A FIVE YEAR COMPARATIVE ANALYSIS OF ANNUAL BASELINE NEUROCOGNITIVE TEST SCORES FOR SOUTH AFRICAN HIGH SCHOOL ATHLETES

A thesis submitted in fulfilment of the
requirements for the degree of

MASTERS OF ARTS
IN PSYCHOLOGY

of

RHODES UNIVERSITY

by

MARCELLE ANTOINETTE REICHLING

ORCID ID http.//orcid.org/ 0000-0002-4392-4647

SUPERVISOR: PROF ANN EDWARDS
CO-SUPERVISOR: DR GARY STEELE

OCTOBER 2019

The financial assistance from the Rhodes University Prestigious Scholarship towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to Rhodes University.

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily attributed to the NRF.

Contribution from National Research Foundation (NRF) incentive funding awarded to my supervisor, Professor Ann Edwards, during 2014.

ABSTRACT

The primary objective of this study was to assess the pattern of change in neurocognitive performance for adolescent athletes on baseline measures of the Immediate Post Concussion Assessment and Cognitive Testing (ImPACT) test, over five consecutive years, with a view to providing an indication of the optimal interval for repeat baseline testing of high school athletes. Participants were non-clinical, predominantly South African high school athletes in the overall age range 13 to 18 years (N = 108), divided into five groups (Grades 8,9, 10, 11 and 12), and tested at five test intervals. Repeated-measures ANOVA analyses examined differences in score performance across the test intervals for each of the five composite scores of the ImPACT test (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, Impulse Control). For the Verbal Memory, Visual Memory, Visual Motor Speed and Reaction Time composites there were significant neurocognitive score changes between several test intervals. Taking these results into account, in conjunction with substantial variability in performance, it is concluded that there is a need for *annual* baseline testing throughout the high school years. The secondary objective was to generate normative tables (Means and Standard Deviations) on the ImPACT test for the five participant groups at each of the five test intervals, including data for: the five composite scores (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, Impulse Control); for the twelve subtest scores test that go to make up the composite scores; and for four additional memory subcomponent scores (Word Memory immediate recall, Word Memory delayed recall, Design Memory immediate recall, Design Memory delayed recall). The results provide a clinical and heuristic normative platform for future use with brain injured individuals, which can be used to facilitate clinical interpretations of postconcussion assessments.

ACKNOWLEDGEMENTS

Firstly, I wish to thank the National Research Foundation (NRF), Rhodes University, and the National Student Financial Aid Scheme (NSFAS) for the financial assistance during the course of this study. Thanks to my primary supervisor, Prof Ann Edwards, for your direction, patience, and perseverance throughout this long journey. I am deeply grateful for your dedication, knowledge and expertise. I would also like to thank my co-supervisor, Dr Gary Steele, for your welcomed guidance and advice, especially through the most difficult parts of this study. My thanks also to Prof Sarah Radloff, for your kind and efficient statistical input and expertise. Finally, to my family for your continuous and unwavering support and encouragement.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

The primary objective of this study (Objective 1) was to assess the pattern of change in neurocognitive performance for adolescent athletes on baseline measures of the ImPACT test, over five consecutive years with test-taking intervals approximately one year apart. Participants were non-clinical predominantly South African high school athletes from Grades 8 to 12, in the overall age range 13 to 18 years, divided into five groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30). An enhanced understanding of the year-by-year neurocognitive score changes experienced by high school athletes across the 13 to 18 year adolescent age range would allow for the development and implementation of more effective recommendations related to sports-related concussion management. In particular, an indication of the optimal interval for the repeat baseline testing of high school athletes.

The secondary objective of this study (Objective 2) was to generate normative tables for the ImPACT tests' five composite scores for the five participant groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30). This was done with the view to establishing normative platforms on which to base interpretations of concussion follow-up testing depending on the grade and associated age stage of the athlete, and the number of baseline tests previously completed. The five normative participant groups are as follows, stratified by Grade as the primary variable, and age mean (SD), age range, and number of baseline tests previously completed as qualifiers: Grade 8 [mean age 13.69 (0.29) years, age range 12-13 years, 0 prior baseline tests]; Grade 9 [mean age 14.62 (0.32) years, age range 13-14 years, 1 prior baseline test]; Grade 10 [mean age 15.58 (0.31) years, age range 14-15 years, 2 prior baseline tests]; Grade 11 [mean age 16.62 (0.31) years, age range 16-17 years, 3 prior baseline tests]; Grade 12 [mean age 17.62 (0.30) years, age range 17-18 years, 4 prior baseline tests]. On this basis, it was considered that normative data would be more focused and accurate, thereby providing a relevant basis for clinicians to refer to when assessing athletes that pertain to a particular grade and age stage, who have completed a specific number of prior baseline tests, thereby aiding clinicians in making more accurate comparisons and clinical diagnoses.

A subsidiary aspect of Objective 2 was to generate normative tables for the twelve subtest scores that make up the five ImPACT composite scores, as well as four additional subcomponent scores that make up the ImPACT Word Memory and Design Memory subtests (i.e., Word Memory immediate recall; Word Memory delayed recall; Design Memory immediate recall; Design Memory delayed recall). The provision of normative data on the twelve subtest scores and the four additional memory subcomponent scores would enrich the clinical interpretation of ImPACT test performance and provide additional interpretative information to aid clinical judgement. Providing normative data on the immediate and delayed recall subcomponents of the two memory subtests would be especially relevant comparative indexes for clinicians managing concussive injuries because delayed recall is a function that is particularly sensitive to brain dysfunction (Lezak, Howieson, Bugler, & Tranel, 2012).

In summary, descriptive data for the five participant groups delineated above will be tabled at each test interval for each of these five composite scores, twelve subtest scores, and four additional memory subcomponent scores. The results will provide a clinical and heuristic normative platform for future use with brain injured individuals, which could be used to facilitate clinical interpretations of postconcussion assessment.

1.1 Overview of Concussion in Sport

Between 1.7 and 3 million sports-and-recreation-related concussions (SRCs) occur each year in the United States (UPMC Sports Medicine, 2019). Whilst a common injury in both adult and youth sport, the prevalence of youth concussion is considered to be at ‘epidemic’ levels (Udall, 2011). From 2001 to 2005, about three in 1000 youths in the United States, aged 14-19, had an emergency department visit for a concussion sustained during an organised team sport (Bakhos, Lockhart, Myers, & Linakis, 2010). Concussion in organised sport is highest amongst high school athletes and is the second highest leading cause of concussion after motor vehicle accidents amongst 15 to 24-year olds (Marar, McIllvain, Fields, & Comstock, 2012). It also accounts for 8.9% of all high school athletic injuries in the United States (Gessel, Fields, Collins, Dick & Comstock, 2007). A review of sports-related concussions (SRCs) in youth reported that high school American football players experience concussions at a rate of 11.2 concussions per 10,000 athletic exposure, which is a rate nearly twice that of college football players (Institute of Medicine and National Research Council, 2013). Possible explanation for adolescent vulnerability to concussion will be discussed in section 1.6.

The incidence of concussion varies greatly depending on the sport, with high contact sports naturally resulting in higher incidence rates. In fact, concussion is believed to be the leading cause of injury in all contact sports (Moser, 2007). Within the sports context, a concussion typically, but not exclusively, is categorised as a mild traumatic brain injury (mTBI) (Alexander, Shuttleworth-Edwards, Kidd, & Malcolm, 2015). In the United States, American football followed by lacrosse and soccer represent the highest prevalence of concussion (Gessel et al., 2007), with around 300 000 concussions occur annually in American Football (UPMC Sports Medicine, 2019). In high school contact sport, two out of 10 athletes suffer at least one concussion annually. However, as it is estimated that five of every 10 concussions go unreported or undetected (UPMC Sports Medicine, 2019), the concussion prevelance rate is likely higher than recorded. According to McCrea, Hammeke, Olsen, Leo, and Guskiewicz (2004), when the frequency of unreported concussions is taken into consideration, it is likely that "close to 15% of high school [American] football players sustain a concussion each season" (p. 15).

Concussion is a common injury in Rugby Union, a high-risk injury sport widely played at school, collegiate, club and professional level in over 100 countries worldwide, including South Africa (Macleod, 1993). The sport involves tackles, scums, mauls and rucks, which results in frequent body-to-body and body-to-ground collissions "and an associated high incidence of multiple concussions as a result of repeated head-jarring trauma" (Alexander et al., 2015, p. 1113). Rugby Union has been shown to be even more susceptible to concussive injuries than American football (see review by Shuttleworth-Edwards, Border, Reid, & Radloff, 2004a), and is classified as one of the most dangerous sports in the world due to the frequent collisions involved in the sport (Hillis, McIntyre, Maclean, Goodwin, & McKenna, 1994; Wekesa, Asembo & Njororai, 1996). In Rugby Union, concussion is one of the three most common rugby injuries, with most incidences occurring during the tackling phase of play (Kemp, Hudson, Brooks, & Fuller, 2008). Additionally, in a survey of three South African schools' top teams, Shuttleworth-Edwards et al. (2004a) recorded an average incidence rate of 2.3 concussions per rugby playing schoolboy compared to 0.4 concussions for an equivalent field-hockey group. In the 1999 Super 12 professional rugby tournament, concussion was the most common injury, with an incidence rate of 20% (Holzhausen, Schwellnus, Jakoet, & Pretorius, 2002). At adult club level, a seasonal incidence rate of 30% was reported in a New Zealand sample (Wills & Leathern, 2001). At high school level, the incidence of concussion in Rugby Union over one season was documented at 18% (Roux, Goedeke, Visser, van Zyl, & Noakes, 1987), and 21.5% (Nathan, Goedeke, & Noakes, 1983). The prevalence of concussion could however be as high as 50% in South African schoolboy rugby players, as the majority of mild brain injuries are often not recognised and reported in this age group (Patricios, Kohler, & Collins, 2010).

It is current consensus that many concussive symptoms go unreported and/or unrecognised due to one or a combination of the following: an athlete's fear of being side-lined and possibly losing their position in the team; an absence of concussion management guidelines; a failure to recognise symptoms due to either a lack of knowledge or as a result of their often rapid onset and spontaneous recovery (McCrea et al., 2004); the pressure placed on talented players to downplay symptoms to ensure team victory and to secure financially lucrative sporting contracts (McCrea et al., 2004; Moser, 2007); and most disturbingly, the misconception held by some athletes, their families as well as coaches that concussion is simply having your 'bell rung' and is not a serious injury. This misconception is heavily ingrained in the culture of contact sport, where athletes are expected to be tough, and where "there is a myth that real men do not suffer concussion" (Nell & Brown, 1991, p. 34). Not reporting and/or recognising concussive injuries and their symptoms is a serious challenge to concussion management and naturally affects the prevalence records. In a study of over 1500 high school football players, it was found that less than half of those who sustained a concussion reported it to a trained professional (McCrea et al., 2004).

1.2 Defining Concussion

Despite decades of research and numerous articles, there is still no universally agreed upon definition for concussion, due mainly to the evolving nature of the injury and its rapidly changing clinical signs and symptoms (McCrory et al, 2013). The first working definition of concussion presented over fifty years ago is that of "a clinical syndrome characterized by the immediate and transient post-traumatic impairment of neural function such as alteration of consciousness, disturbance of vision or equilibrium etc. due to brain stem dysfunction" (Congress of Neurosurgeons, 1968, p. 319). Definitions have evolved throughout the past four decades, as new knowledge of the injury has come to light. Despite this, it became clear that the existing definitions were inadequate to comprehensively define and understand such a complex injury as sports concussions. This, in addition to confounding epidemiological studies, evaluation and management guidelines disputes and a lack of objective data guiding return-to-play decisions (Leclerc, Lassonde, Delaney, Lacroix, & Johnston, 2001) resulted in a need for a more collated approach to brain injury management in sport. In response, the First International Symposium on Concussion was held in Vienna in 2001 (Aubry et al., 2002). Since this date there has been a 2nd (McCrory et al., 2005), 3rd (McCrory et al., 2009), 4th (McCrory et al., 2013), and 5th (McCrory et al., 2017) International Consensus Conference on Concussion in Sport, with each conference making adjustments in keeping with developing scientific research.

