more than an hour but do not require a night's sleep
2	I need a night's sleep before I can continue whatever I am doing
3	I need several days rest in order to recover

Concentration difficulties

Do you find it difficult to gather your thoughts and concentrate?

0	I can concentrate as usual
1	I sometimes lose concentration, for example, when reading or watching TV
2	I find it so difficult to concentrate that I have problems, for example, reading a newspaper or taking part in a conversation with a group of people
3	I always have such difficulty concentrating that it is almost impossible to do anything

Memory problems

Do you forget things more often than before, do you need to take notes or do you have to search for things at home or at work?

0	I have no memory problems
1	I forget slightly more often than I should, but I am able to manage by making notes
2	My poor memory causes frequent problems (for example forgetting important meetings or turning off the cooker)
3	I can hardly remember anything at all

Slowness of thinking

Do you feel slow or sluggish when you think about something? Do you feel that it takes an unusually long time to conclude a train of thought or solve a task that requires mental effort?

0	My thoughts are neither slow nor sluggish when it comes to work involving effort
1	My thoughts are a bit slow one or a few times each day when I have to do something that requires serious mental effort
2	My thoughts often feel slow and sluggish, even when carrying out everyday activities, for example a conversation with a person or when reading a newspaper
3	My thoughts always feel slow and sluggish

Sensitivity to stress

Do you find it difficult to cope with stress, that is doing several things at the same time while under time pressure?

0	I am able to cope with stress in the same way as usual
1	I become more easily stressed, but only in demanding situations that I was previously able to manage
2	I become stressed more easily than before. I feel stressed in situations that previously did not bother me.
3	I become stressed very easily. I feel stressed in unfamiliar or trying situations.

Increased tendency to become emotional

0	I am not more emotional than I used to be
1	I am more emotional than other people but it is something that is natural for me. I start to cry or my eyes fill with tears easily, but only in relation to things that affect me deeply.
2	My emotions are problematic or embarrassing. I sometimes even start to cry about things that mean nothing to me. I try to avoid certain situations because of this.
3	My emotions cause me great problems. They disturb my day-to-day relationship with members of my immediate family and make it difficult for me to cope outside the home.

Irritability or a 'short fuse'

Are you unusually short-tempered or irritable about things that previously did not bother you?

0	I am not more short-tempered or irritable than I used to be
1	I become more easily irritated, but it does not last very long
2	I become irritated very quickly about small things or things that do not bother other people
3	I react with extreme anger or rage, which I find very difficult to control

Sensitivity to light

Are you sensitive to strong light?

0	I have no increased sensitivity to light
1	I sometimes experience problems with strong light such as sunlight reflected by snow, water or glass or strong lights at home, but I am able to cope with it, for example, by wearing sunglasses
2	I am so sensitive to light that I prefer to carry out my daily activities in dim light. I find it difficult to leave the house without sunglasses
3	My sensitivity to light is so strong that I am unable to leave the house without sunglasses. I keep the blinds (or equivalent) drawn at all times.

Sensitivity to noise

Are you sensitive to noise?

0	I do not suffer from increased sensitivity to noise
1	I sometimes have difficulty with loud noise (for example music, noise from the TV, or radio or sudden unexpected sounds), but I can deal with it easily by turning down the volume. My sensitivity does not disturb my everyday life
2	I have a marked over-sensitivity to noise. I have to avoid loud noise or reduce it (for example by means of ear plugs) in order to cope with everyday life.
3	My sensitivity to noise is so great that I find difficult to manage at home despite sound insulation.

Decreased sleep at night

Do you sleep badly at night?

0	I do not sleep less than before (this includes sleeping more)
1	I have slight problems falling asleep or my sleep is shorter, lighter or more restless than before.
2	I sleep at least 2 hours less than before and wake up frequently during the night without anything disturbing me
3	I sleep less than 2-3 hours per night

Increased sleep

Do you sleep longer and/or more deeply than before?

0	I do not sleep more than usual (this includes sleeping less)
1	I sleep longer or deeper, but less than 2 hours more than usual, including naps during the day
2	I sleep longer or deeper. At least 2 hours more than usual, including naps.
3	I sleep longer or deeper. At least 4 hours more than usual and in addition I need to take a nap during the day.

24-hour variations

Do you find that at certain times of the day or night the problems we asked about (for example, tiredness, lack of concentration) are better or worse? In the statements below, 'regularly' means at least 3-4 times of the week.

0	I have not noticed that my problems are regularly better or worse at certain times, or I do not have any specific problems.
1	There is a clear difference between certain times of the day. I can predict that I will feel better at certain times and worse at other times.
2	I feel unwell at all times of the day or night

If you experience 24-hour variations:

When do you feel at your best?	Morning	Afternoon	Evening	Night
When do you feel at your worst?	Morning	Afternoon	Evening	Night

Appendix Six – Dysexecutive Questionnaire

This questionnaire looks at some of the difficulties people sometimes experience. We would like you to read the following statements, and rate them on a 5-point scale according to your own experience.

0	Never
1	Occasionally
2	Sometimes
3	Fairly Often
4	Very Often

1. I have problems understanding what other people mean unless they keep things simple and straightforward.
2. I act without thinking, doing the first thing that comes to mind
3. I sometimes talk about events or details that never actually happened, but I believe did happen
4. I have difficulty thinking ahead or planning for the future
5. I sometimes get over-excited about things and can be a bit 'over the top' at these times
6. I get events mixed up with each other, and get confused about the correct order of events
7. I have difficulty realising the extent of my problems and am unrealistic about the future
8. I am lethargic, or unenthusiastic about things
9. I do or say embarrassing things when in the company of others
10. I really want to do something one minute, but couldn't care less about it the next
11. I have difficulty showing emotion
12. I lose my temper at the slightest thing
13. I am unconcerned about how I should behave in certain situations
14. I find it hard to stop repeating saying or doing things once I've started
15. I tend to be very restless, and 'can't sit still' for any length of time
16. I find it difficult to stop myself from doing something even if I know I shouldn't
17. I will say one thing, but will do something different
18. I find it difficult to keep my mind on something, and am easily distracted
19. I have trouble making decisions, or deciding what I want to do
20. I am unaware of, or unconcerned about, how others feel about my behaviour

Appendix Seven – Cognitive Failures Questionnaire

The following questions are about minor mistakes which everyone makes from time to time, but some of which happen more often than others. We want to know how often these things have happened to you in the past 6 months. Please circle the appropriate number.

4	Very Often
3	Quite Often
2	Occasionally
1	Very rarely
0	Never

1. Do you read something and find you haven't been thinking about it and must read it again?
2. Do you find you forget why you went from one part of the house to the other?
3. Do you fail to notice signposts on the road?
4. Do you find you confuse right and left when giving directions?
5. Do you bump into people?
6. Do you find you forget whether you've turned off a light or a fire or locked a door?
7. Do you fail to listen to people's names when you are meeting them?
8. Do you say something and realise afterwards that it might be taken as insulting?
9. Do you fail to hear people speaking to you when you are doing something else?
10. Do you lose your temper and regret it?
11. Do you leave important letters unanswered for days?
12. Do you find you forget which way to turn on a road you know well but rarely use?
13. Do you fail to see what you want in a supermarket (although it is there)?
14. Do you find yourself suddenly wondering whether you've used a word correctly?

Appendix Eight – Rivermead Post-Concussive Symptoms Scale

Below is a list of difficulties that individuals may experience. I would like you to think about whether you have experience any of these recently. If you have, I would like you to compare how your experience compare to 12 months ago/before your injury*. Please select the closest to your answer.

0	Not experienced at all
1	No more of a problem
2	A mild problem
3	A moderate problem
4	A severe problem

1. Headaches
2. Feelings of dizziness
3. Nausea and/or vomiting
4. Noise sensitivity, easily upset by loud noise
5. Sleep disturbance
6. Fatigue, tiring more easily
7. Being more irritable, easily angered
8. Feeling depressed or tearful
9. Feeling frustrated or impatient
10. Forgetfulness, poor memory
11. Poor concentration
12. Taking longer to think
13. Blurred vision
14. Light sensitivity, easily upset by bright light
15. Double vision
16. Restlessness

Appendix Nine – Digital pen usability survey

During the tasks where you were asked to copy an image then later recall that image you were asked to use a digital pen. The following questions will ask you about how you found using this pen. Please indicate your response by circling which number is most applicable.

1) Did you notice you were not given a standard pen to write with?

Yes/No

2) Ease of use: how did you find the digital pen to draw with compared to a normal pen?

Much improved	Somewhat improved	No difference	Somewhat harder	Much harder
---------------	-------------------	---------------	-----------------	-------------

3) Comfort of pen: how did you find holding the digital pen compared to a normal pen?

Much improved	Somewhat improved	No difference	Somewhat harder	Much harder
---------------	-------------------	---------------	-----------------	-------------

4) Use of paper: how did you find drawing on the specially printed paper compared to drawing on plain paper?

Much improved	Somewhat improved	No difference	Somewhat harder	Much harder
---------------	-------------------	---------------	-----------------	-------------

5) Perception of performance: how did you feel you completed the task with the digital pen compared to if you completed it with a normal pen?

