Online monitoring of brain health in former athletes

Ellen Frances Boucher

Registration Number: 100299159/1

Submitted for the Degree of Doctor of Philosophy (PhD) in Studies Allied to Medicine

University of East Anglia, School of Health Sciences

Date of Submission: June 2024

Word Count: 54,116

None of the material offered has been previously submitted for a degree in any university.

Abstract

Growing evidence suggests that sport-related repetitive head impacts (RHI) increase risk of dementia. Longitudinal, prospective cohort studies are now needed to address remaining research gaps. This thesis aimed to evaluate the acceptability and validity of an online longitudinal study that monitors former athletes' brain health, to examine the study's first dataset, and to explore methods of improving representativeness.

Acceptability was qualitatively explored after one and five assessment repetitions. Participants found the study acceptable due to low burden, high accessibility, and high comfort. The study's feedback process, where participants can monitor their performance, met their expectations of ethicality and promoted motivation to participate.

The online format was compared to in-person testing to examine validity for assessing brain health. Findings from a mixed methods study reported quantitative validity of the online tests, and demonstrated qualitative contributors to validity that could influence results.

Quantitative analysis of the study's first dataset demonstrated significant relationships between sport related RHIs and poorer cognition, mental health, and behavioural outcomes in professional and amateur male athletes. Similar trends were demonstrated in female contact sport athletes, but further investigation is needed to better identify long-term effects in females and potential sex-differences.

To improve the representativeness of the study's dataset, a narrative review explored the definitions of elite female athletes in previous research. This review found considerable variation of definitions and proposes a new approach to unifying definitions of elite and improving inclusion of female athletes in research.

This thesis demonstrates the strengths of a novel approach towards monitoring the brain health of former athletes. Key learnings from this thesis include methods for improving recruitment and retention in longitudinal studies, influences on validity of online brain health assessments, and strategies to facilitate analysis of exposure to contact sport on brain health in male and female, professional and amateur athletes.

Contents

Figures and tables

Abbreviations

AQ: Buss-Perry Aggression Questionnaire

BDNF: Brain-Derived Neurotrophic Factor

BIS-15: Barratt Impulsivity Scale

CLSA: Canadian Longitudinal Study of Ageing

CTE: Chronic Traumatic Encephalopathy

DSB: Digit Span Backwards assessment

FA: Football Association

FAW: Football Association of Wales

GAD-7: General Anxiety Disorder questionnaire

GP: General Practitioner

ICC: Intraclass Correlation Coefficient

IFA: Irish Football Association

MCI: Mild Cognitive Impairment

MoCA: Montral Cognitive Assessment

mTBI: Mild Traumatic Brain Injury

NeurOn: Neuropsychology Online

NHS: National Health Service

NIWFA: Northern Irish Women's Football Association

OSA: Obstructive Sleep Apnoea

PHQ-9: Patient Health Questionnaire

PPI: Patient and Public Involvement

PRH1: Picture Recognition Correct Hits

PRH2: Picture Recognition Correct Hits Delayed Recall

PRSM1: Picture Recognition Source Memory

PRSM2: Picture Recognition Source Memory Delayed Recall

PSQI: Pittsburgh Sleep Quality Index

RHI: Repetitive Head Injury

RT: Reaction Time

SART: Sustained Attention to Response Task

SCORES: Screening for Cognitive Outcomes after Repetitive Head impact Exposure in Sport

SD: Standard Deviation

SNP: Single Nucleotide Polymorphism

SWFA: Scottish Women's Football Association

SWM: Spatial Working Memory assessment

TBI: Traumatic Brain Injury

TES: Traumatic Encephalopathy Syndrome

TFA: Theoretical Framework of Acceptability

TMTA: Trail Making Task A

TMTB: Trail Making Task B

UEFA: Union of European Football Associations

UK: United Kingdom

VSTF: Virtual Supermarket Task Final Position

VSTS: Virtual Supermarket Task Starting Position

WFA: Women's Football Association

WHO: World Health Organisation

WOES: Warwick and Oxford Exploration of Sleep

WSL: Women's Super League

Acknowledgements

This work would not have been possible without the support and dedication of many people who I would like to acknowledge here.

Firstly, I would like to thank the SCORES participants for your commitment to the project and for sharing your voices. I am grateful for your generosity in giving your time to this research. I understand the meaning of the research for many of you, and I hope that through your efforts we can contribute to protecting future generations of athletes.

I would also like to thank the University of East Anglia's Faculty of Medicine and Health Sciences for funding my studentship. I am very thankful for this opportunity to complete this PhD research.

Over the course of the four years of my PhD my supervisory team have been excellent mentors. Your encouragement to learn new approaches and your expertise has helped me to develop as a researcher and uncover valuable understanding throughout this project. I am grateful for your time and dedication to my supervision over the past four years, and I feel very lucky to have worked with you all. I would like to say a big thank you to all three of you for encouraging me to pursue a research topic that I am passionate about, and for creating an environment where I feel I have learned a great deal.

I am also grateful to the wider UEA community, in particular Suzanna Dell who has been excellent in supporting the administration of the ongoing project and data collection for this thesis. I would like to thank Allan Clark, whose pragmatic approach to statistics has helped me to unravel a lot of my confusion during analysis. I am also grateful for input from Julia Robathan, whose consultation was incredibly valuable towards considering how female athletes can be better included in this research.

I am lucky to have had a wonderful support team with me along the way. To Thorpe St Andrews Women's Football Club, thank you for becoming a place for me to step away from my work, and for understanding my avoidance of heading the ball. To my PGR friends, thank you for cultivating a collective wisdom of how to get through a PhD in one piece. Your advice and support are invaluable, and your successes are inspiring. To Ellie, Issy and Laura, thanks for pushing me to do this in the first place! Your friendship and belief mean the world to me.

Throughout this project my family have been a constant source of inspiration in my work by encouraging my interests in sport and neuroscience and inspiring me to advocate for better inclusion of women in research. To Mum, Dad, Joe, Robbie (and Pippin), thank you all for cheering me on and being there for the wins and the hard times too. Your love and support has made this all possible, and I am forever grateful for your part in helping me to "finish strong"!

Thank you all.

Publications and statement of authorship

Publications arising from this thesis:

Boucher E. F., Grey M. J., Hornberger M, Hanson S. Online longitudinal monitoring of brain health in former contact sport athletes: A study of acceptability and ethicality. European Journal of Sport Science. 2024 May;24(5):549-56.

Boucher E. F., Holdaway F., Robathan J., Grey M. J., Hanson, S. (2024). Moving the Goalposts: variance in defining 'elite women's sport' limits inclusion in research. Submitted to Sports Medicine in March 2024. (Under second review)

Boucher E., F., Clark A., Hanson S., Hornberger M., Grey M. J. (2024). Online monitoring of brain health in former contact sport athletes: Early findings from the SCORES Project. (Preparing for submission)

Statement of authorship

The above publications have been prepared with the support of the authorship teams. My contributions to the publications have been in leading the design, data collection and analysis, interpretation of results and drafts of the papers. The authorship teams have supported me by helping in the design of the studies, supporting the data collection and analysis, and editing the drafted papers.

Introduction

Background

Health and physical activity

The positive effects of physical activity on health are well established. Physical inactivity is associated with a higher risk of diseases such as diabetes, heart disease and some cancers, and increasing physical activity is associated with improvements in both mental and physical health (1). These associations are supported by a vast body of literature (2-4) which have contributed to the development of international guidelines. Worldwide, these guidelines are designed to promote physical activity as a means for reducing prevalence of diseases and to combat the socioeconomic burden of ageing populations (1, 5). In the United Kingdom (UK), where 25% of adults are classed as inactive (6), initiatives such as 'Get Active' are designed to increase national levels of physical activity and reduce associated health issues (7). The National Health Service (NHS) also promotes programmes such as 'Couch to 5k' to tackle sedentary behaviour and promote physical activity (8). This association is therefore established beyond academic pursuits and forms the basis of global initiatives to reduce the prevalence and consequences of health problems worldwide.

Physical activity, referring to any "any bodily movement produced by skeletal muscles that results in energy expenditure" (9), includes participation in sport as well as movement for travel, occupation or domestic and leisure activities. However, physical activity is distinct from sport, which refers to physical activities performed individually or as part of a team with a common set of goals, expectations and rules (10). Participation in sport can promote the positive health benefits gained from participation in physical activity, and additional health benefits achieved from increased social structures associated with sport. For example, participation in endurance sports is associated with reduced risk of obesity, Type 2 diabetes and coronary heart disease (11, 12). Team sports are associated with improved cardiovascular fitness (13) and strong evidence suggests an association with improved mental health benefits (14). Evidence therefore supports that participation in sport has many positive effects on health.

The positive effects of physical activity specifically on brain health are also well established. The World Health Organisation defines brain health as, "the state of brain functioning across cognitive, sensory, social-emotional, behavioural and motor domains, allowing a person to realise their full potential over the life course, irrespective of the presence or absence of disorders" (15). In the context of this thesis, the term brain health refers to the overall functioning of the brain as indicated by measures of cognitive, mental and behavioural health. Participation in physical activity is associated with a lower risk of neurodegenerative disease by reducing the incidence of health conditions that increase risk of dementia including diabetes, heart disease, obesity and cancer. Sport and exercise reduce stress and inflammation, and improve cardiovascular factors,

insulin sensitivity and neurotrophic factors, which are all related to risk of dementia (16). Physical activity promotes neuroplasticity and neurogenesis, and is suggested to improve white matter preservation (17). These are important factors towards maintaining long-term brain health and cognitive reserve (18). On an environmental level, participation in sporting activity reduces social isolation and improves social support structures (19), which can reduce dementia risk. Overall, there are many benefits of participation in sport in general on long-term brain health, and promoting participation in physical activity is a key objective in reducing dementia risk worldwide (20).

Risk associated with sport

Whilst strong evidence supports that participation in sport and physical activity is associated with several positive health benefits, participation in sport is also associated with a number of risks. Acceptance of potential harm to physical health has been demonstrated to create a culture of risk in athletes seeking to reach maximum performance in their sports (21). Participation in sport is coupled with an increased risk of injury, particularly musculoskeletal injuries (22). Consequences of musculoskeletal injuries can include absence from work, school, athletic training or competitions, and associated financial costs (23). These injuries can therefore be lifechanging, but not necessarily life ending. However, some sporting injuries also have long-term consequences on health. For example, occurrence of sport-related knee injuries has been found to be associated with increased risk of developing osteoarthritis later in life (24). This demonstrates an increased interest in research that investigates the long-term consequences of sport-related injuries on health.

Of particular concern, is the growing evidence that suggests an association between sport-related injury in contact sport and long-term brain health. For the purpose of this thesis, contact sport refers to any sport where physical contact between competitors, or their environment, is an inherent part of the game. This includes but is not limited to sports such as association football (football), rugby, American football, or hockey. Musculoskeletal injuries occur in both contact and non-contact sports, however studies demonstrate that contact sports pose a higher risk for exposure to repetitive head injuries (RHI) (25, 26). Concerns about the long-term consequences of RHI from participation in contact sport have been recorded in scientific literature for more than 100 years. In 1906, the Harvard University American football team physicians expressed concern about the long-term consequences of exposure to multiple RHIs through sport (27). They called for the conditions of the game, “to be modified as to diminish to a very great degree the number of injuries”. Martland later described “dementia pugilistica” as a long-term consequence of exposure to repetitive head injuries in boxers, manifested as confusion, Parkinsonian symptoms and slowed movements (28). A growing body of evidence now supports this association between participation in contact sport and later neuropsychological consequences (29), based on the mechanism of exposure to sport-related RHIs increasing risk of neurodegenerative disease.

Traumatic brain injury and neurodegenerative disease

Traumatic brain injury (TBI) occurs as the result of force transmitted to the head or body causing damage and dysfunction (29). TBIs are typically categorised as severe, moderate, or mild, referring to symptom outcomes on the Glasgow Coma Scale (30). Mild TBI (mTBI) accounts for 58-88% of all TBIs (31). The terms mTBI and concussion are often used interchangeably in research and clinical contexts, which has implications on the diagnosis and treatment of patients, and the heterogeneity of research samples in trials (32). The definition of concussion has undergone significant modifications and updates in the past 20 years (33). The most recent clinical definition (33) does not provide specific diagnostic criteria (33) and does not distinguish thresholds between concussions and brain injuries that do not present with observable signs of concussions. This is important, because these injuries that do not result in concussive symptoms are thought to accumulate damage in the brain which drives long-term neurodegeneration in the brain (34, 35). In sport, exposure to these injuries occurs, for example, through exposure to heading in football, tackles in rugby, or sparring in boxing. These brain injuries that do not present as a concussion are typically referred to as sub-concussive injuries although this terminology is also considered a misnomer (36).Recent revision of this term highlights that whilst the term sub-concussive suggests that the severity of impacts are less severe than concussions, research finds that approximately 10% of impacts that do not result in concussion are more severe than those that result in concussive symptoms (30). The term RHI is therefore used within this thesis and other research in this field (31) to describe exposure to repetitive head injuries that may include concussive or insults that do not result in concussive symptoms. Regardless of terminology, the study of these accumulative RHIs is important towards understanding the link between contact sport and dementia because it is suggested that exposure to RHIs, rather than concussions alone, is responsible for driving neurodegeneration (31).

In TBI, the primary injury to the brain is damage to the nervous tissue that occurs following exposure to physical forces (38-40). Rapid acceleration-deceleration in the brain produces forces within the nervous tissue that causes axonal injury and impact with the cranial wall (39, 40). The secondary injury in TBI occurs as a response to biochemical changes in the nervous tissue (40-42). These changes occur in the hours or days following injury and consist of a metabolic cascade of mitochondrial dysfunction, increased excitotoxicity and overactivation of microglial inflammatory processes (40). The secondary injury in TBI prolong cellular injury past the initial traumatic event, causing disruption to cellular processes that contribute to long-term brain health (40). Axonal damage occurs through Wallerian degeneration – a process whereby the most distal section of the axon to the injury degenerates 24-48h after injury and ultimately results in the degradation of the axon, myelin sheath and Schwann cells (43). A consequence of the disruption to cellular processes is the accumulation of proteins like amyloid or tau which play key roles in the pathology of neurodegenerative diseases (40). The aggregation of these proteins, as well as prolonged

neuroinflammation, excitotoxicity and oxidative stress are suggested mechanisms for increased incidences of neurodegenerative disease found in people with a history of TBI (40).

TBIs are a recognised modifiable risk factor for neurodegenerative disease (44, 45). Systematic reviews and meta-analyses estimate the increased risk of neurodegenerative disease following exposure to traumatic brain injuries to be between 63% and 96% (46-48). This data presents a clear link between exposure to TBI in general and neurodegenerative risk, but the link between RHIs and neurodegenerative disease is less understood. A reason for this is that moderate and severe TBIs are typically recorded in medical histories, meaning that risk analyses are more feasible. Sport-related concussions have been demonstrated to be underreported in athlete populations (49) and both concussive and sub-threshold injuries cannot be detected using standard clinical imaging (33). These factors suggest that risk analysis is more challenging in concussive and sub-concussive injuries. However, growing evidence (28) suggests a higher prevalence of neurodegenerative disease in athletes exposed to concussive and sub-concussive injury through contact sport.

Neurodegenerative diseases include disorders caused by a progressive loss of neuronal structure or function through neurodegeneration (32). Neurodegenerative diseases are characterised by progressive disruption to neuronal processes, resulting in progressive neuronal atrophy that spreads throughout the nervous system causing widespread disruption and ultimately death (33). The most common neurodegenerative disease is Alzheimer's disease, which contributes to approximately 60-70% of cases worldwide (34). Chronic Traumatic Encephalopathy (CTE) is a neurodegenerative disease thought to be caused by repetitive exposure to RHIs, for example in military contexts, domestic violence or sport (35). Currently, CTE can only be diagnosed post-mortem on a neuropathological level (35), as in vivo biomarkers that can accurately detect CTE pathophysiology are not yet clinically available (36).

The clinical presentation of symptoms of CTE is referred to as Traumatic Encephalopathy Syndrome (TES) (37). Retrospective interviews have largely informed the clinical profile of TES to date. They suggest that TES is associated with impulsivity, depression/suicidality, substance misuse, anxiety, anger, gait instability, motor slowness and cognitive changes (38). However, retrospective interviews are subject to recall bias, these symptoms are also commonly seen in many other dementias, and symptoms such as depression and substance misuse can be caused by many factors. A literature review of the clinical profile of CTE describes distinct behavioural, mood, cognitive and motor symptoms as presented in Table 1 (39).

Table 1: Clinical Features of CTE from Montenigro et al., 2015.

Behavioural features	Mood features	Cognitive features	Motor features
Explosivity	Depression	Dementia	Ataxia
Loss of control	Hopelessness	Memory impairment	Dysarthria
Short fuse	Suicidality	Executive dysfunction	Parkinsonism
Impulsivity	Anxiety	Lack of insight	Gait Disturbance
Aggression	Fearfulness	Perseveration	Tremor
Rage	Irritability	Impaired attention	Masked facies
Physical violence	Labile emotions	Impaired concentration	Rigidity
Verbal violence	Apathy	Language difficulties	Muscle weakness
Inappropriate speech	Loss of interest	Dysgraphia	Spasticity
Boastfulness	Fatigue	Alogia	Clonus
Childish behaviour	Flat affect	Visuospatial difficulties	
Social inappropriateness	Insomnia	General cognitive impairment	
Disinhibited speech	Mania	Reduced intelligence	
Disinhibited behaviour	Euphoria		
Paranoid delusions	Mood swings		
Personality changes	Prolix		
Psychosis			
Social isolation			

Over time, the symptoms of CTE progress to dementia. Currently, the clinical profile of dementia associated with CTE is similar to the clinical profiles of other neurodegenerative diseases including Alzheimer's Disease, and Frontotemporal dementia, which makes diagnosis based on clinical symptoms challenging (58). Because of this, there are currently no clinical criteria for diagnosing CTE in vivo (59). This proposed profile of behaviour, mood, cognitive and motor features is based on retrospective data from neuropathologically confirmed cases of CTE. In these cases, information about the presentation of symptoms is collected from informant interviews. Whilst these interviews offer important information that builds our understanding of the presentation of symptoms, prospective evidence is needed to clarify the clinical profile of TES to a point at which it can be diagnosed whilst the individual is still alive.

Monitoring for declining brain health

Physiological and cognitive decline are hallmarks of "healthy" ageing, observed in people who do not have a neurodegenerative disease. In healthy ageing, neuroanatomical changes such as gradual shrinking of the brain, decline of white matter integrity, increases in the size of the ventricular system, reduced axon myelination, and reduced integrity of the corpus callossum and functionality of visual areas contribute to a progressive decline in cognitive ability and brain health over time (60). These neuroanatomical changes have consequences on cognition, where domains such as memory, processing speed and reasoning are observed to gradually decline over the lifespan (60). Rates of cognitive ageing are established to be caused by many factors including genetics and lifestyle factors like diet, levels of physical activity, or access to social support

networks (61). In general, rates of cognitive ageing are slow and gradual across the lifespan (60). Cognitive ageing in neurodegenerative diseases is also progressive, but what differentiates healthy ageing from cognitive impairments as a consequence of neurodegenerative disease, is that neurodegenerative impairments begin to occur earlier than healthy ageing, and progress at a faster rate (60) leading to dementia.

The term dementia refers to a group of symptoms including impairments in memory, thinking, mood or social abilities as a consequence of neurodegenerative disease. It is estimated that more than 50 million people are living with dementia worldwide, and this figure is predicted to triple by 2050 (40). Dementia is progressive, and is characterised by cognitive, functional and behavioural impairment that worsens over time (41). In 2023, an analysis of age of dementia diagnosis in England found that the median age of diagnosis was between 81 and 86 years, varying by ethnic group (42). However, dementia follows a progressive disease continuum and symptoms can be observed years prior to diagnosis (41). The earliest evidence of neurodegenerative disease is observed in pathological changes up to 20 years before the onset of cognitive symptoms (43). This asymptomatic stage is referred to as preclinical dementia (41). The following stage where cognitive impairments are observed is referred to as mild cognitive impairment (MCI), which is typically characterised by impairments in memory, language, executive function and navigation that begins to affect everyday tasks (41). Over time, impairments may progress to a stage characterised by severe cognitive deficits that interfere with everyday functioning to an extent where assistance is required, at which point a diagnosis of dementia is typically made (41).

Effectively monitoring brain health across the lifespan is considered a valuable approach towards mitigating against the effects of dementia by aiding early detection (66). The importance of early recognition of dementia is that earlier lifestyle and medical interventions can take place to improve quality of life and slow progression of disease (67). Although current medical interventions are not available to reverse the progression of neurodegenerative disease, emerging drugs are found to be more effective in earlier stages of dementia (68, 69). The UK adopts an approach towards screening for dementia whereby at-risk groups are invited to assessments, including people aged over 75 or with vascular conditions (67). These assessments are infrequent, meaning that they may not capture progressive change over time accurately or identify when symptoms appear or worsen. The use of these assessments in research is also limited, given that they are infrequent and costly, which limits the range of participants that can be captured by cohort studies. In order to monitor brain health for signs of MCI, frequent and accessible measures are needed to accurately detect impairments and track changes at regular intervals.

Studies of sport-related neurodegenerative disease

Research to date has largely approached the study of sport-related neurodegenerative disease by examining post-mortem physiology, assessing risk in population studies using health data, or by exploring long-term neuropsychological consequences of exposure to contact sport in former athletes. The following section describes each approach, evaluates their strengths and limitations, and identifies the gaps in literature to date.

Post-mortem studies identify the neuropathological consequences of contact sport

To examine the long-term consequences of exposure to RHI on brain health, studies have analysed the postmortem brain tissues of former contact sport athletes to identify signs of neurodegeneration. In 1973, Corsellis presented early neuropathologic findings following their study of the brains of 15 retired boxers (44). They observed a pattern of cerebral damage and degeneration in the septal regions, in the deep temporal grey matter, and in the cerebellar and nigral pathways. Using data from informant interviews, they also reported a wide variation of observable cognitive and behavioural symptoms. It was not until 2005 that dementia in sport gained widespread lay and scientific attention, when a deceased former American football player exhibited neuropathological evidence of neurodegenerative disease (45). Subsequently, neuropathologically confirmed cases of neurodegenerative disease have been found in male athletes from American football (46, 47), rugby (48), ice hockey (49), boxing (50), and Association Football (football) (51). To date, a single case study reports neuropathologically confirmed neurodegenerative disease in a female Australian rules football athlete (52).

Whilst postmortem studies allow the understanding of pathological consequences on the brain and suggest potential domains of cognition and behaviour that were affected, they offer little information about the manifestation of consequences whilst athletes are still alive. Informant data sheds light on domains of cognition, mental health, and behaviour (57), but reliability of data from retrospective informant interviews is limited. Whilst these interviews offer valuable qualitative information about the scope of symptoms impaired in former athletes, quantitative data is needed to isolate the specific domains of brain health effected and to track their progression over time.

Health record analyses identify the scale of neurodegenerative disease in sport

Retrospective analysis of health records also supports a link between participation in contact sport and long-term brain health problems. In 2019, an analysis of the health records of former professional male association footballers found a decreased risk of heart disease and lung cancer, but an increased risk of neurodegenerative disease 3.5 times the risk in matched controls (53).

These findings were supported by two further studies of former male professional footballers, whereby risk of neurodegenerative disease was higher among football players than controls (54, 55). Another study of former professional male footballers (56) has also demonstrated higher rates of amyotrophic lateral sclerosis, a neurodegenerative disease that effects motor neurons in footballers. These health surveillance studies demonstrate that contact sport players are more likely to die with neurodegenerative disease than control populations, and taken together with pathology studies, support a link between repetitive sport related RHI and neurodegenerative diseases, such as Alzheimer's disease or CTE. However, they offer relatively little information about the manifestation of pathology in observable signs and symptoms whilst athletes are still alive.

Health record studies are limited in their validity, particularly in regards to representing mental health outcomes (57). The manifestation of mental health symptoms is an area of much debate within the literature. Exposure to traumatic brain injury is associated with increased risk of depression and suicidality (58, 59). This suggests that exposure to repetitive sport-related RHI may also influence mental health outcomes (60). Evidence from pathologically confirmed cases of CTE suggests that increased depression and suicidality may be associated with exposure to contact sport (61). However, studies using data from health records find that male former professional footballers are less likely to record mental health outcomes than control populations (62, 63). A narrative review of literature published in 2019 argued that suicidality should therefore not be included in the clinical features of CTE (64). Therefore, further investigation that monitors for mental health symptoms and tracks their progression over time is necessary to better understand this association.

Additionally, previous work suggests that increased media coverage and public awareness of the link between contact sport and dementia may result in former professional athletes being more likely now to seek a medical diagnosis (65). However, it has also been suggested that former athletes may avoid medical assessment because of heightened media coverage of the link between dementia and contact sport (66). Either outcome may influence bias in studies that solely use health record data, and therefore methods that seek further information than health records are now warranted following increased media attention and public awareness of the issue.

Retrospective and post-mortem study designs also make it challenging to understand the relationship of other confounding lifestyle factors or health comorbidities with the long-term sequelae of exposure to contact sport. Previous investigations highlight the importance of considering lifestyle factors and other risk factors associated with developing dementia, and how these might influence risk associated with contact sport (67). Genetic predisposition remains an area that is poorly understood. Genetic risk factors of interest include variation in SNPs related to

Alzheimer's disease such as the presence of APOE4 and decreased levels of Brain Derived Neurotrophic Factor (BDNF) (68-70) as well as genetic variations reportedly linked to chronic traumatic encephalopathy (71). Demographic factors such as ethnicity (72), education (73), sex (74) and socio-economic status (75) are well-evidenced factors that influence risk of dementia as well as performance on brain health measures, and may vary in contact sport participants. Lifestyle factors that may be associated with retirement from professional sport, such as level of physical activity or cardiovascular health (76, 77), may also be important to consider as confounds on long-term brain health. In order to account for these factors, particularly for genetic risk, large scale study populations are required to make meaningful adjustments for potential confounds and to address this gap in the literature.

Neuropsychological studies offer a snapshot of long-term consequences

In order to better understand the profile of neuropsychological outcomes following exposure to contact sport, previous investigations have measured outcomes in retired athletes and older participants. A systematic review of studies investigating long-term cognitive effects of exposure to contact sport found evidence of deficits in memory, executive function, psychomotor function and self-reported cognitive functioning in former contact sport athletes (65). However, their conclusions did not support an association between deficits and concussion history. A review of studies capturing brain health in retired American footballers reports that verbal memory may be worsened in players compared to controls, but findings were limited by small sample sizes and absence of longitudinal data (78).

In retired rugby players, a 2021 study found that subjective self-reporting of cognitive decline did not correlate with performance on neuropsychological tests, and suggests that rather perceived cognitive decline could be a reflection of psychological distress (79). A study of rugby players in the UK found increased mental health disorder prevalence and sleep disruption in elite rugby players and athletes with higher numbers of reported concussions (80).

In football, retired professional males demonstrate worse performance on verbal fluency and verbal learning assessments than controls (66). Further studies have found an association of increased heading with negative cognitive outcomes later in life (81, 82), although these measures of heading are based on self-report data. A study of female footballers found no difference in neurocognitive function, with the exception of verbal fluency, when compared with a control population (83). This sample represented a relatively young cohort, and offers no longitudinal follow up, but is a rare example of the inclusion of female athletes in this field.

Throughout these studies and reviews, it is notable that the relationship between number of concussions and cognitive outcomes is unclear due to validity and reliability of varying methods of reporting concussions (78). The long-term effects on cognition, mental health and behaviour also

vary with study sizes, assessment techniques and control comparisons. Many of these studies were published since the start of this PhD project and contribute to a growing understanding of the long-term consequences of exposure to contact sport on brain health. However, there are still remaining gaps in knowledge and limitations that need to be addressed.

The data available through prior investigations is further limited in terms of representation of the athletes. Most data are from male former professional athletes. Females and amateur or recreational athletes are underrepresented in this field (60, 84). This is problematic for two reasons. Firstly, female athletes are more likely to sustain concussions than male athletes (85), and their symptoms may be more prolonged and severe (86). Females are also more likely to be diagnosed with dementia than males worldwide (87). Although males represent the majority of participants in contact sports, female participation in sports such as football and rugby has shown considerable growth in recent years (6). To date, very little research has been performed using female participants in this field (84). Published in 2020, a cross-section of a cohort of German 1st league female footballers found no differences between footballers and a normative sample on neurocognitive assessments, with the exception of a verbal fluency test where footballers performed worse than the normative sample (83). This investigation also found that depression was associated with headers but not concussions in the football sample. However, this study cohort had a mean age of 37.4 years (SD = 4.8) and therefore represents a relatively young sample. This study offers important information about how the long-term effects of contact sport may manifest in female athletes, however further data from an older study cohort is needed to identify how neurocognitive health is affected as athletes age. Additionally, further data collection is needed to make appropriate comparisons to male athletes in order to identify any sex-differences in ageing as a consequence of exposure to contact sport.

The absence of data from amateur athletes is problematic because amateur and recreational athletes make up the largest participation groups in sport. Their inclusion in research is paramount because non-professional athletes are also exposed to sport related RHI at levels that may have long-term consequences. Reducing the incidence of this preventable risk factor of dementia is therefore a public health issue. A small body of evidence suggests that amateur association footballers, wrestlers and boxers may be at an increased risk of dementia compared to the general population (60). This Finnish study monitors a cohort of male Olympic athletes who competed between 1920 and 1965 and were technically amateur athletes in that they were not paid, although perhaps represent a cohort of athletes with more regular training and higher intensities of exposure to contact through sport than a general population of amateur or recreational athletes. Although this is an important study to understand the risk in non-professional athletes, the inclusion of a wider population of participants would improve the data about the link between contact sport and brain health and improve policies and safeguarding in lower levels of sport.

There are also outcomes of brain health that remain relatively unexplored in the current literature, including sleep health. Emerging evidence suggests that sleep health may be an important domain of brain health to monitor in former contact sport athletes. A recent systematic review found an increased prevalence of obstructive sleep apnoea (OSA) in male former professional contact sport athletes compared to a control population (88). The relevance of this relationship is that OSA is demonstrated to increase risk of cognitive dysfunction and dementia (89-91). OSA has also been shown to develop early in the stages of dementia, before cognitive impairment is observed (92). Therefore, the presence of OSA or changes in sleep may be important markers in monitoring the brain health of athletes that is unexplored in other investigations.

Importantly, these investigations do not monitor change over time. Longitudinal monitoring is a necessity in this research, in order to better capture changes in ageing over time and compare long-term data to baseline results. Dementia is progressive, and cognitive profiles in MCI and dementia differ in presentations in different populations, such as between males and females (119, 120). Therefore, studies that monitor long-term neuropsychological consequences cannot assume that trajectories of decline are homogenous. Repeated measures, that monitor differences in trajectories over time in a representative sample are needed to better understand the long-term effects on cognition, mental health, and behaviour.

Building a more representative data set

Evidence available from pathology, health surveillance and initial long-term neuropsychological studies is already suitable for prioritising the safeguarding of athletes in general through policy changes (29). Already, research in this area has led to policy changes in sport including the lowering of tackle height in community level rugby (93), and the implementation of limits for heading in football training sessions (94). In order to optimise and update safeguarding procedures and policies that are based on the latest evidence it is clear that there are still under researched areas that could better inform our understanding of the link between contact sport and dementia.

To address the gaps in the literature, large scale prospective research is needed to build a database that is diverse in participants and collects information about potential confounds on brain health. This database should aim to meet several requirements. The recruitment of elite and amateur participants is paramount for furthering our understanding of if and how level of exposure to sport may influence long-term brain health. The recruitment of female athletes as well as males, and the collection of detailed sporting history to accurately compare them by level of sport is crucial to understanding possible sex differences in long-term health. Detailed information on potential confounds and lifestyle factors are also needed to clarify their influence

on brain health in athletes. The collection of qualitative, as well as quantitative data may also benefit this database to better understand ageing in a cohort (95).

Longitudinal prospective studies can address these needs by building datasets that are more representative and track progression over time. However, longitudinal studies are costly for researchers and rely on extended commitment from participants. In person assessment burdens participants with travel and time commitments, which may influence their likelihood of retention. Increased perceived burden of participation has been shown to decrease attendance rates and increase drop-out rates in existing longitudinal research studies (96), and therefore the importance of participant burden must be considered in the design of longitudinal data collection methods. The success of longitudinal data collection is also reliant on the success of recruitment and retention of participants (97). Therefore, the assessment of the acceptability of participation in long-term studies of brain health is important towards understanding factors that can influence recruitment and retention from the perspective of participants.

Remote online assessment is an alternative method for data collection that could promote recruitment and retention in longitudinal brain health studies and create a meaningful database. Online assessment potentially enables greater reach of a more diverse set of participants and reduces the burden associated with in-person assessment. These benefits have been demonstrated in existing longitudinal health studies, whereby the inclusion of online assessments in study protocols can reduce attrition rates in comparison to paper and pencil measures (98). However, online participation also requires a certain level of computer literacy and access to suitable devices (98). There are also advantages in online testing by allowing participants to complete assessments in their own environments, although this also comes with concerns surrounding the ability to control research environments. The possible increased acceptability of completing online assessments as part of a longitudinal study needs to be counterbalanced with confirming their validity in comparison to in-person testing. The evaluation of an online study that monitors the brain health of former contact sport athletes is therefore necessary to explore this as an appropriate format for addressing key research questions.

To monitor athlete brain health over time using online methods, we launched the SCORES (Screening for Cognitive Outcomes after Repetitive Head impact Exposure in Sport) project in October 2020 (www.scoresproject.org). The project is designed to monitor the brain health of former athletes as they age and understand the link between exposure to contact sport and brain health in living participants. The overall goals of this longitudinal project were to address key questions that are currently lacking in the literature, such as:

- How do the long-term consequences of exposure to contact sport appear in living participants?

- Are amateur or recreational, as well as professional, athletes at an increased risk of long-term sequelae?
- How do confounding lifestyle factors affect the long-term risks of participation in contact sport?
- Do sex differences in brain health exist in long-term consequences of exposure to contact sport?
- At what age can a difference in brain health between exposure and contact sport athletes be observed?

To address these questions, the project aimed to build a dataset to:

- Compare group means between appropriate exposure and control groups to isolate the long-term effects of contact sport on brain health.
- Monitor a cohort group over time for changes in brain health as they age.

The methods used in the SCORES project are further described in Chapter 2. Data for this thesis was collected from the SCORES project over the course of the first three years of data collection, as well as from studies taking place outside of the project. This thesis therefore sits separately to the SCORES Project, and uses data collected as part of the project as well as data collected outside of the project. I joined the project at the commencement of my PhD time period in July 2020. In order to collect data for my thesis, my contributions to the SCORES project were to aid in the design from the beginning of the project, and to set up the study online. I led the design of recruitment materials and was responsible for the administration of recruiting participants to the study. I contributed to the collection, processing, and analysis of data for the project. I also contributed to the dissemination of research findings by giving conference presentations, writing research papers, and engaging with public outreach events.

Aims and objectives

The overall aim of this thesis is to explore an online longitudinal method (SCORES) as an appropriate approach for monitoring the brain health of former athletes to better understand the link between contact sport and dementia. To assess this method, this thesis will explore the acceptability of participation, early results, and the validity of online methods as a way to address key questions (see Figure 1 for summary of research questions).

To address these questions, my research aims for this thesis are:

1. To explore the acceptability of participation in an online longitudinal study of brain health from the perspective of participants. (Study 1 and 4)

2. To examine the validity of online cognitive assessments in comparison to in-person testing (Study 2).
3. To use initial findings from the longitudinal study to examine differences in long-term cognitive, mental health, behavioural and sleep outcomes between appropriate exposure and control groups (Study 3)
4. To explore methods of promoting and improving the validity of including female athletes in longitudinal cohort studies (Narrative Literature review)

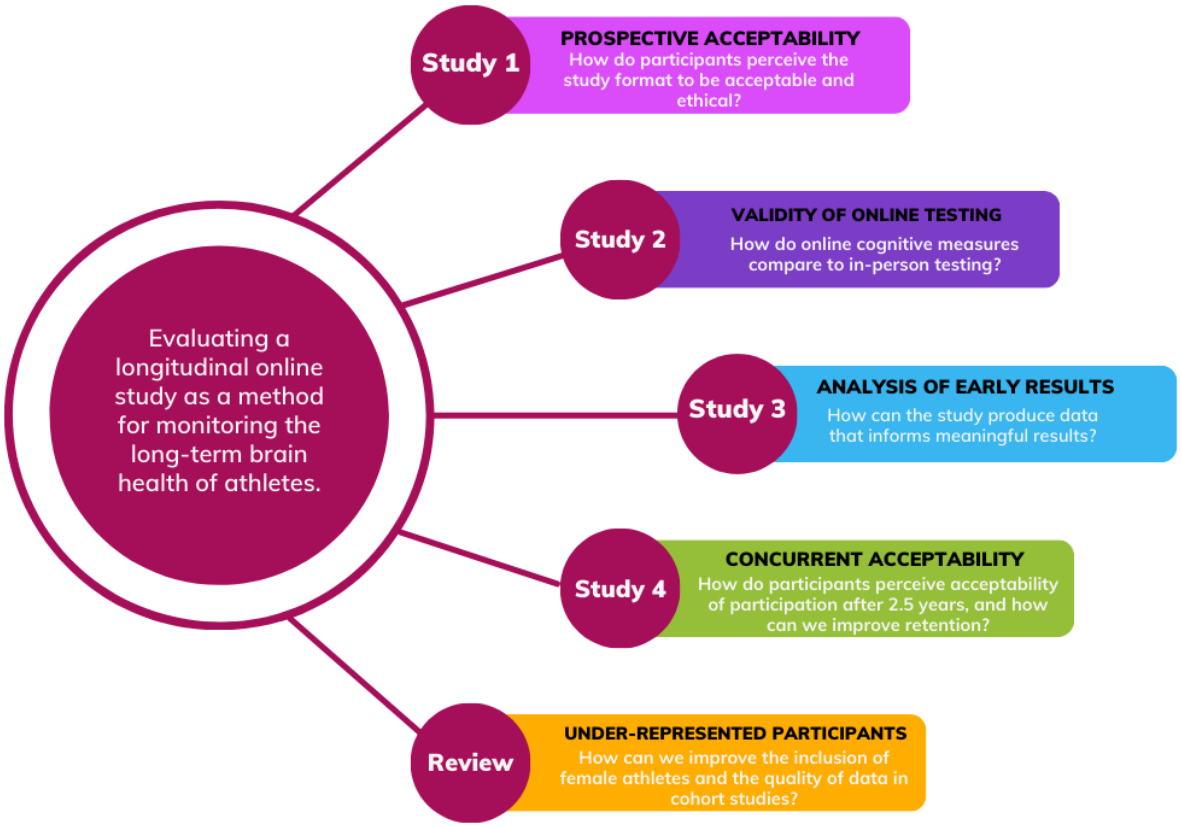


Figure 1: Summary of research questions explored in the order in which they are discussed within this thesis.

Chapter order

This thesis begins with explanations of the overarching methodologies used to address each research question. The methods chapter provides an overview of these approaches, with further detail in each study chapter. The methods chapter also provides a further overview of the SCORES project, and the context in which this thesis fits in relation to the project.

Study 1 assesses the prospective and concurrent acceptability of participation within the SCORES project from the perspective of participants after one round of assessments. This chapter introduces the methods used in the SCORES project and in later chapters of this thesis and

presents key learnings for the development of a longitudinal online study of athletes that were considered in the design of later chapters.

Study 2 uses mixed methods to evaluate the quantitative and qualitative validity of online neuropsychological testing. These results are presented in this chapter in order to establish the validity of online testing in comparison to in-person methods, and to identify qualitative factors that may limit the validity of online testing which were considered in the design of the longitudinal study.

Study 3 presents findings from quantitative analyses of the first set of cognitive, mental health, behavioural and sleep data collected in the SCORES project. For the purpose of this thesis, the presentation of this data establishes if the study format can provide meaningful results that can address key gaps in knowledge within the research field, such as the feasibility of sex difference comparisons.

Study 4 assesses the concurrent and retrospective acceptability of participation in the online longitudinal study from the perspective of participants two and a half years into the study. This study was designed to build upon findings from Study 1, to examine how perceived effectiveness is concurrent with initial perceptions of acceptability. In this thesis, Study 4 is positioned to further examine remaining questions about acceptability of the study format and to propose future avenues for research and development of the online longitudinal study.

Following findings from studies 1-4, it became clear that further work needed to be performed to evaluate how sex difference analyses can be facilitated in a cohort study of former athletes. To address this, a narrative literature review is presented to examine how the definitions of levels of performance within female athlete populations may limit their inclusion in cohort study research. This review was included to build upon results from Study 3, and to guide future research following this PhD project.

This thesis is concluded with a discussion of the main findings of this thesis and a critical analysis of the strengths and limitations of the methods and approaches used in the thesis. The discussion chapter further examines the key learnings from each chapter, highlights the impact of this thesis within existing knowledge, and proposes further avenues for research outside of this thesis.

Methods

Introduction

The focus of this thesis is to explore how an online prospective longitudinal study can monitor the brain health of former athletes as they age, to better understand the link between contact sport and dementia. In summary, the research aims are to investigate the acceptability and validity of a longitudinal athlete cohort study, to analyse the first dataset, and to explore how a representative dataset can be established. To meet these aims, this thesis contains four studies that use quantitative, qualitative, and mixed methods, and a narrative literature review. These methods are described in Figure 2. The purpose of this chapter is to describe the methodological approaches used in my thesis and to discuss my rational behind choosing these approaches. This chapter describes the framework of the SCORES project and where my thesis sits in relation to the SCORES project. This chapter also discusses the process of gaining ethical approval for this research and the SCORES project's brain health feedback process, which provides participants with feedback on their test performance and informed lines of inquiry for Studies 1 and 4.

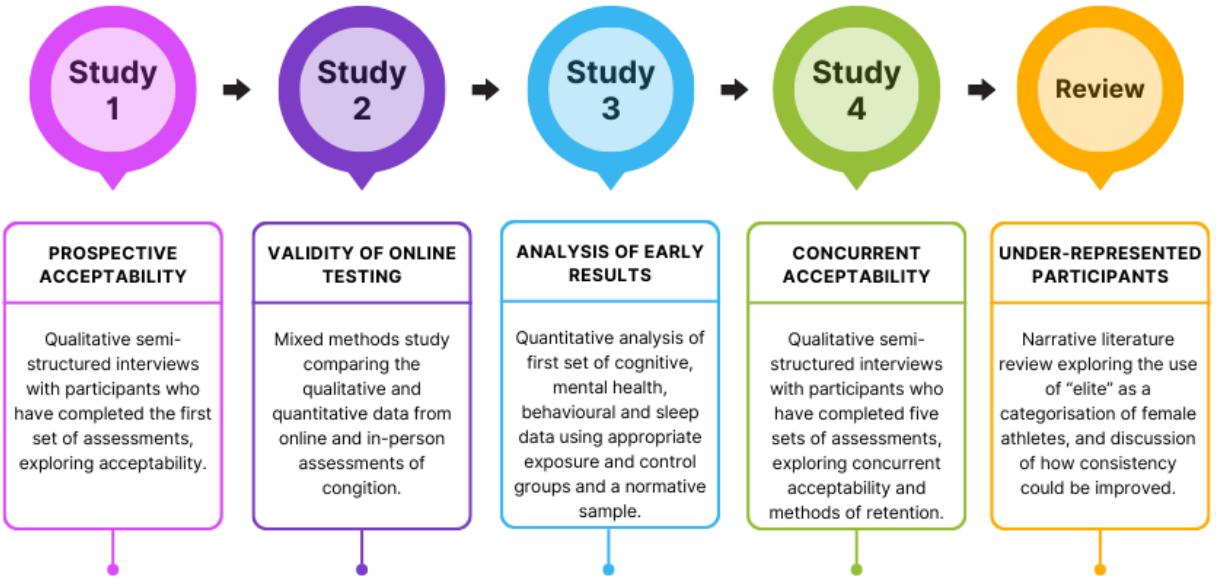


Figure 2: Summary figure of methods used in this thesis in each study.

Overarching research philosophy

The work in this thesis employs different research methods to meet the objectives of my research, including qualitative, quantitative, and mixed methods, as well as a literature review. Individually, the different designs in each study fit into different worldviews. Qualitative methods employed within Studies 1 and 4 were designed to fit within a constructivist worldview. In this paradigm, the researcher seeks to understand the participants' view of the world. For these studies it was important to understand the subjective meanings around the complexities and acceptability of participation (99). Quantitative methods used in Study 3 were designed to take a

postpositivist approach towards research, whereby researchers assess the causes that influence outcomes (99), and the relationships between exposure to contact sport and other risk factors were examined in terms of their influence on long-term brain health outcomes. The choice of mixed methods in Study 2 was intended to examine validity through a pragmatic worldview, whereby researchers emphasize the research problem and use all approaches available to understand the problem in more complex depth than qualitative or quantitative methods alone (99). Finally, the approach used in the literature review was designed to partly follow a postpositivist worldview in which literature was systematically reviewed to test the theory that research definitions vary over time, but also a transformative worldview whereby a feminist lens was employed to examine ways to improve the representativeness of research in sport, to ultimately improve the health and wellbeing of marginalised groups (99). Taken together, the variety of worldviews used to design studies for this thesis and the employment of multiple methods to navigate research problems within this field, reflect an overarching pragmatic research philosophy that guides the work in this thesis.

SCORES Project methods

The studies in this thesis were designed to assess the acceptability and validity of an online longitudinal study and examine the study's first dataset for quantitative results. The SCORES Project is an online longitudinal study of former athletes, that collects cognitive, mental health, behavioural and sleep data at regular intervals. The work in this thesis used the first dataset from the SCORES Project to answer research questions for Study 3, as well as data collected outside of the project to answer research questions in Studies 1, 2 and 4. The following section provides context for the methods used in the project and how they relate to this thesis.

Recruitment and Sampling Strategy of the SCORES Project

The project was first advertised in January 2020. The study advert invited males and females with a history of any sport at any level to register their interest for the study. Participants in the SCORES project are aged 40 and above, live in the UK or Ireland, and do not have a diagnosis of dementia. The youngest age limit of 40 was chosen in order to recruit a wide age range so that results could be compared to other studies (83), but also to monitor signs of cognitive decline across the lifespan. In designing the study, we also needed to strike a balance between the cost of the assessments and recruiting at a young enough age so that we could collect baseline data for each participant. Setting 40 as the youngest age limit was therefore intended to provide baseline data in an age group before cognitive change begins to be detectable. Our inclusion criteria originally stated that participants must live in the UK in order to limit costs and to meet our ethics requirements. We later amended our application so that participants could be recruited from Ireland, and the project team plans to expand recruitment to other nations in the future. Our

inclusion criteria also required that participants did not have a diagnosis of dementia because our objectives are to monitor the stages of brain health decline prior to a diagnosis of dementia in order to better understand the early stages of disease and improve early detection. Although the study is open to participants from any sport, the initial and subsequent study adverts were largely targeted at football participants. The study advertisement was attached to news stories about the cases of dementia in former professional footballers (100) and advertised on social media, on posters in leisure centres, and on the radio. This choice to target footballers was informed by mounting evidence that footballers may be at a higher risk of dementia than controls (53, 54, 56) and because of the interest the study team had in English professional football and the connections that they had already established in the industry.

The participants who registered their interest were invited to participate in the study when the project launched in October 2020. Study recruitment has remained open and new participants have joined the study over the past three years. Out of the total group of participants who completed the demographics questionnaire, and their first set of assessments, appropriate exposure and control groups were selected for analyses. This means that the total database includes a variety of participants in terms of sporting history and level of participation, but not all participants were included in analyses within Study 3 based on the specific research questions.

Data collection method

Upon registration to the study, participants complete detailed demographic, sporting history, health, and brain health questionnaires. Every three months, participants complete 30-minute assessment batteries including measures of cognitive, mental, behavioural and sleep health. The online battery consists of digitised cognitive assessments, as well as standardised mental health, behavioural and sleep questionnaires. Outcomes collected in the demographics and cognitive assessments are summarised in Figure 3.

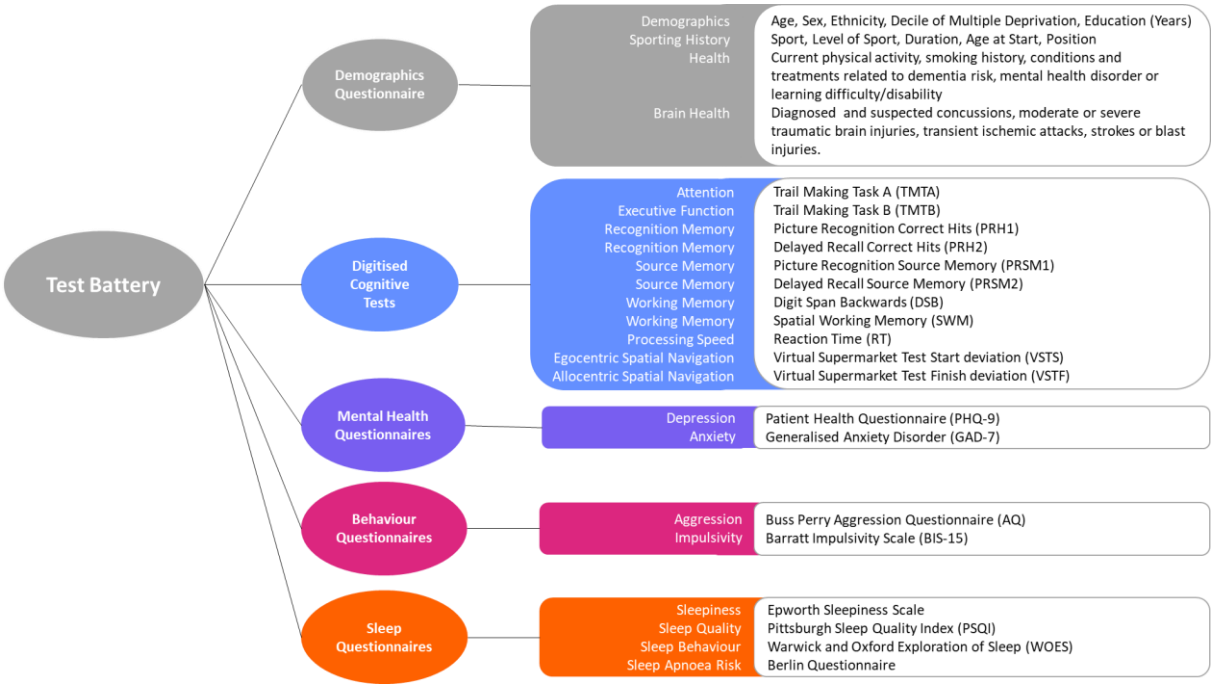


Figure 3: Summary of outcomes collected in the SCORES project.

Demographics questionnaires are completed once on entry to the study. The cognitive, mental health, behavioural and sleep assessments are split between two testing batteries to reduce time burden. Each battery takes roughly 30 minutes to complete. Participants alternate between the two testing batteries, to repeat the same assessments twice per year as described in Figure 4. The time interval of 6 months between repetitions of each assessment was chosen to capture regular intervals of data but allow time between each testing session to reduce practice effects. The influence of practice effects is also monitored by testing test-retest reliability in Study 3, and will continue to be monitored as the SCORES project progresses.