According to the most recent ‘consensus’ definition of sports-related concussion (SRC) based on current knowledge, the injury can be defined as “a traumatic brain injury induced by biomechanical forces” (McCrory et al., 2017, p. 2). Common features that incorporate clinical, pathological and biomechanical injury constructs may be utilized in defining the nature of concussion, and are as follows;

1. “SRC may be caused either by a direct blow to the head, face, neck or elsewhere on the body with an impulsive force transmitted to the head.
2. SRC typically results in the rapid onset of short-lived impairment of neurologic function that resolves spontaneously. However, in some cases symptoms and signs may evolve over a number of minutes to hours.
3. SRC may result in neuropathological changes, but the acute clinical symptoms largely reflect a functional disturbance rather than a structural injury and, as such, no abnormality is seen on standard structural neuroimaging studies.
4. SRC results in a range of clinical symptoms that may or may not involve loss of consciousness. Resolution of the clinical and cognitive symptoms typically follows a sequential course. However, it is important to note that in some cases symptoms may be prolonged.” (McCrory et al., 2017, p. 2).

Whilst this definition has made use of the most current scientific research, the authors note that the management and understanding of concussion in sport is an ongoing process which requires more research (McCrory et al., 2017).

1.3 Biomechanics and Pathophysiology of Sport-related Concussion

Concussion is a closed head injury caused by forces inflicted on the brain by either a direct impact to the head or an impulse (a force to another part of the body that sets the head in motion without directly striking it) (Mihalik, 2012). For example, a direct impact in the sports context would be a head-to-head collision, whereas an impulse would be a tackling collision. The abrupt stopping of the body during a tackle could causes a whiplash effect through the neck and head, which results in the movement of the brain within the skull (Cantu & Hyman, 2012). The forces inflicted by a direct blow or an impulse result in the brain rapidly accelerating (speeding up) or decelerating (slowing down) within the skull, which result in it moving in a linear and/or rotational manner. This linear and/or rotational movement causes significant stress to brain tissue, and may results in axonal tearing, possible cell death and intracranial bleeding (Barth, Freeman, Broshek, & Varney, 2001).

An injury such as this, caused by biomechanical forces, is referred to as a diffuse brain injury as it typically results in widespread disruption of neurological function rather than visible or structural damage (Guskiewicz et al., 2004). This widespread disruption may result in numerous physical signs and symptoms as well as neurobehavioral and neurocognitive deficits (Gardner, Shores, & Batchelor, 2010), with no two injuries ever presenting in the same manner (Broglio & Puetz, 2008). Research based on rodent models indicate that concussion symptoms are produced by metabolic, rather than an anatomic or structural change in the brain. In other words, following concussion there is a "temporary disruption of energy utilization in the brain" (ImPACT Applications, 2007, p. 2). The complex interwoven cellular and vascular changes that occur after a concussion have been described by Giza and Hovda (2001) as a multi-layered neurometabolic cascade, characterized by a 'mismatch' between energy supply and demand.

It is worth noting that the vast majority of concussions that occur in athletes are at the mild end of the mild traumatic brain injury severity continuum, whereby loss of consciousness is typically not present and post-traumatic amnesia is usually brief. Loss of consciousness (LOC) after a concussion is likely the result of damage to the reticular activating system in the brain. This system recovers fairly quickly, thus consciousness is regained shortly after injury (McCrory, Collie, Anderson & Davis, 2004). Unfortunately, the biochemical mismatch lasts significantly longer than damage to the reticular activating system, which makes a loss of consciousness a poor indicator of injury severity (Patricios et al., 2010).

As a result of the non-structural metabolic nature of the injury, brain images on traditional neurodiagnostic techniques (e.g. CT scan, MRI) appear typically normal following a concussion, and are thus of little use in diagnosis of concussion. However, it must be noted that such traditional neurodiagnostic techniques are vital tools when investigating serious effects of head trauma such as cerebral bleeding and skull fracture (Aubry et al., 2002).

1.4 Potential Complications of Concussion

In the last decade there has been an increase in research on the deleterious effects of mild brain injuries in contact sports. Many studies have found that there are not only short-term physiological and neuropsychological (cognitive, emotional, and behavioural) changes that arise post-injury (Belanger & Vanderploeg, 2005), but that some of these issues do not always resolve, resulting in long-term difficulties and chronic sequelae, which for some may be “permanent and disabling” (Lovell & Collins, 1998, p. 6). There is evidence to suggest that repeated brain injury, particularly during the recovery period, may result in more severe, long-lasting, and even life-threatening brain injury (Kelly, 2001). The early complications and late complications of concussion are discussed below.

1.4.1 Early Complications of Concussion

Intracranial space occupying lesions:

Although unusual, concussion may cause damage to cerebral arteries and veins. Bleeding from these vessels may cause epidural, subdural or intracerebral haematomas (Patricios et al., 2010). Any indications of increasing intracranial pressure must be immediately recognised and treated surgically in order to decompress the brain. There may be “considerable overlap between the initial clinical presentation of a concussed athlete and that of a player who has an intracranial bleed” (Patricios et al., 2010, p. 5). This highlights the need for continuous monitoring of the patient in the first 48-72 hours post-injury.

Impact convulsions:

Convulsions (seizures) in contact sports are uncommon but can appear as a dramatic event. They usually occur within two seconds of impact and are not necessarily associated with structural brain damage. Fortunately, they are benign in nature and do not cause long-term cognitive damage nor do they require antiepileptic treatment and prolonged exclusion from play (McCrory, Bladin & Berkovic, 1997).

Second Impact Syndrome:

Second Impact Syndrome (SIS), which occurs mostly in male adolescents (McCrory, 2001), was first reported in American Football players who died after relatively minor brain injury. Albeit extraordinarily rare, a second even mild impact to an already vulnerable brain can result in immediate, catastrophic diffuse cerebral swelling of the brain. Returning an athlete to play whilst their concussion is unresolved or keeping them in play whilst it is undiagnosed places athletes at risk of SIS (McCrory, 2001). The previous impact may result in a cerebral edema and increased vulnerability due to a biomechanical 'mismatch' still being present. A second impact may result in further swelling "followed by loss of the brain's ability to control blood inflow (autoregulation)" (Patricios et al., 2010, p. 5). Cerebral blood flow then increases rapidly, and brain pressure rises uncontrollably, leading to cardiorespitory failure and possible death within only a few minutes after sustaining the second impact. There is minimal time for emergency intervention (Fisher & Vaca, 2004), which explains the near 50% mortality rate, and the long-term brain damage in survivors (Proctor & Cantu, 2000). There have also been documented cases where no second impact was observed, but the athlete still died (McCrory, 2001). As such, there is a debate as to whether a second impact is necessary. It has been suggested that the second blow may be minor and often not observed (i.e. a blow to the chest that indirectly affects the brain) (Cantu, 1998).

Concussion Signs and Symptoms:

There is significant individual variability in the presentation of concussion signs and symptoms, depending on the biomechanical forces involved, the seriousness and location of the injury, the neurological and psychological predisposition of the injured individual, as well as whether in the acute or chronic phase post-injury (ImPACT Applications, 2007; Lovell & Pardini, 2010; Shuttleworth-Edwards & Whitefield, 2007). In some instances, a concussed individual may experience only one concussion symptom, whilst another may report multiple (Fazio, Lovell, Pardini, & Collins, 2007). Concussion symptoms may be apparent immediately following injury whilst in other cases symptoms may only occur in the hours or days later, due to the evolving nature of the injury in the acute phase (Ellemberg, Henry, Macciocchi, Guskiewicz, & Broglio, 2009). This results in a large number of concussions only being identified until 24 hours or more after the injury (McCrory et al., 2009). Such symptom uncertainty and individual variability highlights the importance of rigorously and subjectively assessing a suspected concussion for all common symptoms in the post-concussive period (McCrory et al., 2009).

Initial on-field symptoms include confusion, disorientation, and retrograde (memory loss of events prior to hit or fall) and anterograde (memory loss of events after hit or fall) amnesia (Fazio et al., 2007). Loss of consciousness (LOC) is not a necessary condition of concussion and occurs in only 10% of sports concussions (Bailes, 2009). Furthermore, there is reason to believe that a loss of consciousness may not be an indicator of severe injury, as previously thought. A study by Lovell, Iverson, Collins, McKeag, and Maroon (1999) investigating the relationship between loss of consciousness and post-injury neurocognitive test performance found no difference in test scores between concussed athletes with loss of consciousness and concussed athletes without. Rather the presence of amnesia (retrograde and anterograde) is considered a more accurate and clinically useful measure of the severity of concussion in sport (Cantu, 2001). Headaches are incredibly prevalent and are the most common symptom of concussion. Although seen in over 70% of athletes who sustain a concussion, concussions may occur without headache. They may also develop immediately after injury or minutes to hours later, and should be managed conservatively (ImPACT Applications, 2007).

The acute signs and symptoms of concussions reported within one to seven days post injury usually fall into four categories; physical (somatic), cognitive, emotional (affective) and sleep-related, with individuals experiencing one or more symptoms from one or more categories (see Table 1.1) (Patricios et al., 2010). Cognitive indicators of concussion on objective tests include dysfunctions in memory, learning, processing speed (Erlanger, Kutner, Barth, & Barnes, 1999) as well as deficits in attention, concentration, planning and cognitive flexibility (Erlanger, et al., 2001).

Table 1.1

Concussive Symptoms by Category

Physical	Cognitive	Emotional	Sleep
Headache	Difficulty thinking clearly	Irritability	Sleeping more than usual
Fuzzy or blurred vision	Feeling slowed down	Sadness	Sleeping less than usual
Dizziness	Difficulty concentrating	Feeling more emotional	Trouble falling asleep
Fatigue	Difficulty remembering new information	Nervousness or anxiety	
Drowsiness			
Sensitivity to light	Feeling mentally “foggy”		
Sensitivity to noise	Confusion		
Balance problems			
Nausea or vomiting			

Note. Adapted from Patricios et al. (2010, p. 24).

As concussions are heterogeneous injuries (i.e. there is no specific symptom or cluster of symptoms that are inherent to all sports-related concussions), it is difficult to set specific recovery parameters (McCrory et al., 2004; Patel, Shivdasani, & Baker, 2005). In most cases, a concussion does not appear to produce any permanent injury but is characterized as a "transient, or temporary condition with a brief period of impairment followed by spontaneous recovery" (Schatz & Moser, 2011, p. 4). The majority of concussions (80-90%) resolve without any apparent negative consequences in a short (7-10 day) period, although the time frame may be extended for children and adolescents (see section 1.6.3) (McCrory et al., 2013). In group studies (involving aggregated data) the recovery period was within 2-14 days based on perceived symptoms and neurocognitive test performance (Bleiberg et al., 2004; Lovell, Collins, Iverson, Johnston, & Bradley, 2004; Macciocchi, Barth, Alves, Rimel, & Jane, 1996; McCrea et al., 2003). It must be noted however that "a minority of athletes who remain symptomatic at more than one or two weeks post-injury can be obscured in group analyses" (Moser et al., 2007, p. 911).

A literature review on the impact of concussion on a non-athlete population by Belanger and Vanderploeg (2005) found that whilst the majority of concussive injuries resolved within three months, a small percentage showed cognitive deficits that lingered longer than three months. If symptoms persist beyond three months they are subsequently viewed as relatively chronic and intractable (Reitan & Wolfson, 1999).

1.4.2 Late Complications of Concussion

Post-concussion Syndrome:

According to Jotwani and Harmon (2010), post-concussive syndrome is present if an individual experiences persistent post-concussive symptomology that does not resolve by three months post-injury. The persistent signs and symptoms can last for months or even years post-injury (Patel et al., 2005) and may be exacerbated by poor acute concussion management. Whilst there are differing criteria for diagnosing post-concussion syndrome, according to the DSM-IV diagnostic criteria, at least three of the following symptoms must be present for at least three months post-injury: physical/somatic symptoms of headache, fatigue, disordered sleep, vertigo or dizziness; emotional symptoms of irritability or aggression with little or no provocation, anxiety, depression or affective lability, changes in personality, and apathy or lack of spontaneity (American Psychiatric Association, 2000). This diagnosis unfortunately does not recognise cognitive symptoms. A study by Fazio et al. (2007) found that concussed athletes who denied experiencing symptoms within one week of a concussion were not fully recovered based on neurocognitive testing but experienced some lingering cognitive impairment. This suggests that physical/somatic symptoms resolve within a short period of time, whilst cognitive defects may continue to linger long after injury, often going unnoticed and/or unreported in a small percentage of athletes (Belanger & Vanderploeg, 2005; Fazio et al., 2007). Residual cognitive symptoms include memory and attention deficits (Moser et al., 2007) as well as slowed thinking/reaction time and difficulty problem solving (Duff, 2009).