Much improved	Somewhat improved	No difference	Somewhat harder	Much harder
---------------	-------------------	---------------	-----------------	-------------

References

- Akshoomoff, N. A., & Stiles, J. (1995). Developmental trends in visuospatial analysis and planning: I. Copying a complex figure. *Neuropsychology*, 9(3), 364.
- Akshoomoff, N. A., Feroletto, C. C., Doyle, R. E., & Stiles, J. (2002). The impact of early unilateral brain injury on perceptual organization and visual memory. *Neuropsychologia*, 40(5), 539-561.
- Alegret, M., Espinosa, A., Ortega, G., Pérez-Cordón, A., Sanabria, Á., Hernández, I., ... & Valero, S. (2021). From face-to-face to home-to-home: validity of a teleneuropsychological battery. *Journal of Alzheimer's Disease*, 81(4), 1541-1553.
- Alsalaheen, B., Johns, K., Bean, R. et al. (2019). Women and men use different strategies to stabilize the head in response to impulsive loads: Implications for concussion injury risk. *Journal of Orthopaedic & Sports Physical Therapy*, 49(11), 779-786.
- Anderson, L. B., Jaroh, R., Smith, H., Strong, C. A. H., & Donders, J. (2017). Criterion validity of the D-KEFS color-word and verbal fluency switching paradigms following traumatic brain injury. *Journal of clinical and experimental neuropsychology*, 39(9), 890-899.
- Anderson, P., Anderson, V., & Garth, J. (2001). Assessment and development of organizational ability: The Rey complex figure organizational strategy score (RCF-OSS). *The Clinical Neuropsychologist*, 15(1), 81-94.
- Andronikof-Sanglade, A., Fjellestad-Paulsen, A., Ricard-Malivoir, S., & Evain-Brion, D. (1997). Specific abnormalities in a visual motor psychological test in short children with abnormal growth hormone secretion. *Acta Paediatrica*, 86(2), 154-159.
- Arciniegas, D. B., & Wortzel, H. S. (2014). Emotional and behavioral dyscontrol after traumatic brain injury. *Psychiatric Clinics*, 37(1), 31-53.
- Ariza, M., Pueyo, R., Junqué, C., Mataró, M., Poca, M. A., Mena, M. P., & Sahuquillo, J. (2006). Differences in visual vs. verbal memory impairments as a result of focal temporal lobe damage in patients with traumatic brain injury. *Brain Injury*, 20(10), 1053-1059.
- Ashina, H., Al-Khazali, H. M., Iljazi, A., Ashina, S., Amin, F. M., Lipton, R. B., & Schytz, H. W. (2021). Psychiatric and cognitive comorbidities of persistent post-traumatic headache attributed to mild traumatic brain injury. *The Journal of Headache and Pain*, 22(1), 1-10.
- Balalla, S., Krägeloh, C., Medvedev, O., & Siegert, R. (2020). Is the Rivermead post-concussion symptoms questionnaire a reliable and valid measure to assess long-term symptoms in traumatic brain injury and orthopedic injury patients? A novel investigation using Rasch analysis. *Neurotrauma Reports*, 1(1), 63-72.
- Banja, J., & Johnston, M. V. (1994). Outcomes evaluation in TBI rehabilitation. Part III: Ethical perspectives and social policy. *Archives of physical medicine and rehabilitation*, 75(12 Spec No), SC19-26.

- Bannon, S. M., Fernandes, R. F., Gates, M. V., O'Leary, K. D., & Brandler, E. S. (2022). Short-term changes in romantic relationship satisfaction after mild traumatic brain injury (mTBI): A pilot investigation in an emergency department sample. *Couple and Family Psychology: Research and Practice*. Advance online publication. <https://doi.org/10.1037/cfp0000217>
- Bar-Haim Erez, A., Rothschild, E., Katz, N., Tuchner, M., & Hartman-Maeir, A. (2009). Executive functioning, awareness, and participation in daily life after mild traumatic brain injury: A preliminary study. *The American Journal of Occupational Therapy*, 63(5), 634-640.
- Barker-Collo, S., Jones, K., Theadom, A., Starkey, N., Dowell, A., McPherson, K., ... & BIONIC Research Group. (2015). Neuropsychological outcome and its correlates in the first year after adult mild traumatic brain injury: A population-based New Zealand study. *Brain injury*, 29(13-14), 1604-1616.
- Barker-Collo, S., Theadom, A., Starkey, N. J., Kahan, M., Jones, K., & Feigin, V. (2019). Long-term factor structure of the Rivermead Post-Concussion Symptom Questionnaire in mild traumatic brain injury and normative sample. *Brain injury*, 33(5), 618-622.
- Barker-Collo, S., Theadom, A., Starkey, N., Kahan, M., Jones, K., & Feigin, V. (2018). Factor structure of the Rivermead Post-Concussion Symptoms Questionnaire over the first year following mild traumatic brain injury. *Brain injury*, 32(4), 453-458.
- Barr, W. B. (2014). Mild traumatic brain injury. In *Handbook on the neuropsychology of traumatic brain injury* (pp. 347-369). Springer, New York, NY.
- Basso, M. R., Bornstein, R. A., & Lang, J. M. (1999). Practice effects on commonly used measures of executive function across twelve months. *The Clinical Neuropsychologist*, 13(3), 283-292.
- Beaumont, A., Marmarou, A., Czigner, A., Yamamoto, M., Demetriadou, K., Shirotni, T., ... & Dunbar, J. (1999). The impact-acceleration model of head injury: injury severity predicts motor and cognitive performance after trauma. *Neurological research*, 21(8), 742-754.
- Beaumont, J. G., Kenealy, P., & Rogers, M. (1999). *The Blackwell Dictionary of Neuropsychology*. Wiley.
- Beck, A. T., Steer, R. A., & Brown, G. (1996). Beck depression inventory-II. *Psychological assessment*.
- Becker, J. T., Boller, F., Saxton, J., & McGonigle-Gibson, K. L. (1987). Normal rates of forgetting of verbal and non-verbal material in Alzheimer's disease. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*.
- Belanger, H. G., Curtiss, G., Demery, J. A., Lebowitz, B. K., & Vanderploeg, R. D. (2005). Factors moderating neuropsychological outcomes following mild traumatic brain injury: A meta-analysis. *Journal of the International Neuropsychological Society*, 11(3), 215-227.
- Benjamin, T. B., Nahmias, L., Gibson, D., Chhabria, S., Beaulieu, A., & Taylor, G. (2022). A-169 Utilizing the Rivermead Post Concussion Questionnaire (RPQ) during Post-Covid-19 Patient Follow-up. *Archives of Clinical Neuropsychology*, 37(6), 1324-1324.

- Bennett-Levy, J. (1984). Determinants of performance on the Rey-Osterrieth Complex Figure Test: An analysis, and a new technique for single-case assessment. *British Journal of Clinical Psychology*, 23(2), 109-119.
- Benton, A. L. (1967). Problems of test construction in the field of aphasia. *Cortex: A Journal Devoted to the Study of the Nervous System and Behavior*.
- Berry, D. T., & Carpenter, G. S. (1992). Effect of four different delay periods on recall of the Rey-Osterrieth Complex Figure by older persons. *The Clinical Neuropsychologist*, 6(1), 80-84.
- Berry, D. T., Allen, R. S., & Schmitt, F. A. (1991). Rey-Osterrieth Complex Figure: Psychometric characteristics in a geriatric sample. *The Clinical Neuropsychologist*, 5(2), 143-153.
- Bigler, E. D., Farrer, T. J., Pertab, J. L., James, K., Petrie, J. A., & Hedges, D. W. (2013). Reaffirmed limitations of meta-analytic methods in the study of mild traumatic brain injury: a response to Rohling et al. *The Clinical Neuropsychologist*, 27(2), 176-214.
- Binder, L. M. (1982). Constructional strategies on complex figure drawings after unilateral brain damage. *Journal of Clinical and Experimental Neuropsychology*, 4(1), 51-58.
- Binder, L. M., Rohling, M. L., & Larrabee, G. J. (1997). A review of mild head trauma. Part I: Meta-analytic review of neuropsychological studies. *Journal of clinical and experimental neuropsychology*, 19(3), 421-431.
- Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, 327(8476), 307-310.
- Bledsoe, B. E., Casey, M. J., Feldman, J., Johnson, L., Diel, S., Forred, W., & Gorman, C. (2015). Glasgow Coma Scale scoring is often inaccurate. *Prehospital and disaster medicine*, 30(1), 46-53.
- Bloom, B., Thomas, S., Ahrensberg, J. M., Weaver, R., Fowler, A., Bestwick, J., ... & Pearse, R. (2018). A systematic review and meta-analysis of return to work after mild Traumatic brain injury. *Brain injury*, 32(13-14), 1623-1636.
- Boone, K. B., Lesser, I. M., Hill-Gutierrez, E., Berman, N. G., & D'elia, L. F. (1993). Rey-Osterrieth Complex Figure performance in healthy, older adults: Relationship to age, education, sex, and IQ. *The Clinical Neuropsychologist*, 7(1), 22-28.
- Booth, R. (2006). *Local-global processing and cognitive style in autism spectrum disorders and typical development* [PhD thesis]. King's College London.
- Brandt, J., & Benedict, R. H. (2001). *Hopkins verbal learning test--revised: professional manual*. Psychological Assessment Resources.
- Brazinova, A., Rehorcikova, V., Taylor, M. S., Buckova, V., Majdan, M., Psota, M., ... & Synnot, A. (2021). Epidemiology of traumatic brain injury in Europe: a living systematic review. *Journal of neurotrauma*, 38(10), 1411-1440.
- 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.

- Bretzin, A. C., Covassin, T., Wiebe, D. J., & Stewart, W. (2021). Association of sex with adolescent soccer concussion incidence and characteristics. *JAMA network open*, 4(4), e218191-e218191.
- Bridger, R. S., Johnsen, S. Å. K., & Brasher, K. (2013). Psychometric properties of the cognitive failures questionnaire. *Ergonomics*, 56(10), 1515-1524.
- Bridges, A. J., & Holler, K. A. (2007). How many is enough? Determining optimal sample sizes for normative studies in pediatric neuropsychology. *Child Neuropsychology*, 13(6), 528-538.
- British Psychological Society Division of Neuropsychology (2020). Guidelines to colleagues on the use of Tele-neuropsychology. <https://www.bps.org.uk/sites/www.bps.org.uk/files/Member%20Networks/Divisions/DoN/DON%20guidelines%20on%20the%20use%20of%20tele-neuropsychology%20%28April%202020%29.pdf>
- Broadbent, D. E., Cooper, P. F., FitzGerald, P., & Parkes, K. R. (1982). The cognitive failures questionnaire (CFQ) and its correlates. *British journal of clinical psychology*, 21(1), 1-16.
- Brooks, B. L., Holdnack, J. A., & Iverson, G. L. (2011). Advanced clinical interpretation of the WAIS-IV and WMS-IV: Prevalence of low scores varies by level of intelligence and years of education. *Assessment*, 18(2), 156-167.
- Brooks, B. L., Iverson, G. L., & Holdnack, J. A. (2013). Understanding and using multivariate base rates with the WAIS-IV/WMS-IV. In *WAIS-IV, WMS-IV, and ACS* (pp. 75-102). Academic Press.
- Broshek, D. K., De Marco, A. P., & Freeman, J. R. (2015). A review of post-concussion syndrome and psychological factors associated with concussion. *Brain injury*, 29(2), 228-237.
- Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofacial Science*, 12(1).
- Burgess, P. W. (1996). The Dysexecutive Questionnaire (DEX). En BA Wilson, N. Alderman, PW Burgess, H. Emslie, & JJ Evans. *Behavioral Assessment of the Dysexecutive Syndrome*.
- Caffarra, P., Vezzadini, G., Dieci, F., Zonato, F., & Venneri, A. (2002). Rey-Osterrieth complex figure: normative values in an Italian population sample. *Neurological sciences*, 22(6), 443-447.
- Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist*, 26(4), 543-570.
- Canham, R. O., Smith, S. L., & Tyrrell, A. M. (2000, September). Automated scoring of a neuropsychological test: the rey osterrieth complex figure. In *Proceedings of the 26th Euromicro Conference. EUROMICRO 2000. Informatics: Inventing the Future* (Vol. 2, pp. 406-413). IEEE.