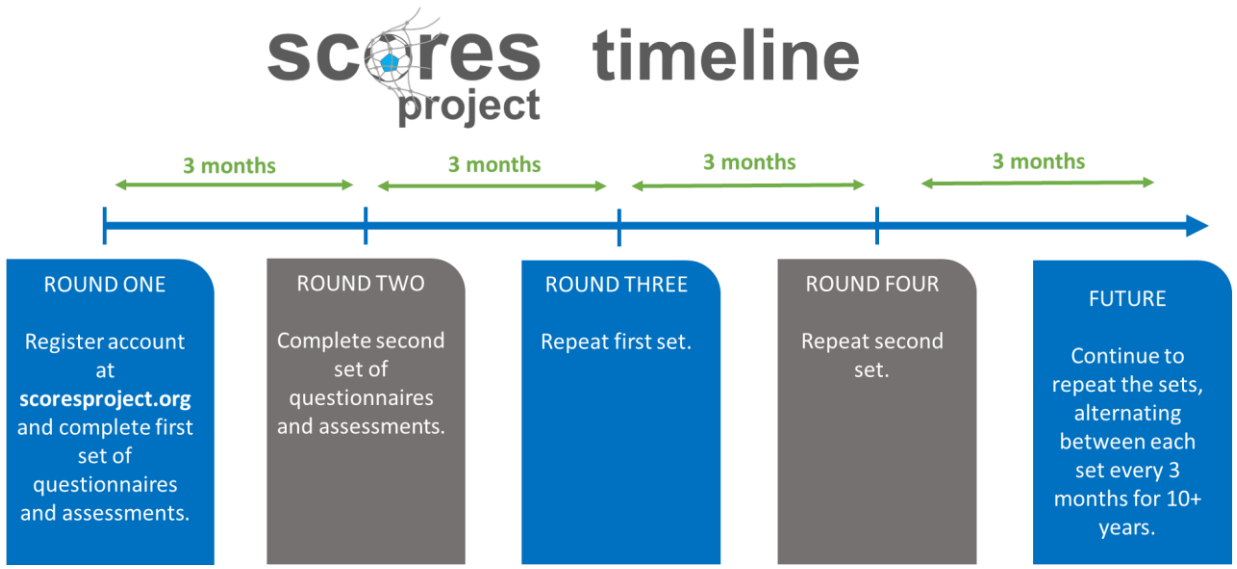


Figure 4: Timeline of participation in the SCORES project

The project currently exists entirely online. Participants complete assessments remotely using their own device (compatible with a laptop, computer, or tablet but not a smartphone). Participants are encouraged to complete their assessments as soon as possible at a time that suits them.

The battery of cognitive assessments in the project includes a set of gold-standard neuropsychology tests that have been digitised by Neuropsychology Online (NeurOn, www.neuropsychology.online/tests). Further details about the assessment battery are described in Study 3. Following each assessment, participants were invited to report whether or not they experienced any distractions, interruptions, or technical issues within their assessments. When an event was reported, the participant was invited to repeat the assessment and only the result of the repetition was included in the dataset. Participants were also asked to repeat assessments if a technical issue was detected, such as the system failing to record results or participants scoring 0 on the Digit Span Backwards task (DSB) and the Spatial Working Memory task (SWM).

Mixed methods approach

Both qualitative and quantitative methods have limitations and strengths as designs for data collection. Mixed methods approaches allow the combination of strengths from qualitative and quantitative study designs to gain deeper further insight (99). Mixed methods are used in Study 2 to assess the validity of the results from online vs. in-person cognitive testing and to compare the experiences of both testing situations from the perspective of participants. The design of Study 2 follows a largely deductive approach, whereby predetermined analyses were performed quantitatively, and predetermined topics were examined qualitatively. The study followed an explanatory sequential mixed methods design (99). This approach involves two phases of data collection in which the researcher firstly collects quantitative data. Based on these results, the researcher then plans and conducts the qualitative data collection (99). In Study 2, the

quantitative data was planned to inform the topics that are discussed in interviews, but the findings from the interviews are also intended to inform the analysis of the quantitative data.

Study 2 took place between November 2021 and June 2023. New participants to the SCORES study were invited to opt in to the validation study to replace their first testing session. Participant selection was limited to those that met the project eligibility criteria, and who lived in Norfolk to reduce travel and time burden. The study was designed so that the initial testing session was counterbalanced for whether participants completed in-person or online assessments first. Further details on the procedure of the study can be found in Study 2.

Mixed methods approaches allow the analysis of both quantitative and qualitative data separately but also in conjunction to provide further meaning to the data and create better insight. In Study 2, quantitative data was used to examine the correlation between results from online and in-person assessments. Analysis of qualitative data was performed to understand the differences and similarities between in-person and online assessments to identify any factors that may influence results. Taken together, findings from quantitative data informed the topics and themes generated from the qualitative data analysis, whilst findings from qualitative data were used to inform further analyses of the quantitative data. The findings from these analyses are presented and discussed in Study 2.

Quantitative approaches

Quantitative methods were used in Study 3 to examine results from the first round of data collection in the longitudinal online study. The first set of data was examined in order to capture the largest groups of participants and to achieve higher effect sizes. Figure 5 summarises the number of assessments completed by males with a history of contact sport, as an example, in the dataset at the time of analysis to demonstrate the number of participants needed to achieve small, medium, and high effect sizes. At the time of analysis, group numbers were too small to perform analysis of change over time and adjust for factors that might influence performance on assessments. Therefore, for the purpose of this thesis, the initial set of quantitative results were analysed to examine the long-term effects of exposure to contact sport on brain health.

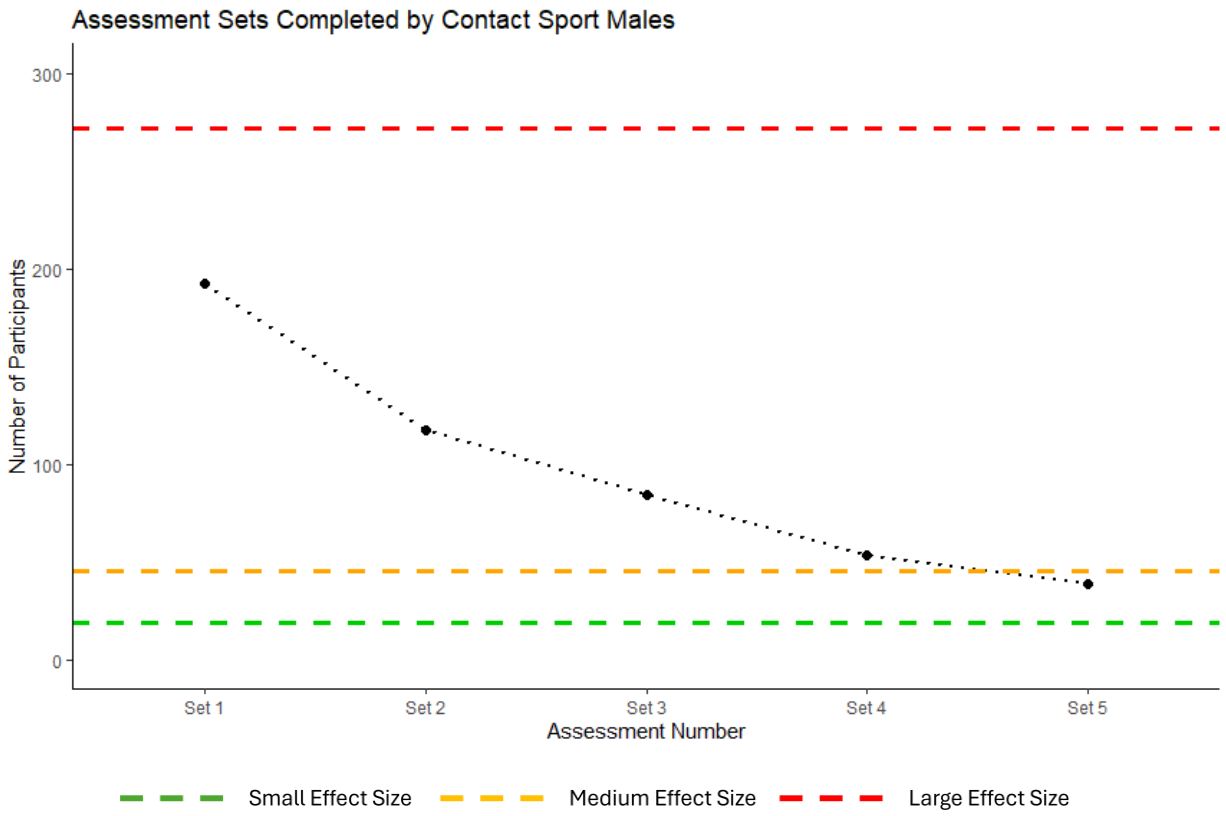


Figure 5: This plot demonstrates the number of males with a history of contact sport who had completed each set of assessments by the time of analysis.

Study 3 was performed using an experimental design, where appropriate exposure and control groups were identified and used to compare outcomes and control for confounding factors. In this study, a cross-section of the data from a longitudinal study was used to analyse data from the first test battery within the study. Control and exposure groups were chosen based on appropriate sample sizes that aimed to isolate the effects of contact sport, or to stratify exposure (for example amateur vs. professional levels of sport).

Participants' raw data were entered into a secure datasheet that was connected to their identifiable information with a unique participant identification code. The data sheet was accessible by only the study team. The demographic information was entered by me and a research administrator. The cognitive, mental health, behavioural and sleep data were only entered by me. The coding of questionnaires was also performed solely by me to avoid discrepancies in how information was coded. To analyse the raw data for Study 3, the relevant data columns were copied into a study datasheet and cleaned for any missing data or results where participants had not yet repeated an assessment. These results were coded to remove them from analysis.

All data analysis was performed in R. The purpose of the analysis was to make appropriate comparisons between exposure and control groups based on the participant groups that were available in the dataset. Comparison selection was based on samples that met minimum sample

size calculations. These were calculated using G*Power and are summarised in Figure 5. The specific analyses performed are also described in Study 3.

Qualitative approaches

Qualitative methods are used in Studies 1 and 4 to explore the prospective, concurrent, and retrospective acceptability of participation in a longitudinal online study. Qualitative approaches towards data collection and analysis aim to understand the meanings that people attach to experiences of the world, and how people make sense of that world (101). A qualitative approach therefore was considered appropriate because the aim of these studies is to understand participants' perceptions of the acceptability and expectations of the study. These studies were designed to use participants' subjective understanding and meaning associated with participation to build an understanding of participants' views. Both Studies 1 and 4 used an acceptability framework for healthcare interventions to inform the design (102). The acceptability studies collected cross-sectional data from participants after one set of assessments (Study 1) and after five sets of assessments (Study 4).

In both studies, participants were recruited from the SCORES study sample. Study 1 was advertised to participants who had completed their first set of assessments as an optional follow-up interview to discuss their thoughts about the study and their opinions on how to give feedback to participants going forward. Out of those that registered their interest, a purposeful sample of eight participants were selected who were varied in sporting history, sex, and age. Study 4 was advertised to participants who had completed five sets of assessments as an optional interview to discuss their experience in the study to date, and to discuss improvements for retention. From participants that registered their interest, a purposeful sample of seven participants were invited to be interviewed who varied in sex, age, level of education, and sporting background.

In both studies, interviews took place over a telephone call or online video call depending on participants' preferences. This method was chosen to reduce geographic constraints or travel burdens to allow recruitment of balanced sample groups. Remote interviewing does raise methodological questions, given that the participants needed to have good computer literacy to use video calls or good service coverage to participate effectively in phone interviews. Previous research assessing the use of online video conferencing calls finds that this method of data collection is generally acceptable to participants, although technical challenges can limit use and internet connection challenges can interrupt flow of interviews (103, 104). However, computer literacy is a prerequisite for participating in the wider project and participants were offered detailed instructions and support for joining video calls, and the connection for phone calls was typically of high quality.

Reflexivity

Reflexivity refers to the awareness of the ways in which characteristics of the researcher and research process may have shaped data (101). These characteristics include the role of prior knowledge, assumptions, and experience which can influence findings (101). To engage in reflexive practice in using qualitative methods I considered my position as a researcher during the design and data collection, and analysis in Studies 1, 3 and 4. Participants that I interviewed had public access to my profile on the SCORES website and researcher profile at my university. Information available to them included my academic background as well as my experience as an amateur football player and coach.

Key characteristics about myself that may influence findings in the qualitative studies in this thesis include that I have a background in football as an amateur player and qualified football coach. This allows me to build rapport with participants in these studies as an insider. Additionally, the study cohort includes former professional athletes and therefore my football background influence rapport building with elite participants as well. Other characteristics that may influence findings and are considered in analyses include that I am white, female and in my mid-20s.

Additionally, given that studies 1, 2 and 4 evaluate the long-term study, it is important to consider how my role as a researcher working on this study may influence discussions of the acceptability and validity of the project. To avoid response bias, participants in these studies are encouraged that their responses will help to improve the quality of the study for themselves and other participants.

Literature review

An objective of this thesis is to explore methods of promoting and improving the validity of including female athletes in longitudinal cohort studies and to widen the representation of the SCORES dataset. This reflects a pressing need in this body of research to better include female athletes in study samples (84). In Study 3, this thesis identifies that a barrier to this may be the way that researchers define elite female athletes, and that a uniform definition framework is necessary. To develop this framework, I chose to perform a narrative literature review to understand the variation in definitions of elite female athletes within sports science research, and to consider how better definitions may improve inclusion in cohort studies. A narrative literature review using systematic methods to identify sources (105) was chosen as the most appropriate approach, in order to capture as much literature as possible and to map the definitions over time.

The approach chosen for this literature review was to map and assess the use of the term ‘elite’ to define female athletes within research, using women’s football research in the UK as an example. This search was also designed to capture literature that discussed the historical development of women’s football in the UK, to identify key milestones that could be matched to the mapped

definitions of elite. A definition framework was then developed based on these milestones, that accounts for developments in the game. The purpose of choosing a literature review was to report a replicable method of how I came to my conclusions for the proposed definition framework that could be used by other researchers in the context of other sports or countries.

Process for obtaining institutional ethical approval

This section explains the process of the institutional ethics application from the University of East Anglia's Faculty of Medicine and Health Sciences Ethics committee.

Patient and Public Involvement (PPI) Work

Prior to applying for ethical approval for the research in this thesis I conducted extensive Patient and Public Involvement (PPI) work. PPI refers to the consultation and inclusion of patients, the public, participants, service users or clients in the design, implementation, analysis, or dissemination of research. The inclusion of PPI work in research allows better quality research that is relevant to participants and meets their expectations (106, 107). PPI work was considered especially important in the design of the research included in this thesis, because of the need for accessibility of online testing and for the study design and feedback forms to meet participants' expectations of ethicality.

The PPI group was made up of former professional footballers and family members of former professional footballers. The PPI group were consulted on the design of the online testing platform to ensure that information and instructions were readable and comprehensible. The group tested the initial assessment battery and informed the choice of measures of mental health and behaviour. Furthermore, the group made considerable contributions to the design of the feedback process, for example by informing the design of the figures and the language used in describing how to interpret assessment performance.

Application Process

The application was approved in August 2020 (Reference: 2019/20-143) to collect demographic and cognitive data for Study 3, and to perform interviews for Study 1. An amendment was accepted in June 2021 (Reference: 2020/21-098) for the addition of the validation study, inclusion of mental health and behaviour questionnaires, and for the inclusion of a feedback process designed to give participants the option of receiving feedback on their brain health. This was a key objective of the project, for participants to be able to monitor their own performance on assessments over time, and to provide resources and signposting to all participants, but especially those that demonstrated signs of worsening brain health.

In addition to consultation with the PPI group, the acceptability of providing feedback was investigated in Study 1, where a more in-depth commentary can be found voicing participants' expectations and perception of anticipatory acceptability. Further consultation with PPI and

findings from Study 1 informed my application for ethical approval to provide feedback, where a format for the feedback was approved. The ethical approval therefore included a process that met both the ethics committee's standards and the expectations of participants within the SCORES cohort.

One feature of the SCORES project is that participants can opt in to receive a feedback form describing their performance on the cognitive tasks after 12 months of participation. This time frame was chosen to allow the accumulation of a cohort large enough to make meaningful comparisons, and to include three sets of data to begin to identify any inconsistencies. The feedback forms were sent to a nominated email. An example of these feedback forms can be found in Appendix A. The feedback forms included a description of how to read and interpret the feedback. Results were presented as grouped performance by domain on executive function, memory, reaction time and spatial navigation assessments. The forms also included signposting to resources related to brain health.

Performance on mental health measures was immediately provided to participants on the completion of the mental health questionnaires. This approach was chosen to give participants real-time feedback and resources that might be more immediately relevant to participants. The decision to provide immediate mental health feedback was informed by findings from Study 1. The acceptability of this feedback procedure is reassessed and discussed in Study 4.

A second amendment was accepted in December 2021 to allow the inclusion of sleep questionnaires in the study battery for Study 3. A third amendment was accepted in February 2023 to include interviews for Study 4.

Concluding summary

The methods included in this thesis were designed to address my research questions using qualitative, quantitative, and mixed methods approaches, as well as a narrative literature review. The combination of these methods allowed me to meet my research aims with a balanced approach and to uncover further understanding about the acceptability, validity, first dataset and methods to improve representation in an ongoing online longitudinal study of athlete brain health. The procedure to gain ethical approval for the methods used in this thesis lead to a further understanding of improving the ethicality of online longitudinal health research and provides guidance for designing a participant performance feedback process that meets the expectations of both a university ethics committee and participants within the study.

Study 1 – Monitoring the brain health of former athletes online: An acceptability study

Abstract

Longitudinal study success is dependent on recruitment and retention. Both are influenced by the perceived acceptability of participation, and therefore it is important to understand participants' perceptions and expectations of acceptability. The longitudinal study evaluated in this thesis also offers regular feedback on assessment performance to participants, which raises ethical issues surrounding duty of care. The study in this chapter was designed to explore acceptability of the SCORES project to improve recruitment, retention and ensure participants expectations of ethicality are met.

Eight participants were purposively sampled and interviewed with a topic guide based on Sekhon's theoretical framework for assessing acceptability. Responses were analysed deductively against this framework.

This study found that promoting altruistic and personal benefits of participation could aid recruitment. Conversely, computer literacy and the possibility of discovering a decline in their brain health was a potential barrier. Participants identified clarity of instructions; regular non-intrusive researcher contact; low assessment burden; emphasis on participation as voluntary and the promotion of a community as avenues towards improving retention. They identified assessment frustration and challenging assessments as possible reasons for attrition. Participants viewed feedback as both necessary and important, and made suggestions for ensuring ethicality.

Findings from this study demonstrate how longitudinal online studies of athletes can be improved to aid recruitment, retention, and ethicality.

Introduction

Dementia, and its cognitive, mental, and behavioural symptoms are progressive and so investigations that seek to understand the progression of these symptoms need to repeatedly measure outcomes. However, many of the investigations in this field only present single measurements of outcomes; one study of the long-term effects on mental health took two measures of depression within a nine-year interval (108) and another took three measures with six-month intervals (109). Currently, the long-term effects of exposure to RHIs on brain health with measurements at regular intervals remains largely unexplored. Retrospective studies (53, 54), which use health records to identify increased risk of neurodegenerative disease in former professional contact sport athletes, can only offer so much information about long-term effects on brain health. It is now necessary for prospective studies to enable us to understand the long-term effects of exposure to sport related RHI over the lifespan.

Whilst longitudinal studies offer valuable information into change over time, cause-and-effect relationships, and the sequence of outcomes, they are also expensive, take time to produce meaningful results, and are time-consuming for both researchers and participants. Longitudinal studies, where large group sizes are needed to monitor group differences and changes over time, rely on the successful recruitment, engagement, and retention of participants in the study. Exploring acceptability is therefore vital within a longitudinal study to promote avenues for improving the study from a participant's point of view, to improve recruitment and engagement, and reduce attrition rates.

Sekhon defines acceptability as a multifaceted construct that reflects the extent to which people delivering or receiving a healthcare intervention consider it to be appropriate, based on anticipated or experiential cognitive and emotional responses to the intervention (102). To summarise the concept, they propose the theoretical framework of acceptability (TFA) which is presented in Figure 6.

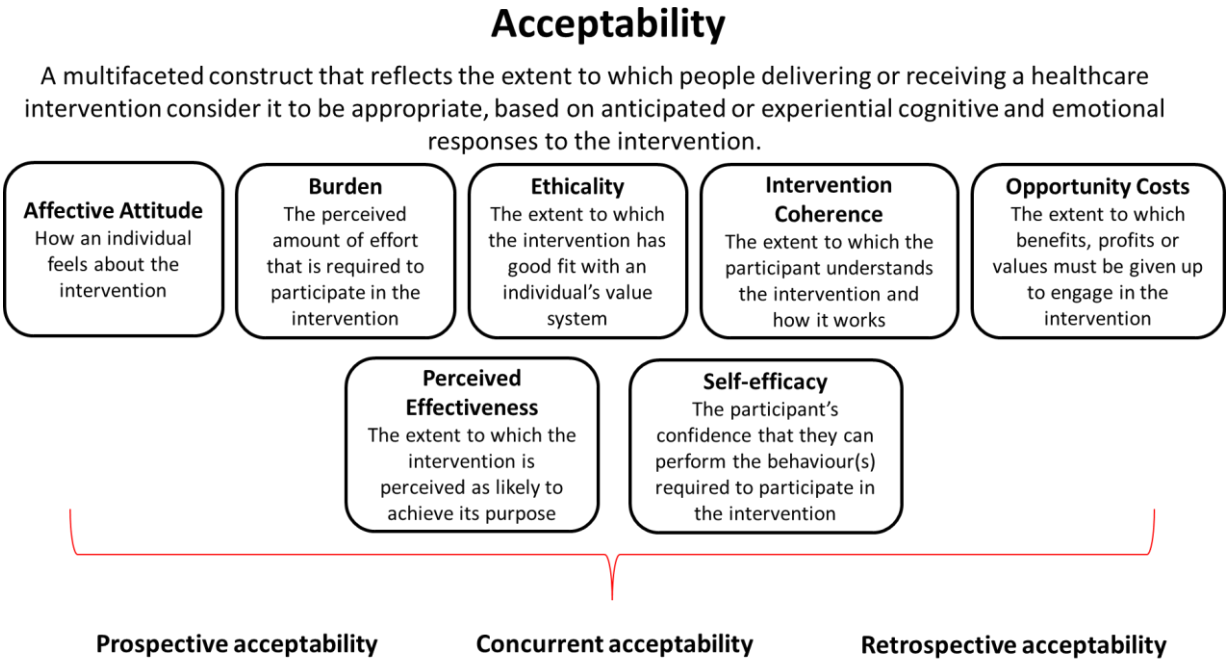


Figure 6: Sekhon et al.'s proposed acceptability framework including seven constructs of acceptability. Constructs can be used to assess acceptability at three time points. Sourced from Sekhon et al., 2017.

Whilst the acceptability framework is typically applied to the assessment of healthcare interventions, previous studies have also used it to assess the acceptability of participation in research, particularly in longitudinal studies. An example of the application of this framework is in the Canadian Longitudinal Study on Ageing (CLSA), a population-based study in Canada that monitors the health of approximately 50,000 participants aged between 45 and 85 as they age (110). In this study, researchers interviewed participants about their experiences in the study, applying the TFA to the design of their interview questions. Participants in the CLSA described factors that may affect the acceptability of participation in the study, such as concerns about the use of their personal data, and the researchers reported that they adapted the study design to meet their expectations. An addition of an acceptability study to this thesis was therefore considered important to understand acceptability from the SCORES participants' perspectives to encourage long-term participation in this study and also develop a method that could have wider applicability.

In developing the SCORES project, we consulted a patient and public involvement (PPI) group made up of former professional footballers, a football manager, and a relative of a former professional footballer with lived experience of dementia. Discussions with this PPI group provided many practical suggestions, for example changing the terminology of "tests" to "assessments", ways to increase participant comfort and highlighted the need for further qualitative work preceding the addition of mental health and behavioural questionnaires and the design of the feedback process. Based on this, an acceptability study was therefore appropriate to understand the expectations and experiences of participants within the study to improve the SCORES protocol, promote recruitment and retention, and to inform the development of the

feedback forms. We were particularly interested in assessing acceptability in terms of motivation to participate, participant experience, feasibility and acceptability of SCORES, procedures for receiving feedback about cognitive health, and the acceptability of the proposed mood and behaviour questionnaires.

This study was designed to investigate the acceptability of the online testing system for assessing the long-term effects of repetitive head injury in former athletes. In particular, the study was designed to explore the TFA constructs of affective attitude, burden, perceived effectiveness, ethicality, participation coherence, opportunity costs and self-efficacy of participation in the SCORES project. This acceptability study took place in the first three months after participants completed their first set of assessments. We were interested in assessing acceptability at this stage to gain understanding about their perceptions of participation to help recruitment, and their early experiences of the online testing system. Findings from this study were intended to improve the design and procedure of the SCORES project.

Methods

Participants in the SCORES project are invited to complete a set of cognitive, mental health and behavioural assessments every three months for at least 10 years. Following each set of assessments participants are invited to provide optional feedback about the study in a short questionnaire, which informed the design of the interview topics. All participants who had completed their first set of online assessments (typically three months into participation) received information about this qualitative acceptability study and were invited to register their interest in taking part in an interview.

Semi-structured interviews were considered to be the most appropriate form of data collection because they allowed the researcher to guide discussion through a range of topics related to the design of the SCORES protocol and also allowed participants to lead discussion about their experience of the project and perception of acceptability (111). Sekhon's TFA (102) informed the development of the acceptability questions in the topic guide around prospective, concurrent and retrospective acceptability of affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness and self-efficacy or participation. For example, questions included "What prompted you to get involved as a participant in the project?" and "How do you feel about receiving feedback about your brain health?". The topic guide is appended (Appendix B). Participants had to be taking part in the SCORES project and therefore met the project inclusion criteria: aged over 40, lived in the UK at time of data collection and did not have a diagnosis of dementia. A purposive sample (n=8) was chosen, from those who expressed interest, to provide a balance of age, sex, and sporting history and all consented to take

part (see Table 2). Purposive sampling by ethnicity was limited because all participants that registered their interest were white British.

Table 2: Demographic characteristics of participants who were interviewed as part of this study. Sporting history was gathered from self-reports of the three highest-level sports that they played over their lifespan.

Participant Number	Age	Sex	Sporting History
1	59	Male	Amateur Football
2	51	Male	Professional Football
3	56	Male	Amateur Football, Recreational Cycling, Recreational Snowsports
4	86	Female	Recreational Swimming
5	59	Female	Recreational Hockey and Recreational Netball
6	79	Male	Amateur Football, Amateur Athletics, Amateur Golf
7	59	Male	Amateur Football
8	75	Male	No Sporting History

I conducted all the interviews online using Zoom. The first interview was used as a pilot interview. Following this interview, a discussion was held with my primary supervisor (SH) to reflect on the interview technique and responses. This reflection led to changing the first question of the interview to ask about what prompted participants to get involved in the research, in order to understand their general motivation and understanding of what the study would involve. Additionally, suggested prompts were added to the interview guide (see Appendix 1) to improve flow of the interview. Social biases were accounted for by encouraging participants to answer openly and honestly because responses would be used to improve the study and feedback process and regular discussions with SH during the design and data collection stages.

Data Management and Analysis

Interviews were audio recorded and saved to a secure OneDrive file and given an anonymous code. This study was designed to answer specific questions related to the design of the SCORES project and acceptability, and therefore the purpose of the analysis was to organise the responses to understand participants feelings towards participation and their understanding of the study as they began participation. Recordings were transcribed by myself and then analysed deductively against predetermined themes based on the TFA (102). To do this, responses were coded as relating to any of the factors of the TFA (affective attitude, burden, ethicality, intervention coherence, opportunity costs, perceived effectiveness, and self-efficacy of participation). Doing so allowed me to breakdown the feedback from participants into key aspects of acceptability. Once the response were categorised by factor, three key topics became apparent which were the perceptions of acceptability of recruitment, retention and ethicality of the feedback process. The coded responses were then integrated within each topic to develop the narrative of participants' collective responses. The findings that follow are a description of these results.

Findings

Following our deductive analysis, based on the TFA, responses from participants were organised into three topics which are recruitment, retention, and the ethicality of the feedback process.

Recruitment

To understand what attracted participants to volunteer their time, participants were asked to describe their original motivation. Some described an awareness of their own brain health as a motivator.

> *"Having played for most of my life… as I saw the information I thought well, it's worth knowing. My father passed away three or four years ago as well and he was sort of a keen amateur footballer, and he had dementia in the end. So, fairly close to my heart." (P1)*

> *"My view is that if we don't volunteer, we can't do this work, and we can't all benefit. I'm of an age now, over 50, where stuff that's going to happen in the next 20, 30 years will probably affect my health. That's a slightly selfish aspect of it." (P3)*

Interview volunteers described the opportunity to receive feedback on their brain health as a key motivation because of their history of sport or concussions, especially following high-profile coverage of former athlete health.

> *"I've sort of been interested with the link to dementia particularly after Jeff Astle and his death which was quite a while ago now. And there seems to be quite a few suffering from this. So, from my own personal perspective I am sort of interested in whether there were any sort of early warning signs or anything like that which might point to something for me." (P2)*

> *"I used to play amateur football and in my position as centre forward I used to head the ball a lot… But also, three and a half years ago I had a pretty bad cycling accident where I went through the back of a car, the back window, and I knocked myself out. I had a severe concussion. …So that concussion element of it for me was also interesting in terms of this particular project." (P3)*

Participants also described an altruistic motivation towards volunteering.

> *"To help in research and hopefully … help prevent (dementia) from happening through sports. To stop head injuries causing dementia." (P5)*

> *"It's in the interest of mankind dare I say." (P6)*

Participants identified key barriers to participation as well including affective attitude towards receiving poor feedback as well as the requirement for good computer literacy.

> *"If anyone is concerned that they have suffered any neurological injury that it might highlight it and bring it home a little bit more than putting it in the back of your mind." (P1)*

> *"Only negatives I would say are for people to partake in this, if they're not good with a computer they'll struggle." (P6)*

Retention

When asked about their expectations of the study, some participants reported having expected the testing to be more comprehensive by including a more detailed assessment of their neurological health, a physical examination or a more in-depth discussion of their sporting history.

> *"I expected it to be a lot more intense. But it was a lot easier than I thought which made me feel a bit more relaxed." (P5)*

> *"I was expecting some sort of physical examination or physical test, and especially a discussion about my sort of activity as a footballer and heading the ball etc." (P3)*

Assessments of mental health and behaviour were not originally included in the test battery to first explore their ethicality prior to their inclusion. At the time of recruitment, participants had an awareness of how the study format worked, and they could consider the acceptability of their inclusion. The participants revealed that assessment of mental health and behaviour were in line with their expectations of the study.

> *"Personally, I'm fine with it. Going into this with my eyes open I sort of anticipated that in some ways." (P1)*

For most participants, the experience of online testing was positive. They described the experience as convenient and easy.

> *"The fact that it's not invasive, it's not time consuming and with a lot of the online stuff you can do it from wherever." (P3)*

> *"The positive is that it is easy, and anyone can do it." (P5)*

However, some participants identified that the online environment and cognitive assessments could be challenging.

> *"I've not always found it easy to sign in and remember passwords or whatever, but actually doing the assessment – I quite enjoyed and I thought it was quite like a game really." (P6)*

> *"I found that the tests for me were too quick." (P7)*

In addition to this, participants suggested that the frustration associated with making mistakes in the assessments could be a negative experience in participation, which could increase attrition.

> *"The feeling of frustration when you're making mistakes, but I think the positive is that you know why you're doing it." (P3)*

When asked how retention could be encouraged in the project, participants identified that the clarity of instructions were a key contributor to finding assessments easy and accessible.

> *"I thought they were very good with the examples as well. …They were very clear and the way that they were spaced meant that I couldn't read them too*

> *quickly. I had to read one, then see the example, then have a practice and then go on." (P3)*

The regularity of contact was described as a positive aspect of participation, and participants expressed that the level of researcher contact could improve comfort in participation.

> *"The contact I have had (has) not been intrusive at all." (P2)*

> *"If you keep people in the loop and keep people up to date then they will stay with you. If you go quiet that's when people go they've lost interest in me." (P3)*

> *"Online I can think oh yeah that's (the researcher), I can call her if I need to know anything or drop her a line. The friendliness is really helpful. It really makes a difference." (P5)*

Participants also described the emphasis on participation as voluntary and the reminder that participants can opt out at any point as an important contributing factor for reducing burden and promoting participation.

> *"I think reassuring people that if they do want to leave… that's reassuring to know that it's their choice. Some people sign up for something and then after a while they think no, I don't want to do this anymore and then they feel guilty and pressured, and that you have given us the choice to leave if we want to." (P5)*

Participants also discussed the importance of developing and emphasizing a community and the scale of the project in retaining participants. An example of how this could be facilitated was suggested to be through the project newsletter:

> *"Also, some form of indication about how many people are taking part because I think again, being part of a community helps. It's not just me, there are other people taking part." (P3)*

> *"I know you're going to do a periodic newsletter as far as progress and news. I think to do that to keep the interest would be key." (P1)*

Ethicality of feedback process

All participants interviewed in this study expressed an interest in receiving feedback on their assessment performance. The feedback process was identified as an opportunity for an external assessment of their brain health, and to give them insight into their brain health and the opportunity to act.

> *"We can't really know how good or bad our brain or how well our brain health is working apart from what we can tell ourselves, but it would be always useful to hear what somebody else generally thinks about it. Especially someone who has got a certain amount of sense about what they're talking about." (P8)*

> *"Even if it was very negative news because I think that would enable me to start action." (P2)*

Although participants highlighted the benefits of receiving feedback, one reflected concern around receiving negative or potentially distressful news.

> *"If I got a poor score there, I'd be quite upset really. Not with you, but thinking I've got a problem and start to worry." (P6)*

Participants were asked to describe a format that they found acceptable for receiving cognitive results. They described a preference for emailed written reports which could be shared with their General Practitioner (GP). Signposting to resources other than the GP were also desirable.

> *"Well just an email saying you've got ill health and that you should see a medical practitioner. I don't see that you need to give huge advice on it." (P7)*

> *"If they have a letter, I can take it to the doctor and if they think I am heading towards slight dementia or something like that then the doctor can take it from there. So, it is evidence as well as information." (P5)*

> *"If there were issues, I'd like to also be sent details of where I can get help or support." (P2)*

Participants also identified that a visual representation of their performance and a reference point to compare their results to other participants who are similar to them would be helpful.

> *"A graph would be good so you can see your ups and downs." (P5)*

> *"I think it needs to be age related, because people of 80 are going to be a little slower than people of 50." (P6)*

The option to discuss results was also emphasised as important to the feedback process.

> *"I guess a combination of written report and if it's not good or it's technical the opportunity to discuss. … I guess (with) someone involved in the study. Not necessarily an academic but maybe someone with experience of what's being studied. Would they have to be medical – I'm not sure as long as I understood the results and could explain the results then that would be fine for me." (P1)*

Participants also considered the ethicality of whether or not feedback should be provided for performance on measures of mental health and behaviour. They emphasised the importance of receiving feedback, particularly about mental health.

> *"Some people don't want to admit that they have got a problem, and it sometimes takes someone else … to say you need help go and get it." (P5)*

> *"If I thought that I had a mental health problem developing, and I wasn't aware of that, I would probably be grateful if privately I was advised by somebody who knew what they were talking about. … Because people, men probably more than women, are always reluctant to get help." (P6)*

> *"I think it would be very useful for your organisation to mention to somebody if you see that they are beginning to suffer from any of these problems, because until a person knows the situation, they can't do anything about it, and if they*

> *want to do something about it then it would be useful to have that information. (P8)*

This discussion was balanced by participants revealing concerns about the outcomes of receiving feedback, particularly surrounding aggressive behaviour.

> *"I think you've also got to think about.... if someone's showing those (aggressive) tendencies to then receive something to say you're test results are showing that those traits... how that could potentially affect people they live with or socialise with." (P2)*

> *"I think if you're telling somebody that they're showing signs of depression it might be a relief that people know. …. On the anger side I don't know whether that could trigger something that you would probably want to make sure that the individual is in a safe environment." (P2)*

Discussion

This acceptability study used Sekhon's Acceptability Framework (102) to explore constructs of participant motivation, experience and perception of ethicality of a longitudinal online study of athlete brain health that provides performance feedback to understand opportunities to improve recruitment, retention and ensure ethicality.

This investigation revealed that motivations to join the study included an awareness of participants' own brain health as a consequence of playing contact sport or a family history of dementia, and the opportunity to receive feedback. Participants also expressed an altruistic motivation towards volunteering for the sake of protecting the next generation of athletes. The importance of altruistic motivations and affective attitude was also found in the acceptability study of the Canadian Longitudinal Study of Ageing (CLSA) (110). Findings from the CLSA suggest that these motivations are important benefits of the study to highlight in recruitment. However, a basic need for computer literacy and concerns around receiving feedback were identified as potential deterrents to the study. This demonstrates a limitation of online studies in a study population that is diverse in age, education, and socio-economic status, and highlights the need for strong technical support in participation. Our findings also suggest that highlighting the opportunity to receive feedback is a good avenue for recruitment, but that participants want this to be an opt-in procedure.

Benefits of participation included a low time burden and an assessment schedule that could be completed when convenient from their own home with a familiar computing device. The inclusion of mental health assessments in an online study format were deemed acceptable and were anticipated by participants given that the context of the project is to monitor brain health. Participants found cognitive assessments to be mixed in terms of ease and suggested that frustration could occur as a result. However, clarity of instructions and the opportunity to contact

researchers by email were identified as helpful in reducing frustration and improving ease of online assessment. Participants highlighted that creating a community feeling was also important to improving the long-term retention of participation in the project, and suggested that this could be promoted by disseminating information through the project newsletter. Interestingly participants also suggested that emphasising the voluntary nature of the project was important to retention, particularly that the consent form includes a statement about being able to drop out at any time, and that rather a mandatory time commitment might deter participants. After running for 2.5 years, the SCORES project has an attrition rate of approximately 20% of the total study population, which is a lower than estimates for other longitudinal studies where attrition rates range between 30% and 70% of the total study populations (112). It is yet to be determined if the long-term attrition rate is in line with other longitudinal studies, but it could be suggested that in these early stages the option to participate for as long as the participant wants, as well as other previously mentioned factors, might increase retention at this stage.

The feedback process in the project is a key benefit of participation, particularly in a study that monitors for signs of dementia where early identification of prodromal symptoms are vital to understanding the disease and for providing early interventions (113). Receiving regular feedback was important to participants, not only as a measure of external assessment but also as a motivator to look after their brain health or seek the advice of a health professional. This is congruent with findings from the CLSA, which identified that the provision of individual results could catalyse long-term involvement (110). The preferences for the format for this feedback were helpful in developing a procedure that was ethical and meaningful for participants. Participants wanted reports with peer comparisons that could be shared with a health professional or their family, the option for a discussion, and signposting to relevant resources. Participants highlighted a need for sensitivity particularly when giving feedback about mental health and behaviour to reduce distress or propagation of behaviours. These findings suggest that inclusion of a feedback procedure within a longitudinal study is acceptable and beneficial to participants, but key considerations need to be made when designing feedback procedures. Feedback provision should be considered in terms of how it might affect performance on future assessments, and the offer of discussions with team members needs to be considered in context of the scale of sample size. However, this study suggests that inclusion of a participant feedback process that meets these needs is important for recruitment and retention and has wider benefits for participants who can monitor their own brain health.

Further suggestions for reducing attrition are revealed by the participants' experience in the SCORES study. The time burden was acceptable, and the online setting was helpful to this as they were able to complete assessments at a time that suited them. However, the online format was also identified as a barrier to participation because a certain level of computer literacy is required for participation. This suggests that offering technical support and increasing accessibility of the

online environment may aid retention and recruitment, but also highlights there are potential participants for whom participation may not be feasible. Participants' appreciation of the clarity of instructions and number of examples demonstrates that these might be two avenues to improve accessibility of the online assessments. Whilst some participants appreciated the familiarity of aspects of the assessments, comparing them to games and expressed enjoyment in completing them, others expressed frustration in making mistakes.

For these people, reminders that the assessments are designed to be taxing might improve their perception of participation. Regular contact was also identified as a key method of increasing participation, but my findings suggest that this should be non-invasive. In summary, recommendations for improving retention based on my findings can be found in Table 3.

Table 3: Summary of recommendations to improve participant experience and reduce attrition based on findings from this study.

Theme	**Suggestion**
Reducing burden	30 minute assessments
	Complete assessments in own time
Increasing accessibility	Offer technical support
	Include video and text instructions
	Include practice assessments
Reducing frustration	Reminder that assessments are designed to be challenging
Promoting altruism	Sustain a community-feel
	Provide regular updates

The interviews revealed that participants expected the inclusion of mental health and behaviour questionnaires and found them acceptable to complete online. Concern around integrating these assessments along with the behavioural measures largely centred around the process of providing feedback on mental health and behaviour, rather than the inclusion of the assessments in the battery or completing them online.

The ethicality of receiving feedback was discussed by the participants in this study. The findings demonstrate that receiving regular feedback on performance in the SCORES assessments was important to the participants, not only as a measure of external assessment but also as a motivator to look after their brain health or seek the advice of a health professional. This is congruent with findings from the CLSA, which identified that the provision of individual results would likely catalyse long-term involvement (110). Answers in the interviews emphasised the importance of external assessment of mental health, with multiple references to times when other's noticing signs of poor mental health helped them to take action. They also referred to the importance of receiving feedback on signs of poor mental health, particularly in males, to reduce fear around the sensitive topic of mental health and to sign-post towards help. Participants' answers reflected that a duty of care was expected from the project, whereby information about their performance should be returned to them in exchange for participation in the project.

However, participants' responses also reflected a recognition of the distress that negative feedback might inflict and advised a caution around delivering such sensitive information. Their responses suggested a conflict between this duty of care and the need to reduce distress.

The preferences for the format for this feedback also reflected this conflict. Participants wanted to receive written reports about their brain health including comparisons to their peers that could be shared with a health professional or their family, but also recognised the value of a discussion with a project member to address questions and increase sensitivity. They expressed that the process should be non-invasive, with the chance to discuss remaining in the study as an option for them to choose rather than requested by a project member. They also demonstrated an expectation that feedback reports would include resources to help support participants in finding help. Participants distinguished between the level of necessary sensitivity required when providing results on cognitive, mental, or behavioural health. They expressed that information about their mental health required more sensitivity than information about their cognitive health, and they raised concerns around distress caused by delivering information about high aggressive behaviours.

Actions taken to improve SCORES based on this study

Following this acceptability study, we implemented measures in the SCORES project based on findings from the interviews summarised in Table 4. Many of the findings regarding methods to improve retention were already in place in the project, but we revised the existing structure to meet recommendations established by this study. Instructions preceding the online assessments were improved in clarity, particularly for the Sustained Attention to Response Task. Expected completion times for each module were added to the homepage of participants' accounts, so that they could estimate how long each battery might take them to complete. Instructions in the more difficult tests, particularly the Virtual Supermarket task, were altered to remind participants that the assessment is designed to be challenging. We also continued to publish the regular newsletter to continue to promote a community feel, including data about how many participants were included in the study and some demographic information.

Table 4: Summary of actions taken within SCORES based on findings from this study. Actions listed are those that we either weren't already doing or that improved upon our protocol.

Actions taken within SCORES
Improved assessment instructions for clarity
Expected completion times added for each module
Difficult assessments include reminder that the assessments are designed to be challenging
Addition of mental health and behaviour questionnaires
Followed NHS protocol for providing mental health feedback
Development of feedback forms for cognitive results

The most significant work that came as a result of this study was the development of the feedback forms. This investigation confirmed the need for the feedback process within the SCORES study and enabled me to propose evidence-based suggestions to improve the ethicality of providing feedback. These suggestions were ultimately adopted and integrated within the feedback procedure. The chosen format (see Appendix A) was a written report summarising the participant's performance across four cognitive domains, with comparisons to their peer group within SCORES and a normative group sample. The report includes an example of how to interpret the results and refers the participant to further resources. The report also clearly explains that the results are not diagnostic but emphasizes that if the participant is worried about their brain health, they should speak to a health professional. The report includes a link to a webpage describing the assessments that clinicians and participants can refer to, to identify which assessments were used in the project. I developed two types of cover letters, one for participants who showed no signs of results that warrant further investigation, and one for participants whose scores warranted further investigation. The latter participants were identified using the Clinically Significant Change Index and the Reliable Change Index to identify scores that were both lower than those of their peers and also demonstrated significant change over time. These are two gold-standard measures used in Clinical Psychology to determine participants who fall outside of normative ranges. Participants who scored outside of these ranges are asked to repeat the assessment in question, and if they continue to score outside of these ranges, they were sent a cover letter that explained that their scores show change that they may want to speak with a health professional about.

For feedback on mental health performance, the decision was made to follow the protocol that the NHS uses on their online self-assessments of mood (114), which use the same mental health measures as used in the SCORES project. Their approach is to provide an automated message based on the score that participant receive, with encouragement to contact their GP if they are worried about their mental health and signposting to resources. An example of this message can be found below:

> *"Based on your responses to questions 1-8, you are experiencing some symptoms seen in depression but only an experienced health professional can tell for sure. You should make an appointment to see your GP for a review or call NHS 111. Both can help you and assess whether you might benefit from extra support or treatment."*

Alternatively, given that there are few examples of providing feedback for the behavioural questionnaires, and that they are not designed as self-assessment tools in the same way that the mental health questionnaires are used by the NHS, the decision was made not to provide feedback on these assessments at this time. This decision was supported by findings from my investigation, whereby participants expressed concerned particularly about providing feedback on

aggressive behaviours in terms of both the ethicality and the consequences of sharing this information.

There are limitations to this study, including the lack of diversity in the study sample in terms of ethnicity and sex. These are both important factors in dementia risk research, (72, 87), and a more diverse sample could have improved the insight gained from this study. Recruitment to this study was limited by the recruitment within the SCORES Project, and this limitation is further considered within the discussion chapter of this thesis. Despite a more representative sample being desirable, a strength of this study was that participants were well balanced in terms of experience levels of contact sport and therefore motivations to participate as well as strategies for recruitment and retention could be explored in depth. The study sample was also well balanced in terms of age relative to the SCORES study sample (see Study 3 for SCORES demographics break down), and therefore the acceptability of an online study format could also be discussed in detail in relation to varying computer literacy levels.

Conclusions

This study aimed to explore the acceptability of participation in the SCORES project, an online longitudinal study that monitors the brain health of participants at regular intervals for at least ten years. This qualitative study found that participants find the protocol of the study to be acceptable in the early stages and makes helpful suggestions for improving participant engagement and retention. The inclusion of mental health and behaviour measures were anticipated by the participants, and they felt comfortable with their addition. Participants emphasised the importance of providing feedback, raised concerns about the ethicality of doing so, and provided suggestions for increasing the acceptability of this process. Findings from this study were used to directly improve the protocol of the SCORES project and serve as recommendations for improving participant experience and reducing attrition in future work.

Study 2 - Mixed methods evaluation of the validity of online cognitive testing

Abstract

The online assessment format is advantageous for addressing many of the key limitations and research gaps within this field of research, but the validity of using an online measure rather than in-person measures needs to be considered. Online digitised cognitive assessments have previously been shown to demonstrate improved ecological validity and suitable construct validity and may improve accessibility and participant experience. Qualitative comparisons between in-person and online assessments remains important to explore to better understand differences that may influence validity.

This study uses a mixed methods approach to examine the validity of an online test battery. This study aimed to assess quantitative validity by testing repeated-measures results from in-person and online assessments for equivalence or correlation. Participants were also interviewed to examine if qualitative experiences of in-person and online assessment differed, and if these differences may have influenced results.

Findings in this study were unable to demonstrate equivalence due to group size limitations but instead demonstrated poor to moderate correlations between in-person and online assessment results. Better performance in either testing condition was found to vary by assessment. Participant responses revealed that participants suggested that the presence of a researcher, verbal or written instructions, the ability to access technical support, and the ability to control their environment may have influenced the differences observed between in-person and online testing. Qualitative results also suggested that the order of completing the in-person condition first may have influenced their comfort of completing subsequent online tests. This informed a re-examination of quantitative results which suggested a possible effect of first testing condition on correlation.

This study concludes that there are key differences in online and in-person assessments that should be considered in designing online cognitive studies. In particular, the design of studies should consider the effects of age, socioeconomic status, and computer literacy of participants on the validity of online remote testing. This investigation also suggests that online and in-person study methods compliment each other, and whilst online methods can reach remote participants, in-person assessments may be appropriate to integrate within study designs to aid initial participation.

Introduction

In the 2023 Consensus Statement on Concussion in Sport (115), the main limiting factor that excluded studies from the review was the inability to adjust for factors that are associated with mental health and neurological outcomes of interest. However, to conduct this analysis, large sample sizes would be needed. Online assessment is advantageous over in-person assessment in that it enables a greater reach of participants, and potentially also a more diverse participant group. There is also advantage from reduced costs and time associated with participation and data collection, which benefits a longitudinal study design. These may be key advantages that enables studies to meet sample size requirements, and to be acknowledged by the consensus review.

There is increasing concern around the representativeness of volunteer based recruitment in epidemiological studies (116). In-person testing limits the recruitment of remote or non-urban participants, or those with little time to travel in to a university (117). In a 2012 review of reasons why eligible participants decline to participate in clinical trials, researchers found that 33% of participants cited inconvenience as the primary barrier to participation (118). Converting in-person neuropsychological testing to online methods could therefore directly reduce burden on participants, increase both the numbers and diversity recruited and aid long-term retention. Computerised cognitive assessment has also been reported to be a more comfortable environment for participants to complete assessments in and less distressing than traditional paper-and-pencil assessment (119). This may benefit longitudinal research where affective attitude towards participation is important to long term retention.

The online study design of the SCORES project therefore offers the potential to increase reach and sample size, increase frequency of testing, improve retention, reduce research costs, and include participants who are unable to travel frequently (120). There were also advantages of online research during the Covid-19 pandemic, where social distancing policies restricted in-person testing and research during this time period. Online research designs could comply with university social distancing policies and maintain data collection during the course of unprecedented changes. Therefore, an online study design offers advantages over in-person testing. Importantly however, the validity of assessments must be unchanged.

Previous quantitative studies have found acceptable construct validity on online assessment and similar ecological validity to in-person assessment (117). A study comparing results from a neuropsychological test battery administered in-person and online demonstrated the reliability and construct validity of online neuropsychological assessment (121). Findings in this study emphasised that ambulatory neuropsychological assessment could improve the ecological validity of assessment, by administering assessments in participants' 'natural' environments. The use of digitised cognitive assessments in research about the pre-clinical stages of Alzheimer's disease

was evaluated by a 2021 review. This review reported emerging evidence that digitised neuropsychological assessments could be used to identify subtle cognitive decline, as found in prodromal dementia, with better sensitivity and reliability than traditional in-person assessment (122). In addition to this benefit, the review reported that digitised assessments could meet the accessibility needs of a wider range of participants in dementia research. The quantitative validity of the cognitive assessments used within the SCORES project was also assessed by another group who established criterion validity in comparison to the Montreal Cognitive Assessment (MoCA) in a population of healthy older people (123). What remains to be explored further is the qualitative experience of completing online assessment vs. in-person assessment, and what affect this might have on validity.

The potential limitations on validity of online, self-administered and unsupervised cognitive testing were outlined by a joint position paper (124) by the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology including examiner contact, testing environment and workstation. In traditional in-person testing the researcher is typically present to deliver instructions and may or may not remain present during the task. An example of this is the MoCA, where the examiner asks participants a series of questions and instructs them to complete short tasks. The examiner is typically present to collect responses and observe completion of the assessments (such as the short trail making task) and is therefore available to clarify vocabulary and repeat and explain instructions. However, the counter argument is that the presence of a researcher may lead to social desirability bias and could make the participant uncomfortable when they might prefer to interpret instructions and navigate instructions by themselves. SCORES participants were encouraged to contact the research team if questions arose. However, this represents a potential barrier for participation depending on their comfort with asking for online support which may affect validity of the findings.

The testing environment in online assessments offers more variability in terms of opportunities for distraction and interruption of the assessments. Although participants report more comfort in completing assessments from home (119), the environment is uncontrolled by the researcher and therefore the validity of the results could be affected. Traditionally, neuropsychological assessments in research are administered in quiet environments with minimal distractions to optimally challenge cognitive performance. However, these environments may also increase test anxiety and discomfort which could influence how well cognitive performance in these situations reflects cognitive performance in real-world environments. SCORES participants were encouraged to access the assessments in a quiet environment with minimal distractions, and to report any distractions or interruptions that might have influenced performance. However, this has not been explored and represents an area which needs to be better understood so that we understand the potential to affect validity.