According to research, persistent post-concussion syndrome can also result in individuals experiencing performance-related problems at work or school. Difficulties include "mental slowness in completing assignments, problems understanding verbal communication in group discussions, greater distractibility and poor concentration, ineffective multitasking, and severe fatigue" (Moser, 2007, p. 700). Experiencing these difficulties long after injury (of which the individual is assumed to be fully recovered) can be misinterpreted by others as laziness and/or attitudinal/behavioural problems. Individuals with no history of academic or behavioural difficulties have shown a gradual decline in test grades, increased incomplete homework assignments, low frustration tolerance, and weak self-confidence as a result of an undiagnosed post-concussion syndrome (Moser, 2007).

Chronic Traumatic Encephalopathy:

Chronic traumatic encephalopathy (CTE) is a "progressive neurogenerative syndrome caused by single, episodic, or repetitive blunt force impacts to the head and transfer of acceleration/ deceleration forces to the brain" (Omalu et al., 2011, p. 173). CTE is characterised by an accumulation of Tau protein, "a substance which serves to stabilize cellular structure in the neurons, but which may become defective and subsequently may cause major interference with the function of the neurons" (Brain Injury Research Institute, 2019, para. 2). It has been found that the primary driver of CTE pathology is the cumulative effect of long-term exposure (over a period of years or decades) to repetitive head impacts (i.e. sub-concussive blows) and not the number of concussions (Huber, Stein, Alosco, & McKee, 2016). A recent study analysing teenage brains that had suffered brain injuries alongside the brains of mice that had suffered brain trauma, found that the process that culminates in CTE does not require the presence of any concussion symptoms. In other words, sub-concussive blows to the brain, which do not cause the sufferer to be knocked unconscious, experience dizziness, or lingering headaches, are enough to cause CTE (Bissell, 2018). These findings explain why approximately 20 percent of athletes with CTE never suffered a diagnosed concussion (Bissell, 2018), and indicate that numerous sub-concussive events caused by player-to-player and player-to-ground collisions and experienced over the course of athletic careers can cause serious injury (Schatz & Moser, 2011).

The symptoms of CTE can be debilitating and disabling, and include loss of memory, difficulty controlling impulsive or erratic behaviour, personality and mood changes, impaired judgement, behavioural disturbances (i.e. aggression and depression), increased suicidality, speech and gait abnormalities, difficulty with balance and a gradual onset of dementia. These symptoms may mistakenly be attributed to the normal process of aging or receive a wrong diagnosis as many symptoms are similar to conditions such as Alzheimer's or Parkinson's disease (Brain Injury Research Institute, 2019). Although the most recent International Symposium of Concussion in Sport (McCrory et al., 2017) notes that "a cause-and-effect relationship has not yet been demonstrated between CTE and SRCs or exposure to contact sport" (p. 7), a recent study of the brains of 202 deceased American football players from a brain donating program produced results to challenge this. Mez et al. (2017) found that 177 players were neuropathologically diagnosed with CTE across all levels of play (87%), including 110 of 111 former National Football League players (99%), 7 of 8 Canadian League players (88%), 9 of 14 seniprofessional players (64%), 48 of 53 college players (91%), 3 of 14 high school players (21%), and 0 of 2 pre-high school players. Whilst usually observed in professional athletes (McKee et al., 2009; McKee et al., 2013; Schatz & Moser, 2011), CTE has been identified in high school, collegiate and semi-professional athletes, with incidence rates and severity of the disease increasing with increased years-of-play (Mez et al., 2017).

1.5 Effects of Multiple Concussions and Sub-Concussive Injuries

Concussion can be considered a cumulative disorder in that the effects of each concussion are additive (Guskiewicz et al., 2003; Iverson, Gaetz, Lovell, & Collins, 2004). Statistically speaking, once concussed an athlete increases his/her risk of sustaining a future concussion. According to Guskiewicz et al. (2003), previously concussed athletes are four to six times more likely to experience a second concussion, even if the second impact is relatively mild. In a two-year prospective study by Zemper (2003) there was a six times greater relative risk of concussion amongst high school and collegiate football players with a history of concussion compared to individuals with no such history. It is not known the extent to which the increased risk is due to a) changes in the brain after an initial concussion that may make it more vulnerable to future impacts as opposed to b) factors that may make an individual more susceptible to concussion in general (e.g. aggressive playing style, position, and musculature) (Institute of Medicine and National Research Council, 2013).

Athletes who have had previous concussions are also more likely to have longer recovery periods than those with no history of concussion (Guskiewicz et al., 2003). In a mixed high school and collegiate sample, a history of concussion was associated with prolonged recovery for subsequent concussions (Iverson et al., 2004). A history of concussion also makes an athlete more vulnerable to sustaining injuries of greater severity in the future. Collins and colleagues (2002) compared on-field concussion presentation between athletes with no history of concussion and athletes who had sustained three or more previously concussions. Results revealed that athletes with a history of concussion were 9.3 times more likely to demonstrate three to four markers of concussive severity (including LOC, anterograde and retrograde amnesia, and confusion) than the non-concussive history group.

Sustaining multiple concussions can put athletes at risk of developing subtle chronic cognitive deficits (Moser et al., 2007), displaying decreased performance on baseline neurocognitive testing (Moser, Schatz, & Jordan, 2005), and reporting more ongoing post-concussion symptoms (Iverson, et al., 2004; Schatz et al., 2011). Research suggests that the effects of a concussive injury are largely undetectable on neurocognitive measures for athletes with a history of only *one or two* prior injuries (Macciocchi, Barth, Littlefield, & Cantu, 2001). However, high school athletes with a history of *three or more* concussions have been found to performed worse on a memory task and demonstrated more concussive symptoms pre-season in comparison to a group of athletes with no history of concussion (Iverson et al., 2004).

In addition to diagnosed concussive injury, it had been suggested that sub-concussive injuries also have cumulative effects. A sub-concussive injury is a microtrauma to the brain that "involves similar smaller impact forces to the concussive injury, but functions below the threshold necessary to produce symptoms" (Alexander et al., 2015, p. 1113). A sub-concussive injury is thought to occur in rugby/football contact sports where repeated head jarring incidents are common. Examples of an impact that could cause a sub-concussive injury include most tackles and collision in rugby and American football, headers in soccer, checks or collisions with the boards or other players in ice hockey, and body checks in lacrosse (Concussion Legacy Foundation, 2018). Additive sub-concussive injuries would continually lower the cerebral reserve of athletes, which would make them increasingly vulnerable to chronic cognitive deficits (Shuttleworth-Edwards & Radloff, 2008; Alexander et al., 2015), especially in the areas of memory and attention (Concussion Legacy Foundation, 2018). In the South African rugby context, it has been suggested that, in addition to the cumulative effects of diagnosed concussive injury, players who are continually exposed to high levels of contact may suffer from sub-concussive injuries, and resulting neurocognitive vulnerability (Alexander et al., 2015; Shuttleworth-Edwards & Radloff, 2008).

1.6 Sports Concussion in the Youth

Young athletes are more vulnerable and susceptible to concussions than older athletes, during both the critical pre-adolescent period and adolescent development periods (Giza & Hovda, 2001). They are also more likely than adults to develop Second Impact Syndrome (SIS) as well as experience long-term negative cumulative consequences of concussion. Ongoing research also suggests that young athletes experience a “more prolonged disturbance of brain function following a concussion that previously believed” (Buzzini & Guskiewicz, 2006, p. 376), and have slower and differential recovery patterns compared to collegiate and professional athletes (Collins, Lovell, Iverson, Ide & Maroon, 2006). Whilst almost all past research has focused on collegiate and professional athletes (Lovell et al., 2003), over the last decade research has begun to focus on how concussions specifically affect children/adolescents, since the injury manifests differently in these age groups compared to adults (Gioia, Schneider, Vaughan, & Isquith, 2009; McCrory et al., 2004).

1.6.1 Physiological Differences between Youths and Adults

Most concussions in organised sports occur in high school adolescent athletes (Powell & Barber-Foss, 1999) in part due to the sheer number of athletes competing at this level (Barr, 2003) and to increased youth vulnerability to concussion compared to adults. Exact reasons for this vulnerability are unknown, however there are “normal changes in brain structure, blood flow, and metabolism that occur with brain development that may influence susceptibility to and prognosis following concussion in youth” (Institute of Medicine and National Research Council, 2013, p. 70).

Possible explanations include the “immaturity of the developing central nervous system, a larger head-to-body ratio, thinner cranial bones, reduced development of neck and shoulder musculature, a larger subarachnoid space in which the brain can move, and differences in cerebral blood volume" (Institute of Medicine and National Research Council, 2013, p. 69). A greater head-to-body ratio (i.e. disproportionately larger heads for their overall size) and weaker necks and shoulder muscles could inhibit a young athletes' ability to sustain a blow to the head (Patoine, 2010). A blow of the same degree will exert more acceleration force to a youth than to an adult due to the relative weakness of the neck and shoulder musculature, which essentially acts as a shock absorber by dissipating the forces applied to the head (Patoine, 2010). Furthermore, gains in weight and mass that occur during adolescent growth spurts may increase force and momentum during collisions without similarly increasing neck strength, which may place the athlete at risk of a concussive injury (Buzzini & Guskiewicz, 2006). Additionally, incomplete myelination of the brain tissue may put the developing brain at risk for shear injury (Cook, Schweer, Shebesta, Hartjes, & Falcone, 2006; Kieslich, Fielder, Heller, Kreuz, & Jacobi, 2002; Ommaya, Goldsmith & Thibault, 2002). Furthermore, as child and adolescent brains are still developing, it is much more vulnerable to the deleterious effects of a concussive injury, which would disrupt this developmental process (Henry & Sandel, 2015).

1.6.2 Neuropsychological Differences between Youths and Adults

According to the Luria's theory of cognitive development (Luria, 1980), cognitive development of primary, secondary, and tertiary brain regions become functional by 12 years of age but continue to develop throughout adolescence and into early adulthood, "most likely reflecting a refinement of the neural networks that support the more specialised aspects of cognition that characterized adulthood" (Cromer, Schembri, Harel, & Maruff, 2015, p. 1). Longitudinal neuroimaging studies also demonstrate that the adolescent brain continues to mature well into the 20's (Johnson, Blum, & Giedd, 2009). Language (receptive and expressive) and sensory/motor areas of functioning are the first areas to mature, with language being developed by 10 years of age. Formal operational thought such as abstract processes and reasoning (i.e., executive functioning) is the last to develop, beginning during adolescence and continuing throughout late adolescence into adulthood (Toledo et al., 2012). Adolescence begins at approximately 12-13 years of age, with entrance into Piaget's Formal Operations stage of cognitive development (i.e., the stage of abstract higher-level thought), and with a corresponding peak in frontal and parietal brain growth that begins with the onset of puberty (Casey, Jones, & Hare, 2008). Adolescence is a critical period for maturation of neurobiological processes that underlie higher cognitive functions and social and emotional behaviour (Yurgelun-Todd, 2007).

According to recent neuroimaging studies, changes in brain structure and function are thought to correlate with improvements in cognitive abilities, behaviours and emotional processes. The brain undergoes substantial structural changes during adolescence, in terms of “grey matter volume, surface area and cortical thickness, as well as white matter volume and microstructure” (Foulkes & Blakemore, 2018, p. 3). White-matter connections between frontal-subcortical networks are important for the development of complex attention and executive functioning (Giedd et al., 1999; Konrad, Firk, & Uhlhaas, 2013).

Neuronal structural changes continue throughout adolescence (Giedd et al., 1999; Konrad et al., 2013), and even into young adulthood (Sowell, Thompson, Holmes, Jernigan, & Toga, 1999). According to longitudinal studies, adolescence is an important period of neuronal organisation, where experience-dependent changes result in new synapses being established, some of which are strengthened over time, while others are left unused and pruned away (Konrad, et al. 2013). For normal, healthy, high-school adolescents significant cognitive and neurodevelopmental changes occur in the frontal-striatal brain regions (i.e. the prefrontal cortex) that are critical for the development components of executive functioning (cognitive processes that regulate, control, and manage other cognitive processes), and social and emotional behaviour (i.e. improvements in affective modulation and discrimination of emotional cues) (Boelema et al., 2014). Executive functioning is strongly associated with the prefrontal cortex, which is one of the slowest-developing brain areas (Benes, 2001, Yurgelun-Todd, 2007) and continues to mature into adulthood, reaching cognitive maturity as late as 25 years of age (Sukel, 2017).