Canham, R. O., Smith, S. L., & Tyrrell, A. M. (2005). Location of structural sections from within a highly distorted complex line drawing. *IEEE Proceedings-Vision, Image and Signal Processing*, 152(6), 741-749.

CANTAB® [Cognitive assessment software]. Cambridge Cognition (2019)

Caplain, S., Blanche, S., Marque, S., Montreuil, M., & Aghakhani, N. (2017). Early detection of poor outcome after mild traumatic brain injury: predictive factors using a multidimensional approach a pilot study. *Frontiers in neurology*, 8, 666.

Carr, E. K., & Lincoln, N. B. (1988). Inter-rater reliability of the Rey figure copying test. *British Journal of Clinical Psychology*, 27(3), 267-268.

Ceban, F., Ling, S., Lui, L. M., Lee, Y., Gill, H., Teopiz, K. M., ... & McIntyre, R. S. (2022). Fatigue and cognitive impairment in Post-COVID-19 Syndrome: A systematic review and meta-analysis. *Brain, behavior, and immunity*, 101, 93-135.

Cermak, L. (2014). *Human Memory and Amnesia (PLE: Memory)*. Psychology Press.

Chancellor, S. E., Franz, E. S., Minaeva, O. V., & Goldstein, L. E. (2019, July). Pathophysiology of concussion. In *Seminars in Pediatric Neurology* (Vol. 30, pp. 14-25).

Chapman, J. E., Gardner, B., Ponsford, J., Cadilhac, D. A., & Stolwyk, R. J. (2021). Comparing performance across in-person and videoconference-based administrations of common neuropsychological measures in community-based survivors of stroke. *Journal of the International Neuropsychological Society*, 27(7), 697-710.

Charter, R. A., Walden, D. K., & Padilla, S. P. (2000). Too many simple clerical scoring errors: The Rey Figure as an example. *Journal of Clinical Psychology*, 56(4), 571-574.

Cherrier, M. M., Mendez, M. F., Dave, M., & Perryman, K. M. (1999). Performance on the Rey-Osterrieth Complex Figure test in Alzheimer disease and vascular dementia. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*.

Chiulli, S. J., Haaland, K. Y., Larue, A., & Garry, P. J. (1995). Impact of age on drawing the Rey-Osterrieth Figure. *The Clinical Neuropsychologist*, 9(3), 219-224.

Cicerone, K. D., & Azulay, J. (2002). Diagnostic utility of attention measures in postconcussion syndrome. *The Clinical Neuropsychologist*, 16(3), 280-289.

Claude Blais, M., & Boisvert, J. M. (2005). Psychological and marital adjustment in couples following a traumatic brain injury (TBI): A critical review. *Brain Injury*, 19(14), 1223-1235.

Coetzer, R., & Balchin, R. (2014). *Working with brain injury: a primer for PSYCHOLOGISTS working in under-resourced settings*. Psychology Press.
Cohen, J. (1988). Set correlation and contingency tables. *Applied psychological measurement*, 12(4), 425-434.

Collie, A., Maruff, P., Makdissi, M., McStephen, M., Darby, D. G., & McCrory, P. (2004). Statistical procedures for determining the extent of cognitive change following concussion. *British Journal of Sports Medicine*, 38(3), 273-278.

- Corwin, J., & Bylsma, F. W. (1993). Psychological examination of traumatic encephalopathy. *The Clinical Neuropsychologist*, 7(1), 3-21.
- Crowe, L., Collie, A., Hearps, S., Dooley, J., Clausen, H., Maddocks, D., ... & Anderson, V. (2016). Cognitive and physical symptoms of concussive injury in children: a detailed longitudinal recovery study. *British journal of sports medicine*, 50(5), 311-316.
- Cullum, C. M., Hynan, L. S., Grosch, M., Parikh, M., & Weiner, M. F. (2014). Teleneuropsychology: Evidence for video teleconference-based neuropsychological assessment. *Journal of the International Neuropsychological Society*, 20(10), 1028-1033.
- Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: clues about the brain from the skull of a famous patient. *Science*, 264(5162), 1102-1105.
- De Guise, E., Bélanger, S., Tinawi, S., Anderson, K., LeBlanc, J., Lamoureux, J., ... & Feys, M. (2016). Usefulness of the rivermead postconcussion symptoms questionnaire and the trail-making test for outcome prediction in patients with mild traumatic brain injury. *Applied Neuropsychology: Adult*, 23(3), 213-222.
- Dean, P. J., O'Neill, D., & Sterr, A. (2012). Post-concussion syndrome: prevalence after mild traumatic brain injury in comparison with a sample without head injury. *Brain injury*, 26(1), 14-26.
- Deckersbach, T., Savage, C. R., Henin, A., Mataix-Cols, D., Otto, M. W., Wilhelm, S., ... & Jenike, M. A. (2000). Reliability and validity of a scoring system for measuring organizational approach in the Complex Figure Test. *Journal of Clinical and Experimental Neuropsychology*, 22(5), 640-648.
- DeLellis, S. M., Kane, S., & Katz, K. (2009). The neurometabolic cascade and implications of mTBI: mitigating risk to the SOF community. *Journal of special operations medicine: a peer reviewed journal for SOF medical professionals*, 9(4), 36-42.
- Denman, S. B. (1984). Manual for the Denman Memory Battery. *Charleston, SC: SB Derman*.
- Donnell, A. J., Belanger, H. G., & Vanderploeg, R. D. (2011). Implications of psychometric measurement for neuropsychological interpretation. *The Clinical Neuropsychologist*, 25(7), 1097-1118.
- Draper, K., & Ponsford, J. (2008). Cognitive functioning ten years following traumatic brain injury and rehabilitation. *Neuropsychology*, 22(5), 618.
- Duff, M. C., Hengst, J. A., Tranel, D., & Cohen, N. J. (2008). Collaborative discourse facilitates efficient communication and new learning in amnesia. *Brain and language*, 106(1), 41-54.
- Duff, M. C., Hengst, J., Tranel, D., & Cohen, N. J. (2006). Development of shared information in communication despite hippocampal amnesia. *Nature neuroscience*, 9(1), 140-146.
- Duley, J. F., Wilkins, J. W., Hamby, S. L., Hopkins, D. G., Burwell, R. D., & Barry, N. S. (1993). Explicit scoring criteria for the Rey-Osterrieth and Taylor complex figures. *The Clinical Neuropsychologist*, 7(1), 29-38.

- Durand, E., Chevignard, M., Ruet, A., Dereix, A., Jourdan, C., & Pradat-Diehl, P. (2017). History of traumatic brain injury in prison populations: A systematic review. *Annals of physical and rehabilitation medicine*, 60(2), 95-101.
- Edginton, T., Loveday, C., Iddon, J. & Hutchinson, P. (2010). Assessing and responding to the cognitive and psychosocial sequelae of Traumatic Brain Injury in a multi-disciplinary neurotrauma out-patient clinic. *Society of British Neurological Surgeons*. London
- Edginton, T., Loveday, C., Iddon, J.I., Borg, N., Seeley, H., & Hutchinson, P.J. (in prep.) Triangulating diagnostic, clinical and self-report measures to assess the cognitive and psychosocial sequelae of Traumatic Brain Injury in a multi-disciplinary neurotrauma out-patient clinic.
- Egeland, J., Løvstad, M., Norup, A., Nybo, T., Persson, B. A., Rivera, D. F., ... & Arango-Lasprilla, J. C. (2016). Following international trends while subject to past traditions: Neuropsychological test use in the Nordic countries. *The Clinical Neuropsychologist*, 30(sup1), 1479-1500.
- 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.
- Elbaum, J. (2019). Acquired brain injury and the family: Challenges and interventions. In *Acquired brain injury* (pp. 335-347). Springer, Cham.
- Ellis, M. J., Boles, S., Derksen, V., Dawyduk, B., Amadu, A., Stelmack, K., ... & Russell, K. (2019). Evaluation of a pilot paediatric concussion telemedicine programme for northern communities in Manitoba. *International journal of circumpolar health*, 78(1), 1573163.
- Evans, J. J. (2010). Basic concepts and principles of neuropsychological assessment. *Handbook of clinical neuropsychology*, 15-26.
- Eyres, S., Carey, A., Gilworth, G., Neumann, V., & Tennant, A. (2005). Construct validity and reliability of the Rivermead post-concussion symptoms questionnaire. *Clinical rehabilitation*, 19(8), 878-887.
- Fastenau, P. S., Bennett, J. M., & Denburg, N. L. (1996). Application of psychometric standards to scoring system evaluation: is "new" necessarily "improved"? *Journal of Clinical and Experimental Neuropsychology*, 18(3), 462-472.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41, 1149-1160.
- Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191
- Ferdinand Pennock, K., McKenzie, B., McClement Steacy, L., & Mainwaring, L. (2020). Under-reporting of sport-related concussions by adolescent athletes: a systematic review. *International Review of Sport and Exercise Psychology*, 1-27

- Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health services research, 48*(6pt2), 2134-2156.
- Fisk, J. D., Ritvo, P. G., Ross, L., Haase, D. A., Marrie, T. J., & Schlech, W. F. (1994). Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. *Clinical Infectious Diseases, 18*(Supplement_1), S79-S83.
- Folbrecht, J. R., Charter, R. A., Walden, D. K., & Dobbs, S. M. (1999). Psychometric properties of the Boston qualitative scoring system for the Rey-Osterrieth complex figure. *The Clinical Neuropsychologist, 13*(4), 442-449.
- Fox-Fuller, J. T., Ngo, J., Pluim, C. F., Kaplan, R. I., Kim, D. H., Anzai, J. A., ... & Quiroz, Y. T. (2021). Initial investigation of test-retest reliability of home-to-home teleneuropsychological assessment in healthy, English-speaking adults. *The Clinical Neuropsychologist, 1*-15.
- Fralick, M., Sy, E., Hassan, A., Burke, M. J., Mostofsky, E., & Karsies, T. (2019). Association of concussion with the risk of suicide: a systematic review and meta-analysis. *JAMA neurology, 76*(2), 144-151.
- Fraser, E. E., Downing, M. G., Biernacki, K., McKenzie, D. P., & Ponsford, J. L. (2019). Cognitive reserve and age predict cognitive recovery after mild to severe traumatic brain injury. *Journal of neurotrauma, 36*(19), 2753-2761.
- Frazier, T. W., Adams, N. L., Strauss, M. E., & Redline, S. (2001). Comparability of the Rey and Mack forms of the Complex Figure Test. *The Clinical Neuropsychologist, 15*(3), 337-344.
- Frenchman, K. A., Fox, A. M., & Maybery, M. T. (2005). Neuropsychological studies of mild traumatic brain injury: A meta-analytic review of research since 1995. *Journal of Clinical and Experimental Neuropsychology, 27*(3), 334-351.
- Gagnon, M., Awad, N., Mertens, V. B., & Messier, C. (2003). Comparing the Rey and Taylor Complex Figures: A test-retest study in young and older adults. *Journal of Clinical and Experimental Neuropsychology, 25*(6), 878-890.
- Gallagher, C., & Burke, T. (2007). Age, gender and IQ effects on the Rey-Osterrieth complex figure test. *British Journal of Clinical Psychology, 46*(1), 35-45.
- Gallagher, V., Kramer, N., Abbott, K. et al. (2018). The effects of sex differences and hormonal contraception on outcomes after collegiate sports-related concussion. *Journal of Neurotrauma, 35*(11), 1242-1247.
- Gallagher, V., Kramer, N., Abbott, K., Alexander, J., Breiter, H., Herrold, A., ... & Reilly, J. (2018). The effects of sex differences and hormonal contraception on outcomes after collegiate sports-related concussion. *Journal of neurotrauma, 35*(11), 1242-1247.
- Gardner, R. C., & Yaffe, K. (2015). Epidemiology of mild traumatic brain injury and neurodegenerative disease. *Molecular and Cellular Neuroscience, 66*, 75-80.
- Gasparini, M., Masciarelli, G., Vanacore, N., Ottaviani, D., Salati, E., Talarico, G., ... & Bruno, G. (2008). A descriptive study on constructional impairment in frontotemporal dementia and Alzheimer's disease. *European Journal of Neurology, 15*(6), 589-597.

- Geary, E. K., Kraus, M. F., Rubin, L. H., Pliskin, N. H., & Little, D. M. (2011). Verbal learning strategy following mild traumatic brain injury. *Journal of the International Neuropsychological Society*, 17(4), 709-719.
- Georgiou, R., Lamnisos, D., & Giannakou, K. (2021). Anticholinergic burden and cognitive performance in patients with schizophrenia: a systematic literature review. *Frontiers in Psychiatry*, 2029.
- Gnassounou, R., Defontaines, B., Denolle, S., Brun, S., Germain, R., Schwartz, D., ... & Maillet, D. (2022). Comparison of neuropsychological assessment by videoconference and face to face. *Journal of the International Neuropsychological Society*, 28(5), 483-493.
- Govender, P., Naidoo, D., & Naidoo, U. (2019). Participant incentives in rehabilitation research: a pilot study of researchers' perspectives. *African Health Sciences*, 19(3), 2778-2783.
- Graff, H. J., Christensen, U., Poulsen, I., & Egerod, I. (2018). Patient perspectives on navigating the field of traumatic brain injury rehabilitation: a qualitative thematic analysis. *Disability and rehabilitation*, 40(8), 926-934.
- Graff, H. J., Deleu, N. W., Christiansen, P., & Rytter, H. M. (2021). Facilitators of and barriers to return to work after mild traumatic brain injury: A thematic analysis. *Neuropsychological rehabilitation*, 31(9), 1349-1373.
- Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate behavioral research*, 26(3), 499-510.
- Guidelines on psychological assessment undertaken remotely. (2020, May 12). In [www.BPS.org.uk](https://www.bps.org.uk). British Psychological Society. <https://www.bps.org.uk/node/2311>
- Guyll, M., Spoth, R., & Redmond, C. (2003). The effects of incentives and research requirements on participation rates for a community-based preventive intervention research study. *Journal of Primary Prevention*, 24(1), 25-41.
- Haenlein, M., & Kaplan, A. M. (2004). A beginner's guide to partial least squares analysis. *Understanding statistics*, 3(4), 283-297.
- Hamby, S. L., Wilkins, J. W., & Barry, N. S. (1993). Organizational quality on the Rey-Osterrieth and Taylor Complex Figure Tests: A new scoring system. *Psychological assessment*, 5(1), 27.
- Hammers, D. B., Stolwyk, R., Harder, L., & Cullum, C. M. (2020). A survey of international clinical teleneuropsychology service provision prior to and in the context of COVID-19. *The Clinical Neuropsychologist*, 34(7-8), 1267-1283.
- Happé, F. (1999). Autism: cognitive deficit or cognitive style?. *Trends in cognitive sciences*, 3(6), 216-222.
- Harder, L., Hernandez, A., Hague, C., Neumann, J., McCreary, M., Cullum, C. M., & Greenberg, B. (2020). Home-based pediatric teleneuropsychology: A validation study. *Archives of Clinical Neuropsychology*, 35(8), 1266-1275.

- Harper, D. (2017). Clinical psychology. In C. Willig, & W. Rogers *The SAGE Handbook of qualitative research in psychology* (pp. 473-495). SAGE Publications Ltd, <https://dx.doi.org/10.4135/9781526405555.n28>
- Hartman, M., & Potter, G. (1998). Sources of age differences on the Rey-Osterrieth complex figure test. *The Clinical Neuropsychologist*, 12(4), 513-524.
- Hawryluk, G. W., & Manley, G. T. (2015). Classification of traumatic brain injury: past, present, and future. *Handbook of clinical neurology*, 127, 15-21.
- Hede, K. (2008). Chemobrain Is Real but May Need New Name. *Journal of the National Cancer Institute*. 100 (3), 162–169, <https://doi.org/10.1093/jnci/djn007>
- Heinrichs, R. W., & Bury, A. (1991). Copying strategies and memory on the Complex Figure Test in psychiatric patients. *Psychological Reports*, 69(1), 223-226.
- Herrmann, N., Rapoport, M. J., Rajaram, R. D., Chan, F., Kiss, A., Ma, A. K., ... & Lanctôt, K. L. (2009). Factor analysis of the Rivermead Post-Concussion Symptoms Questionnaire in mild-to-moderate traumatic brain injury patients. *The Journal of neuropsychiatry and clinical neurosciences*, 21(2), 181-188.
- Holmes, A. D., Copland, D. A., Silburn, P. A., & Chenery, H. J. (2011). Acute nicotine enhances strategy-based semantic processing in Parkinson's disease. *International Journal of Neuropsychopharmacology*, 14(7), 877-885.
- Hsieh, G., & Kocielnik, R. (2016, February). You get who you pay for: The impact of incentives on participation bias. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing* (pp. 823-835).
- Hubley, A. M. (1999). The modified Taylor complex figure: A comparable measure to the Rey-Osterrieth figure. *Archives of Clinical Neuropsychology*, 8(14), 734.
- Hubley, A. M. (2010). Using the Rey–Osterrieth and Modified Taylor Complex Figures with older adults: A preliminary examination of accuracy score comparability. *Archives of Clinical Neuropsychology*, 25(3), 197-203.
- Hubley, A. M., & Jassal, S. (2006). Comparability of the Rey-Osterrieth and Modified Taylor Complex Figures using total scores, completion times, and construct validation. *Journal of Clinical and Experimental Neuropsychology*, 28(8), 1482-1497.
- Hubley, A. M., & Tremblay, D. (2002). Comparability of total score performance on the Rey–Osterrieth complex figure and a modified Taylor complex figure. *Journal of Clinical and Experimental Neuropsychology*, 24(3), 370-382.
- Hyun, G. J., Park, J. W., Kim, J. H., Min, K. J., Lee, Y. S., Kim, S. M., & Han, D. H. (2018). Visuospatial working memory assessment using a digital tablet in adolescents with attention deficit hyperactivity disorder. *Computer Methods and Programs in Biomedicine*, 157, 137-143.
- Iowa Trail Making. (n.d.). Iowa Trail Making . Retrieved September 12, 2022, from <http://apps.usd.edu/coglab/schieber/psyc423/pdf/IowaTrailMaking.pdf>
- Iverson, G. L. (2010). Mild traumatic brain injury meta-analyses can obscure individual differences. *Brain injury*, 24(10), 1246-1255.