Workstation refers to the device and equipment that participants use to access the online neuropsychology battery. The workstation in in-person testing can be set by researchers, and can be consistent in terms of processing speed, internet connection and manual operation. However, in remote self-administered online assessments the workstation that the participants use can vary in terms of whether they use a touch screen or keyboard, whether they use a handheld mouse or touchpad, screen size, internet connections, processing speeds or quality of visuals. These factors are important in assessments where milliseconds can differentiate between how well participants perform on an assessment and are therefore an important consideration in assessing results from the SCORES project.

To explore the validity of the online cognitive assessments used in the SCORES project battery, I compared the quantitative outcomes and qualitative experiences from online and in-person assessment sessions. In this mixed methods study, my first objective was to identify if results from online and in-person cognitive assessments are statistically equivalent. My second objective was to compare the experiences of online and in-person assessments for elements that might influence validity.

Methods

Design

The purpose of this validation study was to compare the quantitative and qualitative outputs from online and in-person assessments. An explanatory concurrent mixed methods design (99) was used whereby both quantitative and qualitative data were collected and analysed simultaneously. Quantitative data consisted of results from a set of both online and in-person cognitive assessments, and qualitative data were collected using semi-structured follow-up interviews where participants were asked to compare testing experiences. Both datasets were collected simultaneously and initially analysed separately. Following this, results from the quantitative analysis were used to inform analysis of qualitative data, and results from the interviews were used to inform additional analyses on the quantitative data.

This study was advertised as an optional addition to the SCORES project for participants that lived in Norfolk and were able to come into the university for in-person testing. The study was advertised on local radio, social media, posters in local leisure centres, through local walking football clubs, and by word of mouth. Study advertisements stated that participants would be offered a £25 retail voucher in return for their time. Interested participants were asked to register their interest on the SCORES entry form, and if eligible they were then contacted with further participant information sheets and invited to sign a consent form. As in the SCORES project, participants were eligible to participate if they were aged above 40 years and did not have a diagnosis of dementia. However, unlike the SCORES project where participants must live in the UK

or Ireland, participants were only eligible for this study if they lived in Norfolk to reduce travel costs and burden for participants.

A sample size calculated was performed using G*Power. At least 23 participants needed to be recruited to the study to have statistical power, based on a high effect size of d = 0.8. The effect size is justified by a similar investigation (125) that found high effect sizes in correlations between digitised tests and their traditional pen-and-paper counterparts. To ensure that the minimum sample size was met, and that participants were diverse by sex and age, we aimed to recruit 30 participants.

Procedure

Participants who registered their interest for the study completed the online demographics questionnaire as described in the methods chapter. This questionnaire collects information about age, sex, ethnicity, level of deprivation (collected using postcodes) and sporting history. These demographics were used to inform the selection of a balanced sample and are reported below. Based on their answers, participants who met the eligibility criteria were invited to take part in this optional validation study.

Participants completed two testing sessions – an in-person and online testing session using the same battery of cognitive tests. The order of testing condition was counterbalanced for each participant to account for learning effect: Group A completed online testing first and paper-based testing second, Group B completed paper-based testing first and online testing second. A time interval of four weeks was chosen based on methods used in previous investigations of online and in-person testing (125). The test battery is described in Table 5. Assessment details are described further in Study 3.

Table 5: Assessments included in the validation study battery.

Assessment	Estimated testing time
Trail Making Task A	3 minutes
Picture Recognition	5 minutes
Trail Making Task B	3 minutes
Picture Recogntion - Delayed Recall	5 minutes
Simple Reaction Task	2 minutes

In the online testing session, participants completed the assessment battery on their personal computing device (laptop, computer, or tablet). In the in-person testing session, participants completed all assessments except for the trail making tasks on a laptop computer provided by the researcher. The trail making tasks were completed in person using traditional paper and pencil versions of the task.

Level of instruction was controlled for by ensuring that participants were offered the same level of instruction in person as they are online as part of the SCORES project. Online, participants were

shown an instructional video and they were given the opportunity to ask questions by emailing us. Thus, in person participants were read the same instructions as are in the videos and were given the chance to ask questions in-person.

Following the second incident of testing, participants were invited to an interview with the researcher to discuss their experience using the online and in-person testing. The interview outline is appended (Appendix C). This short interview was semi-structured and explored how their experience of completing the tests may have influenced their results in either testing session. In particular, this interview explored factors that have been suggested (124) to specifically influence results in either condition (90) including examiner contact, testing environment and workstation. The interview took place over a Zoom call or a telephone call depending on participants' preferences. Typically interviews lasted approximately 30 minutes. The interviews were recorded and stored in a secure OneDrive file.

Analysis

Quantitative results were analysed using R. Online results were plotted against in-person results to visualise the data. If sample size was sufficient and data was normally distributed, a two one-sided t-test would have been performed to assess for equivalence. As sample size was not sufficient (see participant results), an absolute agreement two-way mixed effects intraclass correlation coefficient (ICC) was calculated to identify concurrent validity. This model of ICC was chosen because the quantitative research question asks for the difference between online and in-person conditions (absolute agreement) using the mean of multiple raters (mean of k) who were the only raters of interest (two way mixed effects) (126) .

Qualitative results from the interviews were analysed using framework analysis (127). This type of content analysis is a form of thematic analysis where a matrix is employed to group summaries of data, which provides a structure to analysing and interpreting findings in qualitative studies. It is widely used in health research, and is growing in popularity within research designs with large data sets (127). However, the benefits of framework analysis also include it's use in multidisciplinary research teams (such as within the SCORES project) and where obtaining a descriptive overview of the data is required, such as within this study of factors that may influence validity. The use of framework analysis in this study was also chosen in order to examine the intersecting experiences of individual participants as they related to other participants' responses, as well as in relation to their own answers which was facilitated by the analysis matrix. Limitations of this method include the application of the analysis to highly heterogenous data, where the approach requires data from similar topics or themes in order to categorise the data effectively (127). However, this approach is effective in semi-structured designs where interview topics are predetermined and remains within a particular scope. Overall, use of this method was appropriate for answering my research questions within Study 2, where participants were guided by a semi-structured interview which asked them to reflect on their participation in the project.

Interviews were listened to by me and key quotes were transcribed. I then read the extracted quotes to familiarise with the interviews. Quotes were then coded based on the research questions and findings from the quantitative analysis (researcher presence, testing environment, workstation and preference). Based on the first three transcripts, a working analytical framework was developed and discussed with SH. This consultation was performed in order to sense-check with an experienced qualitative researcher to aid in my development as a qualitative researcher. We chose to discuss only the first three transcripts because all 15 interviews followed a similar structure and even after reviewing all interviews, the framework based on the first three transcripts was appropriate for the remaining interviews. The analytical framework was organised by the four main research question themes (researcher presence, testing environment, workstation and preference) and sub-themes were identified within these groups (i.e. within researcher presence, the sub-themes were the feeling of being observed, the quality of instructions, the availability of support). The analytical framework was then applied to all quotes and the data was charted into a framework matrix. Results from this framework matrix were then interpreted and are described in the findings.

Results

Participants

In total, 31 participants registered interest for the study. Of these, seven were ineligible due to location, eight did not complete the first testing session, and one completed the first session but not the second. In total, 15 participants completed both testing sessions and the final interview.

Demographic characteristics of the 15 participants are described in Table 6. The sample had a mean age of 65.5 (SD = 10.2) and included 3 females and 12 males. The sample had a mean decile of multiple deprivation of 5.8 (SD = 2.5), where 10 indicates the least deprived areas of the country based on government indices of deprivation. The online condition was the first assessment for 8 participants, whilst the in-person condition was the first assessment for 7 participants.

Table 6: Demographic characteristics of participants included in the validation study.

Participant Number	Age	Ethnicity (as on UK Census)	Decile of Multiple Deprivation	Sporting History	First testing condition
1	42	White (English/Welsh/Scottish/ Northern Irish/British)	8	No reported sporting history	Online
2	54	White (English/Welsh/Scottish/ Northern Irish/British)	10	Amateur football Amateur rugby	In person
3	82	White (English/Welsh/Scottish/ Northern Irish/British)	2	Recreational tennis Recreational badminton	In person
4	66	White (English/Welsh/Scottish/ Northern Irish/British)	5	Recreational football Recreational athletics	Online
5	65	White (English/Welsh/Scottish/ Northern Irish/British)	4	Amateur football Amateur cycling Amateur golf	In person
6	63	White (English/Welsh/Scottish/ Northern Irish/British)	4	No reported sporting history	Online
7	66	White (English/Welsh/Scottish/ Northern Irish/British)	10	Amateur football	In person
8	62	White (English/Welsh/Scottish/ Northern Irish/British)	9	Amateur football Recreational walking football	Online
9	74	White (English/Welsh/Scottish/ Northern Irish/British)	7	Amateur football Recreational walking football	Online
10	70	White (English/Welsh/Scottish/ Northern Irish/British)	4	Amateur football Recreational squash Recreational golf	In person
11	61	White (English/Welsh/Scottish/ Northern Irish/British)	5	Amateur cycling Amateur running	Online
12	75	White (English/Welsh/Scottish/ Northern Irish/British)	6	Amateur football Amateur tennis Recreational volleyball	Online
13	79	White (English/Welsh/Scottish/ Northern Irish/British)	6	Recreational table tennis Recreational walking football	Online
14	56	White (other white ethnicity)	3	Amateur rugby Amateur cricket Amateur football	In person
15	67	White (English/Welsh/Scottish/ Northern Irish/British)	4	Recreational tennis Recreational running	In person

Quantitative findings

Table 7 summarises the descriptive statistics and ICC estimates for results from each assessment. Mean results from the Trail Making Task A, the picture recognition correct hits, delayed recall correct hits, and source memory hits, and the reaction time assessment indicate better performance in the in-person condition. Mean results from the Trail Making Task B and the source memory outcome in the picture recognition task indicate better performance in the online condition. The ICC estimates indicate poor correlation in the Trail Making Task B, the picture recognition hits original assessment and delayed recall assessment, and the reaction time assessment. The poor correlation in the picture recognition assessment hits outcomes might reflect the lack of variability in results, as participants tend to score very highly on this assessment (see results from Study 3) which can influence the ICC result. Moderate correlation was found in the Trail Making Task A and the picture recognition source memory original assessment and delayed recall assessment.

Table 7: Summary of Descriptive Statistics and Intraclass Correlation Coefficients for each assessment.

Test	In-person mean (SD)	Online Mean (SD)	ICC (Koo & Li Reliability (123))
Trail Making Task A (s)	30.53 (9.61)	32.27 (6.42)	0.62 (Moderate)
Trail Making Task B (s)	50.20 (19.72)	47.33 (12.69)	0.17 (Poor)
Picture Recognition Hits (%)	95.07 (5.26)	94.20 (9.04)	0.026 (Poor)
Picture Recognition Source Memory (%)	80.87 (18.42)	85.33 (17.62)	0.75 (Moderate)
Picture Recognition Hits (Delayed Recall) (%)	92.73 (7.79)	90.60 (11.46)	0.25 (Poor)
Picture Recognition Source Memory (Delayed Recall) (%)	74.67 (21.35)	74.13 (26.74)	0.69 (Moderate)
Reaction Time (ms)	315.69 (29.56)	335.33 (80.71)	0.39 (Poor)

Figure 7 describes the distribution of results across both testing conditions for each task. Correct hits in the picture recognition task and the delayed recall task demonstrate a ceiling effect and therefore low variability in these scores. The source memory outcome demonstrates greater variability in the results across both the original condition and the delayed recall condition. Participants completing the reaction time assessment show greater variability in the online condition.

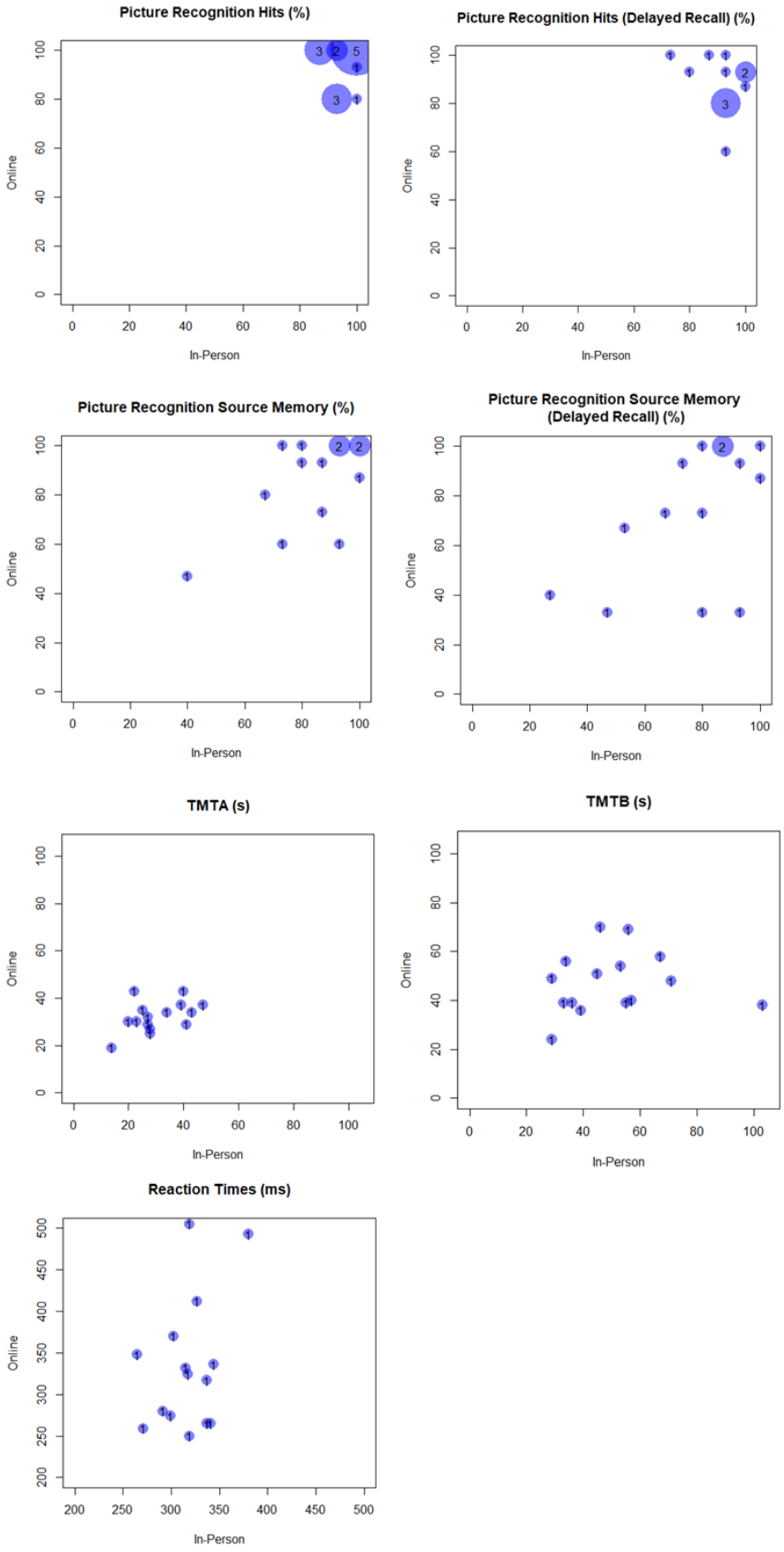


Figure 7: Distribution of results from in-person and online testing conditions for each assessment.

Following the qualitative interviews, of which the findings are summarised later in this chapter, a further variable of interest arose, which was whether the first testing session took place in-person or online. Because of the small sample size further inferential statistics were not feasible, but this variable was plotted to understand the distribution of results in Figure 8. Picture recognition hit accuracy does not show a noticeable difference between first testing conditions due to the ceiling effect and lack of variability. Source memory accuracy results in both the original and delayed recall condition indicate that participants who completed the assessments in-person first were more accurate in both testing conditions. Conversely participants who completed their assessments online first were less accurate in both testing conditions. Participants completing the trail making tasks online or in person first do not show a noticeable a difference between outcomes depending on first condition. In the reaction time assessment, the results indicate that participants who completed the task online first show greater variability in the online testing condition than participants who completed the task in-person first except for one outlier.

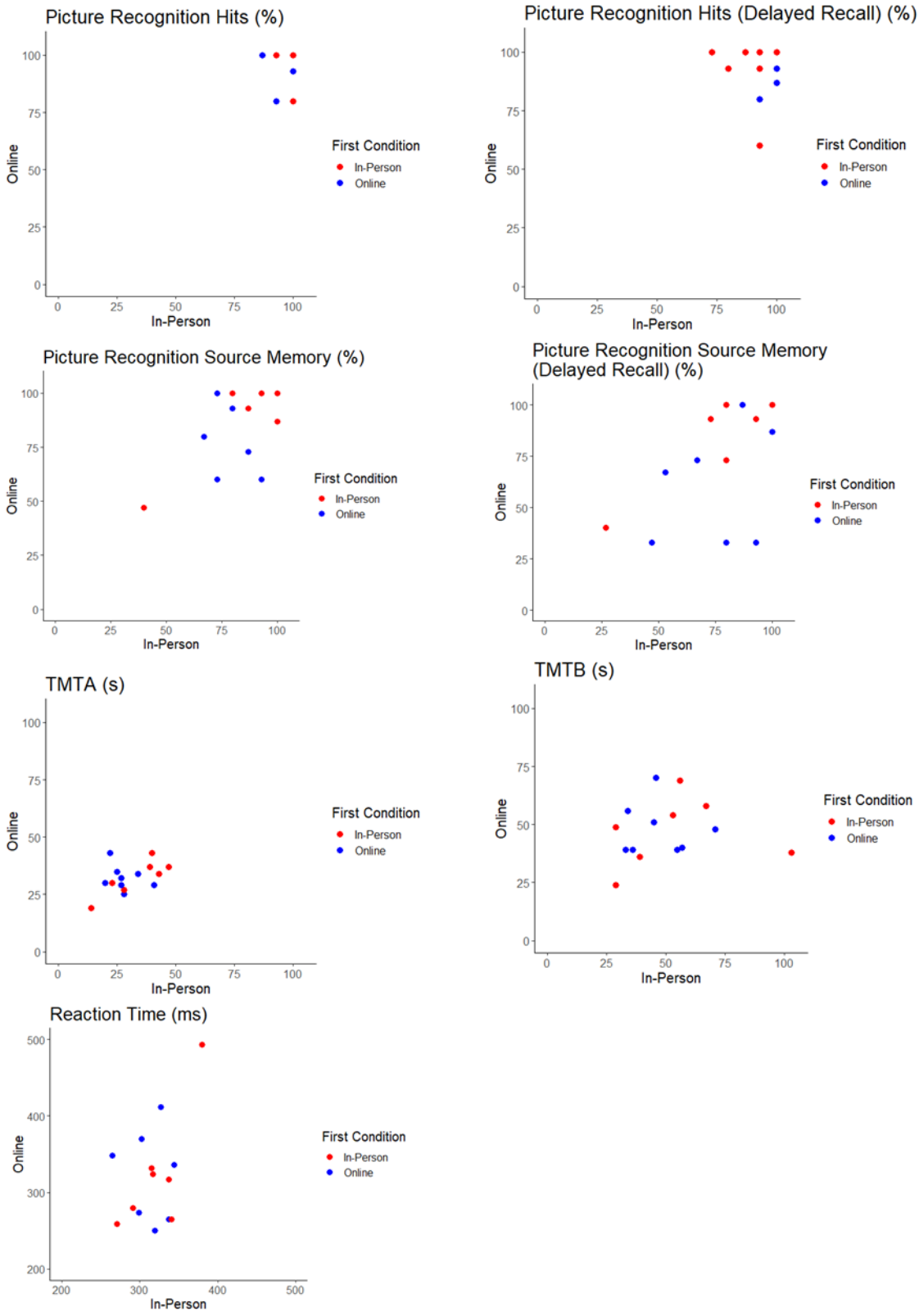


Figure 8: Distribution of online and in-person results by first testing condition

Qualitative findings

Interview responses were analysed using framework analysis, to chart participants' responses and gain an understanding of how their individual experiences identified themes of differences in testing situations. These themes are described below.

Examiner Contact

Participants were asked to discuss the experience of having a researcher present during in-person testing and compare it to the experience of completing the assessments alone at home. A common theme in participants' responses was feeling watched during in-person assessment, which may have led to additional pressure:

> *"You almost feel embarrassed when you get one wrong in front of somebody else." (P7)*
>
> *"You're doing a survey of people who are naturally competitive. So, if you have got somebody sitting there, there is a kind of I want to get this right, I want to impress this person. It's ridiculous, but it's undoubtedly there." (P11)*
>
> *"There's always that bit of test anxiety, and I can't say that the anxiety was any more or any less than with you there or doing it online. I know some people don't like to sit there and have someone watching them." (P4)*

One participant acknowledged that they were also aware of this effect at home:

> *"You're still aware that somebody is analysing the results so whatever you do is going to be seen by somebody." (P7)*

Others acknowledged the feeling of being watched in the in-person testing, but thought that it had a positive effect on performance:

> *"(The researcher was) very quiet and calm on the desk and not interfering, so it didn't trouble me in the least. It was quite reassuring really." (P6)*
>
> *"I would expect it to make it slightly better because it's like being in an exam format with an invigilator. Somebody watching over you just means you're more focused. It didn't make me more nervous, just more focused to get the task done." (P5)*

Participants also acknowledged that the presence of another person in the room with them could be a source of distraction:

> *"There was the temptation to comment or to chat, but I tried to resist that really." (P6)*

Whilst other participants reported that they were used to working in the presence of other people or were used to ignoring distractions:

> *"I'm quite good at blanking out everything because I am used to working in offices." (P8)*
>
> *"I'm a person who can just block out things." (P10)*

Instructions in both the online and in-person assessments were the same in both conditions. Participants reported that instructions were clear in both conditions:

> *"Straightforward. Clear. The instructions online and (in-person) were all ok for me." (P1)*

However, it was identified that verbal instructions were more difficult to understand than written instructions, which have the benefit of being able to read them multiple times.

> *"Surprisingly it took me a little bit more time to absorb what was being said to me than when I did it at home… When I was given written instructions, I could reread it before I did it. Whereas when you spoke to me, I was trying to take it all in and I didn't necessarily actually do that." (P12)*

When prompted about if they knew where to find technical support in the online condition, participants reported a difference in terms of feeling they could ask for help:

> *"I think I would prefer to have someone with me, because if I had any questions at all I had someone to talk to immediately." (P9)*

> *"I think the difference between (in-person testing) and the online one was that there was someone there constantly. Had I had a problem there would have been someone there instantly. You don't quite get that online." (P4)*

Even though asking questions in person was reported to be easier, participants did feel they were able to ask for support in the online condition, but that the quality of this support was important:

> *"I would have emailed and said I don't get this. Given that it was a voluntary and hopefully helpful contribution I wouldn't have fallen at the first hurdle and would have gone well we can make this work." (P11)*

> *"I really wasn't concerned. If there was a problem I would have asked, but if you didn't get clear answers, it would have made the problem a little bit more complicated than it needed to be." (P5)*

Furthermore, the online condition was evaluated to be a more challenging test environment where participants not only had to succeed in completing the assessments but also had to succeed in navigating the online testing platform:

> *"It's probably a subtle thing but if you are doing a test on your own at home you are genuinely on your own. So, the context of the test may be slightly more honest in a way because it's just you reacting to the test and using your mind and your memory and your thoughts without any influence. So, I guess at home it is a more sort of raw environment because there isn't anyone there immediately to help and maybe that subconsciously effects the way you would approach it." (P6)*

In general, these findings suggest that the effect of an absent or present examiner is mixed and may have advantages for some whilst disadvantages for others. However, the ability to ask questions and seek immediate support is a key advantage of the in-person condition.

Testing Environment

Participants were asked to evaluate how comfortable they felt in both testing conditions. Participants reported comfort in both testing conditions and similarity in both settings. It was also noted that the online environment had become more familiar since the Covid-19 pandemic lockdowns.

> *"I felt very comfortable in person." (P10)*
>
> *"It was actually very similar to how it was at home." (P6)*
>
> *"For me, the online bit was a natural environment. It's become a natural environment over the last two years." (P10)*

One of the older participants reported greater comfort in the online session because it was the second testing session, and they were aware of what to expect:

> *"I think I was a bit more relaxed with the second one that I did at home. I was very nervous with the first one that I did at (the university) because I didn't know what to expect." (P3)*

When asked to compare testing conditions, it became clear that many of the participants had an allocated space at home in which they could complete the assessments:

> *"It was like a home office – one of the bedrooms we converted to a study. I was in there on my own with my own desktop computer." (P2)*
>
> *"I have a dedicated home office. It's my office and it's contained and I'm comfortable." (P5)*

These workspaces were reported to be a space where participants felt in control of the environment to minimize distractions and increase comfort whilst completing assessments:

> *"It's a small room at the end of the house. I close the door and it's quiet and there's no interruptions. It's my room." (P12)*
>
> *"At home I closed the door and went into assessment mode." (P4)*
>
> *"If you're doing the online part at home, you're in control to a certain extent." (P11)*

In comparison to online testing, the in-person assessment required an additional mental exercise of "shutting out" the examiner in the room:

> *"It was a conscious thing because I knew you were at the other side of the desk. Sometimes you think you can't help but grimace and you think oh you're opposite someone." (P14)*
>
> *"I have been used to working on a computer around others so even if there are distractions, I can shut those out." (P10)*

Whilst comparing the online and in-person assessment sessions, an emerging theme was the convenience of completing the assessments at home vs. travelling into the university. Recruitment for this study was limited to participants who live within the same county as the university, and the maximum distance travelled by any of the participants was approximately 50 minutes by car. Some participants reported the travel requirement as acceptable:

> *"It was ok, I didn't mind. It's not too far from where I live and easy to get to." (P15)*

> *"It's not the easiest of journeys for me but it was fine. I didn't mind doing it. It's all worth it these things to help out." (P13)*

However, the convenience of completing the assessments at home was also seen as a benefit to online testing and perceived as a less stressful form of testing:

> *"It was more of an effort (to come in) really because you get in the car and drive there. There was a little bit of stress around finding (the building) and parking. So, they're things that wouldn't happen if I did it at home." (P6)*

> *"When you need to be somewhere and achieve an appointment that's a different mental stress than when you can just do it at your leisure at home." (P6)*

In summary, the testing environment did not seem to have a major effect on participants' experiences of the assessments, given that they had allocated work environments in their own homes. However, the convenience of online assessment contributed to participants' affective attitudes towards participation.

Workstation

When asked to compare the workstation used to complete the assessments in both testing conditions, participants described a noted difference in completing the trail making tasks on paper (as they are traditionally completed) compared to on their own device. Participants noted that their performance on the assessment might have been influenced by their visual field being interrupted by their hand in the paper and pencil task:

> *"I think I preferred (doing the trail making tasks) on the laptop because by hand you cover the numbers and on the laptop, you can see them all." (P7)*

> *"When I was taking the test at home, I had the screen in front of me to see everything there. When I was doing it on paper my hand was on top of the paper so perhaps that had an influence on how much information I could see at the time." (P1)*

Participants also noted that the online condition might have been more accessible:

> *"If I had been physically inhibited with that type of drawing action it might have made a difference because on the screen, I am using the mouse to draw the lines." (P10)*

However, some participants reported a preference for the paper and pencil version of the task:

> *"I think doing things on paper and pencil, when you're as old as I am, because that's what you're used to is probably a bit easier." (P13)*

> *"That was easiest to do on a piece of paper rather than on a computer." (P6)*

An emerging theme from comparing workstations was that participants felt more comfortable with their own devices, and that the unfamiliar device in the in-person condition may have impeded on performance. Participants were very comfortable using their own devices to complete the assessments:

> *"Obviously in the home environment I am fairly used to the set up I've got in front of me." (P10)*

However, participants also suggested that the unfamiliarity of the device provided at the in-person session may have had an effect on performance:

> *"There was a difference in what I did at home and what I did in the centre. Partly I think because I don't use a laptop, I use a PC with a mouse. So, it took me a few seconds at the start of each exercise to move the cursor along." (P11)*

> *"Once I got used to the actual mechanics of moving the cursor and clicking, I was fine with it." (P4)*

> *"I haven't used a laptop in a lot of years. It was unfamiliar to me." (P9)*

One participant suggested that the unfamiliarity of the device provided may have added to the stress of testing:

> *"I think in the in-person session I wasn't totally relaxed at the start of each session because I hadn't grasped exactly how to do it manually." (P12)*

In summary, these findings suggest that participants felt they performed better on the online assessments due to differences in manual operation and visual field, particularly for the trail making tasks.

Preference

At the end of the interviews, participants were prompted to suggest their preferred method of assessment and explain their choice. Participants who preferred the online session largely cited the convenience of completing assessments at home:

> *"The online (condition) is more convenient, because I can fit it in around me and from home." (P1)*

> *"I'd rather be at home because it's more convenient, I don't have to go anywhere, and it's quieter." (P2)*
>
> *"The online simply because I could do it when I felt like doing it and it didn't involve any travel." (P4)*
>
> *"I would prefer the online one because there is no reason to come in and there is no difference between them as far as I concerned." (P7)*
>
> *"I would probably say the home one just purely for convenience. Then I could do it in my own time." (P8)*
>
> *"I would prefer to sit in my own home because I don't have to travel." (P13)*
>
> *"Probably online just because it's more convenient. I can do it when it fits in to my day and I don't have to schedule getting there. So, it's easier to manage." (P15)*

One participant also suggested that in-person assessment might not be as accessible as at home testing:

> *"I could see that (coming in) may be a barrier to lots of people if all the assessments were in person." (P4)*

Participants who preferred the in-person session tended to cite the ability to ask questions or for technical support as a key benefit of completing assessments in-person:

> *"I think because of my limited use of the iPad I think I would probably prefer to come to the (university) and do it with you. Then you are there in control of the laptop whereas if something goes wrong with my iPad then I wouldn't know what to do." (P3)*

In summary, the preference of testing condition was determined by a preference for convenience versus a preference for accessing immediate technical support.

Discussion

This explanatory concurrent mixed methods study was designed to compare the quantitative outcomes from the two batteries of tests and to qualitatively better understand participant experiences and insights from online and in-person assessment sessions. The findings are two-fold: (1) equivalence could not be established but variables such as testing condition and first testing condition may influence performance, and (2) that participation in the online assessment condition was considered preferable but mixed in terms of factors that might affect validity.

Quantitative findings indicate that assessments of attention (Trail Making Task A), memory (Picture Recognition Correct Hits and Source Memory) and reaction time produced higher scores in the in-person assessments whilst measures of source memory (Picture Recognition delayed source memory condition) and executive function (Trail Making Task B) produced better scores when completed online. This variation is interesting, because it demonstrates that some

assessments could be better performed online, whilst others produce better results in person. The source memory delayed recall condition is the most challenging condition of the picture recognition outcomes, and the Trail Making Task B is also the more challenging assessment in the trail making tasks so this could reflect the need for more focussed concentration, which the findings from the qualitative study suggest that the online condition can achieve. Conversely the Trail Making Task A, the picture recognition correct hits and original source memory conditions, and reaction time assessment could benefit from the motivation of being "examined" that participants noted in the interview.

The ICC estimates from this analysis indicate moderate correlation for measures of attention (Trail Making Task A) and source memory (Picture recognition source memory task and delayed recall source memory task) and poor correlation for executive function (Trail Making Task B), memory (Picture recognition and delayed recall) and reaction time assessments. However, poor ICC estimates in the picture recognition task may be a result of low variability and the small number of participants (128). Therefore, their interpretation should be done with caution. Combined with the descriptive statistics the assessment outcome that remains convincing is that the Trail Making Task B results may benefit from an online testing environment given the higher mean and poor correlation between assessment sessions. Otherwise, the remaining assessments are difficult to discern between the effect of the testing session or the effect of low variability and sample size.

The lack of variation in the picture recognition hit accuracy results indicate a suggested ceiling effect on performance. This suggests that this part of the assessment may be too easy for a non-clinical population, and that future studies could consider altering the task to make it more difficult. However, the results from the source memory part of the assessment seem to show more variation, which is also reflected in the findings from the SCORES dataset as a whole (see Study 3 results).

The interviews revealed that some participants felt more comfortable completing the online assessments, if they had already completed the in-person session. This prompted interest into if this comfort was reflected in the results, and if participants who completed the first session in-person had an advantage in the second testing session compared to those who completed the assessments online first. The small sample size limited further inferential investigation, but the results were plotted to observe if a trend could be seen. This trend can be observed in the source memory outcomes of the picture recognition assessment, whereby participants who completed the assessment online first tend to have better results in both testing conditions. In addition, greater variability can be observed in the online reaction times of participants that completed the online assessment first. Although there are outliers in both examples, this suggests that initial online assessments might have reliability issues due to participants not knowing what to expect,

whereas an initial in-person assessment might provide a more comfortable testing session to complete assessments for the first time.

The presence of a researcher was noted to be both beneficial and disadvantageous to perceived performance. Key benefits of researcher presence included the comfort of asking questions, reassurance when technical issues occurred, and motivation to focus on the assessments. Key disadvantages included the temptation to engage with the researcher whilst completing the assessments, distraction, and the pressure of "feeling watched" whilst completing the assessments. Previous research suggests that the lack of researcher presence may allow participants to answer personal questions more candidly (129, 130), which may also translate into more comfort with completing cognitive assessments without a researcher presence (120). It may be that completing assessments in the absence or presence of the researcher could be a personal preference, but the key takeaway for designing an online study is that the ease of asking questions and gaining technical support is important to study success and participants' motivation to continue in a longitudinal study. Offering an email address might meet these requirements for some participants. However, for participants who need more reassurance a chat function, video call opportunity or holding the first testing session in-person may benefit participation. Participants also noted a difference between verbal and written instructions based on personal preference and learning style, which suggests that offering online instructions in verbal, written, and visual form might benefit performance and participation.

An interesting finding from this study was the concept of a controlled environment in either home or in-person testing sessions. Traditionally, a university room with controlled noise levels is considered an ideal setting for a controlled testing environment. However, participants emphasized that at home they felt in control of their environment, with the ability to shut the door on noise or distraction, or to ask partners or family to answer the doorbell or keep the dog entertained whilst they complete the assessments. The ability to "close the door and go into assessment mode" may be more accessible following the Covid-19 pandemic, where working from home was normalized for many in the UK and rooms became home offices equipped with suitable working spaces and good internet connection. This notion of improved accessibility in remote testing is also supported in the literature, where remote testing is found to allow improved accessibility for testing in non-clinical populations as well as clinical Alzheimer's Disease populations (122). Importantly though, the ability to "close the door and go into assessment mode" may remain a privilege for those who have greater socio-economic advantage, such as having larger home spaces, a dedicated workspace, and less caring responsibilities.

Discussions of the workstations revealed that differences between completing the assessment by hand and online might have influenced the time to complete the task. Participants reported that they felt they could complete the online trail making tasks faster because their hand wasn't

covering the numbers. This is reflected in the time that it took to complete the Trail Making Task B, but not the Trail Making Task A. However, the difference in field of vision between the digitised assessment and the paper and pencil assessment should be considered when comparing results from online studies and in-person studies. Participants also reported more comfort with completing the assessments on their own device because of the manual familiarity. To control for this, an improvement to this study would be to allow participants the option of a desktop computer, laptop, or tablet or to use a mouse rather than a touchpad when completing the in-person assessments. With a larger sample size, it would have also been interesting to identify if participants using a tablet or computer at home might have affected results in-person. This was not possible with the current sample because only one participant reported using a tablet at home.

Discussion surrounding the preference of testing condition was dominated by a contrast between the convenience of online testing and the availability of support in the in-person testing session. The convenience of online testing is a clear strength of this method, particularly where participants feel comfortable with technology and accessing the assessments independently. Findings from this study that support this are also congruent with previous literature, that suggests high levels of comfort and low levels of distress in online testing environments (119). However, the availability of support is an area where additional measures need to take place in online testing to meet this need. Improved access to support through online chat functions, phone or video calls, or detailed instructional videos with examples might reduce the effect that completing the assessments for the first time online might have on results. The possibility to invite participants to complete their assessments in-person for the first time might also improve validity of the results, and this might be an important consideration for studies using clinical populations. Although, the ability to do so is limited by study resources and geographical reach of the study population. Other investigations also suggest that computer literacy or computer anxiety can be monitored by collecting information on participants' comfort with a computer or their frequency of using the device (120). This may be an important variable to control for in future, particularly in older participant groups and groups where computer literacy varies.

The quantitative findings from this study are limited by small sample size. Recruitment for the present dataset ran from November 2021 until June 2023. The study was advertised on social media, regional radio, a local science public outreach event, posters at a local leisure centre and through word of mouth with walking football groups. The most successful avenue for recruitment was through contact with walking football groups across Norfolk in January 2023. Seven clubs were contacted to build a rapport and advertise the study, which revealed that walking football participants were enthusiastic about the topic of brain safety in sport and represented an older cohort who had a range of footballing experience. Importantly, walking football is a non-contact sport where heading is not just discouraged but classed as a foul and a direct free kick is awarded

to the opposing team (131). Therefore, this group represents a cohort with varying exposure to sport related RHI through their participation in football and walking football and may form an important control group of active participants. In addition, their enthusiasm to promote brain health safety was helpful to the recruitment of participants within this study.

There are limitations to this study. In terms of the demographics, the study population only included three females, and a more representative study population would benefit findings. Females in the SCORES project make up 35% of the study population, and therefore a minimum of five females within this sample would be more representative of the project demographics. The lack of diversity of ethnicity and socio-economic status based on deprivation data also limits findings and reflects a wider limitation of recruitment in the SCORES project to date, which is considered in further detail in the discussion chapter of this thesis. However, the present study population was more balanced in terms of age, decile of multiple deprivation and sporting history.

Conclusion

In conclusion, findings from this study highlight important considerations when assessing the validity of online self-administered neuropsychological assessment. Although one of the initial objectives to assess quantitative equivalence was not possible due to the sample size, ICC comparisons found poor to moderate correlation in scores from both testing environment, where a lack of variation in results on some assessments may explain the poor correlations found. Qualitative findings highlight the benefit of a lack of examiner presence and the ability to control their environment on participants' online results, whilst also highlighting the benefits of being able to ask for immediate help during in-person assessment. The effects of these factors on validity are important to consider when designing online studies, particularly in relation to age, socioeconomic status, and computer literacy of participants. Taken together, findings from this mixed methods study identify key factors that may influence validity of online neuropsychological testing in longitudinal research.

Study 3 – Analysis of brain health data from former contact sport athletes

Abstract

A growing body of evidence suggests that former contact sport athletes are at an increased risk of neurodegenerative disease. To understand this relationship further, longitudinal prospective study of a cohort diverse in age, sex, and sporting level that adjusts for factors that influence brain health is necessary. This study presents the findings from an ongoing online longitudinal study that monitors the cognitive, mental, behavioural and sleep health of former contact sport athletes. The aim of this chapter was to analyse the first data set from the project, to examine if data suggests a long-term consequence of exposure to contact sport on brain health.

In total, 328 participants completed online questionnaires and digitised neuropsychological assessments to collect demographic information and monitor cognitive, mental, behavioural and sleep health. Outcomes were analysed to identify if level of exposure to contact sport had a significant effect on brain health measures. This study also examined the effects of adjusting for potential confounds on brain health. Brain health outcomes were also compared to an age and sex-matched normative mean.

Male former professional contact sport athletes demonstrate worse cognitive scores than male former amateur contact sport athletes, but lower levels of anxiety, depression, impulsivity, and aggression. Females with a history of contact sport did not demonstrate worse cognitive scores than females with a history of non-contact sport, but demonstrated higher levels of depression, anxiety, impulsivity, and aggression. Regressions identified variables that should be adjusted for in future research. Male and female contact sport athletes aged 40-50 years demonstrated better performance on cognitive outcomes than a matched normative sample, but participants aged 65 and above demonstrated worse cognitive scores than the matched sample. A further analysis also examined the validity of results from the Sustained Attention to Response Task and identified potential issues of the measure as a test of executive function.

This investigation identifies domains of cognitive, mental, behavioural and sleep health that may be affected as long-term sequelae of exposure to sport related RHI. This study demonstrates the potential impact of including novel measures of spatial navigation in assessment batteries for monitoring early signs of cognitive decline. Findings from this study suggest a potential sex specific interaction on cognitive and mental health performance across contact athletes, with males exhibiting worse cognitive health but better mental health than females, and females demonstrating the opposite effect. However, this study highlights the need for better recruitment and categorisation of female athletes to further examine the influence of sex on outcomes.

Introduction

Retrospective and post-mortem studies suggest that participation in contact sport increases risk of neurodegenerative disease (29). The threat of dementia on families, healthcare services and the economy is due to increase as worldwide dementia rates are forecasted to triple by 2050 (132). Reducing modifiable risk factors, like exposure to sport-related RHI, is paramount towards mitigating against rising dementia cases.

Fundamental work in this field has identified neurodegenerative disease in the brains of former professional contact sport athletes (133) associated with age (134) and length of career (135). Health record data demonstrates a higher rate of death by neurodegenerative disease in former contact sport athletes (53-56, 136), and retrospective interviews with next of kin has described symptoms of dementia in former contact sport athletes (61, 137). Prospective research that monitors athletes as they age and can adjust for many confounds on brain health is needed to further understand the long-term effects of exposure to sport-related RHI on brain health (115, 138).

Since the commencement of my PhD project in 2020, findings have been published from other research groups studying the long-term brain health of former contact sport athletes. An example of this is the Foot/ankle Osteoarthritis and Cognitive impairment in UK Soccer players (FOCUS) study, which collected brain health data from former professional male footballers in the UK and found an increased prevalence of dementia and worse cognitive performance in footballers than the general population control group (66). However, to date research in this area has largely been limited to studying the long-term brain health of male former professionals (84). The scarcity of data from female participants is problematic, given that females are more likely to be diagnosed with dementia in general (74), and as athletes are more likely to sustain concussions and experience prolonged and more severe concussions than males (85, 139, 140). Studies of amateur athletes are also rare, despite representing the largest participation group in sport. The lack of diverse participants in this field limits the quality of policies that can be applied to protect athletes from exposure to a preventable risk factor of dementia.

The SCORES Project was designed to address these gaps in the literature and to assess a more balanced sample of participants. The aim of this chapter is to identify if initial results from the project demonstrate group differences in cognitive, mental, behavioural and sleep health by levels of exposure to contact sport. The first objective is to compare all outcomes in male former professional contact sport athletes with male former amateur contact sport athletes, as well as to compare female former contact sport athletes with female former non-contact sport athletes. The second objective is to examine the influence of confounding factors on the relationship between exposure to contact sport and neuropsychological outcomes. The third objective is to compare cognitive outcomes in groups exposed to contact sport with a normative sample

matched by age and sex. This study tested the hypothesis that male former professional contact sport athletes would demonstrate significantly poorer scores on assessments of cognitive, mental, behavioural and sleep health than former amateur contact sport athletes. This study also tested the hypothesis that female former contact sport athletes would demonstrate significantly worse brain health outcomes than female former non-contact sport athletes. Furthermore, this investigation tested the hypothesis that male and female contact sport athletes would perform significantly worse on brain health outcomes than an age and sex matched normative sample.

Methods

Participants

Eligible participants (n=328) were aged over 40, lived in the UK or Ireland, and did not have a diagnosis of dementia. Recruitment was open to participants with contact or non-contact sport experience at any level of play. Participants needed to have access to a computer or tablet to complete the assessments. Professional sport was defined as being paid to play, amateur sport was defined as playing competitively but not paid to play, and recreational sport was defined as playing non-competitively.

Procedure

Data for this chapter were collected from the first set of results from the SCORES Project. A demographics questionnaire gathered age, sex, ethnicity, postcode, and years of education. Postcode was collected to estimate socioeconomic status using UK indices of multiple deprivation. Participants described their sporting history including level of sport, years playing the sport, age at start of playing the sport, and playing position if relevant. Participants also provided information about their general health including their current level of physical activity (later categorised by WHO guidelines (1)), smoking history, conditions related to dementia risk (blood pressure conditions, heart conditions, osteoporosis, diabetes or blocked arteries), treatments related to dementia risk (statins or hormone replacement therapies) or conditions that might affect their performance on online assessments (learning difficulties/disabilities, attention deficit hyperactivity disorder, or mental health conditions). Participants were also asked about their general brain health including number of diagnosed concussions, estimation of suspected concussions, and history of moderate or severe brain injuries, transient ischemic attacks, strokes, or blast injuries.

Cognitive health was assessed using digitised neuropsychological assessments designed to measure a range of cognitive domains including attention, executive function, memory, working memory, processing speed and spatial navigation. The choice to study these domains was based on work that identified potential cognitive domains effected in the long-term in former contact sport athletes (38, 39, 137) and domains that are typically effected in the early stages of MCI (41)

such as spatial navigation (141).The battery included 7 assessments with a total of 9 cognitive health outcomes.

To monitor attention, the Trail Making Task A (TMTA) was included in the study battery (see Figure 9). In this assessment, participants are shown a set of numbers and asked to connect them in order as quickly as possible. The primary outcome in this assessment is the time in seconds from participants beginning the task to when they have correctly completed the connections, and a faster time indicates better performance.

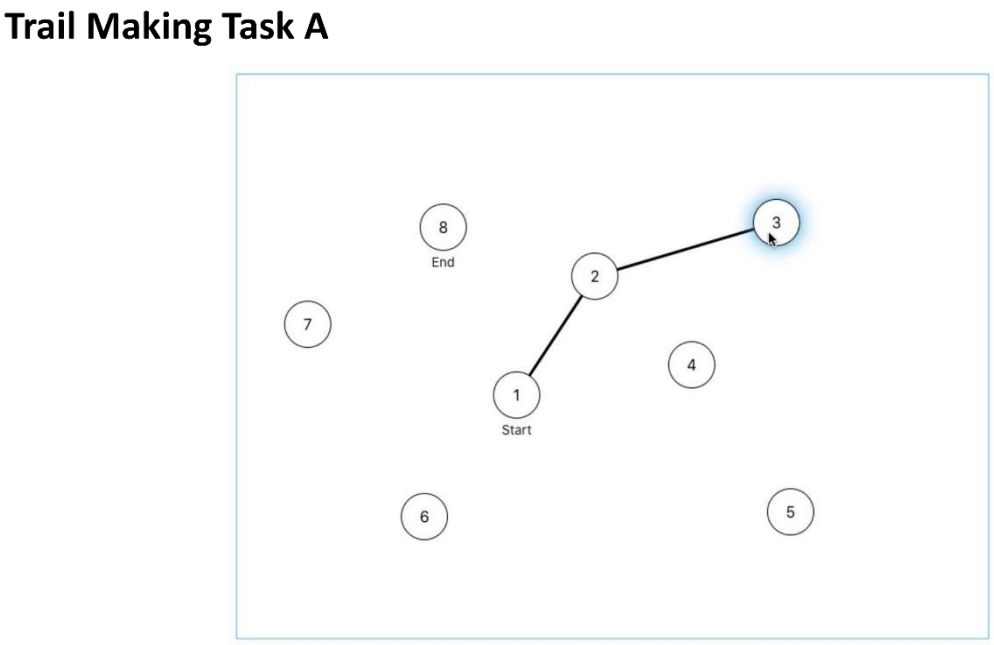


Figure 9: Example of the Trail Making Task A assessment. Participants must connect the numbers in ascending order as fast as possible. This image demonstrates the practice task, but in the real task participants must connect 25 numbers.

In addition to monitoring attention within the TMTA, the Sustained Attention to Response Task (SART) was included to assess sustained attention and inhibition (see Figure 10). In this assessment, participants are tasked with watching a series of numbers between 1-9 appear in the middle of the screen and are asked to press the space bar when they see any number except for the number 3. The assessment lasts approximately 5 minutes and requires sustained concentration. Participants are asked to try to respond as fast and accurately as possible. The primary outcomes of this task are correct commission (correctly not responding to the number 3) percentage, the correct omission (correctly responding to any number except for 3) percentage, the reaction time preceding correct non-target responses and the average reaction time of responses prior to a non-target stimulus. The correct commission percentage was calculated as follows, where target responses refer to participants successfully abstaining from responding to number 3's:

$$Correct\ Commission\ \% = \frac{Correct\ Target\ Responses}{Total\ Targets} * 100$$

The correct omission percentage was calculated as follows, where non-target responses refer to participants successfully responding to numbers other than 3:

$$Correct\ Omission\ \% = \frac{Correct\ Non\ Target\ Responses}{Total\ Non\ Targets} * 100$$

The preceding correct non-target reaction time was calculated as the mean reaction time of responses prior to the presentation of a non-target (number other than 3). The commission error reaction time was calculated as the mean reaction time of incorrect responses to a target (when participants pressed the space bar in response to a number 3). Both reaction time units are milliseconds.

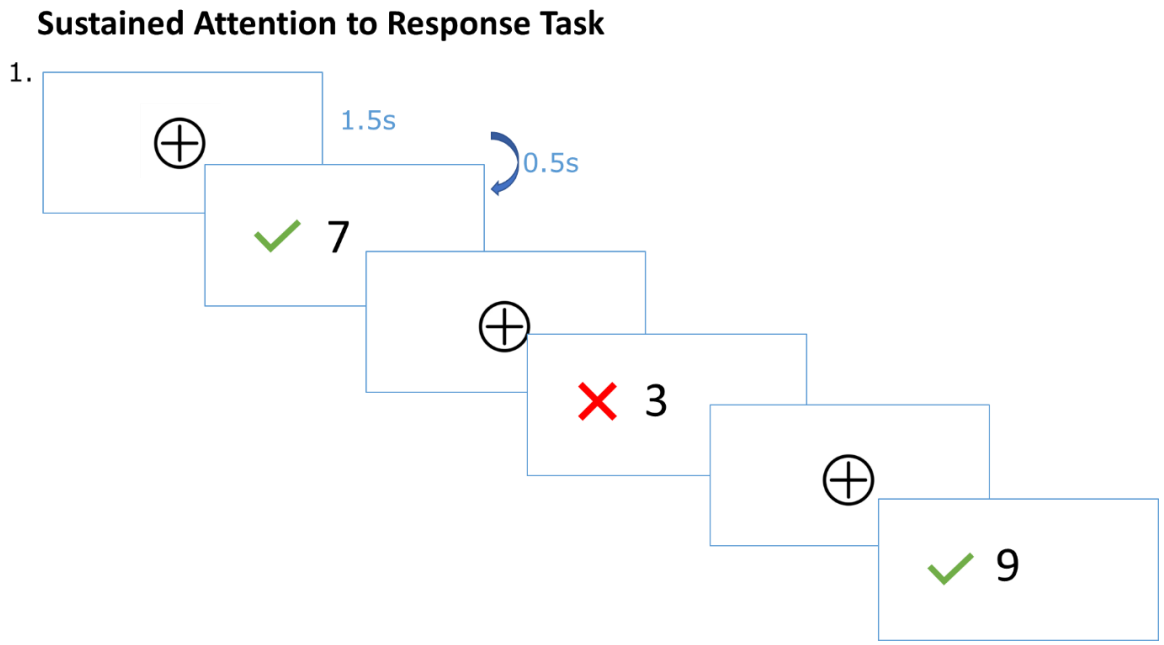


Figure 10: Example of the Sustained Attention to Response Task (SART). Participants view a series of 250 numbers between 1 and 9. Participants must press the space bar when any number except for the number 3 appears. When 3 appears, participants must abstain from pressing the space bar.

Executive function, in particular task switching, was measured using the Trail Making Task B (TMTB) (see Figure 11). In the TMTB, participants must connect a series of numbers and letters in ascending numerical and alphabetical order, alternating between numbers and letters (e.g., 1, A, 2, B, 3, C..). Time to complete this task in seconds is the primary outcome, where a faster time indicates a better performance on the task. The Trail Making Tasks have been used in other studies investigating CTE (142), dementia and the long-term consequences of brain injury (143, 144).