An individual’s age and associated brain maturation at the time of a mild traumatic brain injury may influence injury prognosis. This is because “specific neuronal systems (frontal and parental) that are generally developed at specific ages, may be variably adaptive or resilient post-concussion, and may be modulated by gender and pubertal changes” (Toledo et al., 2012, p. 3). It has been suggested that over-learned or well-developed skills are less vulnerable to disruption following brain injury, whereas skills which are in the process of being developed are most likely to result in deficits (McKinlay, 2010). Considering that higher cognitive functions, social cognitive functions and emotional processing are still undergoing cognitive maturation and development in adolescents and young adults, even a mild brain injury during this period may affect one or more of these abilities, which could cause impairments and create problems in day-to-day functioning.

Furthermore, the areas of the brain that process higher cognitive functions (i.e. forward regions of the prefrontal cortex and temporal lobes) are positioned at the front and sides of the brain. These regions are particularly susceptible to damage from a concussive injury, partly due to where they reside in the brain (Patoine, 2010).

1.6.3 Recovery Time Differences between Youths and Adults

Whilst children, adolescents and adults have shown reduced cognitive performance following a concussion, different cognitive recovery rates have been observed (Field, Collins, Lovell & Maroon., 2003). Research showed that high school athletes (aged 14-18) experienced prolonged and differential cognitive recovery following concussion (i.e. working memory, processing speed, reaction time), in comparison to collegiate (aged 18-25), and professional athletes (Collins et al., 2006; Field et al., 2003). According to McCrory et al. (2004), the clear majority of college football players in a large prospective study (*N* = 1631) recover within five to seven days of injury, whereas most high school athletes in a study by McClincy, Lovell, Pardini, Collins, and Spore (2006) needed 14 days post-concussion to return to their cognitive baseline performance. Collins et al. (2006) reported that the majority of high school football players seemed to recover within one-month post-injury.

Another study by Field et al. (2003) compared the recovery rates of both high school and college athletes to healthy controls. Results showed that collegiate athletes had significant memory impairment only within the first 24 hours after injury, whereas high school athletes had memory impairment up to seven days post-concussion. Even though the college athletes in this sample sustained more severe in-season concussions compared to the high school athletes, it was found that the cognitive performance of college athletes was the same as that of healthy controls at three days post-injury. High school athletes however continued to “perform significantly worse than age-matched controls at seven days post-concussion” (Field et al., 2003, p. 1). These results indicate that a younger age is associated with slower recovery time. High school athletes also take longer to recover from a concussive injury than professional athletes.

Possible reasons for the prolonged recovery experienced by adolescent and younger athletes in comparison to adult athletes is unclear, however the following hypotheses have been proposed. Firstly, there is evidence of concussion causing more diffuse and prolonged cerebral swelling and increased metabolic sensitivities in younger individuals compared to adults (Lang, Teasdale, Macphereson & Lawrence, 1994). These physiological changes may result in more apparent (i.e. severe and persistent) symptoms, which would prolong the recovery period (Institute of Medicine and National Research Council, 2013). According to Bauer and Fritz (2004) possible reasons for this include differences in the expression of glutamate receptors, the expression of aquaporin-4 by microglia, and in brain-water content. Secondly, common sense dictates that younger athletes who are particularly susceptible to concussion and are slow to recover may not advance to higher levels of play. Thus, "the more rapid recovery time in college and professional athletes could, in part, reflect a selection bias" (Iverson, 2011, p. 726).

1.7 Assessment and Management of Sports Concussion

Sports-related concussions were initially assessed using grading scales, which characterised the severity level of the injury. The level of severity was then used to inform return-to-play guidelines (Lovell, 2009). However, as research on the individual differences that exist in concussion incidence, severity, and recovery emerged, as discussed above, assessment moved away from group-based grading scales in the direction of an individualised, multi-method form of assessment and return-to-play protocol (McCrory et al., 2017; Meehan & Bachur, 2009). This approach to concussion assessment and management involves the use of evaluation tools such as self-report symptom scales, postural stability assessment, and neurocognitive testing (Institute of Medicine and National Research Council, 2013).

1.7.1 Neurocognitive Assessment

The importance of neurocognitive testing in sport is highlighted by the International Symposiums on Concussion in Sport (Aubry et al., 2002; McCrory et al., 2005; McCrory et al., 2009; McCrory et al., 2013; McCrory et al., 2017). The most recent Berlin panel (McCrory et al., 2017) states that neurocogntive testing "has clinical value and contributes significant information to SRC evaluation" (p. 4). Reasons for such endorsement of neurocognitive testing rests on its ability to detect the subtle effects of concussion when self-reported symptoms are absent (Van Kampen, Lovell, Pardini, Collins & Fu, 2006).

Much of the decision-making process with regards to concussion diagnosis and management is based on the self-reporting of symptoms. This is due to the lack of biomarkers for diagnosis and recovery of concussion (Institute of Medicine and National Research Council, 2013), as well as the fact that standard neuroimaging techniques, such as magnetic resonance imaging and computer topography scans are not sensitive to concussions (Henry & Sandel, 2015).

Relying solely on symptom self-reporting has been heavily criticised in the literature. The subjective nature of symptom reporting coupled with the possibility of athletes underreporting their symptoms, makes self-reporting potentially unreliable and likely to result in underdiagnoses of concussion. This may in turn "result in premature return of athletes to contact sport, potentially exposing them to additional injury" (Van Kampen et al., 2006, p. 1630). This was demonstrated by Van Kampen et al. (2006), who found that when assessing concussed athletes using only self-reporting of symptoms, 64% of the concussed sample were diagnosed as suffering from concussion, in comparison to 83% who were diagnosed by means of neurocognitive testing. The addition of neurocognitive testing resulted in a sensitivity discrepancy of 19%, indicating that such testing is able to detect the subtle effects of concussion and 'add value' to the concussion diagnostic and return-to-play decision-making process (Van Kampen et al., 2006).

This research, along with others, indicates that even athletes who report being symptom-free may continue to exhibit evidence of cognitive impairment on neurocognitive measures, that they are either unaware of or fail to report (Fazio et al., 2007; Van Kampen et al., 2006). Whilst it has been discussed that some athletes will deliberately fail to report symptoms for various reasons, many of the cognitive changes following a concussion (e.g. mental fogginess, difficulty concentrating) are at times imperceptible to the athletes themselves. Additionally, cognitive recovery may occasionally precede clinical symptom recovery, or in some instances an asymptomatic athlete can experience either a delayed onset of symptoms or a delayed resolution of neurocognitive deficits (McCrory et al., 2017). Fortunately, objective measures of brain function can be used to detect neurocognitive impairment from a concussion and monitor recovery (McCrory et al., 2017). Neurocognitive testing allows for a baseline and post-injury analysis of the subtle aspects of cognitive function likely affected by concussive injury, thus providing objective data to make more informed decisions regarding return to play (Lovell, 2009). As such, the testing of cognitive functions "should be an important component in the overall assessment of concussion and in particular, any return-to-play protocol" (McCrory et al., 2013, p. 91).

It must be noted that neurocognitive testing is only one component of concussion assessment and should not be the sole basis of concussion management decisions. Rather, it should be seen as an aid to the clinical decision-making process of deciding when an athlete is ready to return to play, and should be used "in conjunction with a range of assessments of different clinical domains and investigational results" (McCrory et al., 2013, p. 91), such as clinical examination, symptom reporting, medical/academic/psychological history, concussion history, patient report, observer report, postural stability assessment, athletic exertion testing, as well as type of sport and position played (Meehan & Bachur, 2009; Moser, Schatz & Lichtenstein, 2015).

Using neurocognitive assessment to detect neurocognitive impairment from concussion and track cognitive recovery was first prominently studied in the mid-1980's at the University of Virginia (Barth et al., 1989). In this landmark study, neurocognitive tests were administered to assess changes in cognitive functioning from baseline (i.e. preseason) to subsequent post-injury after a concussion. A number of injuries to several high-profile professional National Football League (NFL) athletes resulted in the implementation of baseline/post-injury neurocognitive assessment by several NFL teams in the mid-1990s (Lovell, 1999). Further, several career-ending injuries in the National Hockey League (NHL) resulted in the NHL mandating baseline neurocognitive testing for all athletes (Lovell & Burke, 2000). The use of neurocognitive testing then gradually spread to other sports, such as automobile racing (Olvey, 2002), soccer (Matser, Kessels, Lezak, Jordan, & Troost, 1999), Australian Rules football (Makdissi et al., 2001) and rugby (Shuttleworth-Edwards & Radloff, 2008; Shuttleworth-Edwards et al., 2008a: Shuttleworth-Edwards et al., 2008b). Today, neurocognitive testing as part of the concussion management programmes is used frequently at professional, collegiate and high school level (Moser et al., 2007).

1.7.2 Traditional versus Computer-Based Neurocognitive Testing

Whilst 'pencil-and-paper' neurocognitive tests were initially used to determine cognitive function after a concussion, the late nineties saw the introduction of *computerized* neurocognitive testing (Lovell, 2006). Traditional 'pencil-and-paper' neurocognitive tests were "not ideal for sporting settings as they were designed to detect gross cognitive impairment at a single assessment, not for mild cognitive deficits on repeated assessments" (Collie, Darby, & Maruff, 2001, p. 297). As such, 'pencil-and-paper' tests were adapted to focus on specific areas believed to be affected by concussion (i.e., attention and concentration, cognitive processing (speed and efficiency), learning and memory, working memory, executive functioning and verbal fluency) (Ellemberg et al., 2009; Guskiewicz et al., 2004), and not on extensive cognitive functioning. Despite these efforts to adapt traditional 'pencil-and-paper' tests, these tests had other limitations which inhibited their suitability for the sports-related concussion context. These included the need for a trained neuropsychologist to administer the tests to each athlete individually, manually score each test and interpret the results, resulting in tests that were resource intensive and time-consuming (Lovell & Solomon, 2011). Also, some of the tests such as Digit Span were found to be insensitive to the subtle effects of concussion (Ellemberg et al., 2009). Finally, as traditional 'pencil-and-paper' neurocognitive tests were not designed for repeated serial assessment they were sensitive to the effects of practice, which was problematic for clinicians evaluating concussion in athletes where serial testing is necessary (Lovell, 2009).

In order to address the limitations of traditional 'pencil-and-paper' neurocognitive tests, computerised neurocognitive tests were specifically developed for the assessment of sports-related concussion (Patel et al., 2005). Computerized assessment allows for a large number of athletes to be baseline tested simultaneously as a group, with ease of administration and a minimal labour/time demand required (Lovell, Iverson, Johnston, & Bradley, 2004). Whilst the interpretation of test scores must always be carried out by a registered neuropsychologist, many computerized tests do not require a registered neuropsychologist to administer as a properly trained technician will suffice (Shuttleworth-Edwards, Whitefield-Alexander, & Radloff, 2013). These factors make computerised testing time and cost-effective, which has played an integral role in neurocognitive testings' newfound prominence in sports-related concussion, and its accessibility "to high schools, organizations, and clubs that do not have a large budget to pay for individual assessments" (Lovell, 2006, p. 195). Computerised neurocognitive tests for concussion management is widely used in US high schools, suggesting an eagerness to enhance safely and prevention efforts (Meehan, d' Hemecourt, Collins, Taylor, Comstock, 2011).

Additional advantages of computerized assessment include the following; (i) a normative comparison standard that presents information in a standard and consistent manner, thus increasing the reliability of neurocognitive tests and allowing for more accurate assessments (ii) increased sensitivity to the subtle neurocognitive effects of concussion compared to traditional forms of neurocognitive testing due to more accurate and reliable measurements of cognitive processes such as reaction time and processing speed. Computer technology allows for millisecond accuracy whereas the traditional tests allow for accuracy of only one to two seconds (Lovell, 2006; Lovell, 2009). Such accuracy in these areas is important as there are reliable differences between concussed and non-concussed athletes that range around only 100 milliseconds (see Bleiberg, Halpern, Reeves, & Daniels, 1998); (iii) randomisation of test stimuli and the generation of alternate forms, thus minimising practice effects in repeated administrations and increasing the reliability of the test (Lovell, 2006; Lovell, 2009). This is a factor contributing to the higher test-retest reliability generally held by computerised neurocognitive tests in comparison to traditional pencil-and-paper tests; (iv) the storage of data on a computer network that is easily assessable at a later date, allowing for efficient clinical evaluation as well as research to be more easily conducted (Lovell, 2006; Lovell, 2009); (v) the fast and reliable scoring of many athletes at one season as well as the immediate generation of clinical reports, which are then used by the neuropsychologist to make clinical decision. This allows for timely feedback of test findings to athletes, coaches and parents (Lovell, 2006; Lovell, 2006). The limitations of computerised neurocognitive assessment are few but include being less flexible and interactive than one-on-one assessment as well as being unable to measure verbal functioning or auditory memory (Schatz & Browndyke, 2002; Schatz & Zillmer, 2003).