- Iverson, G. L., Gardner, A. J., Terry, D. P., Ponsford, J. L., Sills, A. K., Broshek, D. K., & Solomon, G. S. (2017). Predictors of clinical recovery from concussion: a systematic review. *British journal of sports medicine*, 51(12), 941-948.
- Janowsky, J. S., & Thomas-Thrapp, L. J. (1993). Complex figure recall in the elderly: a deficit in memory or constructional strategy?. *Journal of Clinical and Experimental Neuropsychology*, 15(2), 159-169.
- Jenkins, D. G., & Quintana-Ascencio, P. F. (2020). A solution to minimum sample size for regressions. *PloS one*, 15(2), e0229345.
- Johansson, B., & Ronnback, L. (2014). Evaluation of the mental fatigue scale and its relation to cognitive and emotional functioning after traumatic brain injury or stroke. *Int J Phys Med Rehabil*, 2(01).
- Johansson, B., Berglund, P., & Rönnbäck, L. (2009). Mental fatigue and impaired information processing after mild and moderate traumatic brain injury. *Brain injury*, 23(13-14), 1027-1040.
- Jolliffe, T., & Baron Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the Embedded Figures Test?. *Journal of Child Psychology and Psychiatry*, 38 (5), 527-534.
- Jurica, P., Valenzi, S., Struzik, Z. R., & Cichocki, A. (2015). Methods for Transition Toward Computer Assisted Cognitive Examination. *Methods of information in medicine*, 54(03), 256-261.
- Kaber, D., Tupler, L. A., Clamann, M., Gil, G. H., Zhu, B., Swangnetr, M., ... & Lee, Y. S. (2014). Evaluation of an augmented virtual reality and haptic control interface for psychomotor training. *Assistive Technology*, 26(1), 51-60.
- Kander, I. C. S. (2020). *Care Pathways in Traumatic Brain Injury Rehabilitation: Searching for the Missing Pieces* (Doctoral dissertation, University of Warwick).
- Kang, H. (2013). The prevention and handling of the missing data. *Korean journal of anesthesiology*, 64(5), 402-406.
- Karr, J. E., Areshenkoff, C. N., & Garcia-Barrera, M. A. (2014). The neuropsychological outcomes of concussion: A systematic review of meta-analyses on the cognitive sequelae of mild traumatic brain injury. *Neuropsychology*, 28(3), 321.
- Keshavan, M. S., Channabasavanna, S. M., & Reddy, G. N. (1981). Post-traumatic psychiatric disturbances: patterns and predictors of outcome. *The British Journal of Psychiatry*, 138(2), 157-160.
- King, N. (2014). Permanent post-concussion symptoms after mild head injury: A systematic review of age and gender factors. *NeuroRehabilitation* (Reading, Mass.), 34(4), 741-748. <https://doi.org/10.3233/NRE-141072>
- King, N. S., & Kirwilliam, S. (2011). Permanent post-concussion symptoms after mild head injury. *Brain injury*, 25(5), 462-470.
- King, N. S., Crawford, S., Wenden, F. J., Moss, N. E. G., & Wade, D. T. (1995). The Rivermead Post Concussion Symptoms Questionnaire: a measure of symptoms

commonly experienced after head injury and its reliability. *Journal of neurology*, 242(9), 587-592.

Kirkwood, K. T., Peck, D. F., & Bennie, L. (2000). The consistency of neuropsychological assessments performed via telecommunication and face to face. *Journal of Telemedicine and Telecare*, 6(3), 147-151.

Klicpera, C. (1983). Poor planning as a characteristic of problem-solving behavior in dyslexic children. A study with the Rey-Osterrieth complex figure test.

Knight, J., Kaplan, E., & Ireland, L. (1994, February). Survey findings of the Rey-Osterrieth Complex Figure use among the INS membership. In *22nd Annual Meeting of the International Neuropsychological Society, Cincinnati, Ohio*.

Kraus, J. F., & Nourjah, P. (1988). The epidemiology of mild, uncomplicated brain injury. *The Journal of trauma*, 28(12), 1637-1643

Kuehn, S. M., & Snow, W. G. (1992). Are the Rey and Taylor figures equivalent?. *Archives of Clinical Neuropsychology*, 7(5), 445-448.

Lang, K., Roberts, M., Harrison, A., Lopez, C., Goddard, E., Khondoker, M., ... & Tchanturia, K. (2016). Central coherence in eating disorders: a synthesis of studies using the Rey Osterrieth complex figure test. *PloS one*, 11(11), e0165467.

Lannsjö, M., Borg, J., Björklund, G., Af Geijerstam, J. L., & Lundgren-Nilsson, A. (2011). Internal construct validity of the rivermead post-concussion symptoms questionnaire. *Journal of rehabilitation medicine*, 43(11), 997-1002.

Lannsjö, M., Geijerstam, J. L. A., Johansson, U., Bring, J., & Borg, J. (2009). Prevalence and structure of symptoms at 3 months after mild traumatic brain injury in a national cohort. *Brain Injury*, 23(3), 213-219.

LaValle, C. R., Carr, W. S., Egnoto, M. J., Misistia, A. C., Salib, J. E., Ramos, A. N., & Kamimori, G. H. (2019). Neurocognitive performance deficits related to immediate and acute blast overpressure exposure. *Frontiers in neurology*, 10, 949.

Lefevre-Dognin, C., Cogné, M., Perdrieau, V., Granger, A., Heslot, C., & Azouvi, P. (2021). Definition and epidemiology of mild traumatic brain injury. *Neurochirurgie*, 67(3), 218-221.

Levack, W. M., Kayes, N. M., & Fadyl, J. K. (2010). Experience of recovery and outcome following traumatic brain injury: a metasynthesis of qualitative research. *Disability and rehabilitation*, 32(12), 986-999.

Levin HS, High WM, Goethe KE, Sisson RA, Overall JE, Rhoades HM, Eisenberg HM, Kalisky Z, Gary HE. The neurobehavioural rating scale: assessment of the behavioural sequelae of head injury by the clinician. *J Neurol Neurosurg Psychiatry*. 1987 Feb;50(2):183-93. doi: 10.1136/jnnp.50.2.183. PMID: 3572433; PMCID: PMC1031490.

Levin, H. S., Grafman, J., & Eisenberg, H. M. (Eds.). (1987). *Neurobehavioral recovery from head injury*. Oxford University Press, USA.

- Levine, A. J., Miller, E. N., Becker, J. T., Selnes, O. A., & Cohen, B. A. (2004). Normative data for determining significance of test-retest differences on eight common neuropsychological instruments. *The Clinical Neuropsychologist*, 18(3), 373-384.
- Lezak, M. D., Howieson, D. B., Loring, D. W., & Fischer, J. S. (2004). *Neuropsychological assessment*. Oxford University Press, USA.
- Li, Y., Clamann, M., & Kaber, D. B. (2013). Validation of a haptic-based simulation to test complex figure reproduction capability. *IEEE Transactions on Human-Machine Systems*, 43(6), 547-557.
- Li, Y., Kaber, D. B., Tupler, L., & Lee, Y. S. (2011). Haptic-based virtual environment design and modeling of motor skill assessment for brain injury patients rehabilitation. *Computer-Aided Design and Applications*, 8(2), 149-162.
- Lieberman, J., Stewart, W., Seines, O., & Gordon, B. (1994). Rater agreement for the Rey-Osterrieth complex figure test. *Journal of clinical psychology*, 50(4), 615-624.
- Lopez, C., Tchanturia, K., Stahl, D., & Treasure, J. (2009). Weak central coherence in eating disorders: a step towards looking for an endophenotype of eating disorders. *Journal of clinical and experimental neuropsychology*, 31(1), 117-125
- Lopez, C., Tchanturia, K., Stahl, D., Booth, R., Holliday, J., & Treasure, J. (2008). An examination of the concept of central coherence in women with anorexia nervosa. *International Journal of Eating Disorders*, 41(2), 143-152.
- Loring, D. W., Martin, R. C., Meador, K. J., & Lee, G. P. (1990). Psychometric construction of the Rey-Osterrieth complex figure: Methodological considerations and interrater reliability. *Archives of Clinical Neuropsychology*, 5(1), 1-14.
- Loveday, C., Sahakian, B.J., Iddon, J., Seeley, H., Corteen, L., Hutchison, P. (March, 2006) *Clinical outcome in Traumatic Brain Injury: How are cognitive function and quality of life really affected by nature of injury and premorbid status?* Annual Conference of the British Neurosurgical Research Group
- Lumba-Brown, A., Yeates, K. O., Sarmiento, K., Breiding, M. J., Haegerich, T. M., Gioia, G. A., ... & Timmons, S. D. (2018). Centers for Disease Control and Prevention guideline on the diagnosis and management of mild traumatic brain injury among children. *JAMA pediatrics*, 172(11), e182853-e182853.
- Luzzi, S., Pesallaccia, M., Fabi, K., Muti, M., Viticchi, G., Provinciali, L., & Piccirilli, M. (2011). Non-verbal memory measured by Rey-Osterrieth Complex Figure B: normative data. *Neurological Sciences*, 32(6), 1081-1089.
- Malec, J. F., Brown, A. W., Leibson, C. L., Flaada, J. T., Mandrekar, J. N., Diehl, N. N., & Perkins, P. K. (2007). The mayo classification system for traumatic brain injury severity. *Journal of neurotrauma*, 24(9), 1417-1424.
- Mani, K., Cater, B., & Hudlikar, A. (2017). Cognition and return to work after mild/moderate traumatic brain injury: A systematic review. *Work*, 58(1), 51-62.
- Manjaly, Z. M., Harrison, N. A., Critchley, H. D., Do, C. T., Stefanics, G., Wenderoth, N., ... & Stephan, K. E. (2019). Pathophysiological and cognitive mechanisms of fatigue in multiple sclerosis. *Journal of Neurology, Neurosurgery & Psychiatry*, 90(6), 642-651.

- Mansfield, E., Stergiou-Kita, M., Cassidy, J. D., Bayley, M., Mantis, S., Kristman, V., ... & Colantonio, A. (2015). Return-to-work challenges following a work-related mild TBI: the injured worker perspective. *Brain injury*, 29(11), 1362-1369.
- Marra, D. E., Hamlet, K. M., Bauer, R. M., & Bowers, D. (2020). Validity of teleneuropsychology for older adults in response to COVID-19: A systematic and critical review. *The Clinical Neuropsychologist*, 34(7-8), 1411-1452.
- Martin, P., & Velay, J. L. (2012). Do computers improve the drawing of a geometrical figure for 10 year-old children?. *International journal of technology and design education*, 22(1), 13-23.
- Martini, A., Weis, L., Schifano, R., Pistonesi, F., Fiorenzato, E., Antonini, A., & Biundo, R. (2020). Differences in cognitive profiles between Lewy body and Parkinson's disease dementia. *Journal of Neural Transmission*, 127(3), 323-330.
- Maruta, J., Lumba-Brown, A., & Ghajar, J. (2018). Concussion subtype identification with the rivermead post-concussion symptoms questionnaire. *Frontiers in neurology*, 9, 1034.
- Mathuranath, P. S., Nestor, P. J., Berrios, G. E., Rakowicz, W., & Hodges, J. R. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology*, 55(11), 1613-1620.
- McCrea, M., Barr, W. B., Guskiewicz, K., Randolph, C., Marshall, S. W., Cantu, R., ... & Kelly, J. P. (2005). Standard regression-based methods for measuring recovery after sport-related concussion. *Journal of the International Neuropsychological Society*, 11(1), 58-69.
- McGuire, J. L., Ngwenya, L. B., & McCullumsmith, R. E. (2019). Neurotransmitter changes after traumatic brain injury: an update for new treatment strategies. *Molecular Psychiatry*, 24(7), 995-1012.
- McInnes, K., Friesen, C. L., MacKenzie, D. E., Westwood, D. A., & Boe, S. G. (2017). Mild Traumatic Brain Injury (mTBI) and chronic cognitive impairment: A scoping review. *PloS one*, 12(4), e0174847.
- McKinlay, A., Bishop, A. & McLellan, T. (2011). Public knowledge of 'concussion' and the different terminology used to communicate about mild traumatic brain injury (MTBI). *Brain Injury*, 25(7-8), 761-766.
- Meador, K. J., Loring, D. W., Allen, E. Y., Moore, E. E., Abney, O. L., & King, D. W. (1991). Comparative cognitive effects of carbamazepine and phenytoin in healthy adults. *Neurology*, 41(10), 1537-1537.
- Meador, K. J., Moore, E. E., Nichols, M. E., Abney, O. L., Taylor, H. S., Zamrini, E. Y., & Loring, D. W. (1993). The role of cholinergic systems in visuospatial processing and memory. *Journal of Clinical and Experimental Neuropsychology*, 15(5), 832-842.
- Medvedev, O. N., Theadom, A., Barker-Collo, S., & Feigin, V. (2018). Distinguishing between enduring and dynamic concussion symptoms: applying generalisability theory to the Rivermead Post Concussion Symptoms Questionnaire (RPQ). *PeerJ*, 6, e5676.