Trail Making Task B

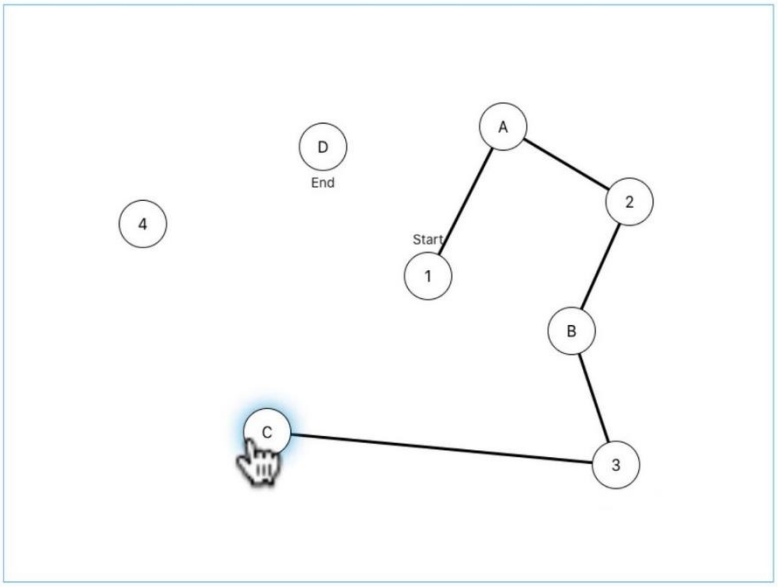


Figure 11: Example of the Trail Making Task B assessment. Participants must connect the numbers and letters in ascending order, switching between numbers and letters. This image demonstrates an example, but in the real task participants must connect a series of 12 numbers and 12 letters.

Short term memory and source memory were measured using the Picture Recognition Task (see Figure 12). In this assessment, participants are shown a series of images that appear in four locations on the screen. Participants are asked to pay attention to the object in the image, as well as where on the screen it appears. Participants then view a second series of images, which includes the original images and new ones. For each image, participants must identify if they saw this image before (recognition) and if so, where on the screen it appeared (source memory). Following the completion of another assessment (the digit span backwards assessment), participants are asked to view a third set of images that contain the original images and new images. Participants must then identify the images they saw before and where they appeared on the screen. This provides a measure of delayed recall. The primary outcome measures of this assessment are therefore the recognition of the original images (PRH1), the source memory of the original images (PRSM1), the delayed recall of the images (PRH2) and the delayed recall of the source (PRSM2). Assessments of visual memory are used in other cognitive batteries like the ImPACT assessment (145). This assessment is less reliant on language than word encoding and memory tasks, and it offers an additional outcome of source memory.

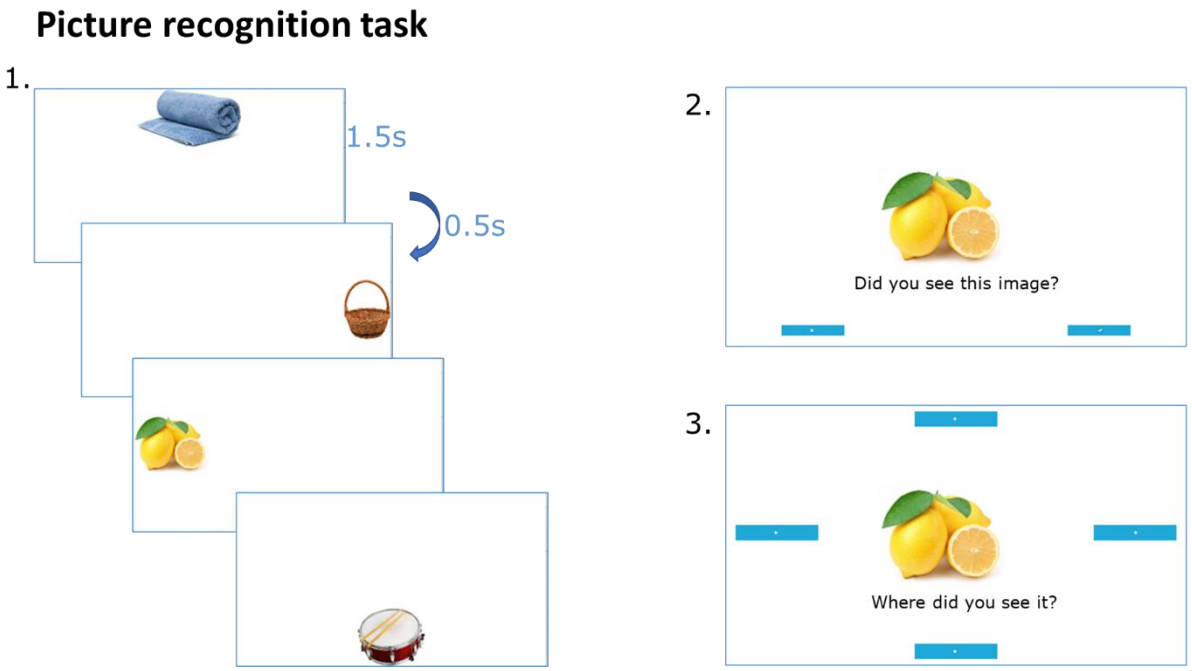


Figure 12: Example of the picture recognition task. Participants view a series of 25 images that appear in one of four locations of the screen. Participants then view a second series and are asked to confirm if the image appeared in the first series. If participants confirm they recognize the image from the first series, participants are asked to identify where on the screen the image appeared.

Working Memory was measured using the Digit Span Backwards task (DSB) (see Figure 13). In this task, participants are presented with a sequence of digits in the reverse order in which they appeared. The length of digits increases until the participant makes two consecutive errors, which ends the test. The primary outcome measure is the maximum digit span. The minimum score on this assessment is 0, and the maximum span is 9. However, if a participant scores 0 the standard procedure was to invite them to repeat the assessment to control for any technical errors. This assessment has been used in studies that investigate the long-term effects of exposure to concussions (142, 143, 146) and in former contact sport players (144).

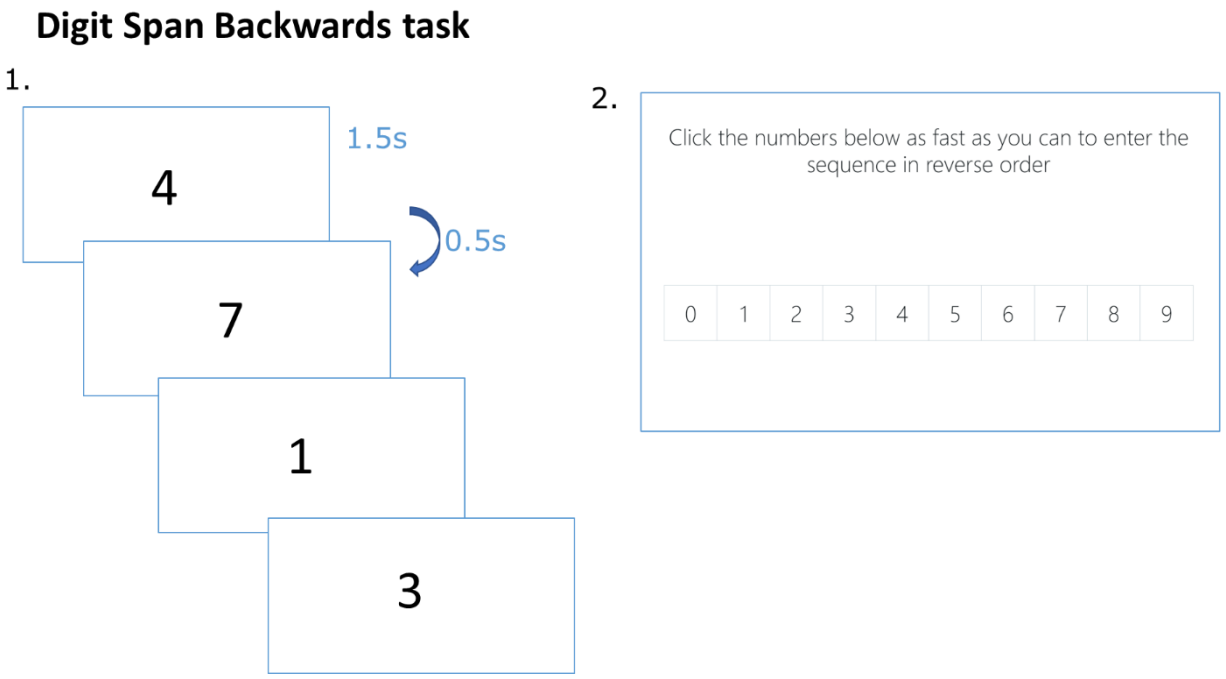


Figure 13: Example of the Digit Span Backwards (DSB) task. Participants view a series of numbers and are then asked to click the numbers in the reverse order that they appeared.

Working Memory was also measured using the Spatial Working Memory assessment (SWM) (see Figure 14). In this assessment, participants are shown a sequence of squares that light up on a screen. Participants are then asked to select the squares in the reverse order in which they appeared. As the task progresses, the sequences become longer until the participant makes two consecutive mistakes. As in the DSB, the minimum span is 0 and the maximum span is 9. Participants who scored 0 were also invited to repeat the assessment to control for technical errors. The inclusion of the SWM as well as the DSB was made to include two measures of working memory that used either digits or visual information.

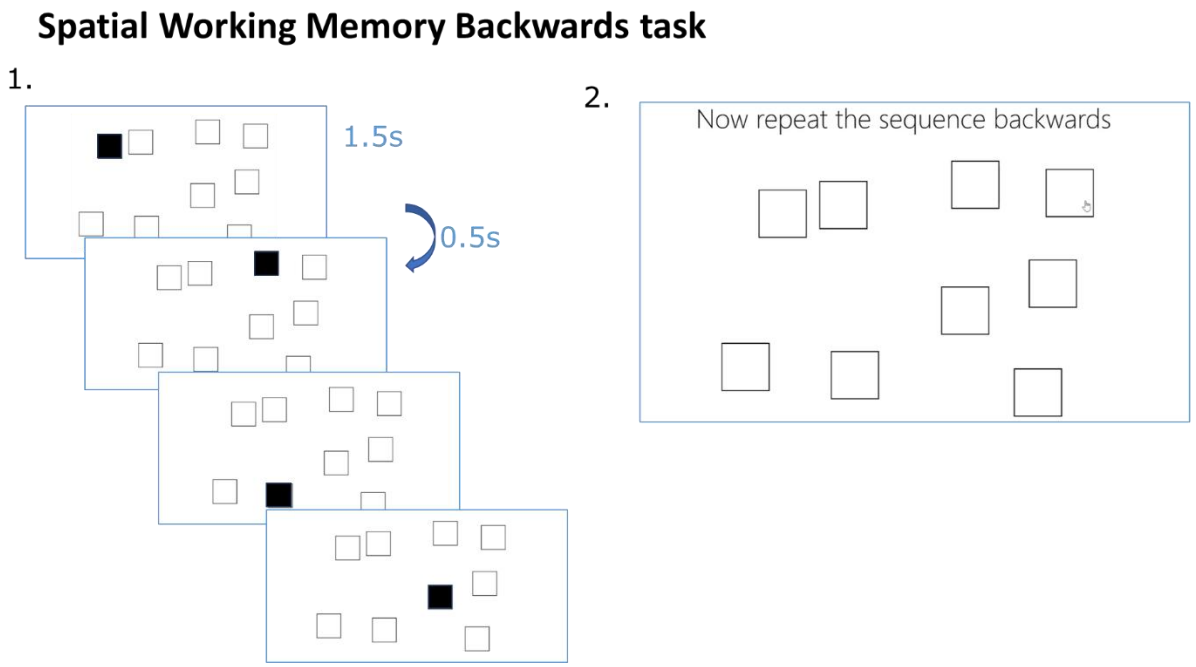


Figure 14: Example of the Spatial Working Memory Backwards (SWM) task. Participants view a series of highlighted boxes and are asked to click the boxes in the reverse order in which they were highlighted.

Processing speed was measured by using the Simple Reaction Time assessment (RT) (see Figure 15). In this assessment, participants are asked to respond to a stimulus as fast as they can. The primary outcome of this measure is the average reaction time (ms) across all trials.

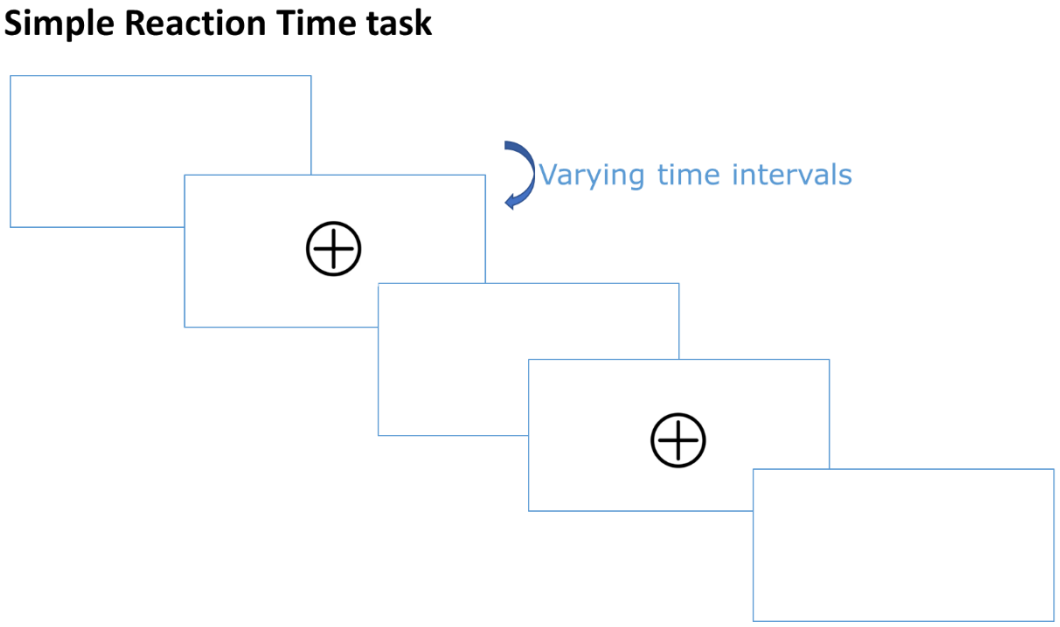


Figure 15: Example of the Simple Reaction Time (RT) task. Participants must respond to a stimulus as fast as possible. Interval times between stimuli vary with each appearance.

Spatial navigation was measured using the Virtual Supermarket task (VST) (see Figure 16). The test consists of a total of 15 videos of a shopping cart that moves through a virtual reality supermarket. The participant is asked to passively watch the videos. After each video, they are asked to indicate the direction of the starting point of the video using an onscreen arrow. This provides a measure of egocentric spatial navigation. They are then asked to trace the path of the trolley on a top-down view of the supermarket layout. Finally, the participant is required to indicate the direction of view at the end point of the video. Spatial navigation is a cognitive domain that has been demonstrated to be impaired in the early stages of dementia (141). Impairments in visuospatial processing have been reported in previous investigations (147), but deficits in spatial navigation ability remains relatively unexplored. The inclusion of this measure represents a novel approach of this thesis.

Virtual Supermarket task

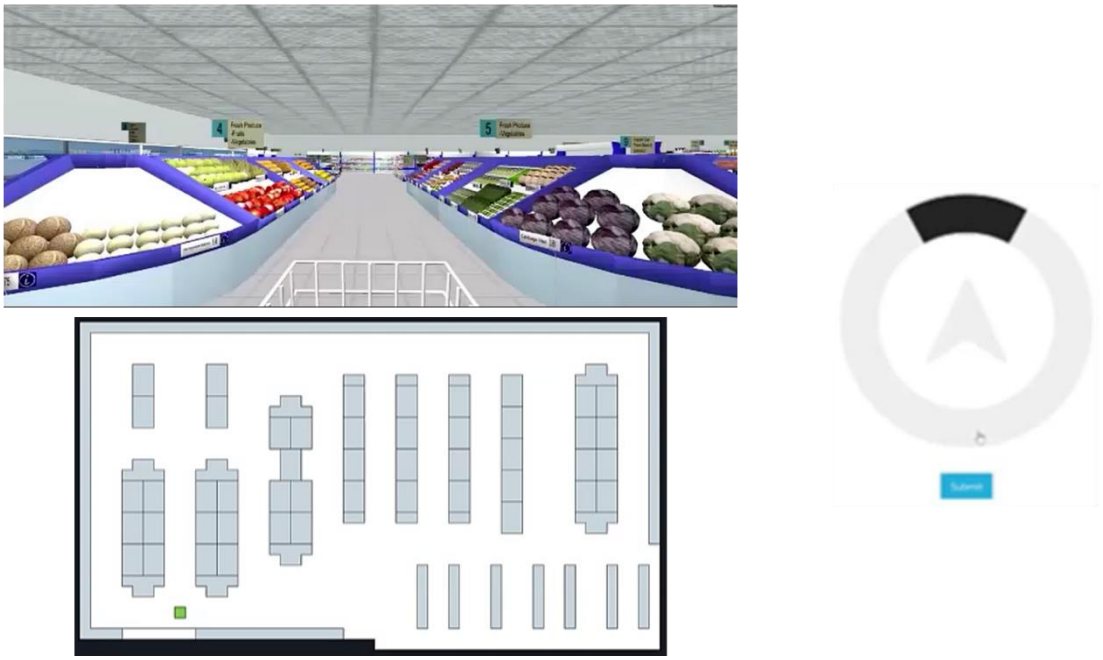


Figure 16: Example of the Virtual Supermarket task. In this assessment, participants watch a video of a shopping cart moving through a supermarket. They are then asked to use the dial on the right to identify where their starting position is in relation to where they are facing. Participants then use the map below to trace their path and identify their final facing position in relation to the top, bottom, left or right of the screen.

Mental health outcomes were measured using a measure of depression and a measure of anxiety. The Patient Health Questionnaire – 9 (PHQ-9) is a short measure of depression (see Appendix D). This measure includes 9 items where participants are asked to indicate how often they experience each item. The Generalised Anxiety Disorder Scale – 7 (GAD-7) is a brief measure for assessing generalised anxiety disorder (see Appendix E). The 7-item questionnaire asks participants to rate how often they experience items. Both scales have been used to assess anxiety in former athletes exposed to concussion and RHIs (148, 149). Both measures were found to have good sensitivity to change over time (150, 151), making them suitable for inclusion in a longitudinal project. The choice of these measures of mental health was also informed by their use by the NHS for online self-evaluations of mental health (114). The format of this online tool was used to inform the design of how this questionnaire appeared as an online assessment, and the feedback that the tool provides on the outcome was also used to inform the feedback that is provided to participants. This choice was made in order to prioritise the acceptability of completing an online measure of mental health, and the acceptability of this choice is further explored in Study 1.

The Aggression Questionnaire (AQ) is a 29-item questionnaire which includes statements about four domains of aggression: physical aggression, verbal aggression, anger and hostility (152) (see Appendix F). Participants rank the statements along a 5-point scale from “extremely uncharacteristic of me” to “extremely characteristic of me”. The measure has been used in previous studies using athlete populations (153, 154) and in athletes who participate in contact sports or who have sustained concussions (155). The Barratt Impulsiveness Scale (BIS-15) consists of 30 items representing non-planning impulsivity, motor impulsivity and attentional impulsivity

(156) (see Appendix G). Items are answered on a 4-point scale ranging from rarely/never to almost always/always. The main outcome is a total score, and three subscale scores can be calculated for each domain of impulsivity. The BIS-15 has been used in athlete populations and to measure the long-term effect of exposure to repetitive head injury in athletes (155, 157). These methods on monitoring behaviour were selected because they provide detailed data on different facets of aggression and impulsivity, and both were demonstrated to have good long-term reliability and construct validity (156, 158, 159).

To measure sleep, a merged questionnaire was developed from the Epworth Sleepiness Scale, the Pittsburgh Sleep Quality Index (PSQI), the Berlin Questionnaire, and the Warwick and Oxford Exploration of Sleep (WOES). The Epworth scale is a measure of daytime sleepiness (160) and asks about how likely participants are to feel sleepy in everyday situations. The PSQI is a measure of sleep quality and disturbance, and produces a global score based on questions about sleep quality, latency, duration, habits, disturbance, medications and subsequent daytime dysfunction (161). The Berlin Questionnaire is a measure used to specifically identify sleep apnoea (162). The primary outcome of this questionnaire is a positive or negative identification of sleep apnoea. The WOES questionnaire is a measure developed by collaborators at the University of Warwick and the University of Oxford Brookes. The questionnaire is designed to measure sleep behaviours and is piloted within the SCORES project. Many of the questions in these questionnaires overlap, so a merged questionnaire was deemed most appropriate to reduce participant burden. This merged questionnaire was designed by our collaborators and implemented within the SCORES study design.

All assessments were completed remotely online without the presence of a researcher, and feedback was collected on interruptions or distractions. Where interruptions or distractions were reported, participants were asked to repeat an assessment. Assessment sessions lasted approximately 30 minutes.

Further validity work

The validity of the online assessment battery was assessed in Study 2. In addition to this, test-retest reliability of all of the cognitive, mental health and behaviour tests were assessed. This comparison used results from the first and second set of data collected in the SCORES project. Results were compared using the ICC. This model of ICC was chosen because the quantitative research question asks for the difference between online and in-person conditions (absolute agreement) using the mean of multiple raters (mean of k) who were the only raters of interest (two way mixed effects) (126) .

Statistical analysis

Statistical analyses were chosen to isolate the effects of participation in contact sport in comparison to appropriate control groups. Exposure groups and control groups were chosen

based on available normative data and sample sizes within the study. Based on a priori power calculations, using G*Power, a minimum sample sizes of 19 was required for one-sample t-tests and 70 for independent t-tests.

Cognitive, mental, behavioural and sleep health results were compared between male former professional contact sport athletes and male former amateur contact sport athletes. Outcomes were also compared between female former athletes with a history of contact sport and a history of non-contact sport. Control groups in these comparisons were chosen based on sample size but also to allow appropriate comparisons where level of exposure to contact sport differs between the two groups. For males in the sample, comparing former professionals to amateurs was appropriate because of the distinctness between amateur and professional careers in terms of training load and exposure to contact. Comparing male contact sport athletes with male non-contact sport athletes was not possible because of the low number or males reporting no contact sport history within the sample. For females in the sample, comparing former professionals to amateurs was not appropriate because of the recentness of professionalisation within women's sport (meaning that females in our sample are unlikely to have been paid to play) and therefore a heterogeneity in terms of level of exposure to contact within the amateur female group. Comparing female contact sport athletes with female non-contact sport athletes was therefore a more appropriate comparison, and allowed me to isolate the effects of contact in sport on female athlete brain health. Therefore, different control group parameters were necessary for comparisons in male and female athletes.

Cognitive results from male and female participants with a history of contact sport were also compared to a normative sample provided by Neuropsychology Online (NeurOn) who developed the digitised cognitive assessments. This normative sample from NeurOn was collected from other studies using the same cognitive assessments in non-clinical populations matched by age and sex.

Data analysis was performed using R. Data from participants with missing data were not analysed. Descriptive statistics were calculated for all groups as means and standard deviations. Group differences in both analyses were compared using inferential statistics. Assumptions of normality were checked using Shapiro Wilk's test and homogeneity of variance was checked using Levene's test. Male and female group comparisons were performed using Student's t-test or Mann-Whitney U tests where data were non-parametric. Linear regression was performed to identify confounds and multiple regression was used to adjust for potential confounds. Comparisons with normative data were performed using one-sample t tests or Wilcoxon tests where data was non-parametric.

Results

Demographics

This investigation reports results collected from the first set of assessments in an ongoing longitudinal study. At the time of analysis, a total of 328 participants had completed at least one full set of assessments. Their demographic information that is relevant to the variables adjusted for in this chapter's analysis is described in Table 8. A table summarising the complete dataset of SCORES participants, including variables that were not included in the current analysis, can be found in Appendix I.

Table 8: Summary of demographic data from participant groups included in this study.

	All	Contact Sport Males	Contact Sport Females	Non-Contact Sport Males	Non-Contact Sport Females	Professional Contact Sport Males	Amateur Contact Sport Males
n	328	192	44	9	49	77	103
Age (mean, SD)	59.01 (9.60)	59.39 (9.61)	52.00 (9.40)	57.67 (8.00)	62.82 (8.33)	59.81 (9.66)	58.74 (9.68)
Ethnicity, n (%)							
White (Eng./Scot./Welsh/N. Irish)	305 (93)	180 (93.75)	41 (93.18)	7 (77.78)	45 (91.84)	72 (93.50)	98 (95.14)
White (Irish)	4 (1.20)	3 (1.56)	1 (2.27)	-	-	1 (1.30)	2 (1.94)
White (Other)	14 (4.30)	5 (2.60)	1 (2.27)	1 (11.11)	4 (8.16)	1 (1.30)	3 (2.91)
Black (Caribbean)	1 (0.30)	1 (0.52)	-	-	-	1 (1.30)	-
Mixed (White and Black African)	1 (0.30)	1 (0.52)	-	-	-	1 (1.30)	-
Mixed (White and Asian)	1 (0.30)	-	-	1 (11.11)	-	-	-
Other Ethnic Group	1 (0.30)	1 (0.52)	-	-	-	-	-
Missing Response	1 (0.30)	1 (0.52)	1 (2.27)	-	-	-	-
Education in Years (mean, SD)	15.52 (3.71)	15.11 (3.41)	16.78 (3.84)	18.78 (5.91)	16.00 (4.04)	14.68 (3.31)	15.16 (3.32)
Deprivation (mean decile, SD)	6.49 (2.41)	6.75 (2.44)	6.20 (2.69)	6.25 (2.38)	5.96 (2.16)	7.25 (2.32)	6.29 (2.50)
Age at start of highest sport level (mean, SD)	14.01 (9.08)	14.59 (5.06)	16.43 (10.19)	16.78 (12.51)	18.50 (13.50)	17.78 (2.90)	12.32 (5.10)
Duration of highest sport level in years (mean, SD)	17.30 (13.27)	18.83 (11.10)	18.89 (11.65)	26.11 (14.12)	19.20 (17.47)	13.54 (6.80)	22.67 (11.57)
Current Physical Activity, n (%)							
Sedentary	45 (13.72)	22 (11.45)	6 (13.64)	2 (22.22)	8 (16.33)	11 (14.29)	10 (9.71)
Within WHO reccomendations	104	58 (30.20)	12 (27.27)	2(22.22)	15 (30.61)	21 (27.27)	33 (32.03)
Above WHO reccomendations	(31.70)	112 (58.33)	26 (59.09)	5 (55.56)	26 (53.06)	45 (58.44)	60 (58.25)
Smoker Status							
Never	221	134 (69.79)	28 (63.63)	7 (77.78)	30 (61.22)	60 (77.92)	64 (62.14)
Past	(67.38)	46 (23.96)	14 (31.81)	2 (22.22)	17 (34.69)	10 (12.99)	35 (33.98)
Current	91 (27.77)	12 (6.25)	2 (4.55)	-	2 (4.08)	7 (9.09)	4 (3.88)
% with health comorbidities	31.80	32.81	25.00	37.50	30.61	36.36	28.16
% with treatments	26.30	26.56	9.09	12.50	34.69	27.27	24.27
% with MHLDD	14.37	13.54	11.36	25.00	16.33	11.69	14.56
Diagnosed Concussions (mean, sd)	1.16 (3.35)	1.57 (4.17)	1.16 (2.28)	0.38 (0.52)	0.43 (0.85)	2.31 (5.80)	1.05 (2.38)
Suspected Concussions, n (%)							
None	161	74 (38.54)	17 (38.64)	5 (55.56)	33 (67.35)	38 (49.35)	32 (31.07)
Less than 5	(49.09)	82 (42.70)	18 (40.90)	4 (44.44)	13 (26.55)	22 (28.57)	54 (52.42)
5 to 9	119	26 (13.54)	7 (15.91)	-	2 (4.08)	12 (15.58)	13 (12.62)
10 to 14	(36.28)	4 (2.08)	2 (4.55)	-	1 (2.04)	2 (2.60)	1 (0.97)
15 to 19	35 (10.67)	1 (0.52)	-	-	-	1 (1.30)	-
20 or more	7 (2.13)	5 (2.60)	-	-	-	2 (2.60)	3 (2.91)
% with other TBI	6.10	7.81	2.27	0.00	4.08	9.09	4.85

The following important demographic variables were found to differ between groups. Demographic information revealed a scarcity of male participants reporting a history of only non-

contact sport, or a history of no participation in sport in comparison to female participants. The average age of female contact sport athletes in the sample is 10 years younger than the average age of female non-contact sport athletes, and the sample includes a small number of females reporting professional participation. The ethnicity of the sample is predominantly white – English/Scottish/Welsh/Northern Irish. A difference in levels of education can be observed between contact and non-contact sport males, and between amateur and professional male contact sport athletes. A difference in level of deprivation is also observed between male amateur and professional contact sport athletes. In terms of exposure to contact sport, male amateur contact sport players reported the longest duration of exposure in years, which was almost 10 years longer than the average exposure of professional contact sport players. More diagnosed and suspected concussions were reported in contact sport athletes than non-contact sport athletes, and in professional contact sport athletes compared to amateur contact sport athletes.

Validity

Before analysis of the data, further validity work was performed to examine the test-retest reliability of results from the SCORES data set and to build on findings from Study 2. To do this, the intraclass correlation coefficient was calculated to determine the correlation between first and second assessment results. The results are summarised in Table 9.

Table 9: Means and ICC estimates comparing first and second assessment results.

Assessment	**Mean at Assessment 1 (Standard Deviation)**	**Mean at Assessment 2 (Standard Deviation**	**ICC**
TMTA	32.71s (9.67)	31.42s (10.09)	0.74 (moderate)
TMTB	48.42s (19.27)	48.20s (22.64)	0.81 (good)
SART Commission	33.12% (29.83)	32.29% (27.92)	0.65 (moderate)
SART Ommission	39.00% (27.01)	37.90% (25.60)	0.62 (moderate)
Picture Recognition Hits	96.15% (6.40)	96.18% (6.95)	0.19 (poor)
Picture Recognition Source Memory	85.75% (15.52)	87.38%(14.46)	0.74 (moderate)
DSB	5.78 (1.76)	5.88 (1.72)	0.65 (moderate)
SWM	5.68 (1.45)	5.76 (1.33)	0.55 (moderate)
Reaction Time	351.81ms (69.92)	342.38ms (70.71)	0.72 (moderate)
VST Starting Deviation	36.18% (13.11)	35.04% (14.17)	0.75 (moderate)
VST Final Deviation	6.44% (1.66)	6.53% (1.43)	0.10 (poor)
PHQ-9	3.34 (4.09)	3.00 (3.81)	0.87 (good)
GAD-7	2.62 (3.64)	2.57 (3.40)	0.86 (good)
AQ	52.12 (15.62)	53.24 (14.70)	0.94 (excellent)
BIS-15	27.52 (5.59)	26.85 (5.67)	0.95 (excellent)

The ICC values demonstrate moderate to excellent test-retest reliability between first and second assessment sessions, with the exception of the Picture Recognition Hits outcome and the Virtual Supermarket Final Deviation outcome. This can be attributed to the low variation in these assessment results, where results demonstrate that participants typically perform very well on these assessments. This low variability subsequently effects the calculation of the ICC.

Additionally, initial examination of the results from the SART assessment identified a distribution of results which suggested that further examination was warranted before including these results within the primary analysis. The decision to examine the data further was based on a higher frequency of participants scoring 0% on assessments than I would have expected from the dataset. This examination is described in further detail in Appendix H. In summary, the findings from these analyses suggest that the data from SART outcomes shows high variability, with the majority of data demonstrating poor performance on the task. Outcomes from the SART show only moderate test-retest reliability, suggesting that poor performance on the first set of assessments may reflect the novelty of the task. Furthermore, performance on the SART shows poor correlation with other assessments of executive function and reaction times. This suggests that results from the SART may more likely reflect participants overemphasising the speed of reactions rather than the accuracy of responses, as supported by other investigations (163). Therefore, the results from the SART assessment in this test battery may not accurately reflect executive function enough to include as a valid measure of this domain. For these reasons, the SART data was not included in the primary analyses in this study.

Comparison of group means

To isolate the effect of level of exposure to contact sport, all outcomes were compared between males who participated in amateur or professional contact sport. Group sizes allowed a comparison between females with a history of contact sport and females with a history of non-contact sport. Findings from this analysis are described in Figure 17.

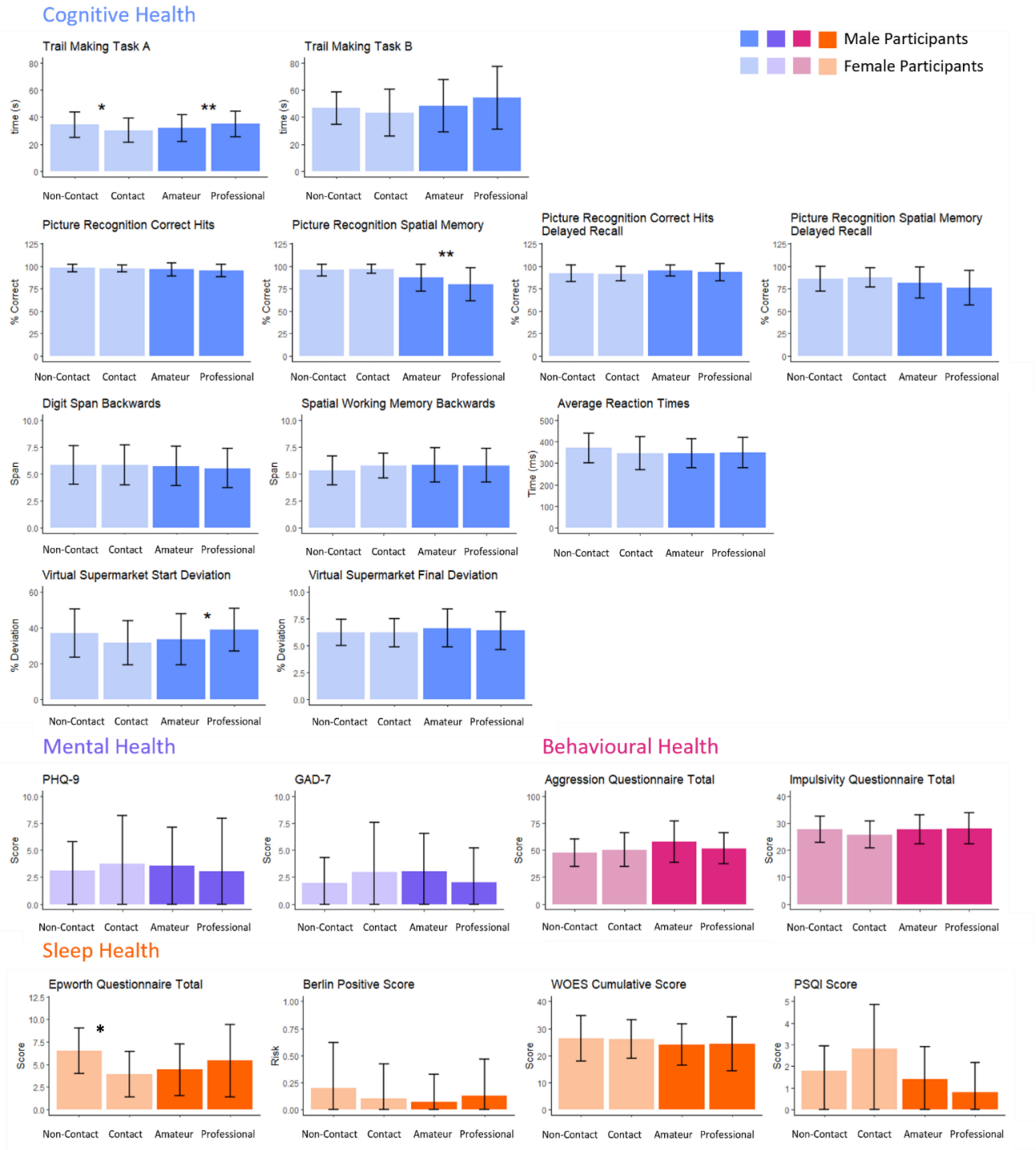


*Figure 17: Group means across cognitive, mental health, behaviour, and sleep outcomes. Lighter coloured bars represent female participants. Darker coloured bars represent male participants. * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001*

Across nearly all cognitive assessments, means reported in Figure 17 suggest that male former professionals performed worse than the male former amateur sample. This difference was statistically significant in measures of attention (TMTA, amateur mean = 31.93s, sd = 9.97, professional mean = 35.60s, sd = 9.45, p = 0.004) source memory (PRSM1, amateur mean = 87.02%, sd = 14.98, professional mean = 79.55%, sd = 18.45, p = 0.006) and spatial navigation (VSTS, amateur mean = 33.48%, sd = 14.18, professional mean = 38.75%, sd = 11.90, p = 0.038). Former professionals showed better mental and behavioural health, by demonstrating lower

scores on measures of depression, anxiety, aggression, and impulsivity. Although sample sizes were small, trends indicate that former professionals demonstrated worse sleep health than former amateurs did, except for in the Pittsburgh outcome.

Female participants with a history of contact sport performed better on most cognitive assessments than non-contact sport females. This was significantly different in the measure of attention (TMTA, contact mean = 30.23s, sd = 8.82, non-contact mean = 34.43s, sd = 9.34, p = 0.025), and demonstrated by non-significant trends in measures of executive function (TMTB, contact mean = 43.22s, sd = 17.25, non-contact mean = 46.67s, sd = 11.88, p = 0.052), working memory (SWM, contact mean = 5.74, sd = 1.16), non-contact mean = 5.33, sd = 1.34, p = 0.37), processing speed (RT, contact mean = 346.00ms, sd = 75.94, non-contact mean = 371.03ms, sd = 68.86, p = 0.067), spatial navigation (VSTS, contact mean = 31.63%, sd = 12.31, non-contact mean = 36.96%, sd = 0.11). Contact sport participants demonstrated higher levels of depression and anxiety, higher levels of aggression and lower levels of impulsivity. Contact sport participants demonstrated significantly better sleep on the Epworth measure contact mean = 3.9, sd = 2.56, non-contact mean = 6.5, sd = 2.55, p = 0.035) and non-significant trends indicated signs of better sleep health in other measures apart from the PSQI outcome.

Adjusting for confounding factors

To determine the influence of other factors on brain health, simple linear regressions were performed. Table 10 describes these regressions.

*Table 10: Summary of regressions and adjustment outcomes. * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001*

Variable	Outcomes signifanctly effected (coefficient)	Adjustment Outcome
Age	TMTA (0.34***) TMTB (0.91***) PRH1 (-0.09*) PRH2 (-0.11*) PRSM1 (-0.21*) PRSM2 (-0.35**) SWM (-0.042***) RT (2.14***) VSTS (0.42***)	When adjusted for, male professionals performed significantly worse on the TMTA and PRSM1 outcomes, and results from the VSTS were no longer significantly different in males. In the female comparison, the difference in TMTA was no longer signficantly different, and results from the PRH1 outcome showed significantly worce performance than contact sport athletes.
Sex	TMTB (4.55*) PRH1 (-1.79*) PRSM1 (-7.13***) PRSM2 (-7.56***) AQ (5.9*) PSQI (-1.16**)	NA - Sex already adjusted for in analysis.
Education	PRH1 (0.25*) PRH2 (0.25*) PRSM1 (0.82**) PRSM2 (0.71*) DSB (0.076*) SWM (0.087***) RT (-4.40***)	No change to significance status of outcomes.
Health Comorbidities	TMTA (2.96*) SWM (-0.51**) AQ (5.83*)	No change to significance status of outcomes.
Treatments	TMTB (6.06*)	No change to significance status of outcomes.
Mental Health Disorder or Learning Difficulty/Disability	TMTA (3.00**) PHQ (4.03***) GAD (3.46***) AQ (12.85***)	No change to significance status of outcomes.
Diagnosed Concussions	TMTA (0.38*) BIS (0.39***) AQ (1.36***)	No change to signifcance status of AQ for females, TMTA or BIS-15 When adjusted for, male professionals scored significantly higher on the AQ than male amateurs.
Suspected Concussions	BIS (1.46***) AQ (5.16***)	No change to significance status of outcomes.
Other Brain Injuries	TMTB (12.93**) VSTS (8.19*)	No change to significance status of outcomes.
Smoking History	PRH1 (-2.01**) PRH2 (-1.94*) BIS (2.05*) AQ (5.91**)	No change to significance status of outcomes.
Length of exposure to sport	SWM (0.016*)	No change to significance status of outcomes.

Variables that were significantly associated with outcomes included age, sex, education, health comorbidities, treatments, the presence of a mental health disorder or learning disability/difficulty, number of concussions, other brain injuries, smoking history, and length of exposure to sport. Variables that were not significantly associated with any outcomes included decile of multiple deprivation, current physical activity, and age of first exposure to main sport. Multiple linear regressions were then performed to adjust for confounds. The outcomes of these adjustments are described in Table 10.

Number of concussions were analysed as potential confounds rather than as primary variables, given that the number of diagnosed and suspected concussions were self-reported rather than validated with medical records and are therefore subject to recall bias (164). Number of diagnosed concussions was significantly associated with results from the TMTA and the impulsivity and aggression questionnaires, whereby a higher number of diagnosed concussions predicted worse performance on the TMTA and higher levels of impulsivity and aggression. After adjusting for the number of diagnosed concussions in these assessments, results in the TMTA remained significantly different for both male and female comparisons. Group differences became significant in the aggression questionnaire when comparing male amateurs and professionals, whereby professionals exhibited significantly lower levels of aggression. Group differences remained non-significant in impulsivity results for male and female comparisons, and in the aggression results for the female comparison. Number of suspected concussions was significantly associated with higher levels of aggression, impulsivity, and poorer sleep on the WOES outcome. After adjusting for number of suspected concussions, group differences in these assessments remained non-significant in both male and female comparisons.

Reporting another brain injury was associated with significantly worse performance on the TMTB and the spatial navigation egocentric outcome. After adjusting for this variable, group differences remained non-significant in both male and female comparisons.

As described in Table 10, age was found to be a significant predictor across the cognitive assessments, except for the digit span backwards assessment. In all other cognitive assessments older age was associated with poorer cognitive performance. As described in Table 10, after adjusting for age in the male comparison, TMTA and PRSM1 results differed significantly. However, results from the spatial navigation egocentric outcome were no longer significantly different between groups. After adjusting for age in the independent t-tests comparing female contact sport athletes with non-contact sport athletes, TMTA results no longer differed significantly between groups. However, results from the picture recognition correct hits outcome became significantly different, whereby contact sport athletes showed worse performance than non-contact sport athletes.

Sex was a significant predictor of results from the TMTB, the correct hits, source memory and delayed recall source memory outcomes from the picture recognition task, and the aggression questionnaire. Regressions showed that males demonstrated worse cognitive performance on these assessments and higher levels of aggression. Given that the analysis already split participants by sex, this variable was not further adjusted for.

Education was found to be a significant predictor of results from the picture recognition assessment outcomes, the digit span backwards assessment, and the spatial working memory assessment. In these assessments, more years of education was associated with better

performance across each assessment. After adjusting for education, the source memory outcome of the picture recognition task remained significantly different between male former professionals and former amateurs, whereby former professionals continued to perform worse on the assessment. In females, no significant differences continued to be observed.

The presence of another comorbidity was found to be a significant predictor of results from the TMTA, the spatial working memory assessment, and the aggression questionnaire whereby this variable predicted worse cognitive performance. When adjusted for in these assessments, group differences remained significant for both male and female comparisons in the TMTA, and non-significant in the spatial working memory assessment and the aggression questionnaire.

Reporting a treatment that might influence dementia risk was found to be a significant predictor in the TMTB, whereby this variable predicted worse performance. When adjusted for in the TMTB, group differences remained non-significant for both male and female comparisons.

The presence of a mental health disorder or learning difficulty/disability was found to be a significant predictor in the TMTA, PHQ-9, GAD-7, and aggression questionnaire. Reporting a mental health disorder or leaning difficulty/difficulty was associated with poorer performance on the TMTA and increased levels of depression, anxiety, and aggression. When adjusted for, group differences remained significant in the TMTA for both male and female comparisons. Group differences remained non-significant in the PHQ-9, GAD-7, and aggression questionnaire, with the exception of the level of aggression in males. In this comparison, male former professionals exhibited significantly lower levels of aggression.

Smoking history was associated with significantly worse performance on the picture recognition correct hits and source memory outcomes. This variable was also associated with a higher level of aggression and impulsivity. After adjusting for smoking history, group differences remained non-significant in both male and female comparisons.

The length of exposure to either contact sport or non-contact sport was associated with significantly better performance in the spatial working memory task, whereby longer exposure predicted a better span. When adjusted for, group differences remained non-significant in both male and female comparisons.

Normative comparison

Cognitive results from contact sport participants were compared to a normative sample matched by age and sex. The findings from male and female participants are described in Figure 18.

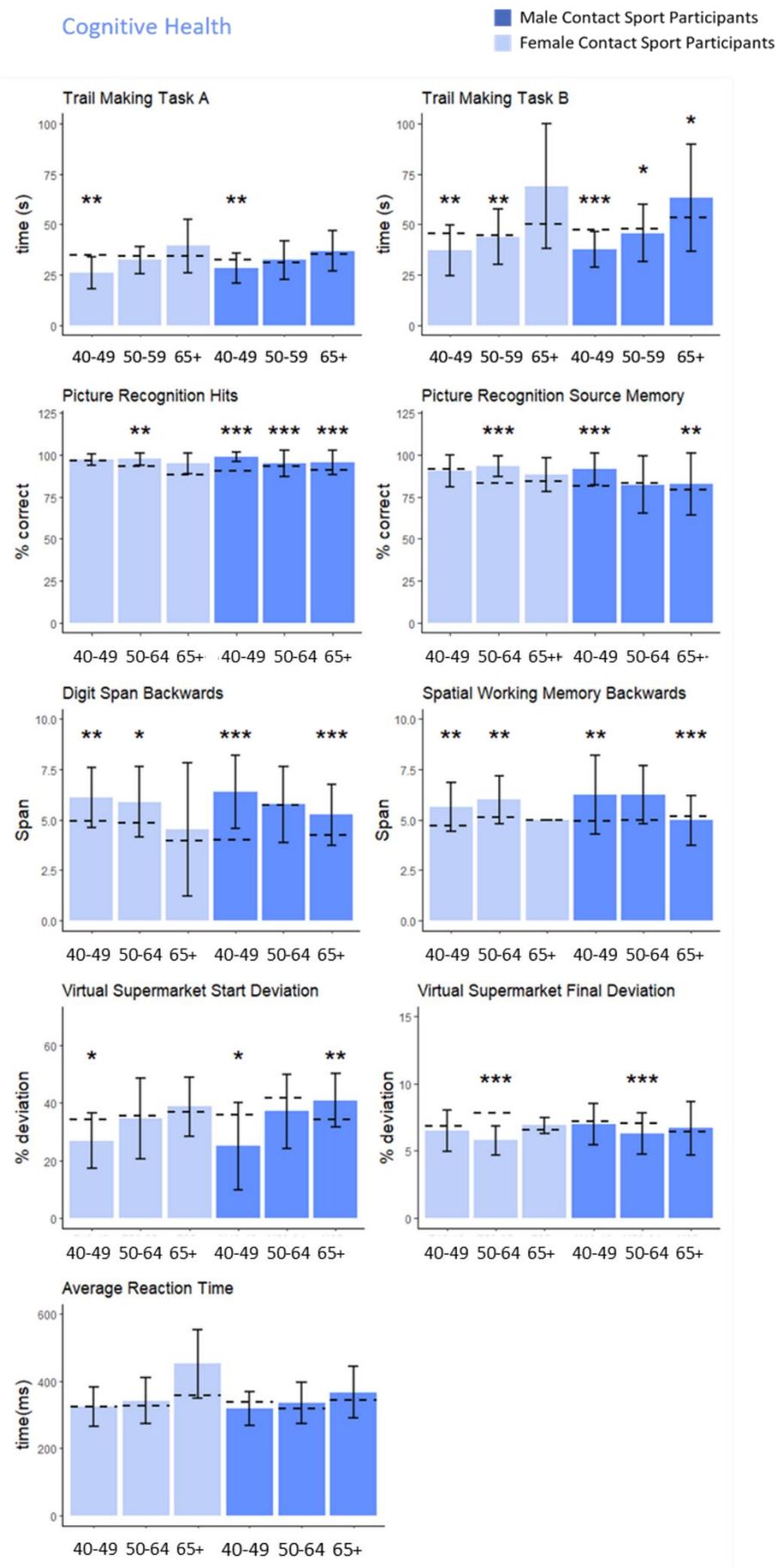


*Figure 18: Results from Normative Comparisons. Plotted bars represent participants with a history of contact sport, segregated by age group. Lighter coloured bars represent females. Darker coloured bars represent males. Dotted lines represent the normative mean. * = p ≤ 0.05, ** = p ≤ 0.01, *** = p ≤ 0.001*

In the youngest group (40-50 years), males performed significantly better in measures of attention (TMTA, male mean = 28.21s, sd = 7.31, normative mean = 32.43s, sd = 15.27, p = 0.005), executive function (TMTB, male mean = 37.56s, sd = 9.04, normative mean = 47.15s, sd = 19.62, p < 0.001), working memory (DSB, male mean = 6.38, sd = 1.83, normative mean = 4.02, sd = 1.83, p < 0.001 and SWM, male mean = 6.24, sd = 1.95, normative mean = 4.91, sd = 1.98, p = 0.0053) and spatial navigation (VSTS, male mean = 24.95%, sd = 15.05, normative mean = 35.75%, sd = 11.08, p = 0.019). In this age group, participants demonstrated a trend of faster reaction times than the normative sample (male mean = 318.33ms, sd = 49.6, normative mean = 337.42ms, sd = 69.34, p = 0.093), although this was not significant. Males in the oldest group (65+) demonstrated a converse trend, performing significantly worse than the normative sample in executive function (TMTB, male mean = 63.26, sd = 26.62, normative mean = 53.44, sd = 16.85, p = 0.024) and the spatial navigation assessment (VSTS, male mean = 40.86%, sd = 9.34, male mean = 34.40%, sd = 13.26, p = 0.001). Males recorded poorer results in measures of attention (TMTA, male mean = 36.88s, sd = 10.03, normative mean = 35.35, sd = 7.39, p = 0.51), working memory (SWM, male mean = 4.96, sd = 1.25, normative mean = 5.17, sd = 1.48, p = 0.12) and reaction time (male mean = 366.48ms, sd = 76.88, normative mean = 344.28, sd = 73.81, p = 0.16), although these differences were not statistically significant. Results from the measure of recognition and source memory assessment demonstrate that males performed significantly better than the normative sample in nearly all age groups.

Similar trends were observed between female participants and an age matched normative sample. In the youngest age group (40-50 years), SCORES females demonstrated significantly better performance in measures of attention (TMTA, female mean = 26.0s, sd = 7.87, normative mean = 34.87s, sd = 7.54, p = 0.002), executive function (TMTB, female mean = 37.21s, sd = 12.47, normative mean = 45.69s, sd = 8.46, p = 0.008), working memory (DSB, female mean = 6.11, sd = 1.50, normative mean = 4.93, sd = 2.41, p = 0.005 and SWM, female mean = 5.64, sd = 1.22, normative mean = 4.72, sd = 1.98, p = 0.006) and spatial navigation (VSTS, female mean = 26.81%, sd = 9.58, normative mean = 34.23%, sd = 13.09, p = 0.049). Sample sizes in the oldest age group were too small for meaningful inferential statistics.

Discussion

This investigation demonstrated domains of cognitive, mental, and behavioural health where groups exposed to sport related RHI through contact sport demonstrated significantly worse outcomes than control groups. Where results were not significant, means demonstrated trends that suggest sequelae of participation in contact sport on brain health outcomes.