Due to their availability, convenience, and psychometric superiority compared to their ‘pencil-and-paper’ counterparts, computerised neurocognitive tests have become a widely used concussion management tool, especially for assessing cognitive functioning in athletes at risk for concussion (Meehan et al., 2011). In a study assessing strategies used by emergency medicine physicians when treating paediatric patients with concussion in the US, more than half (56%) of the 256 physicians involved referred patients with concussion from the Emergency Department for neurocognitive testing (Kinnaman, Mannix, Comstock, & Meehan, 2014).

1.7.3 Baseline/Post-Injury Comparison and Normative Data Comparison

When using neurocognitive testing to identify and manage cognitive impairment after concussion, two options are available; the baseline comparison, also referred to as the baseline/postinjury model (comparison of postconcussion scores to individualised baseline score), and/or the normative comparison (comparison of postconcussion scores to a normative mean) (Schmidt, Register-Mihalik, Mihalik, Kerr, & Guskiewicz, 2012).

The current model of neurocognitive testing for many amateur, collegiate, and professional sports teams is the baseline/post-injury model (McClincy et al., 2006). Baseline testing in and of itself is of no use to concussion management. Instead, “the value of baseline testing lies in its availability for *comparative* purposes with subsequent follow-up testing” (Shuttleworth-Edwards, 2011, p. 391). Pre-season baseline testing allows clinicians to obtain an individualised neurocognitive profile of the athlete, which is then used as a standard against which to compare any of the athlete’s future post-injury scores obtained during the season, to assess for clinically significant cognitive change (Moser et al., 2007). Such serial assessments allow clinicians to closely monitor recovery and track the effectiveness of treatment outcomes. The baseline/postinjury model was designed to "reduce the inherent variability among athletes' cognitive skills" (Henry & Sandel, 2015, p. 267). Within this model, athletes serve as their own controls, thus making managing concussion more individualised and specific to each concussed athlete's own neurocognitive variability, and reducing variation when testing for post-concussion cognitive changes.

Baseline testing is widely regarded as best practice for concussion management (Merritt et al., 2016). Not only can it account for pre-injury differences in individual cognitive profiles (i.e. individual differences in memory, attention, visual motor processing speed), but it can also account for specific and individualistic factors influencing baseline performance such as the presence of attention deficit hyperactivity (ADHD) or learning disorders (LD), history of concussion, cultural/linguistic differences, and psychiatric issues (Barr, 2003; Elbin et al., 2013; Guskiewicz et al., 2003; Moser et al., 2007). Thus, baseline testing should provide a more sensitive approach for evaluating the degree of cognitive change after a concussion. Knowing how an athlete performed prior to a concussive injury allows clinicians to make informed decision regarding the presence or absence of subtle aspects of cognitive impairment following concussion (Lovell, 2009), and to analyse and manage any changes in neurocognitive functioning accordingly. However, despite the benefits of the baseline testing, the model has not escaped criticism.

The primary criticism of the use of baseline testing is that the test-retest reliability of the measure may be too low for clinical utility in cases where there has been a long test-retest interval between baseline and post-injury testing (for a review see Randolph, 2011). However, even in the presence of established test-retest reliabilities over longer time periods, the use of baseline testing may not always be feasible or logistically possible. Significant time, money, resources and personnel are required to administer baseline assessments to all athletes, especially when administering paper-and-pencil measures, and this can be a significant hinderance to the accessibility of baseline assessments (Arnett et al., 2014; Echemendia et al., 2013). Even among programs that routinely utilise baseline testing, the data may not always be available or appropriate to use for various reasons (i.e., athletes not referred for baseline testing, invalid test performance, test administration errors).

In cases where baseline assessments are unavailable or inappropriate for various reasons, post-injury performance can be compared to an established normative database (Henry & Sandel, 2015). The use of normative data are considered a cornerstone of neurocognitive assessment (Merritt et al., 2016). Provided it represents the population for which the measure was designed to test, normative data can be general (the average performance for a diverse sample) or more specific (the average performance of groups divided by age, sex, education, race, sport played etc.) (Strauss, Sherman, & Spreen, 2006). In post-concussion assessment, specific normative data that most closely match an athlete's demographics provides the best comparison method. To illustrate, when evaluating a sports concussion, "a critical question to be addressed by the neuropsychologist is whether it is safe for the athlete to return to play. Given the individualistic nature of this question, applying specific normative data should improve the accuracy of the test results and interpretation of the data" (Merritt et al., 2016, p. 2). With regard to neurocognitive testing for concussion among youths, large samples stratified by grade/age are of particular importance as cognitive maturation occurs dramatically during adolescence (Casey, Giedd, & Thomas, 2000; Yurgelun-Todd, 2007). Gender-based data are also important as men and women perform differently on neurocognitive measures post-injury (Colvin et al., 2009; Covassin, Elbin, Harris, Parker, & Kontos, 2012; Covassin, Elbin, Kontos & Larson, 2010). As such, grade/age-and-gender stratified normative data are essential for determining how and athlete is performing compared to same-grade/age and gendered peers (Henry & Sandel, 2015). Comparing an athlete's post-injury performance with the normative data of same-grade/age and gendered peers allows clincians to identify clinically meaningful cognitive impairment, monitor recovery, or track the effectiveness of treatment outcomes.

According to Hinton-Bayne (2015), a within-subject comparison (a baseline paradigm) will always be more sensitive than a between-subjects comparison (a normative paradigm) when test scores are correlated, and type error rates are comparable. These findings confirm previous research by Iverson, Brooks, Collins, and Lovell (2006), which found that comparing players with their own baseline is more accurate than comparing them with age-and sport-matched normative data. However, it is important to emphasis that even research reporting the superiority of the baseline testing over normative data in identifying impairment after concussion recommends the use of the normative data when baseline testing is unavailable or inappropriate (Hinton-Bayre, 2015; Louey et al., 2014). This supports the utility of normative data in identifying clinically meaningful cognitive impairment after concussion (provided it is appropriate and well-developed).

Whilst the baseline/postinjury model is the preferred model for computerised evaluation within concussion management, it has been suggested that *both* pre-season baseline data and appropriate normative data be used for optimal accuracy in clinical decision making (Shuttleworth-Edwards, Whitefield-Alexander, Radloff, Taylor, & Lovell, 2009). As mentioned above, baseline data should be used as it provides an optimal mode of comparison. It is highly individualised and thereby protects against erroneous interpretations that may arise when there are idiosyncratic variations from the norm in a particular case (Shuttleworth-Edwards et al., 2009). However, normative data should also be used to alert clinicians to significant deviations from the norm at baseline. When an individual's pre-season baseline score is an inaccurate reflection of pre-injury functioning due to factors know to affect neurocognitive performance such as illness, fatigue and low motivation at testing, a post-injury test comparison may produce false positive results (show no injury when one is present). Suspect pre-season baseline scores can be compared to normative data in order to determine whether an individual is deviating significantly from the norm at baseline. In such instances baseline scores will no longer serve as a reliable comparative index, and more emphasis needs to be given to comparing post-injury test results against a demographically appropriate norm (Shuttleworth-Edwards et al., 2009).

1.8 Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT)

There are several computerised neurocognitive tests for use in sports concussion management currently available, such as the Automated Neuropsychological Assessment Metrics (ANAM), CogState Sport, Headminders, CNS Vital Signs, and the ImPACT test. A review of the ANAM, CogState Sport, Headminders and CNS Vital Signs tests are unfortunately beyond the scope of this study. The ImPACT test evaluates multiple aspects of cognitive functioning and uses a baseline/post-injury model, which allows for individualised concussion management (ImPACT Applications, 2004).

The ImPACT test is one of the most commonly used and statistically validated computerised neurocognitive test batteries for concussion management, with a large database of validating clinical research, consisiting of hundreds of peer-reviewed and independent studies on concussion management (ImPACT online, 2019b). The ImPACT test is used internationally at professional, collegiate, high school and primary school levels. Amongst adolescents, the test is perhaps the most utilised concussion management battery (Henry & Sandel, 2015). Professional organisations such as the Major League Baseball (MLB), National Hockey League (NHL), National Football League (NFL), World Wrestling Entertainment (WWE), NASCAR, Formula 1 Racing, UK-based Football Association, as well as over 200 other professional teams and thousands of colleges, universities, high schools, clinics and select military units use the test (ImPACT online, 2019a). It is also the first concussion-specific medical device to receive United States Food and Drug Administration (FDA) *de novo* clearance under Computerised Cognitive Assessment Aid for Concussion, which is a "new category for a device that uses an individual's score(s) on a battery of cognitive tasks to provide an indication of the current level of cognitive function in response to concussion" (ImPACT online, 2019a). The ImPACT test has also received clearance as a concussion-specific device by Health Canada, the Australian Therapeutic Goods Administration (TGA) and other regulatory agencies (ImPACT online, 2019b). The test has been translated into 13 different languages, including Afrikaans, which has increased its applicability in the South African context. In South Africa, it is the only test of its kind that is registered for clinical and research use with the Health Professional Council of South Africa (Health Professions Council of South Africa, 2006). In this country the test is used frequently at professional and school levels, in both the clinical and research setting, mainly to assess the numerous concussions that occur in Rugby Union (Shuttleworth-Edwards et al., 2004a).

ImPACT was created in the United States in the 1990s by Mark Lovell and Joseph Maroon (ImPACT online, 2019a). It was the first computerised test battery developed to assess sports concussion and address the limitations of the traditional 'pencil-and-paper' tests used to assess sports concussion (Lovell, 2006). Since the development of ImPACT 2.0 the test has retained the same basic structure, and refinements have largely been on improving technical aspects of the test, although some alterations have been made to the test items in order to improve the overall reliability and validity of the test (Shuttleworth-Edwards et al., 2013). The test was originally designed for use from ages 12-59, however a paediatric version for individuals aged 5-11 has been FDA-cleared for use (ImPACT online, 2019c).

The ImPACT test has a user-friendly, freestanding Windows-based application that runs on both individual computers or in the Windows network environment and takes about 20-25 minutes to administer. It can be administered by ImPACT trained athletic trainers, clinicians, school nurses or other health care professionals, and does not require administration from a registered neuropsychologist, thus allowing it to be used nationally on a large scale (Shuttleworth-Edwards et al., 2014). The ImPACT battery consists of three parts: (i) demographics, background and concussion history questionnaire (ii) symptom inventory and (iii) neurocognitive test battery. The use of both a neurocognitive test battery and symptom reporting is considered the most effective and accurate method of evaluating concussion (Van Kampen et al., 2006).

The neurocognitive test battery consists of six test modules (Word Memory, Design Memory, X's and O's, Symbol Match, Colour Match, and Three Letter Memory), which together produce five composite scores (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control). The modules represent commonly employed neurocognitive tests. For example, the Colour Match task is similar to the popular Stroop Colour and Word test (Stroop, 1935), and the Symbol Match task is similar to the Digit Symbol Coding task in the Wechsler Adult Intelligence Scale, Fouth Edition (WAIS-IV) (Wechsler, 2008).

A description of the six test modules is as follows:

Word Memory: Utilises a word discrimination paradigm. 12 targeted words are presented for a duration of 750 milliseconds each. Presentation occurs twice to facilitate learning. Participants are then tested for recall via presentation of a 24-word list by choosing 'yes' or 'no' if previously displayed. Following administration of all other test modules, delayed recall is tested by again presenting the 24-word list in a similar manner. This recall is tested without alerting participants at initial presentation that there will be a later recall. This module evaluates attention and verbal recognition memory (ImPACT Applications, 2011).

Design Memory: Utilises a design discrimination paradigm. 12 targeted designs are presented for a duration of 750 milliseconds each. Presentation occurs twice to facilitate learning. Participants are then tested for recall via presentation of 24 designs by choosing 'yes' or 'no' if previously displayed. Following administration of all other test modules, delayed recall is tested by again presenting the 24 designs in a similar manner. This recall is tested without alerting participants at initial presentation that there will be a later recall. This module evaluates attention and visual recognition learning (ImPACT Applications, 2011).