- Meier, T. B., Brummel, B. J., Singh, R., Nerio, C. J., Polanski, D. W., & Bellgowan, P. S. (2015). The underreporting of self-reported symptoms following sports-related concussion. *Journal of science and medicine in sport*, 18(5), 507-511.
- Memisevic, H., & Biscevic, I. (2018). Exploring the link between inhibitory control and cognitive flexibility in preschool children. *Cognition, Brain, Behavior*, 22(1), 1-11.
- Merz, Z.C., Van Patten, R. & Lace, J. (2017). Current public knowledge pertaining to traumatic brain injury: Influence of demographic factors, social trends, and sport concussion experience on the understanding of traumatic brain injury sequelae. *Archives of Clinical Neuropsychology*, 32(2), 155-167.
- Meyers, J. E., & Lange, D. (1994). Recognition subtest for the complex figure. *The Clinical Neuropsychologist*, 8(2), 153-166.
- Meyers, J. E., & Meyers, K. R. (1995). *Rey Complex Figure Test and recognition trial professional manual*. Psychological Assessment Resources.
- Meyers, J. E., & Volbrecht, M. (1999). Detection of malingerers using the Rey Complex Figure and recognition trial. *Applied Neuropsychology*, 6(4), 201-207.
- Mitra, S., Khatri, S. N., Maulik, M., Bult-Ito, A., & Schulte, M. (2020). Allosterism of nicotinic acetylcholine receptors: therapeutic potential for neuroinflammation underlying brain trauma and degenerative disorders. *International Journal of Molecular Sciences*, 21(14), 4918.
- Mitrushina, M., Satz, P., Chervinsky, A., & D'Elia, L. (1991). Performance of four age groups of normal elderly on the Rey Auditory-Verbal Learning Test. *Journal of clinical psychology*, 47(3), 351-357.
- Molitor, S. J., Eadeh, H. M., Bouchtein, E., Smith, Z. R., Green, C. D., & Langberg, J. M. (2019). The rey-osterrieth complex figure: a useful measure of organizational skills for adolescents with ADHD?. *Journal of Pediatric Neuropsychology*, 5(1), 20-30.
- Möller, M. C., Johansson, J., Matuseviciene, G., Pansell, T., & Deboussard, C. N. (2019). An observational study of trait and state fatigue, and their relation to cognitive fatigability and saccade performance. *Concussion*, 4(2), CNC62.
- Moore, H. C. (2014). An overview of chemotherapy-related cognitive dysfunction, or'chemobrain'. *Oncology*, 28(9), 797-797.
- Moser, R. S., Iverson, G. L., Echemendia, R. J., Lovell, M. R., Schatz, P., Webbe, F. M., ... & Planning Committee. (2007). Neuropsychological evaluation in the diagnosis and management of sports-related concussion. *Archives of Clinical Neuropsychology*, 22(8), 909-916.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & 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.
- National Institute for Health and Care Excellence (NICE). 2014. Head Injury: assessment and early management. <https://www.nice.org.uk/guidance/cg176/resources/head-injury-assessment-and-early-management-pdf-35109755595493>

- Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive psychology*, 9(3), 353-383.
- Navon, D. (2003). What does a compound letter tell the psychologist's mind?. *Acta psychologica*, 114(3), 273-309.
- Nelson, H. E., & Willison, J. (1991). *National adult reading test (NART)* (pp. 1-26). Windsor: Nfer-Nelson.
- Newman, M., & Dennis, R. (2021) The resurgence of attention to sports-related concussion, and what this means for our understanding of mild traumatic brain injury. *The Quarterly*. 120. 8-13
- NICE: Head injury: assessment and early management. Clinical guideline [CG176]Published: 22 January 2014 Last updated: 13 September 2019. Accessed 31/05/22
- Oddy, M., Moir, J. F., Fortescue, D., & Chadwick, S. (2012). The prevalence of traumatic brain injury in the homeless community in a UK city. *Brain injury*, 26(9), 1058-1064.
- Office of National Statistics. (2019). Internet access - households and individuals, Great Britain: 2019.
<https://www.ons.gov.uk/peoplepopulationandcommunity/householdcharacteristics/homeinternetandsocialmediausage/bulletins/internetaccesshouseholdsandindividuals/2019>
- Oldenburg, C., Lundin, A., Edman, G., Nygren-de Boussard, C., & Bartfai, A. (2016). Cognitive reserve and persistent post-concussion symptoms—A prospective mild traumatic brain injury (mTBI) cohort study. *Brain injury*, 30(2), 146-155.
- Osterrieth, P. A. (1944). Le test de copie d'une figure complexe; contribution a l'etude de la perception et de la memoire. *Archives de psychologie*.
- Pandor, A., Goodacre, S., Harnan, S. et al. (2011). Diagnostic management strategies for adults and children with minor head injury: A systematic review and an economic evaluation. *Health Technology Assessment*, 15(27), 1–202.
- Parks, A. C., Davis, J., Spresser, C. D., Stroescu, I., & Ecklund-Johnson, E. (2021). Validity of in-home Teleneuropsychological testing in the wake of COVID-19. *Archives of Clinical Neuropsychology*, 36(6), 887-896.
- Parlar, M. E., Spilka, M. J., Wong Gonzalez, D., Ballantyne, E. C., Dool, C., Gojmerac, C., ... & MacKillop, E. (2020). "You can't touch this": Delivery of inpatient neuropsychological assessment in the era of COVID-19 and beyond. *The Clinical Neuropsychologist*, 34(7-8), 1395-1410.
- Peirson, A. R., & Jansen, P. (1997). Comparability of the Rey-Osterrieth and Taylor forms of the complex figure test. *The Clinical Neuropsychologist*, 11(3), 244-248.
- Perez, D. F., Nie, J. X., Arden, C. I., Radhu, N., & Ritvo, P. (2013). Impact of participant incentives and direct and snowball sampling on survey response rate in an ethnically diverse community: Results from a pilot study of physical activity and the built environment. *Journal of Immigrant and Minority Health*, 15(1), 207-214.

Pertab, J. L., James, K. M., & Bigler, E. D. (2009). Limitations of mild traumatic brain injury meta-analyses. *Brain Injury*, 23(6), 498-508.

Peterson RK, Ludwig NN, Jashar DT. A case series illustrating the implementation of a novel tele-neuropsychology service model during COVID-19 for children with complex medical and neurodevelopmental conditions: A companion to Pritchard et al., 2020. *Clin Neuropsychol*. 2021 Jan;35(1):99-114. doi: 10.1080/13854046.2020.1799075. Epub 2020 Aug 5. PMID: 32757708.

Pineau, H., Marchand, A., & Guay, S. (2015). Specificity of cognitive and behavioral complaints in post-traumatic stress disorder and mild traumatic brain injury. *Behavioral Sciences*, 5(1), 43-58.

Ponsford, J., Draper, K., & Schönberger, M. (2008). Functional outcome 10 years after traumatic brain injury: its relationship with demographic, injury severity, and cognitive and emotional status. *Journal of the International Neuropsychological Society*, 14(2), 233-242.

Postal, K. S., Bilder, R. M., Lanca, M., Aase, D. M., Barisa, M., Holland, A. A., ... & Salinas, C. (2021). Inter organizational practice committee guidance/recommendation for models of care during the novel coronavirus pandemic. *Archives of Clinical Neuropsychology*, 36(1), 17-28.

Potter, S., Leigh, E., Wade, D., & Fleminger, S. (2006). The Rivermead post concussion symptoms questionnaire. *Journal of neurology*, 253(12), 1603-1614.

Power, R. K. (2016). The impact of mild traumatic brain injury on romantic relationships. Pownall, M., & Rogers, W. S. (2021). *A Feminist Companion to Social Psychology*. McGraw-Hill Education (UK).

Prange, A., & Sonntag, D. (2019, June). Modeling cognitive status through automatic scoring of a digital version of the clock drawing test. In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization* (pp. 70-77).

Prange, A., & Sonntag, D. (2019, June). Modelling cognitive status through automatic scoring of a digital version of the clock drawing test. In *Proceedings of the 27th ACM Conference on User Modelling, Adaptation and Personalization* (pp. 70-77).

Prange, A., Barz, M., Heimann-Steinert, A., & Sonntag, D. (2021, May). Explainable automatic evaluation of the trail making test for dementia screening. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-9).

Priest, A. (2020). Doing Mixed Methods Research. Combining Outcome Measures with Interviews. In: Bager-Charleson, S., McBeath, A. (eds) *Enjoying Research in Counselling and Psychotherapy*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-030-55127-8_12

Pulsifer, M. B., Grieco, J. A., Burstein, S. M., Parsons, M. W., Gardner, M. M., & Sherman, J. C. (2021). The development and implementation of teleneuropsychology in an academic lifespan neuropsychology center: Lessons learned from the COVID-19 pandemic. *Journal of Clinical and Experimental Neuropsychology*, 43(8), 774-785.