Cognitive Health

In male former contact sport athletes, results demonstrated a trend whereby former professionals performed worse across nearly all cognitive domains than the former amateur participants. Female contact sport athletes were found to perform better than female non-contact sport athletes on most cognitive assessments, however this may reflect the age difference between groups. Comparisons of cognitive results between former contact sport players and a normative sample demonstrate that younger former contact sport athletes performed better than the normative sample on assessments, whereas older contact sport athletes performed significantly worse across some of the cognitive domains. This may reflect the positive benefits of participation in sport on brain health, through improved cardiovascular function (16), reduction of comorbidities (165), promoted neurogenesis and neuroplasticity (18), or improved social networks (19), but also potentially a detrimental consequence of contact sport on later brain health. This analysis identified attention, executive function, spatial navigation, and source memory as cognitive domains impaired in former contact sport players. Executive dysfunction and impaired memory and attention were identified by previous investigations as domains where long-term effects of contact sport participation can be observed (147) Although impairments in visuospatial processing were examined by previous investigations (147), deficits in spatial navigation remain relatively unexplored in contact sport athletes despite representing a cognitive domain effected in the early stages of dementia (141).

Mental Health

In measures of mental health, male former professionals reported lower levels of depression and anxiety. However, female contact sport athletes reported higher levels of depression and anxiety than female non-contact sport players. This difference in outcomes may be influenced by other protective factors on mental health in the demographics of the male former professionals, such as increased socioeconomic status (166). Previous work (62) using health record information found that incidence of hospital admissions for common mental health disorders was lower in male former professional contact sport athletes. However, a review of evidence (167) found an increased prevalence of mental health symptoms in former contact sport athletes compared to appropriate controls. A previous study (80) comparing amateur to elite rugby players has also demonstrated a higher prevalence of symptoms associated with poor mental health in elite level players. My findings in female contact sport athletes supports this, whereby contact sport athletes demonstrated higher levels of mental health symptoms. An analysis of the association of current physical activity and level of deprivation on mental health outcomes was performed, but no significant association was found.

Behavioural Health

Behavioural measures demonstrated higher levels of aggression and impulsivity in amateur males and contact sport females. Aggression and impulsivity have previously been cited as early

symptoms of dementia in contact sport athletes, based on informant reports linked to autopsy-confirmed cases of chronic traumatic encephalopathy (CTE) (168). However, previous work in males has also suggested that levels of aggression are not different to levels within the general population and could affect misdiagnosis of dementia (169). My findings support that aggression and impulsivity may be affected as a consequence of exposure to contact sport, particularly in association with number of lifelong concussions, and therefore should be considered when mapping lifelong neuropsychological consequences.

Sleep Health

Measures of sleep were included at a later stage of recruitment, and for that reason the sleep results are limited by smaller sample sizes than other measures in this study. Sleep health was not reported to be significantly different between group comparisons, except for contact sport females reporting better sleep health according to the WOES measure. Number of suspected concussions was identified to be significantly associated with sleep health as measured by the WOES questionnaire, which is congruent with previous findings surround the influence of exposure to contact sport on sleep health (88). This is an area of brain health that requires further study with sufficient sample sizes.

Confounds and Adjustments

This study identified several predictor variables that are important to consider and adjust for when investigating long-term neuropsychological consequences. As expected, age was negatively associated with most cognitive outcomes. Age and education were associated with correct hit outcomes of the picture recognition assessment where participant results were largely skewed towards scores of 93-100%, as well as the digit span backwards or spatial working memory assessment where there was low variability in results. This suggests that these measures may not be sensitive enough to demonstrate cognitive differences between exposure groups but can act as important markers for identifying participants who score significantly below other participants similar to them in age and level of education. The regression analysis identified sex as a significant predictor of attention, working memory, source memory and aggression which suggests that there may be a difference in neuropsychological in some domains between males and females exposed to contact sport. Further research with appropriate comparison groups is required to explore this relationship further.

Previous research has explored a dose-response relationship of exposure to concussion or sport-related RHI on neurological outcomes (29). My investigation found significant associations between exposure to contact sport and performance on working memory, and between number of concussions and attention, impulsivity, and aggression. However, for the majority of outcomes number of diagnosed or suspected concussions was not associated with performance on assessments. The analysis of concussion data was included in order to consider the influence that number of concussions may have on results, but concussion data was not included as a primary

variable in this investigation. Given that I found little association between concussions and the outcomes, and that the emphasis of this analysis was to assess the long-term consequences of sub-concussive injury in contact sport, the analysis of concussion data was not further investigated in this analysis. However, understanding why concussions aren't related to outcomes in most of the data, but exposure to contact sport is, represents a possible future avenue of investigation.

This investigation did not directly compare male and female athletes within the analysis but did find that sex was a significant predictor of performance on some assessments whereby females demonstrated better performance on assessments of executive function (TMTB), memory (PRH1, PRSM1, PRSM2), lower levels of aggression (AQ) and worse sleep (PSQI) than males. However, this analysis was not broken down by level or type of sport and adjusted for in the analysis. This choice was made because of the group sizes available within the initial sample. The sample included a large number of male professional athletes, but a low number of female professional athletes, and the sample included a large number of female non-contact athletes, but few male non-contact athletes. For this reason, two different exposure vs. control group analyses were used. Had the sample included sufficient non-contact athletes or professional female athletes, a more direct comparison could have been possible to factor investigate the effect of sex on group differences between exposure and control groups.

This limitation also raises an important point about the validity of comparing professional male athletes with professional female athletes. The study sample only included six professional female athletes, compared to 77 professional female athletes, despite females making up over 1/3 of the study population. This discrepancy may be due to 'professional' being an inappropriate categorisation of level of sport for female athletes in this cohort. Professionalisation of women's sport has only occurred recently, with the establishment of the first professional women's football league in the UK occurring in 2018 (170). Given that participants in this study must be aged over 40, the term 'professional' most likely does not capture females who played at the highest level of their sport and would therefore have been exposed to the highest levels of sport-related injury. A better definition is needed to accurately categorise levels of exposure within women's sport for the purpose of research and improve the representativeness of the SCORES dataset. To address this, I performed a literature review of how female athletes have previously been defined in research and proposed a novel definition framework in the review chapter of this thesis.

Further Objectives of the Longitudinal Study

Findings from this study are from initial data of an ongoing longitudinal study, that monitors a cohort of athletes as they age. The future objectives of this study are to closely examine the sex differences between male and female athletes in longitudinal sequelae with appropriate comparison groups, to improve targeted recruitment efforts for a more representative sample, and to continue to monitor brain health outcomes of participants over time. The purpose of this

monitoring will not to be to measure individual change, but rather to use the baseline data from this first dataset to observe changes in groups with higher exposure to RHIs through contact sport.

The findings from this study will inform the methods used to achieve these objectives of the longitudinal study in future. Findings from this study support the appropriateness of this online longitudinal method as an approach to address key research questions within the field, by demonstrating the ability to produce meaningful comparisons. Furthermore, this study supports the validity of assessments used in this study battery by demonstrating moderate to excellent test-retest reliability, and by examining the appropriateness of the SART assessment. Ultimately, the SART was deemed ineligible to include within this testing battery. This was not initially expected, given the inclusion of Go/No-Go tasks in other neuropsychological online batteries (171). However this was ultimately consistent with other assessments of validity of the SART, which advise caution when using the assessment as a measure of sustained attention (163, 172). Specifically, these investigations suggest that the SART measures the decision to respond quickly to stimuli rather than the attention paid to stimuli and non-stimuli. Combined with qualitative feedback from participants that the SART assessment caused frustration, results from this study were used to inform the decision to remove the SART assessment from the study battery.

The virtual supermarket assessment was also considered in terms of appropriateness as part of the study battery following analysis in the study. This assessment takes approximately 20 minutes to complete and consists of 14 trials. Qualitative feedback demonstrated that participants suggested that this assessment was draining and caused frustration. These were important qualitative accounts to consider, given that retention is necessary for the success of longitudinal studies. The approach to monitor spatial navigation is novel within the field of athlete brain health research, but represents an important domain that is effected in the very early stages of dementia (141). Given the importance and novelty of the task, as well as the potential burden on participants, analysis of the results produced in this early dataset was important towards informing the inclusion of the assessment in the battery going forward. Overall, findings from this study demonstrate that the collection of both allocentric and egocentric spatial navigation abilities offers meaningful results that demonstrate differences in brain health in participants by level of exposure to contact sport and supports the inclusion of this measure in the battery. Significant group differences could be observed in the measure of egocentric navigation, which offers valuable insights into the domains effected by exposure to contact sport. Results from the allocentric measure demonstrate a ceiling effect, in a similar nature to the results from the picture recognition tasks, and further considerations about the level of difficulty of the task could be made. However, results from this study overall support the use of this measure as an important and novel approach to achieve the future objectives of the longitudinal study.

Conclusion

Using novel digitised measures of cognitive, mental, behavioural and sleep health, this investigation identifies domains of cognitive, mental, behavioural and sleep health that may be affected later in life as a consequence of participation in contact sport. This study is unique in that it considers long-term sequelae in male and female contact sport athletes, as well as professional and amateur athletes. Findings from this study demonstrate early benefits of participation in sport on brain health, but worse outcomes in older contact sport participants and by levels of exposure to sport. This study emphasises that amateur contact sport players may not be exempt from long-term neuropsychological consequences, and that inclusion of amateurs and female athletes in future research is important.

Study 4 - Concurrent and retrospective acceptability of participation in the SCORES Project after 2 years

Abstract

An aim of the longitudinal study is to collect at least 10 years of data. In order for this to succeed, retention of sufficient participants is vital to analyse long-term trends in brain health. Reducing attrition is a challenge in longitudinal studies, and so it was important to understand participants' perceptions of continued acceptability of the study to inform methods for increasing retention.

In this qualitative study, participants who had completed five repetitions of assessments were invited to register for an optional interview. These semi-structured interviews were designed to explore the experience of participants in the study based on Sekhon's acceptability framework (102). Interviews took place online and responses were analysed using framework analysis.

Seven participants were interviewed. Responses demonstrated good continued acceptability of participation in the study with some areas for improvement in the methods. Participants reflections on motivation were congruent with findings from Study 1, suggesting that promoting altruistic and self-serving (such as the opportunity to receive feedback) benefits to participation could promote retention. Participants continued to consider measures used in the study and the feedback process in line with their expectations of ethicality.

This study concludes that the study format is appropriate for encouraging long-term participation, and that the inclusion of a feedback process is important as a motivator to continue in the study whilst meeting ethical expectations.

Introduction

The success of longitudinal studies in meeting their objectives is dependent on the ability to retain participants for the course of study. Cognitive change in a non-clinical population is slow and subtle, particularly in a cohort with participants as young as 40 years of age. Therefore, the SCORES project aimed to run for at least 10 years to collect a database of participants and monitor them over time. Although the initial acceptability study explored the experience of participants at the beginning of the project, there was more to learn from participants who were further into the study and able to reflect on their experience.

The purpose of assessing concurrent or retrospective acceptability in a longitudinal project is to understand perspectives of participation following some exposure to the assessment and with the intention of further exposure (102). Over the course of the study, mental health and behaviour questionnaires were introduced to the protocol and the feedback process was further developed. Previous studies have evaluated the acceptability of monitoring mental health and behaviour online. A study in 2022 (173) found that the PHQ-9 was assessed to meet satisfactory acceptability, however participants reported that the number of responses were limited and that the questions refer to general characteristics of depression rather than situational characteristics of depression, such as those relevant to a certain disease. An investigation into the acceptability of completing daily online mental health questionnaires online with immediate automated feedback suggested that regular introspection could have a negative effect on affective attitude for some participants, whilst others found it a helpful experience and motivation to improve their current mood (174). These investigations highlight that attitude towards the assessments and how feedback is presented might influence perceived acceptability of participation. Missing from this literature, and relevant to longitudinal studies, is how perceived acceptability of online mental health measures might change over time with longer intervals between assessment.

The initial acceptability study demonstrated that participants considered the opportunity to receive feedback on their brain health an important motivation to participate, and a key positive outcome of the study. However, findings from this study also identified that in providing feedback on brain health the clarity of feedback was important, as well as meeting expectations of duty of care. To meet these expectations, the development of the feedback forms involved PPI members including participants from the initial acceptability study and a lengthy ethical approval process. This process was vital to developing a procedure that was hoped to be acceptable in terms of ethicality at the time. However, this assessment could only estimate what would be acceptable, and therefore an assessment of the concurrent and retrospective acceptability of receiving feedback was of interest.

In addition to these changes to the SCORES protocol, between the timepoint of the initial acceptability study (November 2020) and interviews for the current study (May 2023) social

distancing regulations (due to Covid-19 in 2020 and 2021) were relaxed which may have influenced participants' attitudes towards remote assessment. Therefore, the concurrent and retrospective acceptability of participation warranted further study at this stage.

The aim of this study was therefore to understand the retrospective and concurrent acceptability of participation in a longitudinal online study from the perspective of participants. In particular, the participants experience of online cognitive testing, the acceptability of online mental health and behaviour questionnaires, and the process of receiving feedback on their brain health were explored.

Methods

Design

This qualitative study involved semi-structured interviews with participants who had completed two years of participation in the SCORES project. The interviews were designed to explore experience as a participant in the project based on Sekhon's acceptability framework (102). According to this framework, acceptability is a multifaceted construct that reflects the extent to which people delivering or receiving a healthcare intervention consider it to be appropriate, based on anticipated or experiential cognitive and emotional responses to the intervention. This framework assesses the affective attitude, burden, ethicality, coherence, opportunity costs, perceived effectiveness and self-efficacy of an intervention or study (102).

Potential participants for this study included existing SCORES participants who had completed their fifth round of assessments. When recruitment began, this pool included 81 participants (Male = 52, Female = 29). We aimed to recruit a sample of approximately 8 to 12 participants, varied in sex, age, level of education, ethnicity, and sporting background. Ethical approval was granted by the University of East Anglia Faculty of Medicine and Health Science Ethics Committee in February 2023 (Ref: ETH2223-1392).

The interview protocol is included in the appendix (Appendix J). The interview protocol was developed based on emergent themes from the acceptability study, validity study and findings from the early analysis of SCORES data, as well as the process of applying for ethical approval for the feedback forms. Following the acceptability study (Study One), it was important to understand whether the perceived burden and ethicality of the protocol continued throughout participation, so I included questions about continued perceived burden and asked participants to reflect on their perception of the ethicality of measures they completed and the feedback process. Findings from the validity study about finding assessments easier the second time participants completed them, as well as results from test-retest reliability in Study Three, informed the development of questions about whether participants found assessments easier as they repeated them and whether they felt their performances continued to effectively reflect their abilities on the tests. Furthermore, reflections from the process of applying for ethical approval for the feedback

informed the inclusion of further questions about the feedback process and actions taken within SCORES when participants showed worsening performance. The ethical approval had been based on findings from Study One, and therefore it was important to understand whether participants continued to perceive this process as in line with their expectations of ethicality.

Participants who completed their 5th round of assessments were invited to register their interest in taking part in an optional interview to discuss their experiences in the project. The registration of interest form included a brief description of the purpose of the interviews and what participation would involve, and clearly stated that the purpose of the interviews was not to discuss their results. It was also clearly stated that participation in this interview was not mandatory, and that participants who chose not to participate could continue in the SCORES project as usual.

Participants who registered their interest were asked to read the participant information sheet for this study and sign an online consent form. Upon receiving the signed consent form, an online interview was arranged.

Interviews took place over Zoom or a telephone call depending on the participants preference. Interviews lasted approximately 20 minutes and were recorded using Zoom's recording software or a voice recorder. The recordings were stored in a secure OneDrive file accessible by SH and me. The interviews were listened several times with notes taken, to rather than verbatim transcription, in order to deeply engage with what participants said and also how they said it. Quotations that were relevant to the key topics in the interview topic guide were extracted for analysis.

Analysis

Quotations extracted from the interviews were analysed using framework analysis, which is described in the methods in Study 2. I performed data familiarisation by taking notes on the quotations and revisiting the audio recordings where necessary. I then developed a thematic framework based on a combination of a priori and emergent themes. Specifically, responses were organised by overall themes relating to the TFA that were self-efficacy, perceived effectiveness, burden, opportunity costs, ethicality of mental health and behaviour questionnaires, and ethicality of the feedback process. Within these themes, responses were grouped by sub-themes that emerged (i.e. within self-efficacy, sub-themes were perceived difficult and changes in difficulty over time). Once the framework was identified, the framework was applied to all the quotations which were then charted to index the quotations. I then mapped the indexed data and interpreted it to develop a narrative of the findings. The framework and indexed data were then discussed with SH to further develop the narrative of the findings.

Findings

A total of 16 participants registered their interest in the interviews. From these, three were excluded because they had already taken part in the original acceptability interviews (Study 1), four did not complete the consent form after they registered interest, and two did not respond when invited to arrange the interview. In the end a total of seven participants took part in the interviews. Table 11 summarises their demographic characteristics.

Table 11: Demographic characteristics of participants that took part in this study.

Participant ID	Age at time of interview	Sex	Highest Level of Education	Sporting History
P1	56	Male	O-Levels	Professional Football
P2	55	Male	A-Levels	Amateur Football
P3	81	Male	Doctoral Degree	No Sporting History
P4	61	Female	Masters	Amateur Hockey
P5	55	Male	Masters	Professional Football
P6	64	Male	Masters	Amateur Football
P7	82	Male	O-Levels	Amateur Football

Analysis of responses from participants using framework analysis revealed that six constructs in the acceptability framework were explored including self-efficacy of completing the assessments, perceived effectiveness of the assessments, burden, opportunity costs of participation (how participants might benefit from taking part), and ethicality of the introduction of mental health questionnaires and feedback process.

Self-efficacy

Participants were asked to describe their perception of the difficulty of the assessments. Evaluations of the difficulty were mixed:

> *"I find some of it quite hard it I'm being brutally honest." (*P1*)*
>
> *"It's a mixture of easy and difficult." (P3)*
>
> *"I'm still not sure whether it's because I'm being old and slow or whether they are just tricky to everyone." (P6)*

When prompted to discuss which assessments they found most difficult and why, most responses referred to the virtual supermarket assessment:

> *"The only one I find tricky is the supermarket one. On my charts it says I'm doing alright but if you asked me if I was doing alright, I'd say no." (P4)*
>
> *"I find the shopping one difficult because you're thinking well where did I actually start and what angle am I looking at. So, there's probably more technical thinking in that one but most of them are straightforward." (P5)*

> *"I find the supermarket one difficult. If I had a paper and pencil and could make notes it would be easier, but based on memory it is difficult. More difficult than you realise." (P7)*

Others described the picture recognition assessment, particularly the delayed recall condition, and the digit span backwards assessment as challenging:

> *"The one where the characters appear in the top bottom left or right, and you have to remember those is ok, but when it gets really tricky is when that reappears. That was for me the most difficult one." (P6)*

> *"I do find the memory one where you have to recall the numbers in reverse difficult. I find that probably when I get to halfway, I start to struggle." (P2)*

When discussing the difficulty of the assessments one participant commented on if the assessments felt easier over time:

> *"(Over time) it's easier because you know what's coming, but I don't think it's become easier to do because you have done it more often." (P1)*

Perceived effectiveness

Participants were prompted to reflect on their initial expectations of the assessments and to discuss if their experience met their expectations. For many, their understanding of dementia informed their initial expectations of the types of assessments that they would complete:

> *"My father having had dementia; I can see the reasoning for all the tests. For example, the picture recognition one. He would not have been able to do it early on." (P4)*

> *"Having had my father and my mother-in-law who have been through dementia I sort of understood a little bit about it. So, I understand the logic behind the supermarket test and the numbers really." (P1)*

> *"The mental health one – I was quite surprised when I first got it. But seeing it from both sides like my side and someone who I knew had dementia, they always seemed to diagnose depression first. That's what they did with my dad." (P4)*

> *"The behaviour one is quite an interesting one. I think your behaviour changes so seeing that one as well I think you get more angry at little things. So, I can see why you would want to pick up on those quite early." (P4)*

They also described their perception of the project as a research study rather than a health check which reflects good intervention coherence:

> *"It's not like a health assessment where it feeds to my GP. So, I've always known it is part of a study and it's long-term and whatever information we get that picture will just grow and grow for an individual and as a group." (P5)*

> *"I don't think I was under any allusions that you'll do this for a year and there's going to be a solution at the end." (P5)*

The effectiveness of the assessments was discussed in terms of how participants felt the assessments challenged them:

> *"It's really clear to me why the tests are set up how they are." (P4)*

> *"I would have thought they are good examples of testing memory and I have to say even the supermarket one – now I'm actually focussing on where I am in relation to the start it adds another dimension to the thinking." (P2)*

Participants also reflected on how the feedback reinforced their perception of the effectiveness:

> *"One of them is reaction time and for me that is the worse score that I have. However, I'm using an old laptop. That may well be part of why I have a low score." (P1)*

> *"The tests have highlighted to me, not necessarily deterioration, but perhaps changes in the things I remember or don't remember." (P5)*

Burden

Participants generally described that the time burden associated with completing the assessments was an acceptable burden:

> *"I have recently retired but even when I was working with a busy job, I didn't find it a particular burden." (P2)*

> *"I have no problem with it." (P3)*

> *"I think sometimes finding the time and space to do it in your own house is an occasional challenge…. If I can't find thirty minutes in three months, it's slightly ridiculous." (P5)*

Discussions of the burden and ways to promote retention in the project also revealed an assessment of the acceptability of participation in the context of their original motivations:

> *"The driver for me is my own health in that I think that if something was so ridiculously wrong you would say something to somebody, or I would get feedback saying go and do something…(and) I certainly am quite happy to give 30 minutes of my time every 3 months so that people after me maybe have a better chance of living with (dementia) or not getting it." (P1)*

> *"I joined up because I thought it would be helpful and it's also nice for me to get feedback over time." (P2)*

> *"I intend to stick with it for as long as I am able to and you are running the project to be honest with you, in an attempt to make sure that there is a consistent set of data at least for one person." (P6)*

Opportunity costs

In summary of their experience participating in the project, participants described key positives of participation. These included the ability to monitor their brain health:

> *"There are fairly well known age-related cognitive shifts so there is no harm in me having an anchor on those." (P3)*

> *"It's good that there is some form of generic feedback where you sit against quartiles, and I'm reassured by the fact that if someone says that if you ever show that I'm in desperate trouble they'll tell me to talk to my GP." (P6)*

> *"I know one of the hardest things to do is actually getting a person to say they've got a problem with their memory… but hopefully doing it this way is quite positive really because once people start the early onset of dementia, they don't have the capacity to say they are getting worse." (P4)*

One participant commented that participation in the project had encouraged them to look after their wellbeing.

> *"Having done the tests, I am now more positive about my own wellbeing than when I started, because I've had the feedback and thought do you know what, that's ok. I know going forward that may change, but I am now more positive about it, and I try to do more things that would help me than I did before." (P1)*

Other key positives included the feeling of helping and addressing the issue of dementia in football:

> *"My father who played semi-professional football all through his youth did die of dementia. So, when it first came up, I thought wow someone's actually doing something about it." (P4)*

> *"Key positives I think are that it might take a fair amount of time, but this is hopefully going to lead to changes… Even if it doesn't directly have a positive impact for me, maybe it will for people further down the line who are coming up." (P5)*

> *"For me the positives are that it keeps the profile of the impact of head injuries going on and to show that there is work going on to see if there is any alignment." (P2)*

When prompted to suggest any negatives, participants primarily cited the possibility of receiving poor feedback:

> *"I think maybe when you get that first feedback if you are someone who quite naturally was at the bottom of the pile, that might be a negative for them." (P1)*

> *"If someone sees they are in decline they could be upset by it. Whether or not to be proactive about those is an open question. Depending on the nature of the proactivity the proacting itself could be a negative in itself." (P3)*

Ethicality of mental health and behaviour questionnaires

Participants were also asked to reflect on the addition of mental health and behaviour questionnaires. For many the addition was in line with their expectations of the project:

> *"I've worked with people who had mental health issues and I'm sort of aware of that sort of thing so no problem." (P1)*

> *"They're not uncommon in surveys about what sort of tin of soup you prefer, so I had no reason to be surprised by them." (P3)*

However, for some their addition was unexpected:

> *"I don't know what would have driven that to be added to it. I understand what I call to be the physical components of (dementia), and how that's judged. The mental health side of it, and I know those problems can come if you start getting dementia, but I'm not sure how connected they are or how subjective they are." (P6)*

Participants were prompted to discuss if they felt they were giving an accurate representation of their mental health:

> *"I would think I'm giving probably 90% accuracy. If there was a scale of 0-10, I would find it easier to give a completely honest appraisal of where I am opposed to a (five) point scale." (P6)*

> *"I'll answer honestly, but perhaps because I don't feel that I have particular mental health challenges and maybe for that reason I don't particularly find it challenging. I think if people do, they might find it harder to be open and honest." (P4)*

Participants were also asked to assess the inclusion of immediate feedback for mental health scores and the provision of relevant resources, and if this signposting met their expectations for duty of care. For many this action was in line with their expectations of ethicality:

> *"I think steering people towards where they may need help or where there's some signs of anxiety or stress or depression is very very helpful." (P2)*

> *"I think if there are opportunities to direct to that would be very welcome and I would imagine people seeing that there is a direct consequence of them completing this stuff and a direct opportunity to explore if there is other support." (P4)*

Participants also recognised concerns around duty of care:

> *"Because of the confidential nature of it, I think if there were concerns raised, how would you find that way to go one step further with the individual… It's a difficult balance between drawing it to people's attention vs. trying to intervene." (P2)*

Ethicality of feedback process

In the second part of the interviews, participants discussed the feedback forms. In general, participants found the feedback process useful and understandable:

> *"In general terms it is interesting to see the mapping over time and the statistical significance relating to change. I don't find it alarming." (P2)*

> *"I think it was pretty well explained in the guidance and then I applied that. Through work I get an annual health assessment and interestingly it's very similar." (P5)*

> *"I find it useful, a good medium to do it in, and I guess encouraging that not everything was in the bottom quartile." (P6)*

However, some participants reported difficulty with initially understanding the presentation of results:

> *"They're quite difficult to understand, I think. A long-time ago I had A-level maths and statistics so I'm quite used to looking at graphs and percentiles, but I think possibly they could be quite difficult for other people." (P4)*

> *"Graphs that are clear to people who are always looking at graphs may be less clear to those that aren't." (P3)*

> *"Initially it looked a little bit complicated with the graphs. But from a maths point of view once I studied it, I found it quite simple and straightforward." (P1)*

One participant reported that they found the feedback forms so difficult to understand that they had stopped trying to understand the feedback:

> *"Not very good to be honest… I started looking at it and I gave up because I didn't find it easy to deal with and therefore, I gave up… I don't normally have difficulty understanding things, but I wasn't really understanding what it was telling me and therefore I lost interest." (P7)*

When asked to consider ways that the feedback forms could be improved to make the results easier to understand, this participant's suggestions were similar to suggestions from other participants:

> *"I think it needs to be put in more simple terms basically." (P7)*

> *"I think if you could say you are in this quartile, rather than trying to work it out yourself." (P4)*

Participants were asked to describe their initial feelings when receiving feedback.

> *"I think I was quite excited, looking forward to opening it. Looking for the positives." (P5)*

In this discussion participants considered their possible reactions to negative feedback:

> *"I think if I did get a report that showed a significant move down then the next feedback form, I'd probably be a bit more anxious." (P5)*

> *"I think that the only negative would be if it were a poor outlook. I don't know how that would be handled." (P1)*

Participants suggested that distributing the feedback forms to the participant as well as a nominated next of kin contact could benefit the feedback process:

"I would almost like a next of kin number or point of contact number. Maybe I would like somebody to foreworn my wife for example that this is coming… if I was going to receive negative feedback, I would certainly want to receive it with her as a means of support." (P1)

"I wonder whether you could have someone else that you could call and speak to. Probably if there was a decline the closest person to them would have noticed it." (P4)

"I show the feedback to my wife because she is interested in what I'm doing. So, if someone were to send it to her, I would have no problems with that. Obviously as long as you can opt in or opt out." (P6)

Discussion

The purpose of this study was to understand the retrospective and concurrent acceptability of participation in a longitudinal online study from the perspective of participants. The interview framework was designed to guide participants through discussing different aspects of Sekhon's acceptability framework (102). These include self-efficacy, perceived effectiveness, burden, opportunity, and ethicality. The findings of this study demonstrate good acceptability of continued participation in the study, and some areas for improvement in the methodology.

In their discussion, participants demonstrated that they found the assessments mixed in terms of difficulty, where some were more challenging than others, but did not indicate that they were unable to complete the assessments. This demonstrates good self-efficacy in participation in the study. Participants also reported finding the assessments easier over time. The analysis of test-retest reliability in Study 3 did not support practice effects between the first and second assessment session, but as the project continues it will be important to monitor for practice effects appearing. It may be that participant responses reflect a greater comfort with the assessments, rather than objective improvement, but longitudinal analysis will explore this. Findings from this study also revealed that the assessments met their initial expectations of taking part in the project based on their understanding of dementia and cognitive decline. This understanding was drawn from their own experiences of friends or family with dementia or mild cognitive decline. Participants' reflections on how their relatives might have found the assessments challenging were particularly valuable to the discussion and demonstrate good face validity. The exception for this was the mental health and behaviour questionnaires which one participant expressed didn't fit in line with their expectations of the study based on their perception of dementia. However, this was counterbalanced by participants who had witnessed mental health and behavioural changes in their relatives with dementia and testified for the relevance of the tasks. Therefore, the affective attitude, intervention coherence and perceived effectiveness of participation were influenced by participants' understanding of dementia and cognitive decline.

Similarly, to findings from the initial acceptability study (Study 1), participants expressed that the burden associated with participating was still acceptable to them, even following two years of participation. In addition, their motivations for continuing share similarities to the findings of the initial acceptability study, whereby participants were driven to continue for self-serving reasons (receiving regular feedback) and altruistic reasons (helping the next generation). Similar themes were also identified in participants' examples of the positive outcomes of participation in the study. Participants identified that the ability to receive regular feedback and monitor their own brain health, as well as the feeling of giving back were important positives that they were getting out of participation. One participant also reported that participation in the study lead to a greater awareness of their brain health and had encouraged them to adopt behaviours to look after their brain health. Taken together with findings from the initial acceptability study, this reinforces the finding that promoting these outcomes can attract participants to the study and aid long term retention in a longitudinal health study.

Participants demonstrated general comfort with the addition of mental health and behaviour questionnaires. By the time of interview, the participants would have completed these assessments four times. Participants' reflections on ethicality, affective attitude and intervention coherence were similar to the findings from the initial acceptability study. However, in the current interviews they also revealed a concern around perceived effectiveness of the questionnaires, whereby participants suggested a scale with more options would help them to answer more honestly and that there may be some discomfort with answering such questions online. These are important takeaways from this study and suggest that interpretation of results on these measures might require caution when comparing to in-person assessments or other measures of mental health and behaviour.

Participants' reflections on the feedback process were particularly helpful to understand the acceptability of the feedback forms from their perspective. Prior to the release of the initial feedback forms, the initial acceptability study had gathered perspectives on participants' expectations of the feedback process and the ethicality of this process. We had also conducted PPI work with a small group of SCORES participants. This informed the development of the feedback forms and the efforts in applying for ethical approval to share feedback forms, which is discussed in the methods chapter of this thesis. The current study was important to understand the participants concurrent and retrospective perceptions of the coherence, effectiveness, and ethicality of receiving feedback. In this study we learned that for most the process of receiving feedback was a helpful and positive experience, empowering participants to monitor their own brain health and make lifestyle changes or speak to their GP as they felt appropriate. Most participants felt confident that they could understand the forms. However, the study also revealed that for some the feedback was too difficult to understand, which discouraged them from reading the reports. Although this finding mostly came from one participant, the

implications of this are important and highlight areas where the feedback process can be improved. It was suggested that the forms include a straightforward description of where participants lie in their quartiles and what this means for their brain health. Additionally, participants recommended that an option was added for the form to be sent to a next of kin contact who could help with interpretation or aid participants to seek support or further assessment.

Conclusion

In conclusion, findings from this study suggest that following two years of taking part, participants evaluate the study to have continued acceptability. Participants report similar perceptions of acceptability as was reported in the initial acceptability study, and suggest areas where improvement is necessary to meet their expectations. Participants find the process of receiving feedback an important and helpful aspect of the study and identify ways to improve this experience. Therefore, participants suggest that that study format of the SCORES Project is appropriate for encouraging long-term participation, which is vital to improving the quality of evidence within this research field.

Review: Moving the goalposts: variance in defining ‘elite women’s sport’ limits inclusion in research

Abstract

The paucity of data from female athletes in sports science means that their risk of injury is poorly understood. For example, female athletes have a higher prevalence and severity of sport-related concussions, and a higher risk of neurodegeneration than males, yet they are largely excluded from research in this area. As was found in Study 3, this can be in part attributed to difficulties in defining elite level participation for older cohorts of female former athletes. In this chapter, I review relevant literature to clarify and define the concept of elite women’s football in the UK and present a framework that can be used internationally.

I designed a systematic search approach with narrative synthesis to collate sources that defined participants as elite female footballers in the UK or discussed the development of women’s football. A search of databases from inception to February 27th, 2024, identified 141 eligible articles. Definitions of elite in previous research showed considerable variation at youth and adult levels. Six key factors were identified as promotors or inhibitors: league development, attitudes towards women in sport, recruitment, and retention, professionalisation, the impact of mega-events, and administration development. This identified three key eras of change: before 2000, 2000-2010, and after 2011. The definitions of elite extracted from the literature, key factors and eras of change gives necessary evidence that informed the definition of a framework for elite women’s football in the UK.

This investigation identified the need for a unifying definition of the level of sport participation for older cohorts of female former athletes. This chapter provides a definition framework for elite women’s football in the UK and makes recommendations for framework development in other sports or countries.

Introduction

An objective of the SCORES longitudinal study is to collect female data so that appropriate comparisons can be made between male and female former athletes, to better understand female athletes' risk of dementia. Following the first round of data collection, as described in Study 3, the database contained 44 females who had participated in contact sport, compared to 192 males. Out of these 44, five reported playing at a professional level, 36 played at an amateur level and 3 played at a recreational level. Given that there was no fully professional women's football league until 2018 (170), and the SCORES study cohort are aged 40 and above, findings from Study 3 emphasised that categorising female athletes by professional, amateur and recreational status might not be an appropriate categorisation for ongoing research.

Female athletes are underrepresented in sports science research in general (175). A recent analysis of over 5,000 publications in sport science journals found that 31% of publications included only male participants, whereas only 6% of publications included only female athletes (175). This investigation found that out of participants included across these publications, males represented 66% of participants whereas females represent only 34% of participants. The consequence of this is that data from male athletes is extrapolated to produce policies, exercise guidelines, and injury treatments that are applied to female athletes. Female athletes differ in physiology and experience of accessing sport compared to their male counterparts which can influence their risk of sport related injuries. Therefore, policies that protect safety of female athletes that are based on male data fail to protect the female athlete.

Research that investigates the long-term effects of exposure to sport related RHI on dementia risk is no exception to this paucity. Studies demonstrate that male former professional footballers are at an increased risk of diagnosis of neurodegenerative diseases including Alzheimer's disease (53) and amyotrophic lateral sclerosis (56). Dementia prevalence worldwide is expected to double by 2040 (176), so identifying modifiable risk factors is paramount to understanding how to reduce prevalence. Despite findings that females have a higher prevalence of sport-related concussions (85, 139), longer and more severe symptoms of concussions (86, 140, 177-181), and are at a higher risk of dementia than males (87), this field is predominately limited to studying the male professional athlete (84). Although incidence of concussion are low in women's football in comparison to other injury types (182, 183), the cumulative effects of concussions and RHIs may put female athletes at a greater risk of sport-related dementia than males. As women's sports increases in professionalisation and popularity, female athletes need to be included in research that investigates the safety of their participation in sport.

A possible reason for the paucity of studies that include elite female athletes may be their experimental definition. Using head injury research as an example, studies have defined elite female athletes as professionals, or only those that have played at an Olympic or international

level (83, 184). Given that women's football was not an Olympic event until 1996, and professional contracts have only recently been available to female athletes in most countries, these definitions limit the potential sample group to young and small groups which offer limitations in long-term brain injury research. Future research needs to take elite women's football's long-standing history into account, which, without professional contracts or representation in world-wide elite events, includes a cohort of females who have been exposed to sport-related RHI and therefore may be at an increased risk of dementia.

Previous attempts to define elite sport have discussed the challenges of categorising participants by level of sport (185-187), particularly in females (188). These investigations demonstrate considerable variance in the use of elite to describe participant groups, ranging from Olympic gold-medallists to elite under-nine age groups (189). This confirms the need for a unified definition for research to not only clarify participant groups, but also to increase participation whereby female athletes can understand if their experience in the game is eligible for participation in research studies. Frameworks to categorise participants by level of sport were proposed by some of these investigations, however these classify participants who currently play, or who have recently retired, and may not yet be appropriate to apply historically to a developing sport like women's football, which is also limited by access to evidence of performance standards, rankings, and records of matches. As an extension to these investigations, a specific framework for defining past elite participation in women's football is required to improve the sampling quality of future research.

The objective of this scoping review was to examine available literature to clarify and define the key concept of elite participation in UK women's football. It was designed to answer the following research questions:

- What definitions have previously been used to define elite women's football in the UK?
- What key milestones of development of the game does the literature identify in UK women's football?

The purpose of this review is not to provide an extensive history of the development of women's football in the UK, but rather to use the literature available to researchers to demonstrate how to create a definitive framework for the sport.

To achieve the objective, I reviewed two categories of sources. Research studies that referred to their participants as elite female footballers based in the UK were collected to establish previous definitions of elite in the field (definition records). Literature was gathered that discussed the development of women's football in the UK to map key milestones (history records). By using these two sources, the aim was to synthesize the evidence into a framework for defining elite women's sport across key eras of change in the developmental participation pathway.

Materials and methods

Eligibility criteria

Journal articles, theses, and books written in English were considered eligible for this search, to include academic research and historical material. Definition records were considered eligible if they included female footballers based in the UK, who were referred to as elite within the full text. The elite participation group needed to be explicitly defined in terms of level of participation. There were no date restrictions to enable us to map all definitions of elite across time. History records were considered eligible if they described women's football in the UK, from the perspective of players. Sources needed to specifically discuss development in the UK, which was chosen to allow a homogenous context to compare levels of sport in, given that sporting structures differ internationally. The timespan of eligible history records was from 1921 to February 27th, 2024, to include literature about the Football Association's (FA) ban on women's football in 1921 and its influence on development in the period until the ban was lifted in 1971.

Search strategy

My search was designed to follow a systematic approach as outlined by PRISMA guidance (105). The databases searched were SPORTDiscus, Scopus, and Web of Science Social Science Citations, to include sources from sports science as well as social science and history. After consultation with a medical and health sciences librarian, a comprehensive database search was conducted from inception to March 24th, 2023, using the following approach: TI ((wom#n OR female) N2 (football OR soccer)) AND (England OR Scotland OR Scottish OR Wales OR Welsh OR "Northern Ireland" OR "Northern Irish" OR British OR Britain OR UK OR "United Kingdom"). Appropriate search operators were used for each database. References were also searched for citations of eligible sources.

Selection of sources

Sources were exported into Endnote for review and data charting, and duplicates were removed. Titles of sources were screened for eligibility, followed by a full text screening. Identified sources were discussed by three authors to reach a consensus on eligibility.

Data charting process

From the definition records, we extracted the definition that had been used to describe elite participants within each study. From the history records, we extracted key milestones and factors that were described to either inhibit or promote the development of elite women's football.

Synthesis of results

Definitions of elite women's football were synthesized by level of participation in football and date. Using narrative synthesis (190), key milestones that influenced the development of elite women's football were organized by theme and into a timeline of events. This timeline was used to develop a framework that offers a uniform definition of elite women's football in the UK and considers how the definition may differ across eras. The themes, timeline of events, and definition

framework were reviewed through consultative work with a former professional female footballer, a former football administrator, and a team of researchers, to ensure that findings from this review were reflective of the experience of elite women's footballers in different eras or levels of participation.

Results

Overall, 147 sources were identified as relevant to my review (see Figure 19 for PRISMA diagram and Appendix K for full summary of results). Out of these, 51 were identified as definition sources and 101 were identified as history sources, and out of both groups 5 were identified as both.

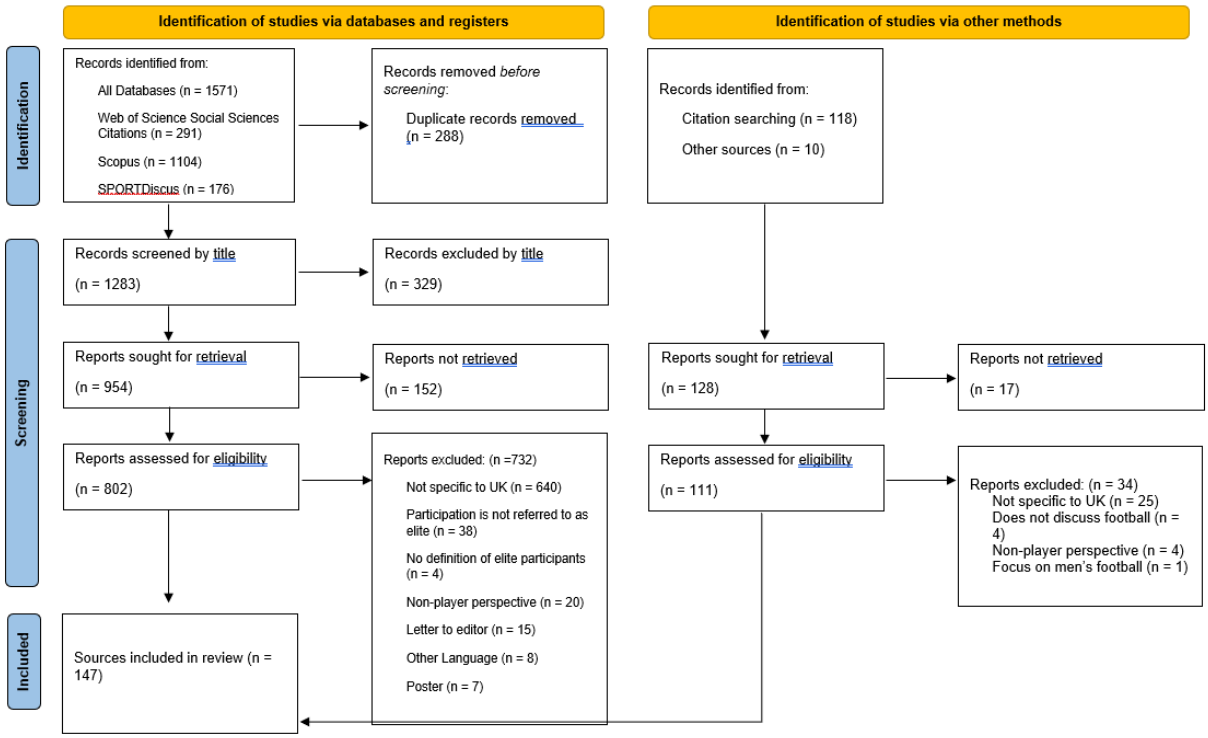


Figure 19: PRISMA diagram depicting search results.

Definition sources

Table 12 describes the 51 papers that met the inclusion criteria for definition sources and their definitions of level of play. This table demonstrates variance in the definition of elite within these sources. Youth definitions vary from regional academies to international level. Some academies described in papers eligible for this review were not described in terms of tier, which is challenging because the term "academy" does not indicate level in the women's football club and may equally refer to grassroot clubs as it can to top level academies. Academies were also described as Tier 1 and 2 academies which is a framework used in men's football. This may reflect efforts to compare male and female youth football in research, by applying a male definition framework to describe female participation in sport. Adult definitions vary between participation

in college international, the fourth highest national league or playing at international level. Early definitions report elite status as participation at international level or in the highest national league. Between 2010 and 2020, the use of highest national leagues is more common for adult participants. Furthermore, from 2020 onwards the use of second highest national leagues or below is more frequent.

Table 12: Summary of definitions extracted from sources eligible for this review.

Authors	Year Published	Definition of Elite
Martin et al.	2006	International Level (Adult)
Price et al.	2009	1st National League & Train >10h/Week (Adult)
Ratna et al.	2010	International Level (Youth & Adult)
Ekstrand et al.	2011	1st National League (Adult)
Munro et al.	2012	2nd National League & Train >10 h/Week (Adult)
Taylor et al.	2013	Academy (Youth)
Bradley et al.	2014	Club level in UEFA Cumulative Club Coefficient Ranking Top 15 (Adult)
Gledhill & Harwood	2014	International, College International, or 1st National League (Youth)
De Ste Croix et al.	2015	Academy (Youth)
Williams et al.	2017	Academy Level (Youth)
Emmonds et al.	2017	Tier 1 Regional Talent Centre (Youth)
Croix et al.	2017	Highest National League Training Academy (Youth)
De Ste Croix et al.	2018	Advanced Coaching Centre (Youth)
Emmonds et al.	2018	Tier 1 Regional Talent Centre (Youth)
Dunn	2018	1st National League and International level (Adult)
Eustace et al.	2019	Highest National League (Adult) or Highest National League Training Academy (Youth)
Emmonds et al.	2019	Highest National League (Adult)
Datson et al.	2020	Elite Performance Camp (Youth)
Sprouse et al.	2020	International Level (Youth & Adult)
Harkness-Armstrong et al.	2020	Regional Talent Centre (Youth)
Emmonds et al.	2020	Tier 1 Regional Talent Centre (Youth)
Emmonds et al.	2020	Tier 1 Regional Talent Centre (Youth)
Pielichaty	2020	Academy Level that requires trials (Youth)
Harrison et al.	2020	Athlete Scholarship Scheme (Adult)
Harkness-Armstrong et al.	2021	Regional Talent Centre (Youth)
Abbott et al.	2021	Top 4 National Leagues (Adult)
Moss et al.	2021	Highest National League (Adult)
Culvin et al.	2021	Highest National League (Adult)
Lucarno et al.	2021	Highest National League (Adult)
Thomas et al.	2021	Highest National League (Adult)
Bishop et al.	2021	Tier 1 Academy (Youth)
Cuthbert et al.	2021	Highest National League (Adult)
Luteberget et al.	2021	Highest National League (Adult)
Simpson et al.	2022	International Level (Adult)
Datson et al.	2022	International Level (Youth & Adult)
Harkness-Armstrong et al.	2022	Regional Talent Centre (Youth)
Fleming et al.	2022	Highest National League (Adult)
Read et al.	2022	Highest National League (Adult)
Myhill et al.	2022	Top 2 National Leagues (Adult)
McHaffie et al.	2022	Top 2 National Leagues (Adult) or International Level (Youth & Adult)
McHaffie et al.	2022	Top 2 National Leagues (Adult) or International Level (Youth & Adult)
Clarkson et al.	2022	Highest National League (Adult)
Parker et al.	2022	Highest National Level & International Level (Adult)
Emmonds et al.	2022	1st & 2nd National League (Adult), Tier 1 Academy (Youth)
Mayhew et al.	2022	1st & 2nd National League (Adult)
Beech et al.	2022	Academy Level (Youth)
Datson et al.	2023	Highest National League & Champions League (Adult)
Emmonds et al.	2023	Top 2 National Leagues (Adult) or Highest National League Training Academy (Youth)
Wheatley et al.	2023	Highest National League (Adult)
Harkness-Armstrong et al.	2023	Tier 1 and Tier 2 Academies (Youth)
McHaffie et al.	2023	International Level (Youth)
Fenton et al.	2023	Top 3 National Leagues (Adult)
Sprouse et al.	2024	International Level (Youth & Adult)

History records

The findings are grouped into six key themes of milestones of development of women's football in the UK: league development, attitudes towards women in sport, recruitment, and retention, professionalisation, mega-events, and administration development. The term mega-events refers to major events in football such as European Championships, World Cups or Olympic games (191).

League Development

The sources identified that league development was a key driver in the development of UK women's football. Development of regionalised leagues and tournaments before 1969 was identified to establish early distinguishments between levels of participation, whereby participation in these leagues and tournaments represented the highest level of football available at the time (192). Formalised leagues in the 1970s were described to contribute to the growth of the game (191). The introduction of the UEFA Women's Champions League in 2001, a policy window after the 2005 European Championships (193), and the formalization of University eleven-a-side leagues in 1988 (England) (192) and 2000 (Scotland) (194) were also credited for driving participation and development of elite women's football in the UK. The establishment of the Women's Super League (WSL) in 2011 was identified as a key moment of change in the development of UK women's football (195, 196), and the subsequent era has been described as a clearly defined period of elevated status for elite women's football in the UK(170, 196). Competitive balance in the WSL was also credited important to development in English women's football (197).

Attitudes towards women in sport

The influence of attitudes towards women in sport were discussed in terms of their impact on the development. These included attitudes towards gender appropriate behaviour (198-200), and how a woman's gender expression (201, 202), sexuality (203, 204), ethnicity (205), religion (206, 207) or class (170) might influence her access to participation levels of sport. Changing attitudes towards women in sport were marked by the impact of the Sex Discrimination Act 1975 (194) and the introduction of policy (193) and social campaigns (208). However, reports suggest that the recent growth in popularity of women's football is also accompanied with a documented rise in misogyny towards women's sports, which may produce barriers towards future generations of players (209). Available literature also noted the resilience shown by women who continued to play in face of such attitudes and societal barriers (210).

Recruitment and Retention

The included sources identified how retention and recruitment opportunities might impact development. These opportunities included the impact of school level participation as an early barrier or promotor towards participation in football sources (211). The Wolfenden Gap (the high

drop-out period of girls playing sport after school age), was also identified as an inhibitor towards development of the game sources (210). The influence of familial and peer support on promoting access to sport and retention of players was also a factor that might impede development (212). The introduction of developmental pathways such as the pyramid organisation of women's football and the introduction of Centres of Excellence in 1998 as the inaugural youth academy structure were identified as promotors of access to participation levels of the sport (193).

Professionalisation

Professionalisation of UK women's football was identified as a factor impacting on the development of the game. Amateur values (211), and the early requirement of national players to retain amateur status before 1993 (213) were noted as impeding on the development of access to women's football at higher levels of participation. The subsequent provision of contracts to English international level players in 2008 marked an improvement in supporting player's access to the higher levels of the game and drove development as players began to be compensated for their time commitment. However, it was also noted that that the delay in constituent countries in securing professional contracts compared to England represents an inequality in development across the UK (201). The impact of professionalisation of the English WSL in 2018 was discussed in terms of training demands and professional expectations (170, 214), as well as barriers that the precarious nature of the work, turbulence for clubs to meet new demands, and lack of policies to meet the needs of female players such as maternity (215, 216), education or post-career planning policies (215).

Mega-events

The impact of mega-events on development was described by sources included in this review. These events included the England men's performance in the 1966 World Cup (191), the Women's Islamic Games in 2001 and 2005 (206, 217), the 2005 Women's European Championships hosted in England (193), the 2012 London Olympics (218) and England women's performance in the 2017 and 2022 Women's European Championships and the 2019 World Cup (170, 219). The effects of the recent 2023 Women's World Cup are not yet represented in the literature identified by this search.

Administration Development

The development of administrative bodies was discussed in the literature to have impacted development. The FA's 1921 ban on women's football was demonstrated to have inhibited the development of the sport across the UK socially, culturally, and economically (209, 211, 220, 221). The ban referred to women being unable to play on the grounds of FA-affiliated clubs, which consequentially dampened the financial sustainability of women's clubs in England and subsequently internationally (221). Although some activity of women playing football after this

ban has been documented (222), the subsequent lifting of the ban in 1971 was described to consequentially formalise the reintroduction of women's football across the UK (223). The inception of the Women's Football Association (England) in 1969 and the Scottish Women's Football Association in 1972 were also recognised to drive organisation of the game (194, 201, 224). Take-overs by the men's football governing body in each UK nation were also cited as key moments of the development of the game in the 1990s (225). Football administration's responses to the Covid-19 pandemic have also been cited as a potential impact factor on the recent development of the women's game, whereby women's football was disproportionately affected in comparison to the men's game (226, 227).