X's and O's: A random assortment of X's and O's are presented for 1.5 seconds, three of which are highlighted yellow. Participants must remember the location of the yellow X's and O's and, following a timed distractor task designed to interfere with memory rehearsal, recall the location of the yellow X's and O's. The timed distractor is a choice reaction time test. Participants are required to press 'Q' if blue square appears and 'P' is red circle appears as quickly as possible.

This module evaluates visual working memory and visual motor (processing) speed (ImPACT Applications, 2011).

Symbol Match: Common symbols (e.g. circle, square, arrow) are presented with a corresponding number (1 to 9). Participants must click on the corresponding number associated with each symbol as quickly as possible and remember these pairings. After 27 trials, the symbols disappear from the grid and the participants must recall the correct pairing by clicking the appropriate number button. This module evaluates visual processing speed and visual learning and memory (ImPACT Applications, 2011).

Colour Match: A word (RED, GREEN, or BLUE) is displayed in either the same colour ink as the word or in a different coloured ink. Participants must click on the word as quickly as possible, but only if the word is presented in matching ink. This module evaluates choice reaction time, impulse control, and response inhibition (ImPACT Applications, 2011).

Three Letter Memory: Consists of 5 separate trials. Three consonant letters are presented. Following 18 seconds of a distractor task, participants must recall the three letters initially displayed. The timed distractor task is a grid consisting of numbers 1-25 randomly placed. Participants must sequentially click as many buttons as possible in backwards order starting from 25. This module evaluates working memory and visual-motor response speed (ImPACT Applications, 2011).

From the six test modules, twelve subtest scores are automatically produced, which contribute to the five composite scores (ImPACT Applications, 2011). Each composite score is made up of two or three subtest scores totalling twelve subtest scores, as can be seen in Table 1.2, which details the composite scores and the formulas used to derive each from their respective subtests. From the table it can also be seen that two of the memory subtest scores (i.e.,Word Memory: Total percent correct and Design Memory: Total percent correct) are in turn made up of separate scores for immediate and delayed recall. Therefore, for norming purposes, it would be useful to have data for these four subcomponents of the memory subtests, in addition to the ‘total percent correct’ scores. The ImPACT neurocognitive composite scores (Verbal Memory, Visual Memory, Visual Motor Speed and Reaction Time) are made up of the average of the contributing subtest scores, whereas the Impulse Control composite score is made up of the sum of the contributing subtest scores (ImPACT Applications, 2004).

Table 1.2

Computation of ImPACT Composite Scores from Twelve Subtest Scores

Composite Scores	Contributing Scores
Verbal Memory	Word Memory: Total percent correct (immediate + delayed/2) Symbol Match: Total correct (hidden)/9 x 100 Three Letters: Percent of total letters correct
Visual Memory	Design Memory: Total percent correct (immediate + delayed/2) X's and O's: Total correct (memory)/12 x 100
Visual Motor Speed	X's and O's: Total correct (interference)/4 Three Letters: Average counted correctly x 3
Reaction Time	X's and O's: Average correct reaction time (interference) Symbol Match: Average correct reaction time (visible)/3 Colour Match: Average correct reaction time
Impulse Control	X's and O's: Total incorrect (interference) Colour Match: Total commissions

Note. Adapted from Henry & Sandel (2015, p. 268).

Individuals who have suffered a concussion may present with different neurocognitive deficits depending on the biomechanics of the injuries, their age, and a variety of other factors (Van Kampen et al., 2006). As no single score can be used to assess the severity of a concussion injury, the ImPACT test has multiple composite scores that evaluate multiple cognitive domains. Four of the ImPACT composite scores measure the broad neurocognitive domains their names suggest: Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, whilst the fifth composite score, Impulse Control, is a validity indicator (ImPACT Applications, 2011). The Verbal Memory composite evaluates attentional processes, learning and memory, and the Visual Memory composite evaluates attention, scanning, learning and memory. The Visual Motor Speed composite (measure of processing speed) evaluates processing, visual-motor response speed, learning and memory (ImPACT Applications, 2011; Wurz, 2016). A higher score on these composites indicates better performance. The Reaction Time composite evaluates average response speed, with lower scores indicating better performance. Rather than an assessment of neurocognitive functioning, the Impulse Control composite serves as an indicator of test validity by calculating the number of errors made in completing the test. Scores above 30 suggest either right-left confusion in completing the test or an unusually high number of errors on several relatively simple tasks. This would indicate that either the individual is seriously confused about the test instructions, or is not putting forth maximum effort. Lower scores indicate better performance and test results with a value greater than 30 should be discarded (ImPACT Applications, 2011). Whilst clinical decisions should not be based on this composite, its inclusion may help in the interpretation of other composites (ImPACT Applications, 2011).

The ImPACT test has alternate forms or stimuli designed to minimise practice effects often found in more traditional neurocognitive tests and improve reliability across serial assessments (ImPACT Applications, 2004). Practice effects are artificial improvements in performance on repeat administrations of a test due to prior exposure and not due to recovery (Duff, 2012). There are five sets of stimuli or alternate forms for the Word and Design Memory test modules. According to Lovell (2006) the high correlations between the five forms indicating they are near equivalent. The stimuli for all other modules are randomly generated by the computer to minimise practice effects. For example, for the Symbol Match test module, the symbol-number pairing change for every administration. Similarly, the number grid for the speed component of the Three Letter Memory test module changes with each administration. All participants are administered the same stimuli of the ImPACT test at baseline (i.e., Form 1) (Lovell, 2006). As such, research comparing baseline-to-baseline scores does not minimise practice effects as the participants are administered the same test at all test intervals, whereas research comparing baseline-to-post-injury scores does minimise practice effects. Practice effects will be discussed in more detail in section 1.9.2.

The ImPACT test has reliable change methodology imbedded into the scoring programme. Post-injury scores that exceed the Reliable Change Index when compared to the individual’s baseline are represented in BOLD RED on the ImPACT combined report (Iverson & Schatz, 2015). Reliable change scores are provided to identify when the difference between a baseline score and a post-injury score fall outside of normal score variation (ImPACT Applications, 2011), and are a means of assessing individual change between baseline and post-concussive functioning in each composite score. If baseline testing is not available, the ImPACT manual recommends that post-injury scores be compared to age and gender specific norms, which will allow for the data to be compared to the individual’s peer group (ImPACT Applications, 2011; Lovell, 2006).

1.8.1 Clinical Use of ImPACT Composite Scores, Subtest Scores, and Additional Memory Subcomponent Scores

The ImPACT composite scores are quantitative summary scores of general cognitive domains (ImPACT Applications, 2011) which are helpful for measuring post-injury cognitive impairment for assessing neurocognitive functioning and for charting progress (Henry & Sandel, 2015). However, in cases that have a slow or complicated recovery course (i.e., where comorbid disorders are present) additional information interpreted from the subtests that make up the ImPACT composite scores may be required (Henry & Sandel, 2015). In these cases, subtest normative data augment the utility of the ImPACT by offering a "deeper level of interpretation that may be missed if one is simply using the ImPACT composite scores" (Henry & Sandel, 2015, p. 272). For example, the Colour Match subtest is comparable to a standard Stroop test, "which is known to be sensitive to attentional deficits (Cohen, Malloy, Jenkins, & Paul, 2006; Ikeda, Okuzumi, & Kokubun, 2013). In ImPACT, the Colour Match subtest is used to formulate the Reaction Time composite score. Therefore, having normative data on the subtests may aid in interpreting the test results of a concussed athlete diagnosed with ADHD" (Henry & Sandel, 2015, p. 272). Normative data on the twelve subtests and four additional memory subcomponents can also provide additional interpretative information to aid clinical judgement, which may be helpful when interpreting post-injury results (Henry & Sandel, 2015; Reynolds, Fazio, Sandel, Schatz, & Henry, 2016). In cases where a composite score indicates a deficit, closer inspection of the subtests norms that make up the relevant composite score (and the additional memory subcomponent norms), could provide significant insight into a particular athlete's performance and "inform return-to-play decisions, especially in such a case where the baseline in either unavailable or otherwise invalid" (Henry & Sandel, 2015, p. 273). It is important to note that subtest norms serve only to supplement the use of the ImPACT test. In cases where the clinical picture is unclear due to comorbid disorders, additional neurocogntive assessment may be necessary. Computerised neurocognitive testing is useful when used in conjunction with, and not as a substitute for, a full clinical neurocognitive assessment (Henry & Sandel, 2015; McCrory et al., 2017).

Normative data on the immediate and delayed recall subcomponents of the Word Memory and Design Memory subtests could have particular clinical utility. These subcomponents make up the 'Word Memory: Total percent correct' and 'Design Memory: Total percent correct' subtests respectively. 'Word Memory: Total percent correct' is one of three subtests that make up the

Verbal Memory composite, and 'Design Memory: Total percent correct' is one of two subtests that make up the Visual Memory composite. Delayed recall on memory tests have been found to be more sensitive to neurocognitive impairment than other concussion-affected cognitive domains (Belanger, Curtiss, Demery, Lebowitz, & Vanderploeg, 2005) including immediate recall (Belanger et al., 2005; Lezak, Howieson, Bugler, & Tranel, 2012). Research indicates that individuals with cognitive impairments caused by brain disorders (Coolidge, Middleton, Griego, & Schmidt, 1996; Belanger et al., 2005; Wurz, 2016) and MTBI (Sim, Terryberry-Spohr, & Wilson, 2008; Shuttleworth-Edwards et al., 2004a) performed significantly worse than healthy controls on delayed recall tasks on memory tests. Delayed recall is an important added component to increase the sensitivity of memory tests to brain dysfunction generally, and accordingly, tests of delayed recall are regularly being included in revisions of memory tasks due to its sensitivity to brain dysfunction that may not appear when testing immediate recall (Lezak et al., 2012).

Therefore, normative data that focuses on delayed recall and immediate recall in isolation might be very useful to clinicians monitoring the recovery of injured athletes. If the delayed recall score is below the norm, but the immediate recall score is 'normal', the athlete may be cognitively impaired from a concussive injury. Normative information on these memory subcomponent scores will allow the clinician to compare the immediate and delayed memory recall results of an injured athlete with what would be expected in a non-clinical athlete of similar demographics, to determine if an athlete is cognitively impaired. Additionally, an athlete may appear 'normal' on the Visual or Verbal Memory composite scores but may be impaired on the delayed recall subcomponent. As the composite scores incorporate both delayed and immediate recall functions, 'normal' immediate recall scores may off-set a deficit on delayed recall, resulting in Visual or Verbal Memory composite scores that appear 'normal' but in fact obscure valuable clinical information. It does not help to have normative data for the Word Memory and Design Memory subtests that go to make up the Verbal Memory and Visual Memory composites, because, as indicated above, these too have been made up of a total percent recall without differentiating between immediate and delayed recall. Accordingly, normative data that provides a further breakdown into the immediate and delayed recall subcomponents of the ImPACT Word Memory and Design Memory subtests (i.e., allowing them to be identified as separate entities) would provide important additional information to be used for clinical purposes.

1.8.2 Normative Data on the ImPACT Test

Current US normative data provided by ImPACT Applications (2011) (presented as percentile scores rather than standardised measures), has banded the ages of 10- to 12- years into one group combining both males and females, whilst the remaining ages are banded into the following age groups and separated by gender: Males, Ages 13-15; 16-18; 19-29; 30-39; 40-49; 50-59; Females, Ages 13-15; 16-18; 19-29; 30-39; 40-49; 50-59 (ImPACT Applications, 2011). The Verbal Memory, Visual Memory, Visual Motor Speed, and Reaction Time composites are normed by the ImPACT developers, whilst the validity composite (Impulse Control) is currently not normed by ImPACT. With regard to the subtest scores, the test developers provide some normative data in the form of percentile cut-offs for the 10th percentile and 2nd percentile to allow clinicians to identify unusually low and impaired subtest scores (see ImPACT Applications, 2004). No other percentile cut-offs are provided, nor are any subtest normative data provided that are based on standardised measures.