- Qin, X., Tupler, L.A., Ma, W., Kaber, D.B., Clamann, M., Hong, Y., & Lee, Y.S. (2015). Computer-based haptic control assessment of 3-D manipulation and drawing skill with analytical application of machining metrics. *Proceedings 19th Triennial Congress of the IEA*. Vol. 9, p.14.
- 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.
- Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology*, 31(3), 206-230.
- Rajabpour Azizi, Z., Akhavan Tafti, M., & Mohsenpour, M. (2021). Dyslexia and the Visual-Spatial Talents: A Critical Review of New Difference-Oriented Research. *J Child Ment Health*, 7(4), 197-214.
- Rakhmanov, O., & Dane, S. (2020). The relationships among gender, handedness, GPA, depression and visual memory in the ROCF test in university students. *J Res Med Dent Sci*, 8, 37-42.
- Rao, R. K., McConnell, D. D., & Litofsky, N. S. (2022). The impact of cigarette smoking and nicotine on traumatic brain injury: a review. *Brain injury*, 36(1), 1-20.
- Rapport, L. J., Charter, R. A., Dutra, R. L., Farchione, T. J., & Kingsley, J. J. (1997). Psychometric properties of the Rey-Osterrieth complex figure: Lezak-Osterrieth versus Denman scoring systems. *The Clinical Neuropsychologist*, 11(1), 46-53.
- Ravindra, V.M., & Hawryluk, G. W. (2020). 3 The classification of traumatic brain injury. *Thieme Connect*. 29-33
- Rawlings, S., Takechi, R., & Lavender, A. P. (2020). Effects of sub-concussion on neuropsychological performance and its potential mechanisms: a narrative review. *Brain Research Bulletin*, 165, 56-62.
- Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique.(Les problems.). *Archives de psychologie*
- Rietdijk, R., Power, E., Brunner, M., & Togher, L. (2017). Reliability of videoconferencing administration of a communication questionnaire to people with traumatic brain injury and their close others. *The Journal of Head Trauma Rehabilitation*, 32(6), E38-E44.
- Riordan, P., Lombardo, T., & Schulenberg, S. E. (2013). Evaluation of a computer-based administration of the rey complex figure test. *Applied Neuropsychology: Adult*, 20(3), 169-178.
- Rivera, D., Perrin, P. B., Morlett-Paredes, A., Galarza-del-Angel, J., Martínez, C., Garza, M. T., ... & Arango-Lasprilla, J. C. (2015). Rey-Osterrieth Complex Figure-copy and immediate recall: Normative data for the Latin American Spanish speaking adult population. *NeuroRehabilitation*, 37(4), 677-698.
- Rochette, A. D., Rahman-Filipiak, A., Spencer, R. J., Marshall, D., & Stelmokas, J. E. (2021). Teleneuropsychology practice survey during COVID-19 within the United States. *Applied Neuropsychology: Adult*, 1-11.

- Røe, C., Sveen, U., Alvsåker, K., & Bautz-Holter, E. (2009). Post-concussion symptoms after mild traumatic brain injury: influence of demographic factors and injury severity in a 1-year cohort study. *Disability and rehabilitation*, 31(15), 1235-1243.
- Roebuck-Spencer, T., Cernich, A. (2014). Epidemiology and Societal Impact of Traumatic Brain Injury. In: Sherer, M., Sander, A. (eds) *Handbook on the Neuropsychology of Traumatic Brain Injury*. Clinical Handbooks in Neuropsychology. Springer, New York, NY. https://doi.org/10.1007/978-1-4939-0784-7_1
- Rohling ML, Binder LM, Demakis GJ, Larrabee GJ, Ploetz DM, Langhinrichsen-Rohling J. (2011) A meta-analysis of neuropsychological outcome after mild traumatic brain injury: re-analyses and reconsiderations of Binder et al. (1997), Frencham et al. (2005), and Pertab et al. (2009). *The Clinical Neuropsychologist*. 25: 608-23. PMID 21512956 DOI: 10.1080/13854046.2011.565076
- Romeu-Mejia, R., Giza, C. C., & Goldman, J. T. (2019). Concussion pathophysiology and injury biomechanics. *Current reviews in musculoskeletal medicine*, 12(2), 105-116.
- Rosselli, M., & Ardila, A. (1991). Effects of age, education, and gender on the Rey-Osterrieth Complex Figure. *The clinical neuropsychologist*, 5(4), 370-376.
- Ruffolo, J. S., Javorsky, D. J., Tremont, G., Westervelt, H. J., & Stern, R. A. (2001). A comparison of administration procedures for the Rey-Osterrieth Complex Figure: Flowcharts versus pen switching. *Psychological assessment*, 13(3), 299
- Rycroft, N., Hutton, S. B., & Rusted, J. M. (2006). The antisaccade task as an index of sustained goal activation in working memory: modulation by nicotine. *Psychopharmacology*, 188(4), 521-529.
- Rycroft, N., Rusted, J. M., & Hutton, S. B. (2005). Acute effects of nicotine on visual search tasks in young adult smokers. *Psychopharmacology*, 181(1), 160-169.
- Sadeh, M., Ariel, R., & Inbar, D. (1996). Rey-osterrieth and Taylor complex figures: Equivalent measures of visual organization and visual memory in ADHD and normal children. *Child Neuropsychology*, 2(1), 63-71.
- Sahakian, B. J. (1988). Cholinergic drugs and human cognitive performance. In *Handbook of psychopharmacology* (pp. 393-424). Springer, Boston, MA.
- Sanada, S., Higa Diez, M., Yamane, D., Kado, Y., Nakano, K., & Ogino, T. (2014). Utilization of the Inkling™ for the Rey-Osterrieth Complex Figure test. *岡山大学大学院教育学研究科研究集録*, 156, 1-5.
- Sariaslan, A., Sharp, D. J., D'Onofrio, B. M., Larsson, H., & Fazel, S. (2016). Long-term outcomes associated with traumatic brain injury in childhood and adolescence: a nationwide Swedish cohort study of a wide range of medical and social outcomes. *PLoS medicine*, 13(8), e1002103.
- Savage, C. R., Baer, L., Keuthen, N. J., Brown, H. D., Rauch, S. L., & Jenike, M. A. (1999). Organizational strategies mediate nonverbal memory impairment in obsessive-compulsive disorder. *Biological psychiatry*, 45(7), 905-916.

- Sawatsky, M. L., Clyde, M., & Meek, F. (2015). Partial least squares regression in the social sciences. *The Quantitative Methods for Psychology*, 11(2), 52-62.
- Sawaya, H., McGonigle-Chalmers, M., & Kusel, I. (2021). Identifying set-switching difficulties in autism spectrum disorder using a rule following task. *International Journal of Developmental Disabilities*, 67(4), 245-255.
- Schoenberg, M. R., Lange, R. T., Brickell, T. A., & Saklofske, D. H. (2007). Estimating premorbid general cognitive functioning for the American WISC-IV: Demographic and current performance approaches. *Journal of Child Neurology*, 22, 379-388.
- Schönberger, M., Ponsford, J., Gould, K. R., & Johnston, L. (2011). The temporal relationship between depression, anxiety, and functional status after traumatic brain injury: a cross-lagged analysis. *Journal of the International Neuropsychological Society*, 17(5), 781-787.
- Schretlen, D. J., & Shapiro, A. M. (2003). A quantitative review of the effects of traumatic brain injury on cognitive functioning. *International review of psychiatry*, 15(4), 341-349.
- Shorr, J. S., Delis, D. C., & Massman, P. J. (1992). Memory for the Rey-Osterrieth Figure: Perceptual clustering, encoding, and storage. *Neuropsychology*, 6(1), 43.
- Sicard, V., Lortie, J. C., Moore, R. D., & Ellemberg, D. (2020). Cognitive testing and exercise to assess the readiness to return to play after a concussion. *Translational Journal of the American College of Sports Medicine*, 5(11), 1-9.
- Smith, S. L., & Hiller, D. L. (1996, April). Image analysis of neuropsychological test responses. In *Medical Imaging 1996: Image Processing* (Vol. 2710, pp. 904-915). SPIE.
- Smith, S. R., Chang, J., Schnoebelen, K. J., Edwards, J. W., Servesko, A. M., & Walker, S. J. (2007). Psychometrics of a simple method for scoring organizational approach on the Rey-Osterrieth complex figure. *Journal of neuropsychology*, 1(1), 39-51.
- Smith-Seemiller, L., Fow, N. R., Kant, R., & Franzen, M. D. (2003). Presence of post-concussion syndrome symptoms in patients with chronic pain vs mild traumatic brain injury. *Brain injury*, 17(3), 199-206.
- Snell, D. L., Martin, R., Surgenor, L. J., Siegert, R. J., & Hay-Smith, E. J. C. (2017). What's wrong with me? Seeking a coherent understanding of recovery after mild traumatic brain injury. *Disability and rehabilitation*, 39(19), 1968-1975.
- Snell, D., Martin, R., Surgenor, L., Hay-Smith, J., Hooper, G., Melzer, T., ... & Anderson, T. (2016). Mild traumatic brain injury and chronic pain: using mixed methods to explore unique and overlapping phenomenology. *Archives of Physical Medicine and Rehabilitation*, 97(10), e2-e3.
- Sonntag, D. (2018). Interactive cognitive assessment tools: a case study on digital pens for the clinical assessment of dementia. *arXiv preprint arXiv:1810.04943*.
- Stern, R. A., Singer, E. A., Duke, L. M., Singer, N. G., Morey, C. E., Daughtrey, E. W., & Kaplan, E. (1994). The Boston qualitative scoring system for the Rey-Osterrieth complex figure: description and interrater reliability. *The Clinical Neuropsychologist*, 8(3), 309-322.