Synthesis of results

Key milestones of development identified by the sources are synthesised in Figure 20. The purpose of this investigation was not to provide an in-depth review of the history of women's football within the UK, but rather to identify milestones of development, that suggest change in the definition of elite as the game has grown.

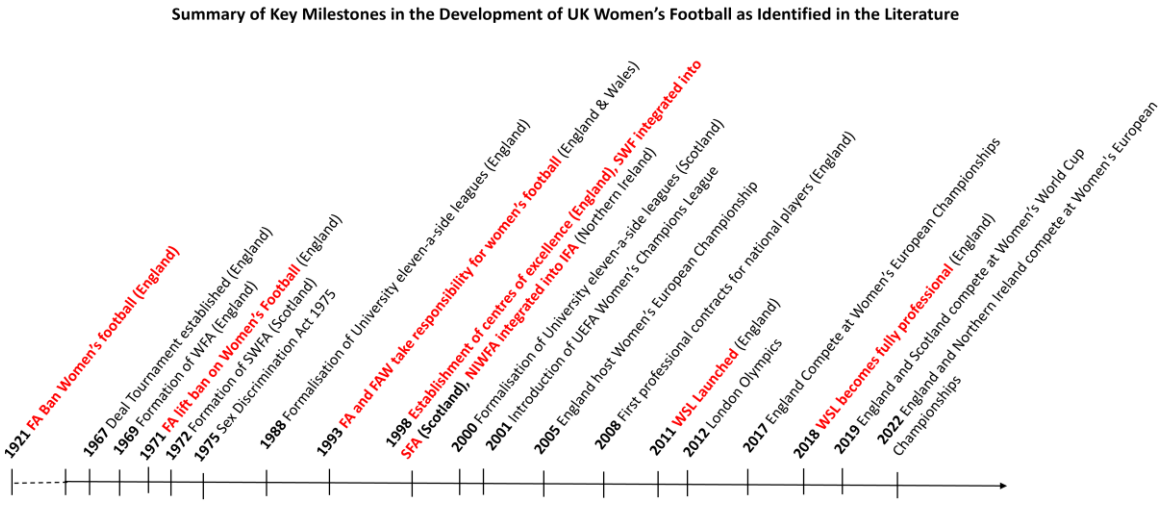


Figure 20: Timeline of key milestones. Key contributors as identified by the available literature are highlighted in red.

From the findings and with the input of the collaborative team, we developed a framework (Figure 21) to enable categorisation of participation, as a proxy for levels of exposure to sport-related injury. This framework defines three key eras of development and levels of participation that fall within each category. Specifically, this framework establishes the era before 2000 as a time where elite participation was accessible at the international level. Between 2000 and 2011, as leagues and centres of excellences developed following affiliation with men's football associations, we see the definition of elite extending to players within the top national leagues, especially with the increased competition driven by the inception of the European Champions League. The third era of development begins with the inception of the WSL. The definition of elite

here is also extended to lower national leagues, which the literature demonstrates as requiring considerable time commitment and exposure to sport (170).

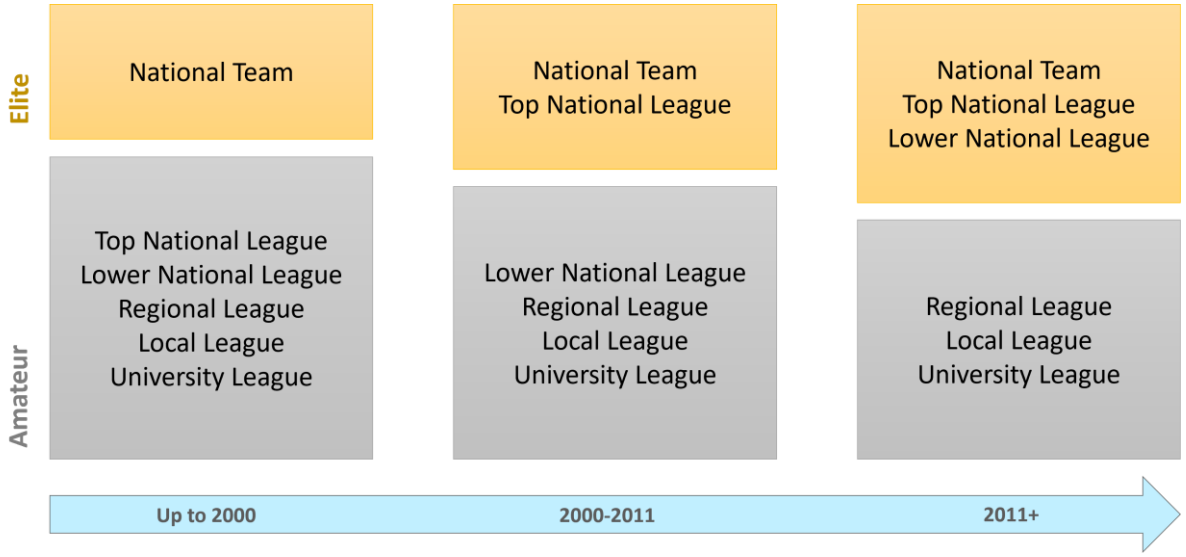


Figure 21: Proposed definition framework to categorize women's football in the UK by level of participation.

Discussion

The objective of this review was to clarify and define the key concept of elite participation in UK women's football. Data was extracted from sources that met an a-priori inclusion criteria to categorise definitions of elite and illuminate promotors and inhibitors in the development of UK women's football. From this, a uniform research definition and a methodology that could be translated for use in defining participation level in other sports or countries were built.

The review found variation in the definition of elite women's football in research, whereby early definitions referred to national team status, whereas later definitions referred to the highest national league or even lower. Some studies specified training time requirements. Whilst this benefits the homogeneity of participant groups that are still playing at time of data collection, including a training time requirement for former players may result in recall bias. Additionally, the available literature produces little information about training time requirements at different levels of sport, and therefore does not support that these times are indicative of playing at a higher level. The search also identified some sources that would otherwise be eligible for inclusion in the review but did not include a definition of the participants that they described as elite. This limits the applicability and generalisability of their findings, as this review demonstrates considerable variation in interpretations of elite. Overall, the variation in definitions found by this review confirms the need for a uniform definition that stratifies levels of exposure to sport and is dynamic across different eras of participation.

Results from the review identified six key factors as influential to the progression of women's football in the UK. League development was reported to have closely mapped progression, with the reorganisation of league structures in the 1990s, the development of enhanced access to regular competition across a wider variety of settings in the early 2000s, and the inception of the

WSL in 2011. The development of recruitment and retention has created a pathway from amateur to elite sport, and milestones under this theme, such as the inauguration of youth academy structures can inform eras of development. The effects of infrastructural changes from administration development support suggests a distinction between the standard of elite women's football before and after the turn of the century in the UK. These three themes therefore largely influenced the selection of eras and levels of elite participation.

The definition framework also considered the themes of attitudes towards women in sport, professionalisation and the impact of mega-events, but these offer fewer clear milestones of development, and their impacts were assessed differently by the included source material. Sources emphasised that attitudes towards women in sport enforced gender-appropriate behaviour on women and created further barriers for some women based on class, sexuality, ethnicity, or religion. However, the sources also described organised participation despite barriers, and that changing attitudes have benefitted women's access to sport. Although events like the Sex Discrimination Act (1975) represent milestone changes in women's access to sport, this factor cannot directly inform the definition of key eras of sport as attitudes change gradually and this review described clear intersectionality in access to sport based on attitudes. Professionalisation has only recently occurred and gradually developed in the UK, and therefore professional status alone cannot define elite status. Sources in this investigation additionally recognised the turbulent process towards professionalisation of the game (170), and the precarious nature of women's football as work. Therefore, professional status may be appropriate to define contemporary and future elite statuses in the UK but is not appropriate for women who previously played elite-level football. This investigation also identified the impact of mega-events as promotors of the game, which opened policy windows for women's football as seen in the Department for Digital, Culture, Media and Sport's investigation following the 2005 European Championships (170, 193), and the 2001 and 2005 Women's Islamic Games (206, 217). However, sources also debate the extent of mega-events as key 'turning points' in the game. Following the record-breaking UEFA Women's European Championship 2021 (in 2022) and World Cup in 2023, a final shift change in the framework for UK women's football could be anticipated and reflected in future published research. Study definitions could reconsider and add to the framework by noting the next significant period of time as 2022 - 2026, allowing a shorter time period for review due to the likely next growth curve. Whilst also remaining cautious about assuming a post European Championship growth (228), a potential rise in sponsorship, development of governance and further growth in participation indicated by the outcome of the virtuous cycle of sport (229) suggests that we could expect 'regional league' to move into the 'elite' level of the framework from 2026 onwards. This would bring the women's game in line with- the male pathway in England for the first time in history.

A strength of this definition framework is that it removes the need to classify participants based on whether or not they were paid to play, but rather by the level of sport that they were exposed to. This allows better categorisation of not just elite athletes, but also amateur level athletes. The benefits of this are exemplified in a recent Finish study of contact sport athletes (60). Although this study only included male participants, participants in the study were eligible if they had represented Finland at an Olympic or international level. Participants in this study were categorised as 'amateur', despite representing an elite level of sport at the time, and therefore the use of 'amateur' requires caution too. Some investigations have used the term 'sub-elite'(230), but there still seems to be some variation into who this refers to (231). The novel uniform definition framework proposed in this thesis allows better categorisation of participants, and clearer comparisons for research.

The themes and milestones identified by this review rely on the quality of the available literature. Most sources discussed the experience of women who were successful in continuing to play sport, despite challenges. Only one study investigated the experience of women who dropped out of football through a negative case analysis (232). Therefore, the extent to which the conclusions of this review can be applied to all women in football is limited by the available evidence. Additionally, most papers included in this review were not written with the primary aim of assessing the development of the sport itself. Although the themes are clear from the data, they may not represent the most prevalent factors for change. As noted by De Bosscher (233), sports policy factors that affect the development of sport are usually guided by the main themes of policy, funding, effective competition, workforce, talent development and marketing/media. The six factors identified by this review are not outside of these themes, but the aim of the review was not to directly assess the factors that affect development in sport but rather to use the available literature to identify promotors or inhibitors of development.

Furthermore, most sources in this review were published in the decade prior to this review. This reflects a recent rise in interest in studying women's football and the history of the game as the involvement of women in the sport has grown internationally. Most sources discussed the game in England, with a minority focusing on Scotland. Investigations in Northern Ireland and Wales are lacking, and this represents a limitation of this review. However, historically women from these constituent countries have accessed elite football in England (225) and therefore this review has relied on evidence from English women's football as an indicator of progress. To address this, key informants were consulted to identify where progress might have been different across the UK, and we recommend consultation for future investigations that do not have access to clear documentation of the sport. Researchers should be careful to ensure a multi-faceted approach is taken, and that several varying roles are recruited in a layered approach to allow multiple perspectives. Suggested informants could include the volunteer workforce, administrators, coaches, officials who should have involvement over a significant period. A second approach to

devising a framework with limited historical information is to select a 'twin' country. Selecting another country that is historically further ahead in the development of the sport can be useful, as timeframes can be compared with assumptions made from the start of initial growth. It can be assumed that the development of sports over time can follow a comparable time frame and learning curve of growth across countries and other sports, ensuring that context is placed at the forefront of consideration to avoid the misunderstanding that there may be one model of elite development that fits all (234). In this light, future research could align to wider framework for comparison, such as the SPLISS framework devised by (233) with a view to provide sports and countries with a comparative set of factors that affect international sporting success and development with a view to provide sports and countries with a comparative set of factors that affect international sporting success and development.

Whilst this framework was designed to identify sporting status of women's football players in the UK, the translation of it to other sports and countries. These may include sports that have seen considerable development over a short period of time, such as other women's sports or Paralympic sports. To adjust the framework for another sport or country, a suggestion following this review strategy is to map how previous literature has defined elite sport, and to identify literature that discusses the development of the sport. To populate the timeline of key events, historical data can be used to understand how extensive participation in the given sport may have been. Once key eras have been established, levels of elite sport can then be identified and supported by a team of key informants. The review strategy relies on literature that is available to researchers, so that definitions can be replicable in research, and validated by the key informants. Translating this definition framework across sports and countries can aid the inclusion of a more representative sample in research.

The purpose of developing this framework is to apply it in future studies to the female athletes within the SCORES cohort. Applying this framework will allow better categorisation of the cohort sample and will enable better validity in comparing elite or amateur male and female samples. Currently, the inclusion criteria of the SCORES project includes a minimum age limit of 40 years. Given that the WSL was only professionalised in 2018, it is unlikely that the SCORES project at this time will contain many female former footballers who were paid full-time contracts. Therefore, the use of the proposed definition framework is important towards the facilitation of accurate stratification by level of sport and therefore level of exposure to injury by proxy. Given the pressing need for female data to better understand their risk (84), this framework represents an important step towards performing sex difference analyses in future. Within the SCORES project, a future aim is to collect further sporting information about the level at which female athletes played to apply this framework and perform this analysis.

Findings from this review have been submitted for publication and are currently undergoing peer review. The paper based on this chapter includes the review of definition sources and presents the variance in definitions used within the literature, but does not include the historical sources. The decision to remove the historical analysis was based on advice from peer review, to improve the rigour of the paper. However, in order to describe the process behind developing the definition framework, the historical analysis was included in this thesis and it will inform further steps taken in the SCORES project to improve the inclusion of female athletes.

Conclusion

The aim of this review was to clarify and define key concepts in women's football in the UK. My search found considerable variation in the description of participant groups and demonstrates a trend towards extending the definition of elite to lower levels of play over time. This investigation supports the need for a uniform definition for participant groups to improve the accurate representation of female athletes within research. The novel definition framework represents a solution to this problem, and the methods used are translatable to other sports or countries. The use of such a framework will improve the categorisation of levels of sport in female athletes, which will ultimately improve their inclusion and greater representation in research and the validity of comparisons with male athletes.

Discussion

The overall aim of this thesis was to evaluate how meaningful results can be determined about the long-term effects of contact sport on brain health by examining the acceptability, validity, and first dataset from an online longitudinal cohort study. To do so, four studies and a narrative literature review were conducted. The main conclusions from each study are summarised in Figure 22.

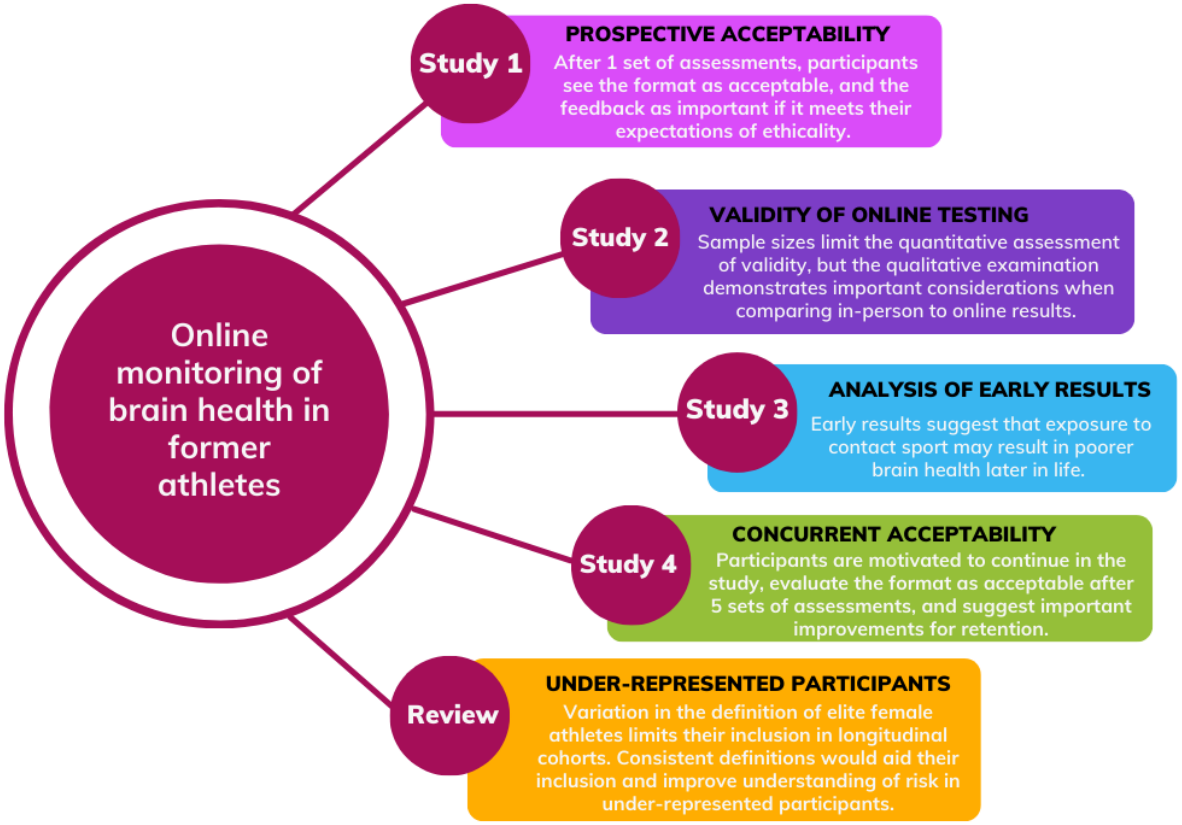


Figure 22: Summary of main conclusions from each study.

Altogether, findings from this thesis suggest that an online longitudinal study of former athletes is an appropriate method of addressing key questions within the research field with many strengths. This approach meets the expectations of acceptability for participants, which promotes recruitment and retention. This investigation supports the overall validity of using online neuropsychological measures and offers considerations for online testing environments, and suggestions to improve the inclusion of female athletes in future research. This approach facilitates the collection of a large and representative dataset that can be used to make meaningful comparisons that contribute to the growing body of evidence that suggests a long-term negative consequence of exposure to sport RHI on brain health.

Acceptability

A key objective of this thesis was to explore the acceptability of the online longitudinal approach to studying former athlete brain health. Studying acceptability was important to this thesis

because the success of this longitudinal approach relies heavily on the participants' perceptions of acceptability. Acceptability of participation was assessed following both the first (Study 1) and fifth (Study 4) set of assessments, to understand both initial and continued perceptions of acceptability as the project developed.

Following the first set of assessments, participants found the online study format to meet their expectations of acceptability and ethicality. Participants reported that they were highly motivated by both self-serving and altruistic reasons. These motivations are important for improving recruitment and retention and can ultimately aid longitudinal studies that include a range of former athletes in terms of level of sport and non-athlete control participants. After one set of assessments, participants perceived the overall study format to meet their expectations of time burden and accessibility and suggest ways to mitigate against the challenges of computer literacy as a requirement for the study and to reduce attrition. Participants highlighted the importance of a feedback process, and described a procedure that meets their expectations of ethicality.

Following five sets of assessments, participant reflections were largely congruent with their perceptions of acceptability found in Study 1. However, Study 4 added important information to this thesis about how participants felt about completing more sensitive assessments, such as mental health and behaviour questionnaires, and importantly their reflections on the process of receiving feedback on their brain health. Study 1 was helpful in initially understanding their expectations of ethicality surrounding the feedback process, but findings from Study 4 demonstrated the lived experiences of participants receiving sensitive feedback surrounding their brain health. Findings from both studies were ultimately used to improve the study and feedback process procedures within SCORES and serve as important suggestions for future study designs that monitor brain health.

Overall, findings in this thesis are largely congruent with the current literature. For example, the CLSA (a population-based health study in Canada) found that participant motivations were largely based on altruistic and self-serving reasons (110), which supports findings from Study 1. They also argue that considering participant motivation is important to recruitment and retention in longitudinal health studies. More specifically, the online study format has also been found to be acceptable by other investigations. Previous investigations support that an online study format increases comfort for participants and reduces distress in completing sensitive measures of brain health (119, 120) which is consistent with findings from Study 4.

In addition to these findings, this thesis contributes new knowledge to the growing field of literature that explores the acceptability of online measures of brain health. Participants' perspectives on the online monitoring of mental health and behaviour, and specifically the duty of care that is expected as part of this protocol, are valuable qualitative findings from this thesis. Furthermore, participants' discussions of ethicality of providing feedback provide novel insight for

designing feedback protocols which meet participants' expectations and encourage recruitment and retention in the study. Overall, the acceptability work in this thesis makes important and novel contributions to the growing field of research that considers the design of longitudinal health research.

Validity

A second objective of this thesis was to examine the validity of the online testing format in comparison to in-person testing. Exploration of the validity of the online testing environment was an important objective of this thesis, to ensure that methods used within the longitudinal online approach were appropriate for monitoring brain health. This was achieved in Study 2, where a mixed methods study design explored both quantitative validity and qualitative contributors to validity.

The original quantitative aim to test equivalence could not be achieved due to sample sizes. However, poor to moderate test-retest reliability was identified between online and in-person assessments, but low variance in the results in some assessments may explain the poor ICC values. A trend indicating that first-testing condition may influence results was also demonstrated. Qualitative data highlighted some factors in both testing conditions that may influence validity and comparisons between in-person and online data, including examiner presence and the ability to control environments remotely and in-person. Findings in Study 2 highlight important considerations that need to be made when designing online studies in relation to age, socio-economic status, and computer literacy of participants. Overall, this chapter supports the validity of the use of online testing batteries in longitudinal studies of former athletes.

This thesis raises important considerations surrounding validity of online measures that are consistent with other assessments made within the field. Quantitatively, the moderate test-retest reliability of in-person and online assessments found in this thesis is reflective of results from another research group using the same digitised assessments in healthy older people (123). Qualitatively, considerations that need to be made surrounding the validity of online assessment as found in this thesis are in line with considerations suggested by other studies. A long-standing body of literature exists surrounding the effect of researcher presence on neuropsychological assessment, particularly in an online testing environment (129, 130). Results from Study 2 support findings in previous literature, suggesting that a lack of researcher presence in online settings may benefit validity of results by removing the "feeling of being watched" or social bias. Furthermore, previous research suggests that online environments create test settings that are high in comfort and low in distress for participants particularly in answering mental health questions(120, 129, 130), which may also improve the validity of results. Results from Study 2 support this in the

context of cognitive tests, but further work could confirm if responses in mental health and behaviour questionnaires are also impacted by the online test environment.

In addition to supporting these findings, results from the validation study also add new understanding to the current literature. Participant responses in this study indicated that although lack of researcher presence in the online condition benefitted their affective attitude towards participating, it also created a challenge in that participants felt unsupported to complete the technical elements of online testing. This distance between researcher and participant may have reduced participants comfort in completing the tasks for the first time, which may have influenced validity of results according to findings from Study 2. This highlights a key concept that although a lack of researcher presence is a key strength of an online testing environment, the ability to quickly reach a researcher is also important towards the success of an online testing platform. This is a particularly important consideration to make when designing research studies that aim to recruit and retain participants with low computer literacy. A further consideration from the validation study is that access to an appropriate remote testing environment may vary by socio-economic status, whereby participants of higher socio-economic status are more likely to be able to access a quiet, uninterrupted environment that is appropriate for online testing, such as a home office space. Therefore, future studies should consider the accessibility of online longitudinal research in the context of lower socio-economic participants. This is particularly important in a study like the SCORES Project, which aims to not only include former professional athletes but also former amateur athletes, and in studies that monitor for signs of dementia, which is associated with socio-economic status (75). Therefore, findings in this thesis emphasise that both the impact of lack of researcher presence on test performance and the impact of socio-economic status on testing environment should be considered in the design of online longitudinal health studies of athletes.

Online monitoring of brain health

A major objective of this thesis was to use initial findings from the longitudinal study to examine differences in contact sport and long-term cognitive, mental health, behavioural and sleep outcomes between appropriate exposure and control groups. This objective was important to meet the overall aim of the thesis, because the online longitudinal approach is proposed to meet many remaining gaps in the literature and to further our understanding of the link between contact sport and long-term brain health. This objective was achieved through Study 3, which examined the first set of results from the ongoing SCORES project.

Overall, the findings from Study 3 that demonstrate better cognitive performance in younger contact sport athletes when compared to a normative sample are consistent with the vast body of evidence that demonstrates a positive effect of physical activity and sport on brain health (16, 18-

20). However, the available data also suggests that male and female athletes who are exposed to contact sport may perform significantly worse on assessments of cognition, behaviour, and mental health later in life. The study presents a trend that professional athletes demonstrate worse brain health than amateurs, but that amateurs may not be exempt from the long-term effects of contact sport on brain health. Findings from Study 3 suggest key domains of cognitive health that should be monitored in further study. These include executive function, source memory and spatial navigation. The latter is a domain that is largely unexplored in this body of research, but this investigation supports growing evidence (141) that spatial navigation is an important domain to consider in the early stages of cognitive impairment. Furthermore, this study suggests a possible sex difference in long-term outcomes, whereby professional contact sport males demonstrate worse cognitive health but better mental health than amateurs, whereas females with a history of contact sport demonstrate better cognitive health but worse mental health than non-contact sport athletes. This thesis also identifies, through Study 3 and the narrative literature review, that further work is needed to accurately make sex-difference comparisons in this field and contributes a novel approach to facilitate comparisons and to benefit future research.

Results from Study 3 contribute to the growing body of literature that demonstrates a link between participation in contact sports and a higher risk of dementia. As outlined in the introduction, the majority of data in this field comes from post-mortem studies and informant data, health record studies, and neuropsychological studies that take single measures of brain health in former athletes. The format of the longitudinal project examined in this thesis is unique because it uses online measures to take repeated measures of brain health, in a population of former professional and amateur, male, and female athletes. Since the start of this PhD project, other projects with similar objectives were launched to try to address the gaps in the literature. In the UK, the FA's FOCUS Study published their first findings in 2023 from questionnaire responses and neuropsychological data from 468 former professional male footballers and 619 general control population males (66). To monitor cognitive decline and health outcomes, the FOCUS Study collected self-reported dementia diagnoses and scores from telephone-administered cognitive tests including the Telephone Interview for Cognitive Status-modified (TICS-m) (a measure of orientation, memory, attention and language), the Verbal Fluency Test (VFT) (a measure of verbal functioning), the Hopkins Verbal Learning Test (HVLT) (a measure of word recall), and Lawton's Instrumental Activities of Daily Living (LADL) to assess general everyday functioning. Compared to methods in the SCORES study, these methods provide more insight into verbal abilities which is a domain demonstrated to be affected in the early stages of dementia (235) and unexplored in the current SCORES protocol. In this sample of males, researchers demonstrated higher prevalence of self-reported neurodegenerative disease and lower performance on measures of verbal literacy and fluency in former professionals than controls.

However, they also found that the overall health profile of the professionals was better than controls. Importantly, the FOCUS study included participants with dementia within their sample, whereas SCORES does not recruit participants with a diagnosis of dementia in order to focus on early detection. Therefore the results of the two studies should be compared with caution. Despite this, and considering that the SCORES battery uses different measures of cognition, the conclusions about overall cognitive health in the FOCUS study are congruent with suggestions from the SCORES results: that former professional athletes are at a higher risk of cognitive impairment than control populations, but also demonstrate other health benefits due to high levels of physical activity. Inclusion of measures of verbal literacy and fluency in the SCORES battery in future may help to further understand this relationship.

Outside of the UK, a German study of 66 former elite female footballers and 45 elite non-contact sport females found no association between participation in football and later neurocognitive performance (83). This study monitored similar cognitive domains to the SCORES project, with the exception of spatial navigation, using computerised and paper and pencil neurocognitive tests. Furthermore, mental health outcomes were not significantly different to non-contact sport controls, but higher estimated exposure to heading was found to be associated with worse mental health outcomes. These findings are not consistent with trends found in the normative analysis within Study 3. However, the study cohort had an average age of 37.4 years, which may explain the discrepancies between study findings. Therefore, findings from this thesis contribute to the current literature by presenting data from an older cohort of female athletes, although further investigation is needed to better understand sex differences in this dataset (236).

Outside of Association football, literature from the USA examines the long-term consequences of participation in American football on brain health. A study (237) of a similarly aged cohort (aged 45-75) to the SCORES cohort found increase impairments in episodic memory, attention, psychomotor speed, and executive function in former professional and college level American football players. These domains are consistent with the domains where impairments were identified in professional male footballers within the SCORES dataset. However, the study of American footballers also identified verbal memory impairment, which may be another domain of interest in future development of the SCORES study.

Measures of mental health, behaviour and sleep are relatively rare within similar large scale long-term study designs, however where these outcomes are studied there are generally mixed findings. A study (80) of former professional and amateur rugby players (105), found higher prevalence of mental health disorders and sleep disruption in players compared to non-contact sport athletes. Interestingly, rugby players with a higher number of concussions were more likely to demonstrate worse mental health outcomes and sleep disruption, as well as higher levels of anger and irritability. These findings were congruent with findings from Study 3, where increased

reported concussions were associated with increased aggression. The primary objective of Study 3 was to examine how exposure to contact sport, rather than total concussions, was associated with outcomes in order to examine the effect of sport related RHIs and to account for reporting bias of concussions. However, it is interesting that this study of rugby players found associations between mental health and behaviour outcomes and estimated heading exposure, as well as number of concussions, and could therefore be considered in future examination of the SCORES dataset. It is however important to note that this study of male rugby players also has a lower average age of participants (46.6 years) and therefore it is possible that lifestyle factors such as proximity to retirement age may influence mental health outcomes (238-240). This is an association that could be further examined through longitudinal, repeated measures of mental health outcomes.

Overall, findings from this thesis support the growing body of evidence suggesting that participation in contact sport has negative consequences on long-term brain health. This thesis began by presenting the positive effects of participation in sport on long-term brain health. These well documented benefits are the reason why increasing physical activity is a common approach towards reducing risk of dementia worldwide and promoting long-term health (20). The growing concern around the increased risk of dementia in contact sports does not challenge this viewpoint, that sport is ultimately good for us, but it does challenge the way in which we participate in sport.

Participation in contact sport is embedded within UK culture, with over 2 million adults and 2.9 million children playing football as an example (6). Recent advances from governing bodies like the Football Association and World Rugby have attempted to reduce the level of exposure to contact through heading restrictions in adults and children (94) and to reduce tackle heights in rugby (93). These changes have been made in response to general research developments in this field, but there is still a lack of studies of the intricacies of this relationship to determine what is ‘safe’. Examination of the consequences of these rule changes will take time, and a longitudinal cohort study can help to track these changes in future, but until then the risk of participation in contact sport has to rely on the policies that are made to protect players. Changing the nature of sports is a challenging position for governing bodies to adopt, but given the increasing evidence of long-term risks and the wide spread participation of non-professional athletes, particularly children (241, 242), this could be necessary (29). Furthermore, the economic incentives of professional football not just in the UK but worldwide are in some cases viewed as a route out of poverty (243-245). Whilst professional footballers represent a minority of participants compared to amateur players, pursuing this career path may put adults and children at a high risk of exposure to brain injuries that have long-term consequences. The solution that participants can be informed of the risks and then choose to participate may be a challenging solution in terms of

ethicality, given that there are such high economic incentives that might pressurise someone into the career path.

Building a representative database

Based on findings from Studies 1-4, a remaining objective of this thesis was to explore methods of promoting and improving the validity of including female athletes in longitudinal cohort studies. This was an important continuation of the work in this thesis, because findings from Studies 1-4 emphasised that females are largely underrepresented in the dataset, and a more representative sample needs to be achieved in future work. Specifically, findings in Study 3 identify that a possible reason for the lack of inclusion of female athletes is the definitions that have been used to describe elite female athletes in past cohort studies. To meet this objective, a narrative literature review examined ways in which female athletes were previously defined in the literature by level of sport and proposed a novel categorisation framework to improve the quality and quantity of female data.

The narrative review demonstrates considerable variation in the definition of elite female participants and identifies that a uniform approach to defining elite in former female athletes would aid their inclusion in future research. This chapter suggests a novel definition framework for use in research. This definition framework will improve the inclusion of female athletes in research, the quality of data, and the validity of comparisons between male and female athletes to examine sex differences in planned future research.

The variability in definitions of elite identified by this review is largely congruent with other attempts to examine the use of this term within sport. Examinations of the use of elite within sports science research identify considerable variation in the use of the term (185-187), including research that focuses on the female athlete (188). In some cases, the use of the term elite was found to range from referring to Olympic level athletes, to under 9 boys football clubs (189). Previous investigations also identify a high number of studies that do not define the participation level they refer to as elite within the full text (185), which is consistent with findings from this thesis. These investigations call for better transparency in reporting athlete characteristics, and many propose their own frameworks for defining participation level in sport (185-188). However, past investigations do not offer a definition framework that classifies past cohorts of athletes, such as female footballers who played in the 1980s or 1990s, which is important in the context of studying brain health in former athletes. The novel approach in this thesis therefore contributes to this field by proposing an evidence-based definition framework that is suitable to categorising past cohorts of athletes and considers change in elite status over time.

The implementation of such a definition framework within this field of research can improve the quality of data and reduce the data gap for female athletes. This will help to facilitate the development of more appropriate policies that are evidence based and protect the female athlete

as well as the male athlete. Findings from this thesis are also novel within this field of research in that they present data from male, female, professional and amateur athletes. Despite the recent publication of a Finnish cohort study (60) that monitors amateur, albeit Olympic and international level, male athletes, findings in this thesis are based on data from a wide range of professional and amateur male and female athletes. The implications of this and further studies from the project are that a wider dataset is made available to improve our understanding of how contact sport effects the brain health of not just an elite minority but instead a wider amateur majority.

Reflections on the methods used in this thesis

A strength of this thesis is the approach using qualitative, quantitative, and mixed methods towards evaluating how an online study format can address key questions within this field of research. Using a multi-disciplinary approach has revealed important information about the strengths and weaknesses of the methodology. The overall process for obtaining ethical approval for the studies in this thesis were ongoing throughout my PhD study period. On reflection, the inclusion of PPI groups and acceptability studies strengthened the applications that navigated more challenging ethical issues, such as the development of the feedback procedure and the inclusion of mental health questionnaires. The evidence-based approach informed by the acceptability studies was beneficial towards representing the views of participants and their own assessments of ethicality, which further supports the value of acceptability studies within research. A qualitative approach towards these studies allowed thoughtful discussion of participants' perceptions of these processes. These discussions were crucial towards the success of the feedback process, as they were used as direct evidence in the procedure for acquiring ethical approval. The mixed methods approach used in Study 2 was also crucial to gaining key findings about factors that may influence validity in online research. In this study, the qualitative element was able to address limitations and gaps in knowledge from the quantitative data. Taking a mixed methods approach was critical in going beyond traditional tests of equivalence and correlation to understand validity.

As a critical part of employing qualitative methods in my study, I made efforts to engage with reflexivity throughout the design, data collection and analysis of Studies 1, 2 and 4. In reflecting on findings from these studies I observed a common theme in participants' responses when discussing their perceived acceptability of the assessments was that participants demonstrated behaviours that could be associated with response bias, either by emphasising that they performed very well on assessments or by overreporting how poorly they thought they performed. This may be a result of my role as a researcher involved in the design of the project. However, my role as a researcher with a background in football may have also benefitted the

quality of findings, particularly in discussions of motivations to participate in the study and when interviewing elite participants.

The online format complied with university social distancing requirements at the time of data collection, but also facilitated the reach of participants across the UK and Ireland which is a major strength of this thesis. However, the online format also limits the inclusion of participants from lower socio-economic demographics and participants that do not have regular reliable access to a computer or internet. Findings in Study 2 suggest that providing university-based equipment to local participants might improve the diversity of the sample in-terms of level of deprivation, but this also limits the geographical benefit of a study completely online. Ideally, a hybrid option could include both strengths but for the purpose of the data collection for this thesis the online format met most needs.

A disadvantage of the questionnaires used in Study 3 is that participants must use scale responses with distinct answers (i.e., “often” to “never”) to answer each question. A frequent comment that participants reported was that they could not decide where their response fit within the predetermined scale. This was also picked up on in Study 4. Although instructions encouraged participants to choose the best fit, this may have limited some responses. This is a limitation of using questionnaires such as the PHQ-9 or GAD-7 that use 5-point scales. To address this, scales could be made numeric rather than categorical, however this would affect the standardisation of the scoring and comparison of results with other studies. This is a common criticism of using questionnaires to measure mental health and behaviour. However, for the purpose of this thesis these questionnaires were appropriate to address the research questions. Furthermore, findings are limited by the accuracy of information from self-report measures. Of particular concern was the validity of data collected surrounding history of concussion, given that bias has been established in previous work (164). Medical records may miss unreported concussions and self-reports may be subject to recall bias. Therefore, a combination of both may be appropriate.

In retrospect, I would also consider redesigning the quantitative aspect of the validation study to meet recruitment needs and to include a wider selection of assessments. The design was limited to only including a selection of the cognitive battery to reduce time burden, but the inclusion of the entire cognitive battery as well as perhaps the mental health, behaviour and sleep questionnaires could help to better understand the validity of the entire assessment battery. Repeating the mental health, behaviour and sleep batteries in-person and online would also help to understand the extent to which online assessment mitigates the effects of researcher presence on responses to compare with findings from other investigations (130).

This cross-section of data provides insight into results as participants begin the study but does not present data from subsequent reassessments. At the time at which the analysis for Study 3 was performed, the study had been collecting data for two and half years. Given that participants are

non-clinical and do not have a diagnosis of dementia, the expectation was that this would be too small of a time frame to see differences in change over time between control and exposure groups. As demonstrated in Figure 5, we also found that at the time of data analysis not enough participant data was available yet to make meaningful comparisons with medium or small effect sizes. Longitudinal analysis is therefore planned as part of the ongoing project. The dataset of participants will continue to be invited to complete assessments, so that a larger data base can be available for longitudinal analysis. This further investigation is important, because it will help to answer remaining research questions about how the brain health of former contact sport athletes might differ from normal ageing.

As described in the methods chapter, the recruitment of footballers was targeted in this study because of growing concern about an increased risk of dementia in former footballers and was supported by established connections that the study team had with English professional football. Football is also the most popular team sport in the UK, with over 1.7 million adult male participants and over 200,000 adult female participants according to the Active Lives Survey (246). Heading the ball and contact with other players are the primary sources of sport-related RHI in football (247). This investigation contributes to the growing body of evidence that suggests that there are long-term consequences of the accumulation of repetitive brain injuries sustained in football. However, there would be merit in extending the targeted recruitment to other contact sports as well as sports where brain injuries are common such as gymnastics, cheerleading, or snow sports (248).

The study sample available was also limited in terms of ethnicity, and represents mostly participants who are socio-economically advantaged. In the dataset used for Study 3, 98.5% of participants identified as white in comparison to 81.7% of English and Welsh residents in the 2021 Census (249). The average level of deprivation using the index of multiple deprivation for the sample was 6.49, which is also skewed towards a less deprived sample. This effected the confounding analysis in Study 3 and had implications on the diversity of views collected in studies 1, 3 and 4. This lack of diversity contradicts the suggestions in the literature that online study designs can attract a diverse and widespread participation sample (122). To understand why this recruitment problem has occurred, it is important to consider that firstly the study design recruits from sporting samples. According to the Sport England's Active Lives Survey, participation in physical activity is shown to vary by ethnicity, with Mixed and White Other ethnicities reporting the highest levels of physical activity, whilst South Asian, Black, and other ethnicities recorded the lowest levels of physical activity (6). Similar to how there is an overwhelming need to study female athletes even though they make up a smaller population of contact sport athletes than males in the UK, there is also a need to collect better data on how brain health might vary in a more diverse population of former contact sport players.

The diversity problem in the dataset may reflect a more general research recruitment problem, rather than one associated with testing modality. The recruitment of the project relies on its interest to potential participants who volunteer their time. As Studies 1 and 4 identified, motivation is a key factor towards participation in longitudinal research. Participants who feel connected to the study, either through a motivation to learn more about their brain health or to contribute to research that helps the next generation, may be more likely to volunteer their time. This sentiment is echoed in other forms of health research as well (118). Therefore, a control population, for example non-contact sport athletes, may be less likely to participate. Currently, participants are not financially renumerated for their time, and instead are offered regular feedback reports on their brain health. There may be a bias in the participants that volunteer their time to the study without a financial incentive. Further efforts should be made to reduce these discrepancies in population samples, particularly given that ethnicity and socio-economic status are important risk factors in dementia (72). Targeted recruitment and offering alternative methods to access and participate in the study (such as using assessments compatible with smartphones or allowing participants to complete assessments at the university) should be considered as potential ways to improve the diversity of the database.

Despite these limitations, the overall sample size of 328 participants is relatively large compared to other similar approaches in the literature, particularly for a study that until recently has been independent with no backing from sporting government bodies. The overall sample size has allowed the primary aims of Study 3 to be met, and although a larger sample size in Study 2 would have allowed better testing of equivalence, quantitative comparison was still possible between in-person and online testing. The dataset also contains a sample size of both amateur and professional male contact sport athletes that were sufficient to make meaningful comparisons. As the longitudinal study progresses, targeted recruitment will aim to reach participants in key demographic groups, such as females with a history of contact sport and males with a history of non-contact sport. Targeted recruitment will also aim to diversify the study population to allow better examination of how ethnicity and socio-economic status can influence findings in the study.

Suggestions for future research

This thesis highlights the need for prospective, longitudinal cohort studies to further understand the link between contact sport and consequences on brain health. This thesis examines the suitability of an online longitudinal study to address key questions in the literature. Through a combination of qualitative, quantitative, and mixed methods approaches, this thesis presents key learnings for ongoing or future prospective longitudinal studies. These key learnings are summarised in Figure 23.

Recruitment and Retention

Recruitment and retention of longitudinal studies drive the success of the study (97). This thesis demonstrates that participant motivation is a key factor in promoting recruitment. Both studies 1 and 4 support that the main motivations for participation were altruistic (to contribute to the research and help the next generation of footballers) and self-serving (to receive feedback to monitor brain health). This is a valuable insight from this work, because it demonstrates that there is a widespread concern surrounding the link between participation in contact sport at both amateur and professional levels and that taking part in research that addresses this is valuable to participants. Further research should consider meeting these expectations of participants to promote recruitment.

This thesis also demonstrates the importance of recruiting sufficient and appropriate control group participants. This is challenging, given that participant motivations are linked to their own history of playing contact sport, but a priority to better isolate the effects of contact sport and gather a sample size suitable to adjust for confounding factors. This thesis also demonstrates that in online studies, the ability to quickly access technical support aids retention. A prerequisite of online research is a certain level of computer literacy, and even though participants in Study 2 favoured the condition where a researcher was not present, they emphasised that technical support was key to their attitudes towards the study and therefore retention. This is a particularly important point to consider in research that includes older participants or participants who may be less computer literate.

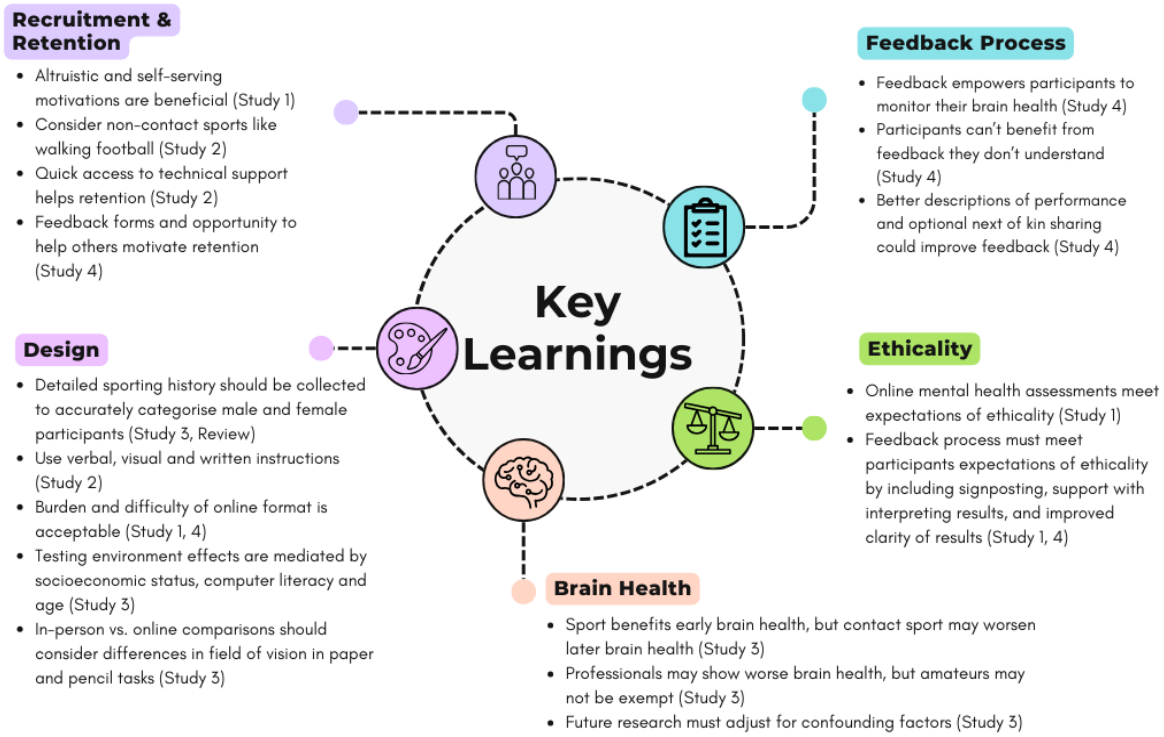


Figure 23: Summary of key learnings from this thesis

The continuation of the project will allow further data collection that will enable the tracking of participants' brain health over time. This research will aim to identify group differences between exposure and control groups to identify how and when brain health might differ between the groups. For this to happen, further recruitment of non-contact sport athletes to the project is likely needed. Given that this thesis identifies that motivation is important, promoting altruistic motivations to take part in the study and the opportunity to receive feedback on brain health would be important in recruiting these participants. Study 4 identified that walking football clubs might be a good avenue for recruiting participants, given that this is a non-contact sport and playing association football is not typically a pre-requisite of participation. This would give access to active adults who may be motivated to contribute to this body of research.

Efforts to better understand sex-differences in the long-term brain health of athletes is a priority. Women's sport is growing globally, and the data gap in male and female research participants means that policies are not optimised to protect the female athlete (84). To enable this research within this longitudinal project, the current participants should be invited to provide further information on their playing careers, so that participants can be categorised using the definition framework approach introduced in the review chapter. Dependent on response rate to this inquiry, further recruitment of female athletes may be necessary. This recruitment could be improved by working with gatekeepers of women's sport like club networks or professional organisations.

Furthermore, efforts to understand the effects of factors like ethnicity that are underrepresented in the current SCORES database will allow the analysis of how important risk factors of dementia influence outcomes in former contact sport athletes. Future planned methods to engage with a more diverse sample in terms of ethnicity will include targeting recruitment to metropolitan areas where populations are more diverse in terms of ethnicity. An example of how this will be achieved is by connecting with veteran football clubs (where participants are usually aged over 35 or 40) and walking football clubs in these areas. Working together with key gatekeepers in the elite field, such as players unions, will also facilitate more targeted recruitment.

Additionally, accessing an appropriate population of non-contact sport athletes will benefit future analyses within the SCORES project. For female athletes, where participation in contact sports is still less common than for male athletes (246), this should be relatively straight forward to facilitate by targeting recruitment to non-contact sporting clubs, such as athletics clubs or veteran hockey or netball clubs. However for male athletes, finding suitable volunteers with no history of contact sport, or very little history of contact sport, but who have lead active lives may be more challenging given the popularity of contact sport for boys within school programmes (246). Therefore, methods to facilitate targeted recruitment of male athletes with a history of non-contact sport will include engaging with amateur non-contact sport clubs.

Study Design

This investigation highlighted that the online format itself was acceptable to participants, and that the convenience of completing testing at home was important to recruitment and retention. However, this investigation also identified that the validity of online research may be influenced by whether or not a participant has access to a controlled space and is sufficiently computer literate. These may vary with age, socio-economic status, caring responsibilities, or education, and therefore studies that employ an online study format to study ageing should consider these factors in the design of their studies. Another consideration in the design of validity studies of online assessments is that if using paper and pencil assessments in person, the difference in field of vision between a paper assessment and a computer assessment should be considered as this likely reduces the effectiveness of participants on a task.

This thesis highlights that the design of the longitudinal study is important towards the acceptability and validity of the study as well as the quality of results. The demographics questionnaire already collects detailed information about sporting history, however findings from Study 3 demonstrates that further detail is needed about level of participation in sport to enable better sex-difference comparisons in future. Within the brain health assessments, the use of verbal, written and visual instructions are a priority for participants, and should be included in longitudinal studies to increase the accessibility of research particularly for the initial online testing session.

Furthermore, this thesis demonstrates the benefits of examining the acceptability of participation in longitudinal studies at regular intervals to understand retrospective, concurrent, and prospective acceptability. Findings from studies 1 and 4 were used to directly improve study protocol and participant instructions, and to inform the development of the feedback process. Assessing ongoing acceptability in future rounds of data collection will allow the researchers to introduce methods to reduce attrition, which will promote the success of meeting research outcomes.

Brain Health

In general, findings from this thesis support the growing evidence that participation in contact sport may have a long-term consequence on brain health. Findings from Study 3 suggest that attention, executive function, processing speed and spatial navigation are important domains of cognition that future research should explore further. However, analysis in Study 3 was limited by sample size and therefore demonstrates the importance of recruiting sufficient sample sizes to better understand sleep outcomes and to adjust for confounds. Although, this investigation did identify potential confounds that should be considered by future research as well. In male

athletes, it was also demonstrated that professionals may demonstrate worse brain health than amateurs, but that amateurs may also not be exempt from long-term consequences in comparison with a normative sample. This highlights the importance of including amateur athletes in future research.

Feedback Process and Ethicality

This thesis demonstrates how the addition of a feedback process to a longitudinal study of health can improve recruitment and retention. This thesis also demonstrates participants' expectations of the ethicality of a feedback process, particularly when presenting cognitive or mental health data. A feedback process that is clear to understand and includes appropriate signposting empowers participants to monitor their own brain health. This thesis also highlighted that including the option for feedback to be shared with next of kin might be appropriate, and this might particularly aid studies using clinical populations.

Concluding comments

This thesis finds that an online longitudinal study format has many strengths that can aid the production of meaningful results about the long-term effects of contact sport on brain health. This conclusion was reached by examining the acceptability, validity and early findings from an online longitudinal study using both qualitative, quantitative, and mixed methods.

The study format is demonstrated to meet expectations of acceptability not just for new participants to the study but also after five repetitions of assessments. This highlights the online study format as appropriate for encouraging recruitment and retention. These are valuable qualities of such a study in a field where large scale, prospective research is needed to address key gaps in knowledge. This thesis also makes suggestions for how the quality and validity of these results can be improved to drive future research in sex-differences and the study of amateur participants.

This thesis demonstrates that the study format can produce meaningful findings that contribute to the body of knowledge within this research field. Specifically, initial findings demonstrate evidence of long-term negative consequences of contact sport on brain health in amateur and professional athletes.

In conclusion, this thesis demonstrates that the online longitudinal study format represents a novel approach within this research field, that contributes to the existing literature about the long-term effects of contact sport on brain health. To build on these findings, future research should aim to examine the longitudinal trends in data collected in this study and aim to improve the inclusion of underrepresented athletes in this research.