Whilst interpretation of the ImPACT subtests can potentially produce useful clinical information, limited normative research on subtests existed until recently, with the publication of two studies by Henry and Sandel (2015) and Reynolds et al. (2016) respectively. A US study by Henry and Sandel (2015) provided normative data on ImPACT composites and the subtests used to generate the composites for 4,500 male and female adolescent and young-adult athletes with the aim of augmenting the ability of health care providers to interpret the ImPACT test. The normative data was separated by gender for athletes aged 13 to 21 and stratified by the norm age stages already employed by ImPACT (13-15; 16-18; 19-21 years-old). The Reynolds et al. (2016) study mentioned above also provided normative data on each of the subtests used to calculate the composite scores for 10-, 11- and 12-year-olds respectively, in addition to normative data on the composites themselves. As there were consistent age differences but a lack of gender differences on the ImPACT composite scores, normative data were split across age groups but combined on gender. Both the Henry and Sandel (2015) and Reynolds et al. (2016) studies presented the subtest normative information as standardized measures (*z* scores) rather than percentiles scores as provided by ImPACT Applications (2004).

1.8.3 Local Normative Data on the ImPACT Test

Whilst the English version of the test for young adult male and female US athletes has been normed, even when the test is administered in English, “the cultural equivalence of performance on neuropsychological tests cannot be assumed” (Shuttleworth-Edwards et al., 2009, p. 46). Local normative data must be derived to ensure diagnostic accuracy (Lezak et al., 2012; Shuttleworth-Edwards et al, 2004b). Some local normative studies of South African adolescent athletes on the ImPACT test have taken place (Shuttleworth-Edwards et al., 2009; Salman-Godlo, 2006; Horsman, 2010).

A study by Horsman (2010) provided normative data on a cohort of Afrikaans first language adolescent rugby players with Model C education. When comparing this cohort to existing data for South African English first language adolescent rugby players with Private/Model C schooling, it was found that the Afrikaans first language sample generally demonstrated poorer performance, which was attributed to poorer quality of education.

A study by Salmon-Godlo (2006) provided normative data for semi-rural Xhosa-speaking schoolboys with an advantaged education (ages 14, 16, 18) on the English version of the ImPACT test. It was found that there were no significant differences between the age groups on the Xhosa-speaking sample. When compared with existing US norms, the South African sample consistently scored lower than the US sample, thus prompting the author to recommend the use of local norms on South African semi-rural Xhosa-speaking educationally advantaged athletes, although this requires further investigation (Salmon-Godlo, 2006, p. i).

In the Shuttleworth-Edwards et al. (2009) study the validity of the US ImPACT normative data within the South African context was investigated using age and educationally matched groups: ages 11-13, 14-16, 17-21. The comparison between the US normative data and the South African sample of neurocognitive scores was “remarkably similar” (Shuttleworth-Edwards et al., 2009, p. 48). As the South African normative sample consisted of English first language athletes from relatively advantaged educational backgrounds, the results indicated that the US normative data are appropriate for use on South African athletes who have English as a first language and who come from relatively advantaged educational backgrounds (Shuttleworth-Edwards et al., 2009).

1.8.4 The Influence of Repeat Baseline Testing on Normative Data

The data generated by existing normative studies on the ImPACT test for the adolescent population (ImPACT Applications, 2011; Henry & Sandel, 2015; Shuttleworth-Edwards et al., 2009) may not have controlled for the number of baseline tests previously administered, and thus would not have controlled for exposure to practice effects from repeat testing. It can be extrapolated that the existing ImPACT normative studies for the adolescent population (which use the cross-sectional between-group design) were either not influenced by practice effects (in that all participants in each norm group were first-time baseline test-takers), or were influenced by practice effects to some degree (in that some participants in each norm group were first-time baseline test-takers whilst others were repeat-baseline test-takers). Norming participants with different rates of baseline test exposure in the same stratified group would result in at least some of the participants having elevated scores due to practice effects from completing the test previously, whilst others having scores free of practice effects. Therefore, application of such norms to the current clinical situation for high school athletes would potentially be unreliable due to the unknown history of prior baseline tests influencing the norming process.

To the author's knowledge, most high schools (including those in South Africa) that conduct computerised baseline testing do so on an annual pre-season basis, resulting in most athletes having completed a specific number of prior baseline tests at each grade or annual testing interval. For example, a 15-year old Grade 10 learner would have completed two prior baseline tests (one in Grade 8 and one in Grade 9). It can be assumed that clinicians using the existing ImPACT norms are comparing the concussion post-injury scores of athletes with a specific number of prior baseline tests to age-appropriate '*no* practice effects' norms or '*some* practice effects' norms, instead of comparing the post-injury scores to age-appropriate norms that accurately match with the individual athlete's specific number of prior baseline tests.

Therefore, existing normative information may not be the most accurate comparative indexes to use in the repeat testing situation, or in the current clinical situation where athletes are exposed to two, three, or four consecutive test administrations on the same test. Rather, group norms that control for the number of baseline tests previously completed, and thus control for exposure to practice effects by ensuring *all* participants have the *same* number of prior baseline tests at each grade or testing interval, would give clinicians a more accurate indication of what is 'normal' neurocognitive performance following two, three, and four consecutive test-taking occasions on the same test, thereby allowing for a more accurate interpretation of results and clinical diagnosis.

1.8.5 Reliability and Validity of the ImPACT Test

Test-retest reliability “is the extent to which the test produces consistent results across multiple administrations in the same individual” (Cole et al., 2013, p. 733). In other words, it is the extent to which scores on a battery remain stable over time. For a measure to be useful and reliable for assessing neurocognitive change in concussed individuals, it needs to be shown that the scores on the assessment are stable over time in healthy individuals (Randolph, McCrea, Barr, Macciocchi, 2005). For computerized neurocognitive tests, good test-retest reliability is critical for there to be diagnostically accurate comparisons between post-injury and baseline neurocognitive performance (Echemendia et al., 2013). If a test has low or marginal reliability, “any differences between the baseline score and post-injury score may reflect the error of retesting rather than any true change in the athlete” (MacDonald & Duerson, 2015, p. 368).

The ImPACT test-retest intervals have shown a level of test-retest reliability higher or comparable to similar traditional neurocognitive tests measuring similar constructs (Iverson, Lovell & Collins, 2003). Numerous test-retest reliability studies on the ImPACT have produced inconsistent and mixed test-retest reliability results, with low (e.g., intraclass correlation coefficients [ICCs] = .21 for verbal memory) (Tsushima, Siu, Pearce, Zhang, & Oshiro, 2016) to high (e.g., ICC = .88 for visual motor speed) (Schatz & Ferris, 2013) reliability coefficients being reported. ImPACT has demonstrated a stable range of test-retest reliability across various intervals, ranging from 7 days (Iverson et al., 2003a), 1 month (Schatz & Ferris, 2013), 45-50 days (Nakayama, Covassin, Schatz, Nogle, & Kovan, 2014), 1 year (Brett, Smyk, Solomon, Baughman, Schatz, 2016; Elbin, Schatz, & Covassin, 2011), 2 years (Brett et al., 2016; Schatz, 2010) and 3 years (Brett et al., 2016). Even a study by Miller, Adamson, Pink, & Sweet (2007) that used repeated measures analyses, and did not provide reliability coefficients, found the ImPACT to be stable over a 4-month time period in football players who were exposed to repetitive contact over the course of one season. In contrast, other investigations present contradictory results, demonstrating low-to-moderate reliability findings for ImPACT composite scores and symptom scale, ranging from intervals of 1-2-3 days, (Register-Mihalik et al., 2012), 30 days (Cole et al., 2013), 45-50 days (Broglio, Ferrara, Macciocchi, Baumgartner, & Elliott, 2007; Resch et al., 2013a), 1 year (Bruce, Echemendia, Meeuwisse, Comper, & Sisco, 2014) and 2 years (Tsushima et al., 2016).

On the ImPACT test, the Visual Motor Speed and Reaction Time composites have demonstrated greater test-retest reliability than the Verbal Memory and Visual Memory composite (Bruce et al., 2014; Elbin, Schatz, & Covassin, 2011; Nakayama et al., 2014; Resch et al., 2013a; Resch, Schneider, Cullum, 2018; Schatz, 2010; Schatz & Ferris, 2013; Womble, Reynolds Schatz, Shah, & Kontos, 2016). In studies that compared repeat testing score change for reliability purposes, the actual amount of score change was quite small for Visual Motor Speed and Reaction Time, and greater for the Verbal Memory and Visual Memory composites (Nakayama et al., 2014; Register-Mihalik et al., 2012; Schatz, 2010; Schatz & Ferris, 2013). Whilst the ImPACT Verbal and Visual Memory composites have repeatedly recorded low (> 0.59) to marginal (0.60 – 0.69) test-retest reliability (Brett et al., 2016; Cole et al., 2013; Register-Mihalik et al., 2012; Resch et al., 2013a; Resch et al., 2018; Schatz, 2010; Tsushima et al., 2016), the Reaction Time composite has repeatedly recorded marginal (0.60 – 0.69) to adequate (0.70 – 0.79) test-retest reliability (Brett et al., 2016; Elbin et al., 2011; Nakayama et al., 2014; Register-Mihalik et al., 2012; Resch et al., 2013a; Schatz, 2010; Schatz & Ferris, 2013; Womble et al., 2016), and the Visual Motor Speed composite has consistently recorded adequate (0.70 – 0.79) to high (< 0.80) test-retest reliability (Brett et al., 2016; Cole et al., 2013; Elbin et al., 2011; Nakayama et al., 2014; Register-Mihalik et al., 2012; Resch et al., 2013a; Resch et al., 2018; Schatz, 2010; Schatz & Ferris et al., 2013; Tsushima et al., 2016; Womble et al., 2016). This is not only the case on the ImPACT test but has also been observed on other computerised neurocognitive tests. A study by Resch et al. (2018) assessing the test-retest reliability of the ANAM, CVS Vital Signs, and ImPACT computerised neurocognitive tests across three short-term test intervals (day 1, day 45, day 50) in healthy college-age participants found that digitised measures of memory across the three batteries recorded the lowest reliability coefficients (ICC = 0.19 to 0.63) and measures of reaction time (ICC = 0.18 to 0.85) and information processing speed (ICC = 0.28 to 0.89) recorded the highest values across platforms. Whilst it is difficult to make direct comparisons among CNTs due to differences in composite score calculations, computerised measures of working memory and short-term memory have routinely been found to be the least reliable measures when compared to timed measures, such as reaction time and information processing speed (Cole et al., 2013; Littleton, Register-Mihalik, & Guskiewicz, 2015; Nakayama et al., 2014; Nelson et al., 2016; Randolph et al., 2005; Resch et al., 2013a; Resch, McCrea, & Cullum, 2013b; Resch et al., 2018;).

Most test manuals and test-retest reliability studies, including the ImPACT test, report on relatively short retest intervals (e.g. days to weeks) where tests may be expected to be more stable. Whilst these studies are valuable in that they report short-term test-retest reliability, they are far shorter than most clinically relevant scenarios, as the time intervals between baseline and post-injury testing can stretch several months to years (Kirkwood, Randolph, & Yeates, 2009). There has however been an influx of research in recent years on the long-term test-retest reliability of the ImPACT test (Brett et al., 2016; Bruce et al., 2014; Elbin et al., 2011; Schatz, 2010; Tsushima et al., 2016; Womble et al., 2016). This is likely in response to the call to establish test-retest reliability over time intervals that are clinically relevant for comparing baseline and post-injury data (Schatz, 2010), and in response to the need for more empirical evidence for how often baseline testing should be repeated in certain populations. The long-term test-retest reliability or test stability of a measure is one way of providing such recommendations because if a measure has good test-retest reliability or stability between test and retest over a period, there is no need to repeat baseline test as frequently in that population.

The ImPACT test has demonstrated good construct validity when examined in relation to common neurocognitive tests. A study by Iverson et al. (2003a) found good construct validity between the ImPACT Memory and Processing Speed composite scores and other traditional neurocognitive measures including Trail Making Tests A & B, Brief Visuospatial Memory Test (BVMT), and the Symbol Digit Modalities Test (SDMT). Research has also found good convergent validity between the Visual Motor Speed and Reaction Time composites of the ImPACT test and a traditional neurocognitive measure – the SDMT (Iverson, Lovell, & Collins, 2005). The ImPACT test has also been found to be excellent in detecting a concussion and discriminating between non-concussed and concussed athletes. A study by Schatz, Pardini, Lovell, Collins, and Podell (2006) found that the combined sensitivity (i.e. the probability that the test result will be positive when a concussion is present) of the ImPACT composites and the symptom score was 81.9%. The specificity (i.e., the probability that a test result will be negative when a concussion is not present) was very high at 89.4%.