- Stika, M. M., Riordan, P., Aaronson, A., Herrold, A. A., Ellison, R. L., Kletzel, S., ... & Smith, B. (2021). Cognition and other predictors of functional disability among veterans with mild traumatic brain injury and posttraumatic stress disorder. *The Journal of head trauma rehabilitation*, 36(1), 44.
- Strauss, E., & Spreen, O. (1990). A comparison of the Rey and Taylor figures. *Archives of Clinical Neuropsychology*, 5(4), 417-420.
- Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. American chemical society.
- Strong, C. A. H., Tiesma, D., & Donders, J. (2010). Criterion validity of the Delis-Kaplan Executive Function System (D-KEFS) fluency subtests after traumatic brain injury. *Journal of the International Neuropsychological Society*, 17(2), 230-237.
- Sufrinko, A.M., Mucha, A., Covassin, T. et al. (2017). Sex differences in vestibular/ocular and neurocognitive outcomes following sport-related concussion. *Clinical Journal of Sport Medicine: Official Journal of the Canadian Academy of Sport Medicine*, 27(2), 133–138.
- Sullivan, E. V., Mathalon, D. H., Ha, C. N., Zipursky, R. B., & Pfefferbaum, A. (1992). The contribution of constructional accuracy and organizational strategy to nonverbal recall in schizophrenia and chronic alcoholism. *Biological Psychiatry*, 32(4), 312-333.
- Sussman, E. S., Pendharkar, A. V., Ho, A. L., & Ghajar, J. (2018). Mild traumatic brain injury and concussion: terminology and classification. *Handbook of clinical neurology*, 158, 21-24.
- Sveen, U., Ostensjo, S., Laxe, S., & Soberg, H. L. (2013). Problems in functioning after a mild traumatic brain injury within the ICF framework: the patient perspective using focus groups. *Disability and Rehabilitation*, 35(9), 749-757.
- Syrjälä, I. (2020). The association between childhood ADHD symptoms and adult executive functioning measured with the Rey-Osterrieth Complex Figure in a perinatal risk cohort.
- Tailby, C., Collins, A. J., Vaughan, D. N., Abbott, D. F., O'Shea, M., Helmstaedter, C., & Jackson, G. D. (2020). Teleneuropsychology in the time of COVID-19: the experience of The Australian Epilepsy Project. *Seizure*, 83, 89-97.
- Taylor, E. M. (1959). Psychological appraisal of children with cerebral defects.
- Taylor, L. B. (1969). Localisation of cerebral lesions by psychological testing. *Neurosurgery*, 16(CN_suppl_1), 269-287.
- Teasdale, G., & Jennett, B. (1976). Assessment and prognosis of coma after head injury. *Acta neurochirurgica*, 34(1), 45-55.
- Tenconi, E., Santonastaso, P., Degortes, D., Bosello, R., Titton, F., Mapelli, D., & Favaro, A. (2010). Set-shifting abilities, central coherence, and handedness in anorexia nervosa patients, their unaffected siblings and healthy controls: exploring putative endophenotypes. *The World Journal of Biological Psychiatry*, 11(6), 813-823.
- The Commons Library, 2022. *Higher education student numbers*. London: Commons Library.

- Theodoros, D., Hill, A., Russell, T., Ward, E., & Wootton, R. (2008). Assessing acquired language disorders in adults via the Internet. *Telemedicine*
- Theodoros, D., Russell, T. G., Hill, A., Cahill, L., & Clark, K. (2003). Assessment of motor speech disorders online: a pilot study. *Journal of telemedicine and telecare*, 9(2_suppl), 66-68.
- Thompson, C., Davies, P., Herrmann, L., Summers, M., & Potter, S. (2016, January). Approaches to establishing validated cut-off scores on the Rivermead post-concussion symptoms questionnaire (RPQ). In *Brain Injury* (Vol. 30, No. 5-6, pp. 770-770). 530
- Todd, Z., Nerlich, B., McKeown, S., & Clarke, D. (2004). Mixing methods in psychology. *The integration of qualitative*.
- Tombaugh, T. N., & Hubley, A. M. (1991). Four studies comparing the Rey-Osterrieth and Taylor complex figures. *Journal of Clinical and Experimental Neuropsychology*, 13(4), 587-599.
- Tremblay, M. P., Potvin, O., Callahan, B. L., Belleville, S., Gagnon, J. F., Caza, N., ... & Macoir, J. (2015). Normative data for the Rey–Osterrieth and the Taylor complex figure tests in Quebec-French people. *Archives of Clinical Neuropsychology*, 30(1), 78-87.
- Troyer, A. K., & Wishart, H. A. (1997). A comparison of qualitative scoring systems for the Rey-Osterrieth Complex Figure Test. *The Clinical Neuropsychologist*, 11(4), 381-390.
- Tsatali, M., Emmanouel, A., Gialaouzidis, M., Avdikou, K., Stefanatos, C., Diamantidou, A., ... & Tsolaki, M. (2022). Rey complex figure test (RCFT): norms for the Greek older adult population. *Applied Neuropsychology: Adult*, 29(5), 958-966.
- Tupler, L. A., Welsh, K. A., Asare-Aboagye, Y., & Dawson, D. V. (1995). Reliability of the Rey-Osterrieth Complex Figure in use with memory-impaired patients. *Journal of clinical and experimental neuropsychology*, 17(4), 566-579.
- Uiterwijk, D., Stargatt, R., Humphrey, S., & Crowe, S. F. (2021). The relationship between cognitive functioning and symptoms of depression, anxiety, and post-traumatic stress disorder in adults with a traumatic brain injury: a meta-analysis. *Neuropsychology review*, 1-49.
- Umile, E. M., Sandel, M. E., Alavi, A., Terry, C. M., & Plotkin, R. C. (2002). Dynamic imaging in mild traumatic brain injury: support for the theory of medial temporal vulnerability. *Archives of physical medicine and rehabilitation*, 83(11), 1506-1513.
- van Ierssel, J., O'Neil, J., Sveistrup, H., Marshall, S., & Graham, I. (2021). A qualitative study of persons with persistent postconcussion symptoms and clinicians with concussion expertise to inform the development of a concussion-specific questionnaire. *Disability and Rehabilitation*, 43(23), 3365-3376.
- Vandborg, S. K., Hartmann, T. B., Bennedsen, B. E., Pedersen, A. D., & Thomsen, P. H. (2014). Memory and executive functions in patients with obsessive-compulsive disorder. *Cognitive and Behavioral Neurology*, 27(1), 8-16.
- Veeramuthu, V., Narayanan, V., Ramli, N., Hernowo, A., Waran, V., Bondi, M. W., ... & Ganesan, D. (2017). Neuropsychological outcomes in patients with complicated versus

- uncomplicated mild traumatic brain injury: 6-month follow-up. *World neurosurgery*, 97, 416-423.
- Vingerhoets, G., Lannoo, E., & Wolters, M. (1998). Comparing the Rey-Osterrieth and Taylor Complex Figures: Empirical data and meta-analysis. *Psychologica belgica*.
- Visser, R. S. H. (1973). *Manual of the Complex Figure Test: CFT: a Test Involving Restructuring Behaviour Developed for the Assessment of Brain Damage*. Swets en Zeitlinger.
- Vogel, A., Stokholm, J., & Jørgensen, K. (2012). Performances on Rey Auditory Verbal Learning Test and Rey Complex Figure Test in a healthy, elderly Danish sample—reference data and validity issues. *Scandinavian journal of psychology*, 53(1), 26-31.
- Vos, L., Whiteneck, G. G., Ngan, E., Leon-Novelo, L., & Sherer, M. (2019). Rasch analysis of postconcussive symptoms: development of crosswalks and the brain injury symptom scale. *Archives of physical medicine and rehabilitation*, 100(10), 1844-1852.
- Waber, D. P., & Holmes, J. M. (1986). Assessing children's memory productions of the Rey-Osterrieth Complex Figure. *Journal of Clinical and Experimental Neuropsychology*, 8(5), 563-580.
- Wallesch, C. W., Curio, N., Kutz, S., Jost, S., Bartels, C., & Synowitz, H. (2001). Outcome after mild-to-moderate blunt head injury: effects of focal lesions and diffuse axonal injury. *Brain injury*, 15(5), 401-412.
- Weider, S., Indredavik, M. S., Lydersen, S., & Hestad, K. (2016). Central coherence, visuoconstruction and visual memory in patients with eating disorders as measured by different scoring methods of the rey complex figure test. *European Eating Disorders Review*, 24(2), 106-113.
- Weinstein, C. S., Kaplan, E., Casey, M. B., & Hurwitz, I. (1990). Delineation of female performance on the Rey-Osterrieth Complex Figure. *Neuropsychology*, 4(2), 117.
- Wilde, E.A., Hunter, J.V., Bigler, E.D. (2014). Neuroimaging in Traumatic Brain Injury. In: Sherer, M., Sander, A. (eds) Handbook on the Neuropsychology of Traumatic Brain Injury. Clinical Handbooks in Neuropsychology. Springer, New York, NY.
- Wilson, N. A., & Batchelor, J. (2015). Examining Rey complex figure test organization in healthy adults. *Journal of Clinical and Experimental Neuropsychology*, 37(10), 1052-1061.
- Wylie, G. R., & Flashman, L. A. (2017). Understanding the interplay between mild traumatic brain injury and cognitive fatigue: models and treatments. *Concussion*, 2(4),
- Yamashita, H. (2015). Effects of the immediate recall trial on delayed recall performance in the Rey complex figure test in young and older adults. *Applied Neuropsychology: Adult*, 22(3), 197-203.
- Yamashita, H., & Yasugi, M. (2008). Comparison of the Medical College of Georgia Complex Figures and the Rey-Osterrieth Complex Figure tests in a normal sample of Japanese university students. *Perceptual and motor skills*, 107(1), 45-50.
- Yardley, L., & Bishop, F. (2017). Mixing qualitative and quantitative methods: A pragmatic approach. *The Sage handbook of qualitative research in psychology*, 398-413.

Young-Bernier, M., Kamil, Y., Tremblay, F., & Davidson, P. S. (2012). Associations between a neurophysiological marker of central cholinergic activity and cognitive functions in young and older adults. *Behavioral and Brain Functions*, 8(1), 1-9.

Zappalá, G., & Trexler, L. E. (1992). Quantitative and qualitative aspects of memory performance after minor head injury. *Archives of Clinical Neuropsychology*, 7(2), 145-154.

Zargar, F., Mohammadi, A., Shafiei, E., & Fakharian, E. (2015). Comparing cognitive failures and metacognitive beliefs in mild traumatic brain injured patients and normal controls in Kashan. *Archives of trauma research*, 4(2).

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale. *Acta psychiatrica scandinavica*, 67(6), 361-370.