References

1. Bull FC, Al-Ansari SS, Biddle S, Borodulin K, Buman MP, Cardon G, et al. World Health Organization 2020 guidelines on physical activity and sedentary behaviour. British Journal of Sports Medicine. 2020;54(24):1451-62.
2. Posadzki P, Pieper D, Bajpai R, Makaruk H, Könsgen N, Neuhaus AL, Semwal M. Exercise/physical activity and health outcomes: an overview of Cochrane systematic reviews. BMC Public Health. 2020;20:1-12.
3. Reiner M, Niermann C, Jekauc D, Woll A. Long-term health benefits of physical activity–a systematic review of longitudinal studies. BMC Public Health. 2013;13:1-9.
4. Lee I-M, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. The Lancet. 2012;380(9838):219-29.
5. Organization WH. Global action plan on physical activity 2018-2030: more active people for a healthier world: World Health Organization; 2019.
6. England S. Active Lives Adult Survey November 2022. 2022.
7. Department for Culture MaS. Get Active: a strategy for the future of sport and physical activity. 2023.
8. Service NH. Get running with Couch to 5K 2023 [Available from: https://www.nhs.uk/live-well/exercise/get-running-with-couch-to-5k/.
9. Caspersen CJ, Powell KE, Christenson GM. Physical activity, exercise, and physical fitness: definitions and distinctions for health-related research. Public Health Reports. 1985;100(2):126.
10. Khan KM, Thompson AM, Blair SN, Sallis JF, Powell KE, Bull FC, Bauman AE. Sport and exercise as contributors to the health of nations. The Lancet. 2012;380(9836):59-64.
11. Kujala UM, Marti P, Kaprio J, Hernelahti M, Tikkanen H, Sarna S. Occurrence of chronic disease in former top-level athletes: predominance of benefits, risks or selection effects? Sports Medicine. 2003;33:553-61.
12. Oja P, Titze S, Kokko S, Kujala UM, Heinonen A, Kelly P, et al. Health benefits of different sport disciplines for adults: systematic review of observational and intervention studies with meta-analysis. British Journal of Sports Medicine. 2015;49(7):434-40.
13. Castagna C, Krustrup P, Póvoas S. Cardiovascular fitness and health effects of various types of team sports for adult and elderly inactive individuals-a brief narrative review. Progress in Cardiovascular Diseases. 2020;63(6):709-22.
14. Andersen MH, Ottesen L, Thing LF. The social and psychological health outcomes of team sport participation in adults: An integrative review of research. Scandinavian Journal of Public Health. 2019;47(8):832-50.
15. Gorelick PB, Sorond FA. What is brain health? Cerebral Circulation-Cognition and Behavior. 2024;6:100190.
16. Kennedy G, Hardman RJ, Macpherson H, Scholey AB, Pipingas A. How does exercise reduce the rate of age-associated cognitive decline? A review of potential mechanisms. Journal of Alzheimer's Disease. 2017;55(1):1-18.
17. Sexton CE, Betts JF, Demnitz N, Dawes H, Ebmeier KP, Johansen-Berg H. A systematic review of MRI studies examining the relationship between physical fitness and activity and the white matter of the ageing brain. Neuroimage. 2016;131:81-90.
18. LLorens-Martín M, Torres-Alemán I, Trejo JL. Exercise modulates insulin-like growth factor 1-dependent and-independent effects on adult hippocampal neurogenesis and behaviour. Molecular and Cellular Neuroscience. 2010;44(2):109-17.
19. Sutin AR, Stephan Y, Luchetti M, Terracciano A. Loneliness and risk of dementia. The Journals of Gerontology: Series B. 2020;75(7):1414-22.
20. Brasure M, Desai P, Davila H, Nelson VA, Calvert C, Jutkowitz E, et al. Physical activity interventions in preventing cognitive decline and Alzheimer-type dementia: a systematic review. Annals of Internal Medicine. 2018;168(1):30-8.
21. Schnell A, Mayer J, Diehl K, Zipfel S, Thiel A. Giving everything for athletic success!–Sports-specific risk acceptance of elite adolescent athletes. Psychology of sport and exercise. 2014;15(2):165-72.
22. Verhagen E, van Mechelen W. Sport for all, injury prevention for all. British Association of Sport and Excercise Medicine; 2010. p. 158-.

23. Grice A, Kingsbury S, Conaghan P. Nonelite exercise-related injuries: Participant reported frequency, management and perceptions of their consequences. Scandinavian Journal of Medicine & Science in Sports. 2014;24(2):e86-e92.
24. Muthuri S, McWilliams D, Doherty M, Zhang W. History of knee injuries and knee osteoarthritis: a meta-analysis of observational studies. Osteoarthritis and Cartilage. 2011;19(11):1286-93.
25. Kerr ZY, Chandran A, Nedimyer AK, Arakkal A, Pierpoint LA, Zuckerman SL. Concussion incidence and trends in 20 high school sports. Pediatrics. 2019;144(5).
26. Prien A, Grafe A, Rössler R, Junge A, Verhagen E. Epidemiology of head injuries focusing on concussions in team contact sports: a systematic review. Sports Medicine. 2018;48:953-69.
27. Nichols EH, Smith HB. The physical aspect of American football. The Boston Medical and Surgical Journal. 1906;154(1):1-8.
28. Martland HS. Punch drunk. Journal of the American Medical Association. 1928;91(15):1103-7.
29. Nowinski CJ, Bureau SC, Buckland ME, Curtis MA, Daneshvar DH, Faull RL, et al. Applying the Bradford Hill criteria for causation to repetitive head impacts and chronic traumatic encephalopathy. Frontiers in Neurology. 2022;13:938163.
30. Nowinski CJ, Rhim HC, McKee AC, Zafonte RD, Dodick DW, Cantu RC, Daneshvar DH. 'Subconcussive'is a dangerous misnomer: hits of greater magnitude than concussive impacts may not cause symptoms. BMJ Publishing Group Ltd and British Association of Sport and Exercise Medicine; 2024.
31. McKee AC, Alosco ML, Huber BR. Repetitive head impacts and chronic traumatic encephalopathy. Neurosurgery Clinics of North America. 2016;27(4):529-35.
32. Erkkinen MG, Kim M-O, Geschwind MD. Clinical neurology and epidemiology of the major neurodegenerative diseases. Cold Spring Harbor Perspectives in Biology. 2018;10(4):a033118.
33. Dugger BN, Dickson DW. Pathology of neurodegenerative diseases. Cold Spring Harbor Perspectives in Biology. 2017;9(7):a028035.
34. Organisation WH. Dementia 2023 [Available from: https://www.who.int/news-room/fact-sheets/detail/dementia.
35. McKee AC, Stein TD, Huber BR, Crary JF, Bieniek K, Dickson D, et al. Chronic traumatic encephalopathy (CTE): criteria for neuropathological diagnosis and relationship to repetitive head impacts. Acta Neuropathologica. 2023;145(4):371-94.
36. Mariani M, Alosco ML, Mez J, Stern RA, editors. Clinical presentation of chronic traumatic encephalopathy. Seminars in Neurology; 2020: Thieme Medical Publishers 333 Seventh Avenue, New York, NY 10001, USA.
37. Shively SB, Priemer DS, Stein MB, Perl DP. Pathophysiology of traumatic brain injury, chronic traumatic encephalopathy, and neuropsychiatric clinical expression. Psychiatric Clinics of North America. 2021;44(3):443-58.
38. Mez J, Daneshvar DH, Kiernan PT, Abdolmohammadi B, Alvarez VE, Huber BR, et al. Clinicopathological evaluation of chronic traumatic encephalopathy in players of American football. Journal of the American Medical Association. 2017;318(4):360-70.
39. Montenigro PH, Baugh CM, Daneshvar DH, Mez J, Budson AE, Au R, et al. Clinical subtypes of chronic traumatic encephalopathy: literature review and proposed research diagnostic criteria for traumatic encephalopathy syndrome. Alzheimer's Research & Therapy. 2014;6(5):1-17.
40. International AsD. World Alzheimer Report 2019: Attitudes to dementia 2019 2019 [Available from: https://www.alz.co.uk/research/WorldAlzheimerReport2019.pdf.
41. Porsteinsson AP, Isaacson R, Knox S, Sabbagh MN, Rubino I. Diagnosis of early Alzheimer's disease: clinical practice in 2021. The Journal of Prevention of Alzheimer's Disease. 2021;8:371-86.
42. Mukadam N, Marston L, Lewis G, Mathur R, Rait G, Livingston G. Incidence, age at diagnosis and survival with dementia across ethnic groups in England: A longitudinal study using electronic health records. Alzheimer's & Dementia. 2023;19(4):1300-7.
43. Bateman RJ, Xiong C, Benzinger TL, Fagan AM, Goate A, Fox NC, et al. Clinical and biomarker changes in dominantly inherited Alzheimer's disease. New England Journal of Medicine. 2012;367(9):795-804.
44. Corsellis J. Boxing and the brain. BMJ: British Medical Journal. 1989;298(6666):105.

45. Omalu BI, Hamilton RL, Kamboh IM, DeKosky ST, Bailes J. Chronic traumatic encephalopathy (CTE) in a National Football League Player: Case report and emerging medicolegal practice questions. Journal of Forensic Nursing. 2010;6(1):40-6.
46. Mez J, Solomon TM, Daneshvar DH, Stein TD, McKee AC. Pathologically confirmed chronic traumatic encephalopathy in a 25-year-old former college football player. JAMA Neurology. 2016;73(3):353-5.
47. Omalu B, Small GW, Bailes J, Ercoli LM, Merrill DA, Wong K-P, et al. Postmortem autopsy-confirmation of antemortem [F-18] FDDNP-PET scans in a football player with chronic traumatic encephalopathy. Neurosurgery. 2018;82(2):237.
48. Stewart W, McNamara P, Lawlor B, Hutchinson S, Farrell M. Chronic traumatic encephalopathy: a potential late and under recognized consequence of rugby union? QJM: An International Journal of Medicine. 2016;109(1):11-5.
49. Gavett BE, Cantu RC, Shenton M, Lin AP, Nowinski CJ, McKee AC, Stern RA. Clinical appraisal of chronic traumatic encephalopathy: current perspectives and future directions. Current Opinion in Neurology. 2011;24(6):525-31.
50. Ling H, Kara E, Revesz T, Lees AJ, Plant GT, Martino D, et al. Concomitant progressive supranuclear palsy and chronic traumatic encephalopathy in a boxer. Acta Neuropathologica Communications. 2014;2:1-11.
51. Grinberg LT, Anghinah R, Nascimento CF, Amaro Jr E, Leite RP, Martin MdGM, et al. Chronic traumatic encephalopathy presenting as Alzheimer's disease in a retired soccer player. Journal of Alzheimer's Disease. 2016;54(1):169-74.
52. Suter CM, Affleck AJ, Pearce AJ, Junckerstorff R, Lee M, Buckland ME. Chronic traumatic encephalopathy in a female ex-professional Australian rules footballer. Acta neuropathologica. 2023;146(3):547-9.
53. Mackay DF, Russell ER, Stewart K, MacLean JA, Pell JP, Stewart W. Neurodegenerative disease mortality among former professional soccer players. New England Journal of Medicine. 2019;381(19):1801-8.
54. Ueda P, Pasternak B, Lim C-E, Neovius M, Kader M, Forssblad M, et al. Neurodegenerative disease among male elite football (soccer) players in Sweden: a cohort study. The Lancet Public Health. 2023;8(4):e256-e65.
55. Orhant E, Carling C, Chapellier JF, Marchand JL, Pradat PF, Elbaz A, et al. A retrospective analysis of all-cause and cause-specific mortality rates in French male professional footballers. Scandinavian Journal of Medicine & Science in Sports. 2022;32(9):1389-99.
56. Pupillo E, Bianchi E, Vanacore N, Montalto C, Ricca G, Robustelli Della Cuna FS, et al. Increased risk and early onset of ALS in professional players from Italian Soccer Teams. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration. 2020;21(5-6):403-9.
57. Kariotis TC, Prictor M, Chang S, Gray K. Impact of electronic health records on information practices in mental health contexts: scoping review. Journal of Medical Internet Research. 2022;24(5):e30405.
58. Madsen T, Erlangsen A, Orlovska S, Mofaddy R, Nordentoft M, Benros ME. Association between traumatic brain injury and risk of suicide. Journal of the American Medical Association. 2018;320(6):580-8.
59. Perry DC, Sturm VE, Peterson MJ, Pieper CF, Bullock T, Boeve BF, et al. Association of traumatic brain injury with subsequent neurological and psychiatric disease: a meta-analysis. Journal of Neurosurgery. 2016;124(2):511-26.
60. Batty GD, Frank P, Kujala UM, Sarna SJ, Valencia-Hernández CA, Kaprio J. Dementia in former amateur and professional contact sports participants: population-based cohort study, systematic review, and meta-analysis. EClinicalMedicine. 2023;61.
61. McKee AC, Stein TD, Nowinski CJ, Stern RA, Daneshvar DH, Alvarez VE, et al. The spectrum of disease in chronic traumatic encephalopathy. Brain. 2013;136(1):43-64.
62. Russell ER, McCabe T, Mackay DF, Stewart K, MacLean JA, Pell JP, Stewart W. Mental health and suicide in former professional soccer players. Journal of Neurology, Neurosurgery & Psychiatry. 2020;91(12):1256-60.
63. Batty GD, Frank P, Kujala UM, Sarna SJ, Kaprio J. Suicide and depression in former contact sports participants: population-based cohort study, systematic review, and meta-analysis. EClinicalMedicine. 2023;60.

64. Iverson GL. Retired national football league players are not at greater risk for suicide. Archives of Clinical Neuropsychology. 2020;35(3):332-41.
65. Cunningham J, Broglio SP, O'Grady M, Wilson F. History of sport-related concussion and long-term clinical cognitive health outcomes in retired athletes: a systematic review. Journal of Athletic Training. 2020;55(2):132-58.
66. Macnab T-MP, Espahbodi S, Hogervorst E, Thanoon A, Fernandes GS, Millar B, et al. Cognitive Impairment and Self-Reported Dementia in UK Retired Professional Soccer Players: A Cross Sectional Comparative Study. Sports Medicine-Open. 2023;9(1):43.
67. Asken BM, Sullan MJ, Snyder AR, Houck ZM, Bryant VE, Hizel LP, et al. Factors influencing clinical correlates of chronic traumatic encephalopathy (CTE): a review. Neuropsychology Review. 2016;26:340-63.
68. Allan CL, Ebmeier KP. The influence of ApoE4 on clinical progression of dementia: a meta-analysis. International Journal of Geriatric Psychiatry. 2011;26(5):520-6.
69. Vasilevskaya A, Taghdiri F, Burke C, Tarazi A, Naeimi SA, Khodadadi M, et al. Interaction of APOE4 alleles and PET tau imaging in former contact sport athletes. NeuroImage: Clinical. 2020;26:102212.
70. Ng TKS, Ho CSH, Tam WWS, Kua EH, Ho RC-M. Decreased serum brain-derived neurotrophic factor (BDNF) levels in patients with Alzheimer's disease (AD): a systematic review and meta-analysis. International Journal of Molecular Sciences. 2019;20(2):257.
71. Cherry JD, Mez J, Crary JF, Tripodis Y, Alvarez VE, Mahar I, et al. Variation in TMEM106B in chronic traumatic encephalopathy. Acta Neuropathologica Communications. 2018;6:1-9.
72. Shiekh SI, Cadogan SL, Lin L-Y, Mathur R, Smeeth L, Warren-Gash C. Ethnic differences in dementia risk: a systematic review and meta-analysis. Journal of Alzheimer's Disease. 2021;80(1):337-55.
73. Peters R, Booth A, Rockwood K, Peters J, D'Este C, Anstey KJ. Combining modifiable risk factors and risk of dementia: a systematic review and meta-analysis. BMJ Open. 2019;9(1):e022846.
74. Mielke MM, Vemuri P, Rocca WA. Clinical epidemiology of Alzheimer's disease: assessing sex and gender differences. Clinical Epidemiology. 2014:37-48.
75. Wang A-Y, Hu H-Y, Ou Y-N, Wang Z-T, Ma Y-H, Tan L, Yu J-T. Socioeconomic status and risks of cognitive impairment and dementia: a systematic review and meta-analysis of 39 prospective studies. The Journal of Prevention of Alzheimer's Disease. 2023;10(1):83-94.
76. Yao P-L, Laurencelle L, Trudeau F. Former athletes' lifestyle and self-definition changes after retirement from sports. Journal of Sport and Health Science. 2020;9(4):376-83.
77. McHugh C, Hind K, Davey D, Wilson F. Cardiovascular health of retired field-based athletes: a systematic review and meta-analysis. Orthopaedic Journal of Sports Medicine. 2019;7(8):2325967119862750.
78. Schaffert J, LoBue C, Fields L, Wilmoth K, Didehbani N, Hart Jr J, Cullum CM. Neuropsychological functioning in ageing retired NFL players: a critical review. International Review of Psychiatry. 2020;32(1):71-88.
79. Van Patten R, Iverson GL, Terry DP, Levi CR, Gardner AJ. Predictors and correlates of perceived cognitive decline in retired professional rugby league players. Frontiers in Neurology. 2021;12:676762.
80. Hind K, Konerth N, Entwistle I, Hume P, Theadom A, Lewis G, et al. Mental health and wellbeing of retired elite and amateur rugby players and non-contact athletes and associations with sports-related concussion: the UK Rugby Health Project. Sports Medicine. 2021:1-13.
81. Bruno D, Rutherford A. Cognitive ability in former professional football (soccer) players is associated with estimated heading frequency. Journal of Neuropsychology. 2022;16(2):434-43.
82. Espahbodi S, Hogervorst E, Macnab T-MP, Thanoon A, Fernandes GS, Millar B, et al. Heading frequency and risk of cognitive impairment in retired male professional soccer players. JAMA Network Open. 2023;6(7):e2323822-e.
83. Prien A, Feddermann-Demont N, Verhagen E, Twisk J, Junge A. Neurocognitive performance and mental health of retired female football players compared to non-contact sport athletes. BMJ Open Sport & Exercise Medicine. 2020;6(1):e000952.
84. D'Lauro C, Jones ER, Swope LM, Anderson MN, Broglio S, Schmidt JD. Under-representation of female athletes in research informing influential concussion consensus and

position statements: an evidence review and synthesis. British Journal of Sports Medicine. 2022;56(17):981-7.
85. Bretzin AC, Covassin T, Wiebe DJ, Stewart W. Association of sex with adolescent soccer concussion incidence and characteristics. JAMA Network Open. 2021;4(4):e218191-e.
86. Ellemberg D, Leclerc S, Couture S, Daigle C. Prolonged neuropsychological impairments following a first concussion in female university soccer athletes. Clinical Journal of Sport Medicine. 2007;17(5):369-74.
87. Berry L. Worldwide research finds dementia affects women more than men. Nursing Older People (2014+). 2015;27(6):9.
88. Howarth NE, White AJ, Pearce AJ, Nowinski C, Cantu R, Ji C, Miller MA. Obstructive Sport Apnea (OSA) and contact sports: A systematic review and meta-analysis. Sleep Epidemiology. 2022:100036.
89. Guay-Gagnon M, Vat S, Forget MF, Tremblay-Gravel M, Ducharme S, Nguyen QD, Desmarais P. Sleep apnea and the risk of dementia: A systematic review and meta-analysis. Journal of Sleep Research. 2022;31(5):e13589.
90. Bucks RS, Olaithe M, Eastwood P. Neurocognitive function in obstructive sleep apnoea: A meta-review. Respirology. 2013;18(1):61-70.
91. Leng Y, McEvoy CT, Allen IE, Yaffe K. Association of sleep-disordered breathing with cognitive function and risk of cognitive impairment: a systematic review and meta-analysis. JAMA Neurology. 2017;74(10):1237-45.
92. Yaffe K, Laffan AM, Harrison SL, Redline S, Spira AP, Ensrud KE, et al. Sleep-disordered breathing, hypoxia, and risk of mild cognitive impairment and dementia in older women. Journal of the American Medical Association. 2011;306(6):613-9.
93. Rugby W. World Rugby confirms lower tackle height law trials for community rugby World Rugby2023 [Available from: https://www.world.rugby/news/808734/world-rugby-confirms-lower-tackle-height-law-trials-for-community-rugby.
94. Football E. Heading in Football England Football2022 [Available from: https://www.englandfootball.com/participate/learn/Brain-Health/Heading-in-Football.
95. Wenger GC. Advantages gained by combining qualitative and quantitative data in a longitudinal study. Journal of Aging Studies. 1999;13(4):369-76.
96. Gabel M, Bollinger RM, Knox M, Coble DW, Grill JD, Edwards DF, et al. Perceptions of research burden and retention among participants in ADRC cohorts. Alzheimer Disease & Associated Disorders. 2022;36(4):281-7.
97. Caruana EJ, Roman M, Hernández-Sánchez J, Solli P. Longitudinal studies. Journal of Thoracic Disease. 2015;7(11):E537.
98. Fitzgerald D, Hockey R, Jones M, Mishra G, Waller M, Dobson A. Use of online or paper surveys by Australian women: longitudinal study of users, devices, and cohort retention. Journal of medical Internet research. 2019;21(3):e10672.
99. Creswell JW, Creswell JD. Research design: Qualitative, quantitative, and mixed methods approaches: Sage publications; 2017.
100. Ames N. UEA begins study with former players into link between football and dementia: The Guardian; 2020 [Available from: https://www.theguardian.com/football/2020/jan/07/university-east-anglia-players-link-football-dementia.
101. Pope C, Mays N. Qualitative research in health care. Fourth edition. ed: John Wiley and Sons, Inc.; 2020.
102. Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Services Research. 2017;17(1):1-13.
103. Archibald MM, Ambagtsheer RC, Casey MG, Lawless M. Using zoom videoconferencing for qualitative data collection: perceptions and experiences of researchers and participants. International Journal of Qualitative Methods. 2019;18:1609406919874596.
104. Boland J, Banks S, Krabbe R, Lawrence S, Murray T, Henning T, Vandenberg M. A COVID-19-era rapid review: using Zoom and Skype for qualitative group research. Public Health Research & Practice. 2022;32(2).

105. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. International Journal of Surgery. 2021;88:105906.
106. Brett J, Staniszewska S, Mockford C, Herron-Marx S, Hughes J, Tysall C, Suleman R. Mapping the impact of patient and public involvement on health and social care research: a systematic review. Health Expectations. 2014;17(5):637-50.
107. Entwistle VA, Renfrew MJ, Yearley S, Forrester J, Lamont T. Lay perspectives: advantages for health research. British Medical Journal. 1998;316(7129):463-6.
108. Kerr ZY, Marshall SW, Harding Jr HP, Guskiewicz KM. Nine-year risk of depression diagnosis increases with increasing self-reported concussions in retired professional football players. The American Journal of Sports Medicine. 2012;40(10):2206-12.
109. Kilic Ö, Hopley P, Kerkhoffs GM, Lambert M, Verhagen E, Viljoen W, et al. Impact of concussion and severe musculoskeletal injuries on the onset of mental health symptoms in male professional rugby players: a 12-month study. BMJ Open Sport—Exercise Medicine. 2019;5(1).
110. Kirkland SA, Raina PS, Wolfson C, Strople G, Kits O, Dukeshire S, et al. Exploring the acceptability and feasibility of conducting a large longitudinal population-based study in Canada. Canadian Journal on Aging. 2009;28(3):231-42.
111. King N, Horrocks C, Brooks JM. Interviews in qualitative research. 2nd edition. ed: Sage Publications Ltd; 2019.
112. Gustavson K, von Soest T, Karevold E, Røysamb E. Attrition and generalizability in longitudinal studies: findings from a 15-year population-based study and a Monte Carlo simulation study. BMC public health. 2012;12:1-11.
113. Rasmussen J, Langerman H. Alzheimer's disease–why we need early diagnosis. Degenerative Neurological and Neuromuscular Disease. 2019:123-30.
114. Service NH. Mood Assessment [Available from: https://assets.nhs.uk/tools/self-assessments/index.mob.html?cookie_consent=true&variant=44.
115. Patricios JS, Schneider KJ, Dvorak J, Ahmed OH, Blauwet C, Cantu RC, et al. Consensus statement on concussion in sport: the 6th International Conference on Concussion in Sport–Amsterdam, October 2022. British Journal of Sports Medicine. 2023;57(11):695-711.
116. Schoeler T, Speed D, Porcu E, Pirastu N, Pingault J-B, Kutalik Z. Participation bias in the UK Biobank distorts genetic associations and downstream analyses. Nature Human Behaviour. 2023:1-12.
117. Chaytor NS, Barbosa-Leiker C, Germine LT, Fonseca LM, McPherson SM, Tuttle KR. Construct validity, ecological validity and acceptance of self-administered online neuropsychological assessment in adults. The Clinical Neuropsychologist. 2021;35(1):148-64.
118. Brintnall-Karabelas J, Sung S, Cadman ME, Squires C, Whorton K, Pao M. Improving recruitment in clinical trials: why eligible participants decline. Journal of Empirical Research on Human Research Ethics. 2011;6(1):69-74.
119. Collerton J, Collerton D, Arai Y, Barrass K, Eccles M, Jagger C, et al. A comparison of computerized and pencil-and-paper tasks in assessing cognitive function in community-dwelling older people in the Newcastle 85+ Pilot Study. Journal of the American Geriatrics Society. 2007;55(10):1630-5.
120. Feenstra HE, Vermeulen IE, Murre JM, Schagen SB. Online cognition: factors facilitating reliable online neuropsychological test results. The Clinical Neuropsychologist. 2017;31(1):59-84.
121. Sliwinski MJ, Mogle JA, Hyun J, Munoz E, Smyth JM, Lipton RB. Reliability and validity of ambulatory cognitive assessments. Assessment. 2018;25(1):14-30.
122. Öhman F, Hassenstab J, Berron D, Schöll M, Papp KV. Current advances in digital cognitive assessment for preclinical Alzheimer's disease. Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring. 2021;13(1):e12217.
123. Morrissey S, Gillings R, Hornberger M. Feasibility and reliability of online vs in-person cognitive testing in healthy older people. medRxiv. 2023:2023.07. 05.23292229.
124. Bauer RM, Iverson GL, Cernich AN, Binder LM, Ruff RM, Naugle RI. Computerized neuropsychological assessment devices: joint position paper of the American Academy of Clinical Neuropsychology and the National Academy of Neuropsychology. Archives of Clinical Neuropsychology. 2012;27(3):362-73.

125. Björngrim S, van den Hurk W, Betancort M, Machado A, Lindau M. Comparing traditional and digitized cognitive tests used in standard clinical evaluation–a study of the digital application minnemera. Frontiers in Psychology. 2019:2327.
126. Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine. 2016;15(2):155-63.
127. Gale NK, Heath G, Cameron E, Rashid S, Redwood S. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. BMC Medical Research Methodology. 2013;13(1):1-8.
128. Lee KM, Lee J, Chung CY, Ahn S, Sung KH, Kim TW, et al. Pitfalls and important issues in testing reliability using intraclass correlation coefficients in orthopaedic research. Clinics in Orthopedic Surgery. 2012;4(2):149-55.
129. Gunter B, Nicholas D, Huntington P, Williams P, editors. Online versus offline research: implications for evaluating digital media. Aslib Proceedings; 2002: MCB UP Ltd.
130. Booth-Kewley S, Larson GE, Miyoshi DK. Social desirability effects on computerized and paper-and-pencil questionnaires. Computers in Human Behavior. 2007;23(1):463-77.
131. Association TF. Walking Football: Laws of the Game 2016.
132. Nichols E, Steinmetz JD, Vollset SE, Fukutaki K, Chalek J, Abd-Allah F, et al. Estimation of the global prevalence of dementia in 2019 and forecasted prevalence in 2050: an analysis for the Global Burden of Disease Study 2019. The Lancet Public Health. 2022;7(2):e105-e25.
133. McKee AC, Cantu RC, Nowinski CJ, Hedley-Whyte ET, Gavett BE, Budson AE, et al. Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. Journal of Neuropathology & Experimental Neurology. 2009;68(7):709-35.
134. Alosco ML, Mez J, Tripodis Y, Kiernan PT, Abdolmohammadi B, Murphy L, et al. Age of first exposure to tackle football and chronic traumatic encephalopathy. Annals of Neurology. 2018;83(5):886-901.
135. Mez J, Daneshvar DH, Abdolmohammadi B, Chua AS, Alosco ML, Kiernan PT, et al. Duration of American football play and chronic traumatic encephalopathy. Annals of Neurology. 2020;87(1):116-31.
136. Phelps A, Alosco ML, Baucom Z, Hartlage K, Palmisano JN, Weuve J, et al. Association of playing college American football with long-term health outcomes and mortality. JAMA Network Open. 2022;5(4):e228775-e.
137. Stern RA, Daneshvar DH, Baugh CM, Seichepine DR, Montenigro PH, Riley DO, et al. Clinical presentation of chronic traumatic encephalopathy. Neurology. 2013;81(13):1122-9.
138. Iverson GL, Castellani RJ, Cassidy JD, Schneider GM, Schneider KJ, Echemendia RJ, et al. Examining later-in-life health risks associated with sport-related concussion and repetitive head impacts: a systematic review of case-control and cohort studies. British Journal of Sports Medicine. 2023;57(12):810-21.
139. O'Connor KL, Baker MM, Dalton SL, Dompier TP, Broglio SP, Kerr ZY. Epidemiology of sport-related concussions in high school athletes: National Athletic Treatment, Injury and Outcomes Network (NATION), 2011–2012 through 2013–2014. Journal of Athletic Training. 2017;52(3):175-85.
140. Broshek DK, Kaushik T, Freeman JR, Erlanger D, Webbe F, Barth JT. Sex differences in outcome following sports-related concussion. Journal of Neurosurgery. 2005;102(5):856-63.
141. Coughlan G, Laczó J, Hort J, Minihane A-M, Hornberger M. Spatial navigation deficits—overlooked cognitive marker for preclinical Alzheimer disease? Nature Reviews Neurology. 2018;14(8):496-506.
142. LoBue C, Schaffert J, Cullum CM, Peters ME, Didehbani N, Hart J, White CL. Clinical and neuropsychological profile of patients with dementia and chronic traumatic encephalopathy. Journal of Neurology, Neurosurgery & Psychiatry. 2020;91(6):586-92.
143. Konrad C, Geburek AJ, Rist F, Blumenroth H, Fischer B, Husstedt I, et al. Long-term cognitive and emotional consequences of mild traumatic brain injury. Psychological Medicine. 2011;41(6):1197-211.
144. Terpstra AR, Vasquez BP, Colella B, Tartaglia MC, Tator CH, Mikulis D, et al. Comprehensive neuropsychiatric and cognitive characterization of former professional football players: implications for neurorehabilitation. Frontiers in Neurology. 2019;10:712.

145. Majerske CW, Mihalik JP, Ren D, Collins MW, Reddy CC, Lovell MR, Wagner AK. Concussion in sports: postconcussive activity levels, symptoms, and neurocognitive performance. Journal of Athletic Training. 2008;43(3):265-74.
146. Tsushima WT, Lum M, Geling O. Sex differences in the long-term neuropsychological outcome of mild traumatic brain injury. Brain Injury. 2009;23(10):809-14.
147. Gallo V, Motley K, Kemp SP, Mian S, Patel T, James L, et al. Concussion and long-term cognitive impairment among professional or elite sport-persons: a systematic review. Journal of Neurology, Neurosurgery & Psychiatry. 2020.
148. Byrd M, Dixon C, Lucke-Wold B. Examining the correlation between acute behavioral manifestations of concussion and the underlying pathophysiology of chronic traumatic encephalopathy: a pilot study. Journal of Neurology and Psychology. 2018;6(1).
149. Carson JD, Healey KE, Frémont P. Using the PHQ-9 to identify and manage depressive symptoms in patients with sport-related concussion. Canadian Family Physician. 2021;67(3):183-4.
150. Löwe B, Kroenke K, Herzog W, Gräfe K. Measuring depression outcome with a brief self-report instrument: sensitivity to change of the Patient Health Questionnaire (PHQ-9). Journal of Affective Disorders. 2004;81(1):61-6.
151. Toussaint A, Hüsing P, Gumz A, Wingenfeld K, Härter M, Schramm E, Löwe B. Sensitivity to change and minimal clinically important difference of the 7-item Generalized Anxiety Disorder Questionnaire (GAD-7). Journal of Affective Disorders. 2020;265:395-401.
152. Buss AH, Perry M. The aggression questionnaire. Journal of Personality and Social Psychology. 1992;63(3):452.
153. Lemieux P, McKelvie SJ, Stout D. Self-reported hostile aggression in contact athletes, no contact athletes and non-athletes. Athletic Insight. 2002;4(3):42-56.
154. Boostani MA, Boostani MH. Investigation and comparing aggression in athletes in non-contact (swimming), limited contact (karate) and contactable (kickboxing) sport fields. Journal of Combat Sports and Martial Arts. 2012;2(2):87-9.
155. Kerr ZY, Evenson KR, Rosamond WD, Mihalik JP, Guskiewicz KM, Marshall SW. Association between concussion and mental health in former collegiate athletes. Injury Epidemiology. 2014;1(1):1-10.
156. Spinella M. Normative data and a short form of the Barratt Impulsiveness Scale. International Journal of Neuroscience. 2007;117(3):359-68.
157. Goswami R, Dufort P, Tartaglia M, Green R, Crawley A, Tator C, et al. Frontotemporal correlates of impulsivity and machine learning in retired professional athletes with a history of multiple concussions. Brain Structure and Function. 2016;221:1911-25.
158. Harris JA. A further evaluation of the Aggression Questionnaire: Issues of validity and reliability. Behaviour Research and Therapy. 1997;35(11):1047-53.
159. Meule A, Mayerhofer M, Gründel T, Berker J, Beck Teran C, Platte P. Half-year retest-reliability of the Barratt Impulsiveness Scale–short form (BIS-15). Sage Open. 2015;5(1):2158244015576548.
160. Johns MW. A new method for measuring daytime sleepiness: the Epworth sleepiness scale. Sleep. 1991;14(6):540-5.
161. Carpenter JS, Andrykowski MA. Psychometric evaluation of the Pittsburgh sleep quality index. Journal of Psychosomatic Research. 1998;45(1):5-13.
162. Netzer NC, Stoohs RA, Netzer CM, Clark K, Strohl KP. Using the Berlin Questionnaire to identify patients at risk for the sleep apnea syndrome. Annals of Internal Medicine. 1999;131(7):485-91.
163. Dang JS, Figueroa IJ, Helton WS. You are measuring the decision to be fast, not inattention: the Sustained Attention to Response Task does not measure sustained attention. Experimental Brain Research. 2018;236:2255-62.
164. Kerr ZY, Mihalik JP, Guskiewicz KM, Rosamond WD, Evenson KR, Marshall SW. Agreement between athlete-recalled and clinically documented concussion histories in former collegiate athletes. The American Journal of Sports Medicine. 2015;43(3):606-13.
165. Warburton DE, Nicol CW, Bredin SS. Health benefits of physical activity: the evidence. Canadian Medical Association Journal. 2006;174(6):801-9.
166. Fone DL, Dunstan F. Mental health, places and people: a multilevel analysis of economic inactivity and social deprivation. Health & Place. 2006;12(3):332-44.

167. Rice SM, Parker AG, Rosenbaum S, Bailey A, Mawren D, Purcell R. Sport-related concussion and mental health outcomes in elite athletes: a systematic review. Sports Medicine. 2018;48:447-65.
168. Mez J, Stern RA, McKee AC. Chronic traumatic encephalopathy: where are we and where are we going? Current neurology and neuroscience reports. 2013;13:1-12.
169. Iverson GL, Terry DP, Luz M, Zafonte R, McCrory P, Solomon GS, Gardner AJ. Anger and depression in middle-aged men: implications for a clinical diagnosis of chronic traumatic encephalopathy. The Journal of Neuropsychiatry and Clinical Neurosciences. 2019;31(4):328-36.
170. Culvin A. Football as work: The new realities of professional women footballers in England: University of Central Lancashire; 2019.
171. Levaux M-N, Potvin S, Sepehry AA, Sablier J, Mendrek A, Stip E. Computerized assessment of cognition in schizophrenia: promises and pitfalls of CANTAB. European Psychiatry. 2007;22(2):104-15.
172. Seli P, Jonker TR, Solman GJ, Cheyne JA, Smilek D. A methodological note on evaluating performance in a sustained-attention-to-response task. Behavior Research Methods. 2013;45:355-63.
173. Beswick E, Quigley S, Macdonald P, Patrick S, Colville S, Chandran S, Connick P. The Patient Health Questionnaire (PHQ-9) as a tool to screen for depression in people with multiple sclerosis: a cross-sectional validation study. BMC Psychology. 2022;10(1):281.
174. Drake G, Csipke E, Wykes T. Assessing your mood online: acceptability and use of Moodscope. Psychological Medicine. 2013;43(7):1455-64.
175. Cowley ES, Olenick AA, McNulty KL, Ross EZ. “Invisible sportswomen”: the sex data gap in sport and exercise science research. Women in Sport and Physical Activity Journal. 2021;29(2):146-51.
176. Wittenberg R, Hu B, Barraza-Araiza L, Rehill A. Projections of older people with dementia and costs of dementia care in the United Kingdom, 2019–2040. London: London School of Economics. 2019.
177. Chiang Colvin A, Mullen J, Lovell MR, Vereeke West R, Collins MW, Groh M. The role of concussion history and gender in recovery from soccer-related concussion. The American Journal of Sports Medicine. 2009;37(9):1699-704.
178. Chamard E, Lassonde M, Henry L, Tremblay J, Boulanger Y, De Beaumont L, Théoret H. Neurometabolic and microstructural alterations following a sports-related concussion in female athletes. Brain injury. 2013;27(9):1038-46.
179. Sollmann N, Echlin PS, Schultz V, Viher PV, Lyall AE, Tripodis Y, et al. Sex differences in white matter alterations following repetitive subconcussive head impacts in collegiate ice hockey players. Neuroimage: Clinical. 2018;17:642-9.
180. Rubin TG, Catenaccio E, Fleysher R, Hunter LE, Lubin N, Stewart WF, et al. MRI-defined white matter microstructural alteration associated with soccer heading is more extensive in women than men. Radiology. 2018;289(2):478-86.
181. Wright DK, Symons GF, O’Brien WT, McDonald SJ, Zamani A, Major B, et al. Diffusion imaging reveals sex differences in the white matter following sports-related concussion. Cerebral Cortex. 2021;31(10):4411-9.
182. Mayhew L, Johnson MI, Francis P, Lutter C, Alali A, Jones G. Incidence of injury in adult elite women’s football: a systematic review and meta-analysis. BMJ Open Sport—Exercise Medicine. 2021;7(3).
183. Beech J, Jones B, King R, Bennett P, Young S, Williams S, et al. The incidence and burden of injuries in elite English youth female soccer players. Science and Medicine in Football. 2022;6(5):605-15.
184. Gouttebarge V, Castaldelli-Maia JM, Gorczynski P, Hainline B, Hitchcock ME, Kerkhoffs GM, et al. Occurrence of mental health symptoms and disorders in current and former elite athletes: a systematic review and meta-analysis. British Journal of Sports Medicine. 2019;53(11):700-6.
185. McAuley AB, Baker J, Kelly AL. Defining “elite” status in sport: from chaos to clarity. German Journal of Exercise and Sport Research. 2022;52(1):193-7.

186. De Pauw K, Roelands B, Cheung SS, De Geus B, Rietjens G, Meeusen R. Guidelines to classify subject groups in sport-science research. International Journal of Sports Physiology and Performance. 2013;8(2):111-22.
187. McKay AK, Stellingwerff T, Smith ES, Martin DT, Mujika I, Goosey-Tolfrey VL, et al. Defining Training and Performance Caliber: A Participant Classification Framework. International Journal of Sports Physiology and Performance. 2022;17(2):317-31.
188. Decroix L, De Pauw K, Foster C, Meeusen R. Guidelines to classify female subject groups in sport-science research. International Journal of Sports Physiology and Performance. 2016;11(2):204-13.
189. Kirkland A, O'Sullivan M. There is no such thing as an international elite under-9 soccer player. Journal of Sports Science & Medicine. 2018;17(4):686.
190. Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. A product from the ESRC methods programme Version. 2006;1(1):b92.
191. Gozillon A, Neys OH. The impact of (mega)sporting events? Comparative study of the institutionalisation of 'women's' football. International Journal of Sport Policy and Politics. 2021;13(3):425-42.
192. Williams J. A game for rough girls?: a history of women's football in Britain: Routledge; 2013.
193. Bell B. Women's Euro 2005 a 'watershed' for women's football in England and a new era for the game? Sport in History. 2019;39(4):445-61.
194. Fraser K. Sisters doing it for themselves: the rich history of women's football in Scotland from the 1960s to 2020. Sport in History. 2020;40(4):456-81.
195. Woodhouse D, Fielding-Lloyd B, Sequerra R. Big brother's little sister: The ideological construction of women's super league. Sport in Society. 2019;22(12):2006-23.
196. Themen K. Female football players in England: examining the emergence of third-space narratives. Soccer & Society. 2016;17(4):433-49.
197. Mondal S. She kicks: The state of competitive balance in the top five women's football leagues in Europe. Journal of Global Sport Management. 2023;8(1):432-54.
198. Harris J. Doing gender on and off the pitch: The world of female football players. Sociological Research Online. 2007;12(1).
199. Grice B, Mackintosh C, Oldfield S, Durden-Myers E. From performing gender to symbolic violence in English women's football:"Women are not supposed to be viewed in this way". Managing Sport and Leisure. 2023:1-22.
200. Pielichaty H. Pleasure and the sanctuary paradox: Experiences of girls and women playing soccer. International Review for the Sociology of Sport. 2020;55(6):788-806.
201. Fleming M, Parker A. Chaplaincy and Scottish women's football: sport, gender, and pastoral care. Journal of Contemporary Religion. 2022;37(1):145-64.
202. Devonport TJ, Russell K, Leflay K, Conway J. Gendered performances and identity construction among UK female soccer players and netballers: a comparative study. Sport in Society. 2019;22(7):1131-47.
203. Caudwell J. 'Out on the field of play: women's experiences of gender and sexuality in football contexts'. British Football and Social Exclusion: Taylor and Francis Group; 2004. p. 127-46.
204. Harris J. Doing gender on and off the pitch: The world of female football players. Sociological Research online. 2007;12(1):140-51.
205. Ratna A. 'Who wants to make aloo gobi when you can bend it like Beckham?'British Asian females and their racialised experiences of gender and identity in women's football. Soccer & Society. 2011;12(3):382-401.
206. Ahmad A. British football: where are the Muslim female footballers? Exploring the connections between gender, ethnicity and Islam. Soccer & Society. 2011;12(3):443-56.
207. Ratna A. 'Taking the power back!' The politics of British–Asian female football players. Young. 2010;18(2):117-32.
208. Byrne H. Where are we now? A review of research on the history of women's soccer in Ireland. Sport in History. 2019;39(2):166-86.

209. Cooper C. A reflection on contemporary myths of women's football: a historical analysis. Accounting, Auditing & Accountability Journal. 2023(ahead-of-print).
210. Pielichaty H. Identity salience and the football self: a critical ethnographic study of women and girls in football. Qualitative Research in Sport, Exercise and Health. 2019;11(4):527-42.
211. Williams J. The fastest growing sport? women's football in England. Soccer & Society. 2003;4(2-3):112-27.
212. Themen K, van Hooff J. Kicking against tradition: women's football, negotiating friendships and social spaces. Leisure Studies. 2017;36(4):542-52.
213. Williams J. Globalising Women's Football: Europe, Migration and Professionalisation: Peter Lang; 2013.
214. McGreary M, Morris R, Eubank M. Retrospective and concurrent perspectives of the transition into senior professional female football within the United Kingdom. Psychology of Sport and Exercise. 2021;53:101855.
215. Culvin A, Bowes A. The Incompatibility of Motherhood and Professional Women's Football in England. FRONTIERS IN SPORTS AND ACTIVE LIVING. 2021;3.
216. Forsyth J, Sams L, Blackett A, Ellis N, Abouna M-S. Menstrual cycle, hormonal contraception and pregnancy in women's football: perceptions of players, coaches and managers. Sport in Society. 2023;26(7):1280-95.
217. Ahmad A. British Muslim Female Experiences in Football: Islam, Identity and the Hijab. Race Ethnicity and Football: Persisting Debates and Emergent Issues2012. p. 101-14.
218. Dunn C. Elite footballers as role models: Promoting young women's football participation. Soccer & Society. 2016;17(6):843-55.
219. Simpson D, MARTINDALE RLJ, Travlos A, Souglis A, Andronikos G. An investigation of the talent development pathway in Scottish female football. International Journal of Sport Psychology. 2022;53:218-41.
220. Skillen F, Byrne H, Carrier J, James G. 'The game of football is quite unsuitable for females and ought not to be encouraged': a comparative analysis of the 1921 English Football Association ban on women's football in Britain and Ireland. Sport in History. 2022;42(1):49-75.
221. Jenkel L. The FA's ban of women's football 1921 in the contemporary press–a historical discourse analysis. Sport in History. 2021;41(2):239-59.
222. Williams J. 'We're the lassies from Lancashire': Manchester Corinthians Ladies FC and the use of overseas tours to defy the FA ban on women's football. Sport in History. 2019;39(4):395-417.
223. Welford J. Globalising Women's Football: Europe, Migration and Professionalization. INTERNATIONAL JOURNAL OF THE HISTORY OF SPORT. 2015;32(5):726-8.
224. Macbeth J. The development of women's football in Scotland. Sports Historian. 2002;22(2):149-63.
225. Lopez S. Women on the ball: a guide to women's football: Scarlet Press; 1997.
226. Clarkson BG, Culvin A, Pope S, Parry KD. Covid-19: Reflections on threat and uncertainty for the future of elite women's football in England. Managing Sport and Leisure. 2022;27(1-2):50-61.
227. Clarkson BG, Parry KD, Culvin A, Pope S. An institutional analysis of gender (in) equalities, COVID-19 and governance of elite women's football in Australia, England and the USA. Sport, Business and Management: An International Journal. 2022;12(4):459-78.
228. Pope S, Allison R. Euro 2022 can be a catalyst for women's football in England – here's how The Conversation2022 [Available from: https://theconversation.com/euro-2022-can-be-a-catalyst-for-womens-football-in-england-heres-how-187874.
229. Grix J, Carmichael F. Why do governments invest in elite sport? A polemic. International Journal of Sport Policy and Politics. 2012;4(1):73-90.
230. Ruiz-Pérez I, López-Valenciano A, Jiménez-Loaisa A, Elvira JL, Croix MDS, Ayala F. Injury incidence, characteristics and burden among female sub-elite futsal players: a prospective study with three-year follow-up. PeerJ. 2019;7:e7989.
231. Mallo J, González P, Veiga S, Navarro E. Injury incidence in a Spanish sub-elite professional football team: A prospective study during four consecutive seasons. Journal of Sports Science & Medicine. 2011;10(4):731.

232. Gledhill A, Harwood C. A holistic perspective on career development in UK female soccer players: A negative case analysis. Psychology of Sport and Exercise. 2015;21:65-77.
233. De Bosscher V, De Knop P, Van Bottenburg M, Shibli S. A conceptual framework for analysing sports policy factors leading to international sporting success. European Sport Management Quarterly. 2006;6(2):185-215.
234. Houlihan B, Green M. Comparative elite sport development. Comparative elite sport development: Systems, structures and public policy: Routledge; 2007. p. 1-25.
235. Gomez RG, White DA. Using verbal fluency to detect very mild dementia of the Alzheimer type. Archives of Clinical Neuropsychology. 2006;21(8):771-5.
236. Gallo V, Seghezzo G, Basinas I, Williamson E, van Hoecke Y, Davoren D, et al. Cognitive function among former professional male soccer players–the HEADING study. medRxiv. 2024:2024.03. 26.24304885.
237. Alosco ML, Barr WB, Banks SJ, Wethe JV, Miller JB, Pulukuri SV, et al. Neuropsychological test performance of former American football players. Alzheimer's Research & Therapy. 2023;15(1):1.
238. Murray L, Vuoskoski P, Wellman J, Hebron C. ‘’It was the end of the world”–The lifeworld of elite male rugby union players living with injury. An interpretative phenomenological analysis. Physiotherapy Theory and Practice. 2022;38(9):1219-32.
239. Brown JC, Kerkhoffs G, Lambert MI, Gouttebarge V. Forced retirement from professional rugby union is associated with symptoms of distress. International Journal of Sports Medicine. 2017;38(08):582-7.
240. Arvinen-Barrow M, Hurley D, Ruiz MC. Transitioning out of professional sport: The psychosocial impact of career-ending injuries among elite Irish rugby football union players. Journal of Clinical Sport Psychology. 2017;11(1):67-84.
241. White AJ, Batten J, Robinson S, Anderson E, Burns A, Batey J, et al. Tackling in physical education rugby: an unnecessary risk? Injury Prevention. 2018;24(2):114-5.
242. Batten J, White AJ, Anderson E. Preventing penalty corner injuries and head trauma in field hockey: time to consider the power play? : BMJ Publishing Group Ltd and British Association of Sport and Exercise Medicine; 2016. p. 639-40.
243. Darby P, Esson J, Ungruhe C. ‘Becoming a somebody’through football. African Football Migration: Manchester University Press; 2022. p. 104-23.
244. Manzo K. Development through football in Africa: Neoliberal and postcolonial models of community development. Geoforum. 2012;43(3):551-60.
245. Darby P. Gains versus drains: Football academies and the export of highly skilled football labor. The Brown Journal of World Affairs. 2012;18(2):265-77.
246. England S. Active Lives Adult Nov 21-22 Tables 6-8 Types of Activity. Sport England2022.
247. Kolodziej MA, Koblitz S, Nimsky C, Hellwig D. Mechanisms and consequences of head injuries in soccer: a study of 451 patients. Neurosurgical Focus. 2011;31(5):E1.
248. Daneshvar DH, Nowinski CJ, McKee AC, Cantu RC. The epidemiology of sport-related concussion. Clinics in Sports Medicine. 2011;30(1):1-17.
249. (ONS) OfNS. Ethnic group, England and Wales: Census 2021. 2021.
250. Cheyne JA, Solman GJ, Carriere JS, Smilek D. Anatomy of an error: A bidirectional state model of task engagement/disengagement and attention-related errors. Cognition. 2009;111(1):98-113.

Appendices

Appendix A: Feedback Form

SCORES Feedback Form

01 June, 2022

How to interpret your results

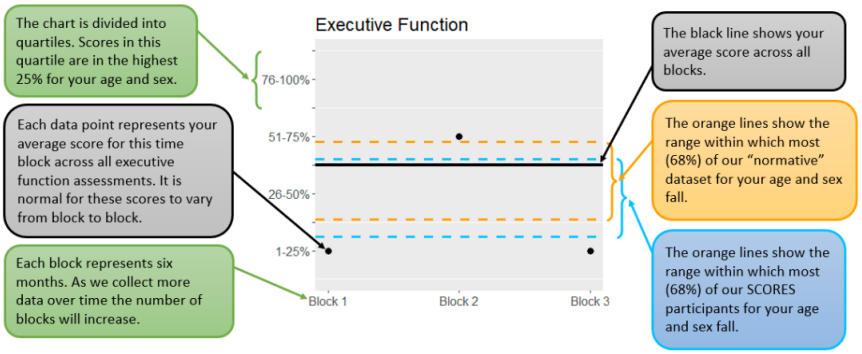
Introduction to the chart.

Your results are grouped into different brain functions. On the next page you will see your results for executive function, Memory, Reaction Time and Spatial Navigation. You can read more about how we assess these brain functions by logging into your SCORES account and opening the module titled “Information About Our Assessments”.

In the example chart below we show the average result from all the assessments that measured executive function. The black dots are this person’s executive function score at each six month block. Block one is the first time this person completed assessments and block two is the second time they were completed them, six months later. While we assess you every three months, we use different tasks so they repeat every six months.

For comparison, the scores are plotted in groups against people of similar age and sex. The four groups on the left side of the chart indicate the quartile. The highest 25% of scores for your age/sex group are at the top of the chart followed by 51-75%, 26-50% and then 1-25%.

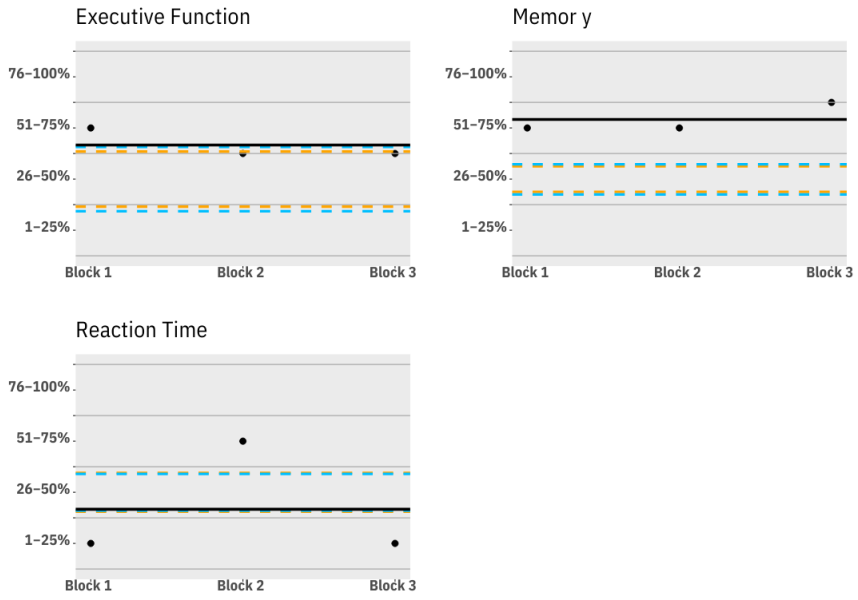
The majority of the participants in SCORES fall within the two blue lines. For comparison, we also show results from data calculated from thousands of people from other studies who performed these same assessments. The orange lines show the range in which most of these people fall. For the technically minded, these bands represent 1 standard deviation from the mean. This means 68% of the people who were assessed in this age/sex group had a score within these lines. The black line indicates the average score for this person across each of the blocks.



What does this mean?

In the example above, the result from the first block of assessments is 26%, falling within the second quartile (26-50%) of the scores for that person's age and sex. It is also fits within the normal band of scores for both participants in the SCORES study and our large dataset of normative values. The results from Block 2 is 51%, falling within the third quartile (51-75%) when compared with people of similar age and sex. This result is also higher than the SCORES group mean and normative mean. Variations such as this are normal and likely to happen over time. Variation can be caused by numerous factors including technical issues, how the person felt on the day, and whether or not there were any distractions. For any individual we need data over a longer period of time before we can see meaningful changes. As you continue in the study we will add more data to these plots and this will generate a clearer picture of your brain health over time.

Your Performance:



For an explanation of each assessment please visit www.scoresproject.org and log into your SCORES account. **We will contact you if we believe your results warrant further investigation by your GP.** Don't be alarmed if some of your scores are in red, that's normal and could be caused by a range of factors including technical issues, feeling poorly on the day or distractions.

What happens next?

We encourage you to continue completing the brain health assessments every three months. Once per year we will contact you to ask you to provide your consent to participate in the study and to let us know if you would like to continue to receive the results of your assessment. We will continue to update you on news about the project through the SCORES newsletter, which can be found on our website at www.scoresproject.org.

If you are concerned about your results, or about your brain health, please arrange to speak to your GP. If you would like to speak with a member of our team about your results or the assessments, please contact us at scoresproject@uea.ac.uk to arrange a discussion. They can explain the assessments that you completed, and help you to seek further support or assessment.

If you are looking for more information about brain health, you may find the following links to resources helpful:

Mind

Sane

Age UK

PFA - Wellbeing

Alzheimer's Society

Dementia UK

Headway

Appendix B: Study 1 interview topic guide

Topic	**List of Questions**
Reason for participation	1. What prompted you to get involved as a participant in the SCORES project? 2. How did you hear about the study? 3. From how it was described, was it what you expected? (Prompt: If not, why not)
Experience in SCORES as a participant	4. Please describe your experience as a SCORES participant. Prompt: a. How did you find completing the online questionnaire? b. How did you get on with completing the brain health assessments? c. Was there an assessment that was particularly difficult/easy? In contrast, was there one that you found easier? d. Did you contact us for support in using the website, or completing any of the assessments and questionnaires? If so, how was the support that you received? Were you able to complete the tests OK after you got the support? 5. The feedback form revealed that many of our participants found the Sustained Attention to Response Task difficult (remind them which one). Was this the case for you? If yes, what made this test difficult to complete? 6. Would you recommend participating in this study to a friend? If yes, what would you say are the positives and negatives about taking part in this study? If no, why not?
Feasibility and Acceptability	7. Having completed the first round of SCORES, how do you feel about completing further rounds of assessments and questionnaires in the future? 8. The feedback form revealed that most participants took roughly 30 minutes to complete the first round of SCORES.

	Future rounds of SCORES will be roughly the same length. How do you feel about this commitment of time? 9. The SCORES project intends to run for at least 10 years. Do you think participants would be able to commit to regular participation for at least 10 years? (Prompt: what would feel reasonable to you? What leads you to think that?) Prompt: a. What can we do to support that participation in the long term?
Receiving feedback about brain health	In the feedback form we asked whether you would like to receive feedback about your brain health. We are currently in the process of making a decision about how to do this in a sensitive but helpful way. It is a bit tricky, as we know that many of our participants would like to receive feedback, but we need to work out how to do this in a way that is sensitive, and also in a way where we can deal with the situation where someone shows signs of poor brain health. So, to help us to make this decision, we would like to ask you as a participant, how would you feel about receiving feedback about your brain health? 10. How do you feel about receiving feedback about your brain health? 11. How would you prefer to receive feedback about your brain health? 12. If a participant shows signs of poor brain health, how should this information be shared with them?
Mood and Behaviour Questionnaires	The SCORES project aims to study the brain health of people exposed to repetitive head injury. In some cases, people who are exposed to repetitive head injury may go on to develop a specific type of dementia called chronic traumatic encephalopathy (CTE). Not only does CTE effect brain health, but it can also lead to depression, anxiety, aggression, or impulsivity. In the future the SCORES project would like to measure these changes using specific questionnaires.

	13. How do you feel about such questionnaires asking about your mental health or behaviour? 14. How would you feel about answering questions about your mental health in the future? 15. If a participant is showing signs of depression or anxiety, how should this information be communicated with them? 16. How would you feel about answering questions about aggressive or impulsive behaviour? 17. If a participant is showing high levels of aggression, how should this information be communicated with them? 18. If a participant is showing high levels of impulsivity, how should this information be communicated with them?

Appendix C: Study 2 interview topic guide

Examiner Contact	In-Person Assessment 1. How did you feel about the level of instructions in the in-person session? 2. Did you feel that you had the opportunity to ask questions about the assessments? 3. Did you ask the researcher any questions about the assessments? a. If so, how did you feel the ability to ask questions effected your experience? b. If no, how did you feel not asking questions may have affected your experience? i.Why didn't you ask any questions? 4. How did the presence of a researcher effect your experience of the in-person assessments? Online Assessment 1. How did you feel about the level of instructions in the online session? 2. Did you feel that you had the opportunity to ask questions about the assessments? 3. Did you contact the research team to ask any questions? a. If so, how might asking questions have influenced your experience? b. If no, how might not asking questions have influenced your experience? i.Why didn't you ask any questions? Both 1. Were there any other differences between the two sessions in terms of contact with the researcher that might have influenced your performance or experience?
Testing Environment	In-Person Assessment 1. Please could you describe the setting of the in-person session? (prompt about noise, comfort)

	a. How do you think this setting might have affected your performance? 2. Did you experience any distractions during the in-person testing? a. How might these have affected performance? Online Assessment 3. Please could you describe the setting of where you completed the online assessments (prompt about noise, comfort) a. How do you think this setting might have affected your performance? 4. Did you experience any distractions during the online assessments? a. How might these have affected performance? Both 5. Were there any other differences between the two sessions in terms of testing environment that might have influenced your performance or experience? Balance between work and volunteering?
Workstation	In-Person Assessment 1. Please could you describe the device and workstation that you used to complete the in-person assessments? a. How did you get on using this device? b. How do you think this workstation may have influenced your performance? Online Assessment 2. Please could you describe the device and workstation that you used to complete the online assessments? a. How did you get on using this device? b. How do you think this workstation may have influenced your performance? Both

	1. Were there any other differences between the two workstations that may have influenced your performance? (internet speed, hardware, software, or processing speed? Paper and pencil
Other	1. Were there any other differences between the two assessment sessions that you think may have influenced your performance in either session? 2. Preference?

Appendix D: Patient Health Questionnaire (PHQ-9)

Over the last 2 weeks, how often have you been bothered by any of the following problems?

1. Little interest or pleasure in doing things
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
2. Feeling down, depressed, or hopeless
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
3. Trouble falling or staying asleep, or sleeping too much
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
4. Feeling tired or having little energy
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
5. Poor appetite or overeating
    a. Not at all
    b. Several days
    c. More than half the days

   d. Nearly every day

6. Feeling bad about yourself – or that you are a failure or have let yourself or your family down

   a. Not at all

   b. Several days

   c. More than half the days

   d. Nearly every day

7. Trouble concentrating on things, such as reading a newspaper or watching television

   a. Not at all

   b. Several days

   c. More than half the days

   d. Nearly every day

8. Moving or speaking so slowly that other people could have noticed

   a. Not at all

   b. Several days

   c. More than half the days

   d. Nearly every day

9. Thoughts that you would be better off dead or of hurting yourself in some way

   a. Not at all

   b. Several days

   c. More than half the days

   d. Nearly every day

Appendix E: General Anxiety Disorder Questionnaire (GAD-7)

Over the last two weeks, how often have you been bothered by the following problems?

1. Feeling nervous, anxious or on edge
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
2. Not being able to stop or control worrying
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
3. Worrying too much about different things
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
4. Trouble relaxing
    a. Not at all
    b. Several days
    c. More than half the days
    d. Nearly every day
5. Being so restless that it is hard to sit still
    a. Not at all
    b. Several days
    c. More than half the days

   d. Nearly every day

6. Becoming easily annoyed or irritable

   a. Not at all

   b. Several days

   c. More than half the days

   d. Nearly every day

7. Feeling afraid as if something awful might happen

   a. Not at all

   b. Several days

   c. More than half the days

   d. Nearly every day

Appendix F: Buss-Perry Aggression Questionnaire (AQ)

"Please rate each of these items in terms of how characteristic they are for you. Put an X in the box that applies best to you."

Factor		Extremely Uncharacteristic	Uncharacteristic	Neither Uncharacteristic or Characteristic	Characteristic	Extremely Characteristic
Physical Aggression						
1	Once in a while I can't control the urge to strike another person.					
2	Given enough provocation, I may hit another person.					
3	If somebody hits me, I hit back.					
4	I get into fights a little more than the average person.					
5	If I have to resort to violence to protect my rights, I will.					
6	There are people who pushed me so far that we came to blows.					
7	I can think of no good reason for ever hitting a person.*					
8	I have threatened people I know.					
9	I have become so mad that I have broken things.					
Verbal Aggression						
1	I tell my friends openly when I disagree with them.					
2	I often find myself disagreeing with people.					
3	When people annoy me, I may tell them what I think of them.					
4	I can't help getting into arguments when people disagree with me.					
5	My friends say that I'm argumentative.					
Anger						
1	I flare up quickly but get over it quickly.					
2	When frustrated, I let my irritation show.					
3	I sometimes feel like a powder keg ready to explode.					
4	I am an even tempered person. *					
5	Some of my friends think I'm a hothead.					
6	Sometimes I fly off the handle for no good reason.					
7	I have trouble controlling my temper.					
Hostility						
1	I am sometimes eaten up with jealousy					
2	At times I feel I have gotten a raw deal of out life.					
3	Other people always seem to get the breaks.					
4	I wonder why sometimes I feel so bitter about things.					
5	I know that "friends" talk about me behind my back.					
6	I am suspicious of overly friendly strangers.					
7	I sometimes feel that people are laughing at me behind my back					
8	When people are especially nice, I wonder what they want.					

Appendix G: Barratt Impulsiveness Scale (BIS-15)

"Read each statement and select how often this applies to you."

Factor		Rarely/Never	Occasionally	Often	Almost Always/Always
Motor Impulsivity					
1	I act on impulse.				
2	I act on the spur of the moment.				
3	I do things without thinking.				
4	I say things without thinking.				
5	I buy things on impulse.				
Non-planning					
1	I plan for job security (inverted).				
2	I plan for the future (inverted).				
3	I save regularly (inverted).				
4	I plan tasks carefully (inverted).				
5	I am a careful thinker (inverted)				
Attention Impulsivity					
1	I am restless at lectures or talks.				
2	I squirm at plays or lectures.				
3	I concentrate easily (inverted).				
4	I don't pay attention.				
5	I am easily bored solving thought problems.				

Appendix H: Further analysis of results from the Sustained Attention to Response Task

The SCORES Study battery originally contained an assessment called the Sustained Attention to Response Task (SART) which assesses sustained attention and inhibition. In this assessment, participants are tasked with watching a series of numbers between 1-9 appear in the middle of the screen and are asked to press the space bar when they see any number except for the number 3. The assessment lasts approximately 5 minutes and requires sustained concentration. Participants must respond as fast as possible, but participant responses less than 200ms are deemed ineligible according to traditional data processing protocol (250). A shorter response time than 200ms may indicate an anticipatory response, rather than processing the new stimulus. However, the majority of participants in their first set of assessments recorded mean responses less than 200ms. In a spot check of 10% of the data, only one participant had a mean response time of more than 200ms. Based on this, the original decision was not to include the SART data within the aggregated statistics due to concerns over response times. However, given that this problem was so widespread within the data, we considered if the issue may lie with the online recording of the responses, and if we could still produce meaningful data from the SART assessments.

To investigate this, the following outcomes were collected from the SART data. The correct commission percentage was calculated as follows, where target responses refer to participants successfully abstaining from responding to number 3's:

$$Correct\ Commission\ \% = \frac{Correct\ Target\ Responses}{Total\ Targets} * 100$$

The correct omission percentage was calculated as follows, where non-target responses refer to participants successfully responding to numbers other than 3:

$$Correct\ Omission\ \% = \frac{Correct\ Non\ Target\ Responses}{Total\ Non\ Targets} * 100$$

The preceding correct non-target reaction time was calculated as the mean reaction time of responses prior to the presentation of a non-target (number other than 3). The commission error reaction time was calculated as the mean reaction time of incorrect responses to a target (when participants pressed the space bar in response to a number 3). Both reaction time units are milliseconds.

In order to test the validity of the SART data, three steps were performed. First, SART outcomes were plotted to examine distribution of results. Secondly, the intraclass correlation coefficient was calculated for first and second SART assessments to examine test-retest reliability. Finally, performance on SART outcomes was then correlated with performance on tasks of executive function and reaction time to understand construct validity.

SART Results

Figure 24 demonstrate the distribution of results for the four main outcomes of the SART assessment. The red and blue plots present the distribution of results from the Commission and Omission outcomes. These outcomes represent the percentage of correct commission and omission responses out of all targets and non-targets, where a higher percentage correct indicates better performance on the task. Figure 24 demonstrates that the medians of the data lie below 50% in both outcomes, indicating that as a whole participants tend to get less than 50% of responses correct. As demonstrated by the violin plot, the majority of scores are relatively low in these outcomes. The grey and gold plots present the mean reaction times for an incorrect commission and preceding a correct omission. Both medians lie below 250ms, and the majority of responses are relatively fast.

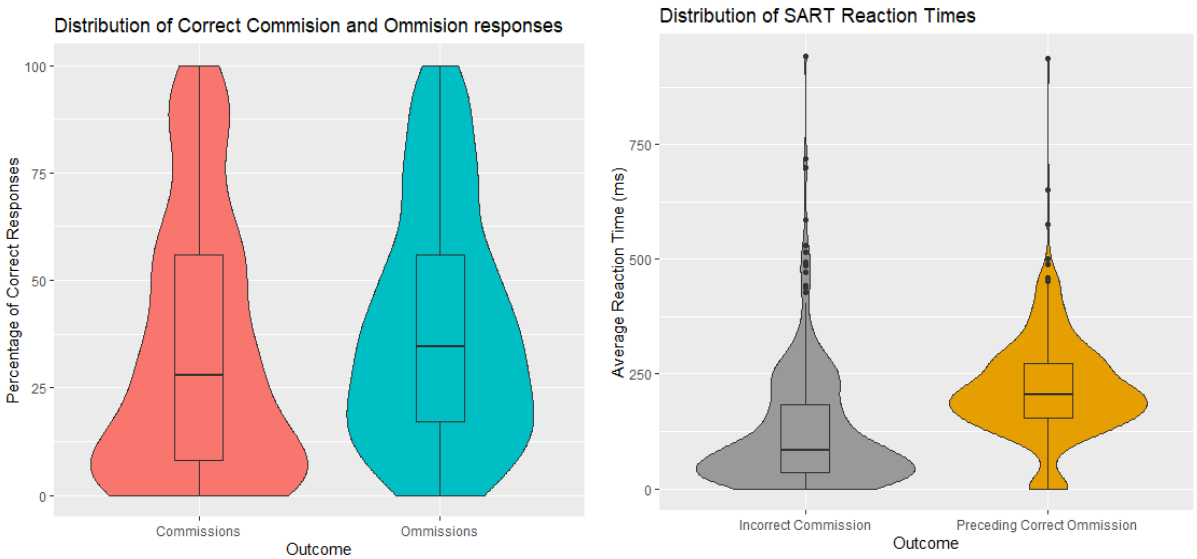


Figure 24: Violin plots demonstrating the distribution of data for each SART outcome.

In order to examine if these lower scores were related to the novelty of the task for participants, a test-retest analysis was performed to determine the correlation between scores from the first and second set of assessments. Using the intraclass correlation coefficient (ICC), outcomes from Assessment Set 1 and Assessment Set 2 were compared and the correlation of results is presented in Figure 25. The ICC of correct commissions was 0.65, representing moderate correlation between assessment 1 and 2. The ICC of correct omissions was 0.62, representing moderate correlation between assessments 1 and 2. The ICC of reaction times for incorrect

commissions was 0.50, representing a moderate correlation between assessments 1 and 2. The ICC of reaction times preceding a correct omission was 0.369, representing a poor correlation between assessments 1 and 2. Across all outcomes, Figure 25 demonstrate participants who had zero correct responses in either outcome in assessment 1, but then scored much higher in assessment 2.

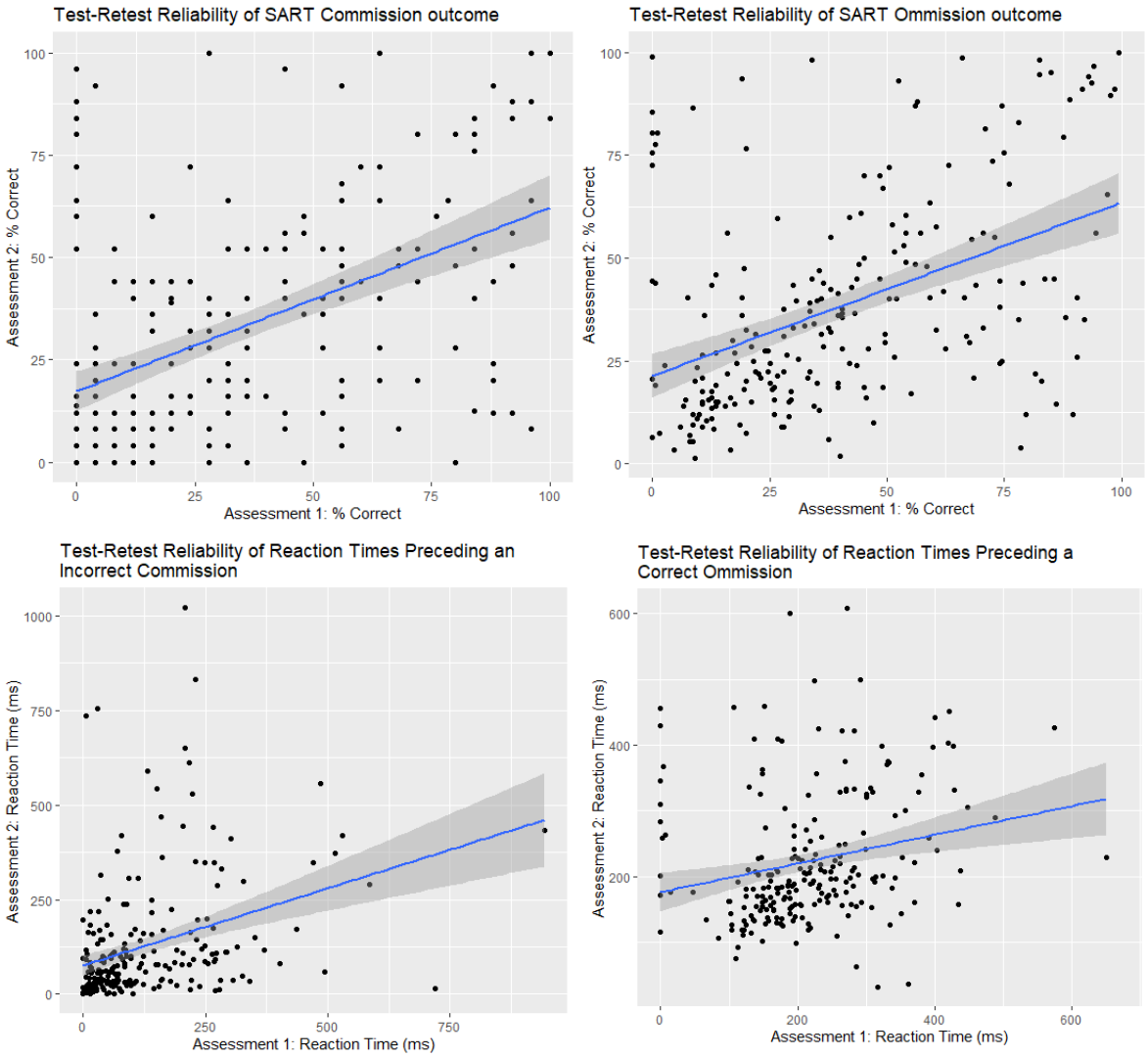


Figure 25: Test-retest reliability of data for each SART outcome.

Results from the SART were also compared to performance on other executive function and reaction time tasks, to assess the validity of the measure. Figure 26 demonstrate the nature of these relationships. These plots suggest a low correlation between performance on the TMTB and SART commission outcomes, whereby higher scores on the SART outcomes are slightly associated with worse performance on the TMTB. A possible reason for this may be that participants who are taking longer to complete the TMTB may be taking longer to respond on the SART and are therefore more likely to respond correctly to targets and non-targets. Findings also demonstrate that reaction times are in general quicker in the SART assessment outcomes than in the simple reaction time assessment, which also supports the suggestion that participants are overemphasising speed of responses rather than accuracy.

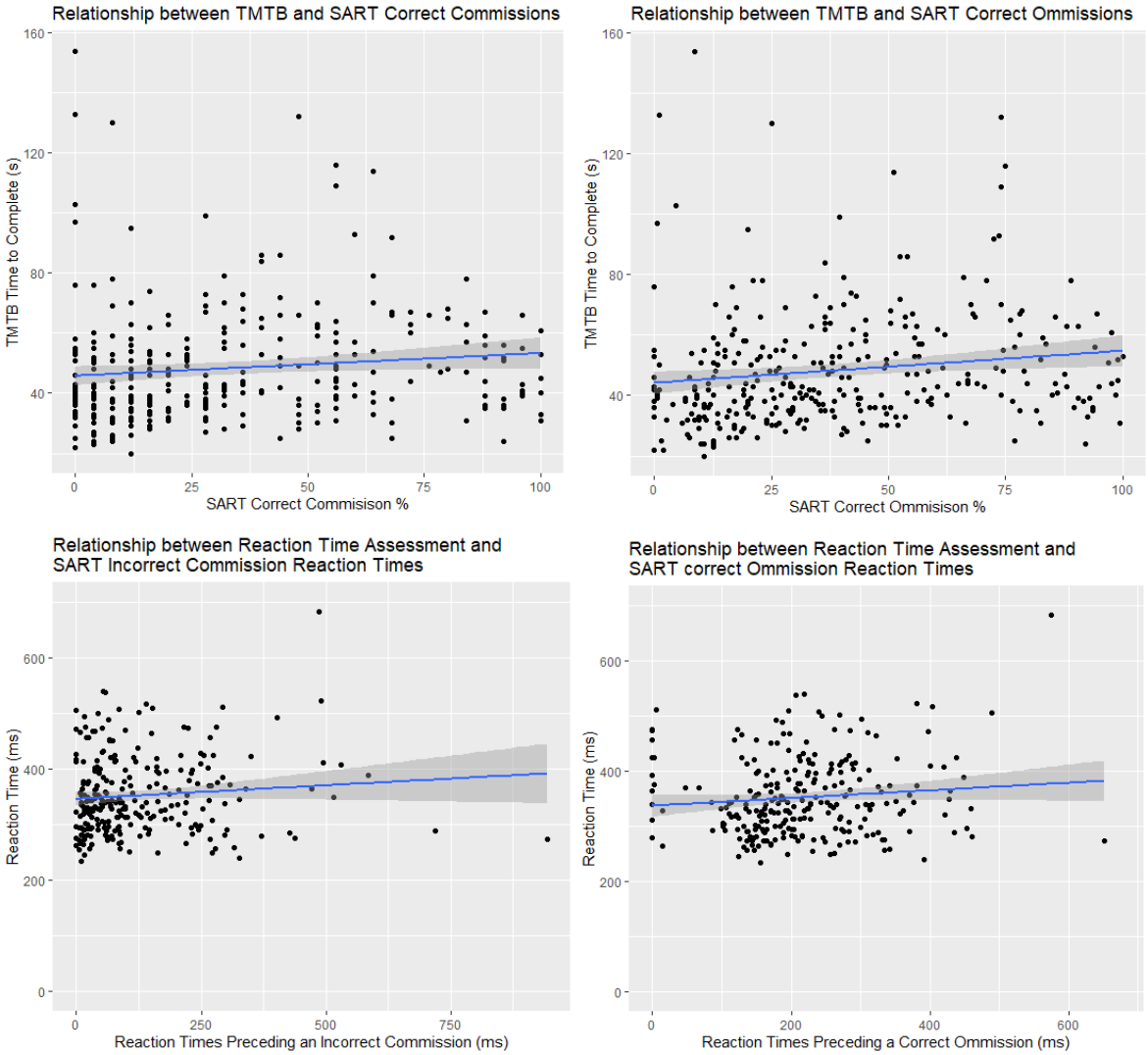


Figure 26: Scatterplots demonstrating the nature of the relationship between SART outcomes and other measures of executive function and reaction time in the SCORES dataset.

The findings from these analyses suggest that the data from SART outcomes shows high variability, with the majority of data demonstrating poor performance on the task. Outcomes from the SART show only moderate test-retest reliability, suggesting that poor performance on the first set of assessments may reflect the novelty of the task. Furthermore, performance on the SART shows poor correlation with other assessments of executive function and reaction times. This suggests that results from the SART may more likely reflect participants overemphasising the speed of reactions rather than the accuracy of responses, as supported by other investigations (163). Therefore, the results from the SART assessment in this test battery may not accurately reflect executive function enough to include as an accurate measure of this domain.

	All	Contact Sport Males	Contact Sport Females	Non-Contact Sport Males	Non-Contact Sport Females	No Sport Males	No Sport Females	Professional Contact Sport Males	Professional Contact Sport Females	Amateur Contact Sport Males	Amateur Contact Sport Females	Recreational Contact Sport Males	Recreational Contact Sport Females	Professional Non-Contact Sport Males	Professional Non-Contact Sport Females	Amateur Non-Contact Sport Males	Amateur Non-Contact Sport Females	Recreational Non-Contact Sport Males	Recreational Non-Contact Sport Females
n	328	192	44	9	49	9	25	77	5	103	36	12	3	0	1	5	22	4	26
Age	59.01	59.39	52.00	57.67 (8.00)	62.82 (8.33)	59.89	61.28 (4.79)	59.81	46.4	58.74 (9.68)	52.94 (9.39)	62.25 (8.70)	50.00	NA	53.00	55.2 (9.07)	62.59 (8.24)	60.75 (6.18)	63.38 (8.48)
(mean, SD)	(9.60)	(9.61)	(9.40)			(12.49)		(9.66)	4.28)				(14.80)		NA				
Ethnicity, n (%)																			
White (Eng./Scot./Welsh/N. Irish)	305 (93)	180 (93.8)	41 (93.2)	7 (77.8)	45 (91.8)	8 (88.8)	24 (96)	72 (93.5)	3 (60)	98 (95.1)	35 (97.2)	10 (83.3)	3 (100)	NA	-	4 (80)	20 (90.9)	3 (75)	25 (96.2)
White (Irish)	4 (1.2)	3 (1.6)	1 (2.3)	-	-	-	-	1 (1.3)	0	2 (1.9)	1 (2.8)	-	-		-	-	-	-	-
White (Other)	14 (4.3)	5 (2.6)	1 (2.3)	1 (11.1)	4 (8.2)	1 (11.1)	1 (4)	1 (1.3)	1 (20)	3 (2.9)	-	1 (8.3)	-		1 (100)	-	2 (9.1)	1 (25)	1 (3.8)
Black (Caribbean)	1 (0.3)	1 (0.5)	-	-	-	-	-	1 (1.3)	-	-	-	-	-		-	-	-	-	-
Mixed (White and Black African)	1 (0.3)	1 (0.5)	-	-	-	-	-	1 (1.3)	-	-	-	-	-		-	-	-	-	-
Mixed (White and Asian)	1 (0.3)	-	-	1 (11.1)	-	-	-	-	-	-	-	-	-		-	1 (20)	-	-	-
Other Ethnic Group	1 (0.3)	1 (0.5)	-	-	-	-	-	-	-	-	-	1 (8.3)	-		-	-	-	-	-
Missing Response	1 (0.3)	1 (0.5)	1 (2.3)	-	-	-	-		1 (20)	-	-	-	-		-	-	-	-	-
Handedness, n (%)																			
Right	291 (88.7)	169 (88)	36 (81.8)	7 (77.8)	46 (93.9)	9 (100)	24 (96)	68 (88.3)	5 (100)	90 (87.4)	28 (77.8)	11 (91.7)	3 (100)	NA	1 (100)	4 (80)	22 (100)	3 (75)	23 (88.5)
Left	28 (8.5)	18 (9.4)	7 (15.9)	1 (11.1)	2 (4.1)	-	-	6 (7.8)	-	11 (10.7)	7 (19.4)	1 (8.3)	-		-	1 (20)	-	-	2 (7.7)
Ambidextrous	8 (2.4)	4 (2)	1 (2.3)	1 (11.1)	1 (2)	-	1 (4)	3 (3.9)	-	1 (1)	1 (2.8)	-	-		-	-	-	1 (25)	1 (3.8)
Missing Response	1 (0.3)	1 (0.5)	-	-	-	-	-	-	-	1 (1)	-	-	-		-	-	-	-	-
Education in Years	15.52	15.11	16.78	18.78 (5.91)	16.00 (4.04)	14.00	14.88 (3.31)	14.68	17.2	15.16 (3.32)	16.88 (3.52)	17.42	15.00	NA	18.00	16.60 (2.07)	16.19 (5.28)	21.50 (8.35)	15.77 (2.85)
(mean, SD)	(3.71)	(3.41)	(3.84)			(3.00)		(3.31)	(0.84)			(3.99)	(9.54)		NA				
Deprivation	6.49	6.75 (2.44)	6.20 (2.69)	6.25 (2.38)	5.96 (2.16)	6.38	6.12 (2.39)	7.25	5.50	6.29 (2.50)	6.29 (2.70)	7.33 (1.97)	6.00	NA	5.00	5.60 (2.88)	6.45 (2.16)	7.33 (0.58)	5.62 (2.16)
(mean decile, SD)	(2.41)					(1.60)		(2.32)	(1.73)				(5.66)		NA				
Age at start of highest sport level	14.01	14.59	16.43	16.78	18.50	NA	NA	17.78	30.4	12.32 (5.10)	14.92 (9.58)	13.67 (5.03)	11.33 (0.58)	NA	25.00	18.20	15.71 (10.10)	15.00 (5.35)	20.50 (15.77)
(mean, SD)	(9.08)	(5.06)	(10.19)	(12.51)	(13.50)			(2.90)	(5.32)						NA	(16.90)			
Duration of highest sport level in years	17.30	18.83	18.89	26.11	19.20	NA	NA	13.54	3.60	22.67	21.14 (11.17)	19.42 (15.45)	17.33	NA	12.00	26.00	17.27 (13.23)	26.25 (18.41)	21.15 (20.74)
(mean, SD)	(13.27)	(11.10)	(11.65)	(14.12)	(17.47)			(6.80)	(2.70)	(11.57)			(6.43)		NA	(12.02)			
Age of first exposure to contact sport	10.55	10.46	10.79	NA	NA	NA	NA	10.35 (4.37)	5.80 (0.84)	9.99 (4.86)	11.61 (10.12)	15.17 (11.56)	9.33 (3.79)	NA	NA	NA	NA	NA	NA
(mean, SD)	(6.28)	(5.43)	(9.36)																
Duration of contact sport exposure	13.65	19.19	18.18	NA	NA	NA	NA	13.54	3.60	22.83	20.22 (9.74)	23.67 (19.46)	18.00 (5.29)	NA	NA	NA	NA	NA	NA
(mean, SD)	(13.06)	(11.94)	(10.36)					(6.80)	(2.70)	(12.29)									
Current Physical Activity, n (%)																			
Sedentary	45 (13.7)	22 (11.5)	6 (13.6)	2 (22.2)	8 (16.3)	1 (11.1)	6 (24)	11 (14.3)	-	10 (9.7)	4 (11.1)	1 (8.3)	1 (33.3)	NA	-	-	3 (13.6)	2 (50)	5 (19.2)
Within WHO reccomendations	104 (31.7)	58 (30.2)	12 (27.3)	2(22.2)	15 (30.6)	5 (55.6)	12 (48)	21 (27.3)	-	33 (32)	12 (33.3)	4 (33.3)	-		1 (100)	2 (40)	7 (31.8)	-	7 (26.9)
Above WHO reccomendations	179 (54.6)	112 (58.3)	26 (59.1)	5 (55.6)	26 (53.1)	3 (3.33)	7 (28)	45 (58.4)	5 (100)	60 (58.3)	20 (55.6)	7 (58.3)	2 (66.7)		-	3 (60)	12 (54.5)	2 (50)	14 (53.8)
Smoker Status																			
Never	221 (67.4)	134 (69.8)	28 (63.6)	7 (77.8)	30 (61.2)	4 (44.4)	18 (72)	60 (77.9)	4 (80)	64 (62.1)	23 (63.9)	10 (83.3)	1 (33.3)	NA	-	3 (60)	15 (68.2)	4 (100)	15 (57.7)
Past	91 (27.7)	46 (24)	14 (31.8)	2 (22.2)	17 (34.7)	5 (55.6)	7 (28)	10 (13)	1 (20)	35 (34)	11 (30.6)	1 (8.3)	2 (66.7)		-	2 (40)	6 (27.3)	-	11 (42.3)
Current	16 (4.9)	12 (6.3)	2 (4.5)	-	2 (4.1)	-	-	7 (9.1)	-	4 (3.9)	2 (5.6)	1 (8.3)	-		1 (100)	-	1 (4.5)	-	-
% with health comorbidities	31.80	32.81	25.00	37.50	30.61	44.44	32.00	36.36	40.00	28.16	19.44	50.00	66.67	NA	0.00	25.00	36.36	50.00	26.92
% with treatments	26.30	26.56	9.09	12.50	34.69	11.11	48.00	27.27	20.00	24.27	5.56	41.67	33.33	NA	0.00	0.00	31.82	25.00	38.46
% with MHLDD	14.37	13.54	11.36	25.00	16.33	22.22	16.00	11.69	20.00	14.56	8.33	16.67	33.33	NA	100.00	0.00	9.09	50.00	19.23
Diagnosed Concussions	1.16	1.57	1.16	0.38	0.43	0	0.16	2.31 (5.80)	0.2 (0.45)	1.05 (2.38)	1.14 (2.10)	1.09 (1.76)	3.00 (5.20)	NA	1.00 NA	0.00	0.70 (1.13)	1.00 (0.00)	0.19 (0.49)
(mean, sd)	(3.35)	(4.17)	(2.28)	(0.52)	(0.85)	(0.00)	(0.37)									(0.00)			
Suspected Concussions, n (%)														NA					
None	161 (49.1)	74 (38.5)	17 (38.6)	5 (55.6)	33 (67.3)	7 (7.78)	25 (100)	38 (49.3)	-	32 (31.1)	17 (47.2)	4 (16)	-		1 (100)	3 (60)	11 (50)	2 (50)	21 (80.8)
Less than 5	119 (36.3)	82 (42.7)	18 (40.9)	4 (44.4)	13 (26.5)	2 (2.22)	-	22 (28.6)	2 (40)	54 (52.4)	14 (38.9)	6 (24)	2 (66.7)		-	2 (40)	8 (36.3)	2 (50)	5 (19.2)
5 to 9	35 (10.7)	26 (13.5)	7 (15.9)	-	2 (4.1)	-	-	12 (15.6)	2 (40)	13 (12.6)	4 (11.1)	1 (4.0)	1 (33.3)		-	-	2 (9.1)	-	-
10 to 14	7 (2.13)	4 (2.1)	2 (4.5)	-	1 (2)	-	-	2 (2.6)	1 (20)	1 (0.9)	1 (2.8)	1 (4.0)	-		-	-	1 (4.5)	-	-
15 to 19	1 (0.3)	1 (0.5)	-	-	-	-	-	1 (1.3)	-	-	-	-	-		-	-	-	-	-
20 or more	5 (1.5)	5 (2.6)	-	-	-	-	-	2 (2.6)	-	3 (2.9)	-	-	-		-	-	-	-	-
% with other TBI	6.10	7.81	2.27	0.00	4.08	11.11	4.00	9.09	0.00	4.85	2.78	25.00	0.00	NA	0.00	0.00	4.55	0.00	3.85

Appendix J: Study 4 Interview Topic Guide

Preamble	How are you getting on with your assessments?
Participation in SCORES	**Assessments** *Affective Attitude* • How are you finding participating in SCORES? *Self-efficacy* • What do you think about the difficulty level of the assessments within the project? • How are you finding accessing the assessments online? *Intervention Coherence* • How does your experience in the SCORES project fit in line with your initial expectations of participation? *Perceived Effectiveness* • How effective do you feel the assessments and questionnaires are at measuring your brain health? • (relevant) *Burden* • How do you feel about completing 30-minuteassessments every 3 months? • The SCORES project has been running since October 2020, and intends to run for at least 10 years in total. What do you think about the time commitment for a further 7-8 years? *Ethicality* • How did you feel about answering questions about mental health and behaviour? • What did you think about the resources and information that we provide after the mental health and behaviour questionnaires? *Opportunity Costs*

	• What are the positives of taking part in the SCORES project for you? • What are the negatives of taking part in the project? **Feedback Process** *Affective Attitude* • How have you found the process of receiving feedback? • How did you feel when you first received feedback about your brain health?

	Ethicality • To what extent do the feedback forms meet your expectations in terms of our duty of care for providing you with feedback about your brain health? *Intervention Coherence* • What do you think about the feedback forms in terms of how easy they are to understand? *Opportunity Costs* • What do you think the benefits of receiving feedback might be? • What do you think any negatives of receiving feedback might be?
Quality of Sporting History (Football Females only)	What was the highest level that you played at? Roughly what age were you when you played at that level? When did you play at that level? For how long did you play at that level? At that time, where would the line be between amateur and elite level women's football? Considering this, would you consider the highest level you played at as amateur or elite? (Show Framework) How does this framework fit in with what you would consider as amateur or elite?

Appendix K: Review Papers

Author	Year	Journal or Book Title	Source Type	History Source	Definition Source	Definition of Elite
Williams & Woodhouse	1991	Sport in History	Journal Article	x		
Williamson	1991	The Belles of the Ball	Book	x		
Fletcher	1993	Women First: The Female Tradition in English Physical Education	Book	x		
Davies	1996	I Lost My Heart to the Belles	Book	x		
Lopez	1997	Women on the Ball: a Guide to Women's Football	Book	x		
Newsham	1997	In a League of their Own: The Dick, Kerr Ladies 1917-1965	Journal Article	x		
Harris	1998		Doctoral Thesis	x		
Caudwell	1999	Journal of Sport and Social Issues	Journal Article	x		
Melling	1999		Doctoral Thesis	x		
Pfister et al.	1999	Sport in Europe	Book Chapter	x		
Scraton et al.	1999	International Review for the Sociology of Sport	Journal Article	x		
Woodhouse & Williams	1999	Offside? The Position of Women in Football	Book	x		
Harris	2001	World Leisure Journal	Journal Article	x		
Caudwell	2002	Football Studies	Journal Article	x		

Hargreaves	2002	Sporting Females: Critical Issues in the History and Sociology of Women's Sport	Book	x
Macbeth	2002	Sports Historian	Journal Article	x
Williams	2002		Doctoral Thesis	x
Caudwell	2003	Sociology of Sport Journal	Journal Article	x
Williams	2003	A Game for Rough Girls? A History of Women's Football in Britain	Book	x
Williams	2003	Soccer & Society	Journal Article	x
Caudwell	2004	British Football & Social Exclusion	Book Chapter	x
Caudwell	2004	NA	Doctoral Thesis	x
Harris	2004	British Football & Social Exclusion	Book Chapter	x
Hong & Mangan	2004	Soccer, Women, Sexual Liberation: Kicking Off a New Era	Book	x
Reid	2004	Scottish Affairs	Journal Article	x
Harris	2005	Journal of Sport and Social Issues	Journal Article	x
Jeanes	2005	Sport, Active Leisure and Youth Cultures	Journal Article	x
Owen	2005	Kicking Against Tradition; A Career in Women's Football	Book	x

Scraton et al.	2005	International Review for the Sociology of Sport	Journal Article	x		
Bell	2006	Sporting events and event tourism: Impacts, plans and opportunities	Book Chapter	x		
Caudwell	2006	Soccer & Society	Journal Article	x		
Caudwell	2006	Sport, Sexualities and Queer Theory	Book Chapter	x		
Macbeth	2006	Football Studies	Journal Article	x		
Martin et al.	2006	Journal of Sports Sciences & Medicine	Journal Article		x	International level (Adult)
Bell & Blakey	2007	International Journal of Sport Management and Marketing	Journal Article	x		
Caudwell	2007	Gender, Place & Culture	Journal Article	x		
Clark	2007	Sport, Education and Society	Journal Article	x		
Harris	2007	Sociological Research Online	Journal Article	x		
Williams	2007	A Beautiful Game: International Perspectives on Women's Football	Book	x		
Caudwell	2008	Women, Football and Europe: Histories, Equity and Experience,	Book	x		
Macbeth	2008	Scottish Affairs	Journal Article	x		
Ratna	2008		Doctoral Thesis	x		

Welford & Kay	2008	Women, Football and Europe: Histories Equity and Experiences	Book	x		
Price et al.	2009	Applied Physiology, Nutrition and Metabolism	Journal Article		x	1st National League & Train >10h/Week (Adult)
Ratna	2010	Young	Journal Article	x	x	Academy (Youth)
Ahmad	2011	Race, Ethnicity and Football	Book Chapter	x		
Ahmad	2011	Soccer & Society	Journal Article	x		
Caudwell	2011	Soccer & Society	Journal Article	x		
Drury	2011	Soccer & Society	Journal Article	x		
Ekstrand et al.	2011	Scandinavian Journal of Medicine & Science in Sports	Journal Article		x	1st National League (Adult)
Ratna	2011	Soccer & Society	Journal Article	x		
Stirling	2011	Choregia	Journal Article	x		
Williams	2011	Women's Football, Europe and Professionalisation 1971-2011	Book	x		
Bell	2012	Sport in Society	Journal Article	x		
Munro et al.	2012	Physical Therapy in Sport	Journal Article		x	2nd National League & Train >10 h/Week (Adult)

Tate	2013	Girls with Balls: The Secret History of Women's Football	Book	x		
Taylor et al.	2013				x	Academy (Youth)
Bradley et al.,	2014	Human Movement Science			x	Club level in UEFA Cumulative Club Coefficient Ranking Top 15
Dunn & Welford	2014	Football and the Fa Women's Super League: Structure, Governance and Impact	Book	x		
Gledhill & Harwood	2014	International Journal of Sport and Exercise Psychology	Journal Article		x	International, College International, or 1st National League (Youth)
Williams	2014	A Contemporary History of Women's Sport-Part One: Sporting women, 1850-1960	Book	x		
De Ste Croix et al.	2015	Scandinavian Journal of Medicine & Science in Sports	Journal Article		x	Academy (Youth)
Gledhill & Harwood	2015	Psychology of Sport and Exercise	Journal Article	x		
Pielichaty	2015	Sport & Society	Journal Article	x		
Welford	2015	Globalising Women's Football: Europe, Migration and Professionalization.	Book Chapter	x		

Williams	2015	The International Journal of the History of Sport	Journal Article	x		
Dunn	2016	Routhledge Handbook of Football Studies	Book Chapter	x		
Themen	2016	Soccer & Society	Journal Article	x		
Croix et al.	2017				x	Highest National League Training Academy (Youth)
Emmonds et al.	2017	Science and Medicine in Football	Journal Article		x	Tier 1 Academy (Youth)
Themen & van Hooff	2017	Leisure Studies	Journal Article	x		
Williams	2017	Football and the Boundaries of History: Critical Studies in Soccer	Book Chapter	x		
Williams et al.	2017	Science and Medicine in Football	Journal Article		x	Academy (Youth)
De Ste Croix et al.	2018	The American Journal of Sports Medicine	Journal Article		x	Advanced Coaching Centre (Youth)
Dunn	2018	Soccer & Society	Journal Article	x	x	1st National League and International level (Adult)
Emmonds et al.	2018	International Journal of Sports Science & Coaching	Journal Article		x	Tier 1 Academy (Youth)
Klein	2018	Female Football Players and Fans	Book Chapter	x		
Welford	2018	Female Football Players and Fans	Book Chapter	x		

Woodward	2018	Sport in Society	Journal Article	x		
Bell	2019	Sport in History	Journal Article	x		
Byrne	2019	Sport in History	Journal Article	x		
Day & Roberts	2019	Sport in History	Journal Article	x		
Devonport et al.	2019	Sport in Society	Journal Article	x		
Emmonds et al.	2019	The Journal of Strength & Conditioning Research	Journal Article		x	1st National League (Adult)
Eustace et al.	2019	Physical Therapy in Sport	Journal Article		x	1st National league and train >10 hours/week (Adult)
Pielichaty	2019	Qualitative Research in Sport, Exercise and Health	Journal Article	x		
Stride et al.	2019	Sport, Education and Society	Journal Article	x		
Williams et al.	2019	Sport in History	Journal Article	x		
Williams	2019	Sport in History	Journal Article	x		
Woodhouse et al.	2019	Sport in Society	Journal Article	x		
Bowes et al.	2020	Managing Sport and Leisure	Journal Article	x		
Clarkson et al.	2020	Managing Sport and Leisure	Journal Article	x		
Culvin	2020		Doctoral Thesis	x		

Datson et al.	2020	Soccer & Society	Journal Article		x	National Training Camp (Youth)
Emmonds et al.	2020	The Journal of Strength & Conditioning Research	Journal Article		x	Tier 1 Academy (Youth)
Emmonds et al.	2020	The Journal of Strength & Conditioning Research	Journal Article		x	Tier 1 Academy (Youth)
Fielding-Lloyd et al.	2020	Soccer & Society	Journal Article	x		
Ford et al.	2020	Journal of Sports Sciences	Journal Article	x		
Fraser	2020	Sport in History	Journal Article	x		
Harkness-Armstrong et al.	2020	International Journal of Performance Analysis in Sport	Journal Article		x	Academy (Youth)
Harrison et al.	2020	Journal of Applied Sport Psychology	Journal Article		x	Athlete Scholarship Scheme (Adult)
Pielichaty	2020	International Review for the Sociology of Sport	Journal Article	x	x	Academy (Youth)
Sprouse et al.	2020	Sports Medicine	Journal Article		x	International Youth Level (Youth)
Themen	2020	Soccer & Society	Journal Article	x		
Abbott et al.	2021	Eating and Weight Disorders - Studies on Anorexia, Bulimia and Obesity	Journal Article		x	Highest Four Leagues (Adult)

Bishop et al.	2021	The Journal of Strength & Conditioning Research	Journal Article		x	Tier 1 Academy (Youth)
Clarkson et al.	2021	The Professionalization of Women's Sport	Book Chapter	x		
Clarkson et al.	2021	Sport, Business and Management: An International Journal	Journal Article	x	x	1st National League (Adult)
Culvin & Bowes	2021	Frontiers in Sports and Active Living	Journal Article	x	x	1st National League (Adult)
Culvin	2021	Managing Sport and Leisure	Journal Article	x		
Cuthbert et al.	2021	Journal of Sports Sciences	Journal Article		x	1st National League (Adult)
Gozillon & Neys	2021	International Journal of Sport Policy and Politics	Journal Article	x		
Harkness-Armstrong et al.	2021	Science and Medicine in Football	Journal Article		x	Academy (Youth)
Jenkel	2021	Sport in History	Journal Article	x		
Lucarno et al.	2021				x	1st National League (Adult)
Luteberget et al.	2021	Frontiers in Sports and Active Living	Journal Article		x	1st National League (Adult)
McGreary et al.	2021	Psychology of Sport and Exercise	Journal Article	x		
Moss et al.	2021	European Journal of Sport Science	Journal Article		x	1st National League & Train >10 h/Week (Adult)
Okholm Kryger et al.	2021	Science and Medicine in Football	Journal Article	x		

Scelles	2021	International Journal of Sport Policy and Politics	Journal Article	x		
Thomas et al.	2021	International Journal of Sports Physiology and Performance	Journal Article		x	1st National League (Adult)
Beech et al.	2022	Science and Medicine in Football	Journal Article		x	National Training Camp (Youth)
Clarksonet al.	2022	Managing Sport and Leisure	Journal Article	x		
Datson et al.	2022				x	International Level (Youth & Adult)
Emmonds et al.	2022	Science and Medicine in Football	Journal Article		x	1st & 2nd National League (Adult), Tier 1 Academy (Youth)
Fleming & Parker	2022	Journal of Contemporary Religion	Journal Article	x	x	International level (Adult), International Youth Level (Youth)
Harkness-Armstrong et al.	2022				x	Regional Talent Centre
Mayhew et al.	2022	Journal of Elite Sport Performance	Journal Article		x	1st & 2nd National League (Adult)
McHaffie et al.	2022	Science and Medicine in Football	Journal Article		x	International Level, 1st & 2nd National League (Adult), International Youth Level (Youth)

McHaffie et al.	2022	Science and Medicine in Football	Journal Article		x	International Level, 1st & 2nd National League (Adult), International Youth Level (Youth)
Myhill et al.	2022	Science and Medicine in Football	Journal Article		x	1st & 2nd National Leagues (Adult)
Nassis et al.	2022	Scandinavian Journal of Medicine & Science in Sports	Journal Article	X		
Parker et al.	2022	Science and Medicine in Football	Journal Article		x	1st National League (Adult)
Read et al.	2022	Science and Medicine in Football	Journal Article		x	1st National League (Adult)
Simpson et al.	2022	International Journal of Sport Psychology	Journal Article	x		
Skillen	2022	Sport in History	Journal Article	x		
Cooper	2023	Accounting, Auditing and Accountability Journal	Journal Article	x		
Datson et al.	2023	Science and Medicine in Football	Journal Article		x	National Training Camp (Youth)
Emmonds et al.	2023				x	Top 2 National Leagues (Adult) or Highest National League Training Academy (Youth)
Fenton et al.	2023				x	Top 3 National Leagues (Adult)

Forsyth et al.	2023	Sport in Society	Journal Article	x		
Grice et al.	2023	Managing Sport and Leisure	Journal Article	x		
Harkness-Armstrong et al.	2023				x	Tier 1 and Tier 2 Academies (Youth)
McHaffie et al.	2023				x	International Level (Youth)
Mondal	2023	Journal of Global Sport Management	Journal Article	x		
Wheatley et al.	2023				x	1st National League (Adult)
Sprouse et al.	2024				x	International Level (Youth & Adult)