1.9 The Stability of Neurocognitive Scores over Time

When assessing the stability of neurocognitive scores between two or more test intervals on healthy non-clinical athletes, two primary factors could cause changes between test intervals, thus negatively affecting the stability of neurocognitive scores over time. These factors are practice effects and cognitive maturation.

1.9.1 Practice Effects (the Effect of Repeat Testing)

Practice effects can cause scores to change between test administrations. What is considered a 'practice effect' can extend beyond simply remembering the answers to certain items (i.e., declarative memory of the content or recall) but could also include procedural learning (i.e., remembering how to do the test or familiarity with the testing procedure) (Strauss et al., 2006). Practice effects can be minimised by the randomisation of test items and the generation of alternate forms of a test (Collie et al., 2001). Whilst a near infinite number of random test items and alternate forms integrated into most computerised neurocognitive tests (including ImPACT) minimise practice effects caused by the recall of specific content, they do not control for practice effects caused by procedural learning or familiarity with the testing procedure (i.e., remembering how to do the test) (Hinton-Bayre & Geffer, 2005). According to Lezak et al. (2012), most variance in retesting is felt to be due to familiarity with the procedure as opposed to the specific content.

Research has found that the amount of time between tests has a negative influence on the size of the practice effect (Calamia, Markon, Tranel, 2012; Salthouse, Schroeder, & Ferrer, 2004). When the time interval between test and retest is relatively short, practice effects increase. Conversely, when the time interval between test and retest is relatively long, practice effects decrease. This is because the longer the test-retest interval, the less test-specific skills (e.g. strategies, rules of the test, correct answers to items) that can be recalled. A longer test-retest interval may also lead to a decrease in familiarity and rule comprehension, "because participants are less likely to remember the test in order to feel familiar with it, its rules and their comprehension thereof" (Scharfen, Peters, & Holling, 2018, p. 47).

Whilst longer retest intervals can diminish practice effects, they do not necessarily eliminate them (Duff, 2012), as practice effects have been reported on cognitive tests years after initial testing (e.g., 2.5 years) (Salthouse, 2010). It has also been noted that "with longer intervals, practice effects become more difficult to distinguish from actual changes and within-person variability" (Calamia et al., 2012, p. 547). This has been reiterated by Salthouse and Tucker-Drob (2008) which state that in longitudinal comparisons that involve repeated testing of the same individuals, it is possible that at least some of the observed within-person change in performance is due to the effects of prior test experience (i.e. practice effects) rather than solely to influences related to maturation.

Several studies on the effect that multiple test administrations have on practice effects have reported a large score gain from the first to the second test, with practice effects becoming smaller with the number of tests (Albers & Hoeft, 2009; Bartels, Wegrzyn, Wiedl, Ackermann, & Ehrenreich 2010; Dunlop, Morrison & Cordery, 2011; Puddey, Mercer, Andrich, & Styles, 2014). This non-linear progression aligns with the theoretical view of the *power law of practice*. This theory "describes the assumption that learning curves show diminishing gains over time", and after a phase of improvement no further gains are observed (Scharfen et al., 2018, p. 45). In other words, retest effects are largest between the first and second test administrations but decrease with the number of test administrations, eventually reaching a plateau or levelling off after a certain number of tests (Scharfen et al., 2018). According to Newell and Rosenbloom (1981), the *power law of practice* should hold for "all types of mental ability" (p. 33) and can thus be applied to practice effects in cognitive ability tests when there are multiple test administrations. The number of test administrations that would need to occur before there are no further gains from practice effects is plausibly dependent on numerous factors, such as the measure used and the length of the test-retest intervals. On some measures, practice effects have been found to influence test performance up to four serial assessments (Maerlender et al., 2016; McCaffrey, Duff, & Westervelt, 2000).

1.9.2 Cognitive Development and Maturation

General cognitive maturation or development will cause scores to change between two or more test administrations if the time interval between test and retest is sufficiently long, as people may cognitively change between test and retest. Generally, shorter test-retest intervals lead to higher score stability, whilst longer retest intervals lead to lower score stability. Even on very stable measures, long intervals between test and retest (longer than about six months) make the study vulnerable to the effects of cognitive maturation (Bless, Higson-Smith, & Kagee, 2006).

Cognitive maturation and development vary depending on the examinee's subjective characteristics, as some groups are intrinsically more variable over time than others (Strauss et al., 2006). Characteristics likely to influence the extent to which scores fluctuate over time are age (i.e., examinee developmental level), neurological status, education level, intellect, gender and physical activity (Duff, 2012; Littleton et al., 2015). Age-related differences are a factor shown to influence performance on neurocognitive tests (Littleton et al., 2015). For example, normal pre-schoolers could be expected to show more test-retest variability than adults on neurocognitive tests (Strauss et al., 2006). According to Luria's theory of cognitive development discussed in section 1.6.2, most cognitive units are functional by age 12 years but continue to develop through early adulthood (Luria, 1980). Whilst much of cognitive development is present at 12 years of age, brain development and underlying neuro maturational processes continue throughout adolescence and into early adulthood (Luria, 1980; Casey et al., 2000; Yurgelun-Todd, 2007; Toledo et al., 2012), which may explain the age-related differences in neurocognitive test performance found in some studies (Hunt & Ferrara, 2009; Register-Mihalik et al., 2012).

A between-groups cross-sectional study by Hunt and Ferrara (2009) examined differences at baseline testing among grade levels in US high school athletes on pencil-and-paper neurocognitive tests used to manage sports-related concussion (Grade 9, 10, 11, 12, with mean ages respectively of 14.15 (SD = 0.68), 15.22 (SD = 0.42), 16.25 (SD = 0.44), 17.09 (SD = 0.61). As this study used different participants at each grade level, practice effects did not influence the findings as participants did not undergo repeat testing (i.e. all participants were tested on one occasion only). There were significant differences on tests of information processing (Trail Making Test), attention (Symbol Digit Modality Test), and motor speed (finger tap test) between Grade 9 and Grade 11 and Grade 9 and Grade 12 athletes, with neurocognitive test scores in these domains differing at 15.22 years of age, or Grade 10. The significant differences between Grade 9 participants and those in later grades, and the fact that older participants seemed to perform better than younger participants on tests of information processing, attention, and motor speed, indicates that cognitive development of these domains continues through high school. Furthermore, the normative test data on both traditional pencil-and-paper and publically available computerised neuropsychological tests for information processing, attention, and motor ability presents a distinct separation in score performance among age groups, with performance improving as age increases (Bornstein, 1985; Kumar Sharma et al., 2014; Smith, 1991; Yeudall, Reddon, Gill, & Stefanyk, 1987). All normative data for the significant tests in these studies showed a difference between those 15 years and older, and younger adolescents (Bornstein, 1985; Smith, 1991; Yeudall, Reddon, Gill, & Stefanyk, 1987). As in the Hunt and Ferrara (2009) study, different participants at each grade or age level were used in these existing normative studies.

In the Hunt and Ferrata (2009) study, no significant differences among grade levels in high school athletes on tests of cognitive function (Standardized Assessment of Concussion), memory/verbal learning (Hopkins Verbal Learning Test), concentration and attention (Digit Span Test (forward)), and immediate memory and concentration (Digit Span Test (backward)) were found. Several theories may explain why test results in these domains were not different. It could be that "these domains obtain cognitive maturity before 9th grade, reliability of the neuropsychological tests was poor, or the test did not stress the specific circumstances under which a deficit within the domain would be demonstrated" (Hunt & Ferrara, 2009, p. 407).

A subsequent study by Register-Mihalik et al. (2012) support the findings of Hunt and Ferrara (2009) that adolescents displayed age effects for processing speed. One of the purposes of this study was to determine if age affects neurocognitive test performance on both pencil-and-paper (Hopkins Verbal Learning Test, Brief Visual-Spatial Memory Test scores, Trail Making Test B total time, Symbol Digit Modalities Test, Stroop Test total score) and computerised neurocognitive tests (the five composite scores of the ImPACT). In this study, two separate groups, one consisting of collegiate student-athletes (mean age 20.00 ± 0.79 years) and the other consisting of high school student-athletes (mean age 16.00 ± 0.86 years) completed various neurocognitive tests. It was found that the collegiate student-athletes performed significantly better than high school student-athletes on the Visual Motor Speed composite of the ImPACT test across all 3 test sessions approximately 24 hours apart, indicating that continued improvements in visual motor (processing) speed are expected between adolescence and early adulthood. Of all the tests used in this study, age-related differences between high school and collegiate athletes were found on only the ImPACT Visual Motor (processing) Speed composite. There were no significant differences in neurocognitive performance between high school and collegiate athletes on the Verbal Memory, Visual Memory, Reaction Time or Impulse Control composites. As the mean age of the high-school athletes in this study was 16 years old, it is possible cognitive maturity on these domains was obtained before this time, during the period of late childhood or early adolescence.

The significant age effects for visual motor (processing) speed found in the Hunt and Ferrara (2009) and Register-Mihalik et al. (2012) studies indicate that significant cognitive maturation and development in this domain continues through adolescence and even into early adulthood, with improvement as age increases. Further corroborating these findings is a subsequent cross-sectional between-group study by Kumar Sharma et al. (2014), which presented year-wise normative data for age groups from 12 years onward to 17 years. This study found consistent improvement in performance with increasing age in adolescents on pencil-and-paper measures of visual motor processing speed (e.g. Trail Making Test A and B and Two Target Letter Cancellation Task).

Such developmental changes are less expected in the domains of verbal memory, where Hunt and Ferrara (2009) found no significant differences between grade groups during adolescence on the Hopkins Verbal Learning Test [Grade 9, 10, 11, 12, with mean ages respectively of 14.15 (SD = 0.68), 15.22 (SD = 0.42), 16.25 (SD = 0.44), 17.09 (SD = 0.61)] and Register-Mihalik et al. (2012) found no significant difference between high school athletes (age 16) and collegiate athletes (age 20) on the Hopkins Verbal Learning Test and the ImPACT Verbal Memory Composite. It may be the case that in the domain of verbal memory, cognitive maturity is largely complete by early adolescence or before late adolescence, although this is not to say non-significant improvements do not continue through adolescence.

In the domain of visual memory, Register-Mihalik et al. (2012) found no significant difference between high school athletes (age 16) and collegiate athletes (age 20) on the Brief Visual-Spatial Memory Test and the ImPACT Visual Memory composite. Research on recognition memory found that recognition memory begins to plateau around 13 years of age (Flin, 1985). These studies indicate that in the domain of visual memory (consisting of both visual working memory and visual recognition memory), cognitive maturity may largely be complete by early adolescence or before late adolescence, although this is not to say non-significant improvements do not continue through adolescence. There is however research on visual working memory that contradicts this. According to Isbell, Fukuda, Neville, and Vogel (2015) visual working memory continued to develop throughout adolescence on an adaptation of the visual change detection task, and does not reach adult levels even in 16-year-olds.

In the domain of reaction time, Register-Mihalik et al. (2012) found no significant difference between high school (age 16) and collegiate athletes (age 20) on the ImPACT Reaction Time composite. In a between-group cross-sectional study of 180 participants using the CogSport neurocognitive test, there were substantial improvement in performance between the ages of 8 and 18 years on tests on simple and choice reaction time (as well as working memory and new learning). However, the largest test performance improvements were recorded between the ages of 8 and 15, “with minimal changes after this age paralleling adult performance” (McCrory et al., 2004, p. 516). These studies indicate that in the domain of reaction time, cognitive maturity may largely be complete by early adolescence or before late adolescence, with there being minimal changes after this time.

Conceptually, in a within-subjects repeated-measures study using cognitively developing participants with a long period of time between test-retest intervals, practice effects could plausibly be a contributing factor towards any observed test-retest change. Unfortunately, it is not possible to differentiate what portion of the observed change is due to naturally occurring cognitive development and what is due to practice effects from repeat testing. Any observed change in neurocognitive scores between test intervals could plausibly be a combination of both of these factors.

It is relevant to note that the studies above were able to determine cognitive maturation or development between grades/ages by using between-group cross-sectional research designs. In other words, the cognitive performance of different groups of participants were compared at different grades/ages (i.e., the cognitive performance of a group of 13-year-olds were compared with the cognitive performance of a different group of 15-year-olds, and any improvement between these age-groups was considered cognitive maturation or development by the relevant authors). Practice effects were not considered a contributing factor to any observed change in the above mentioned studies due to the fact that no within-subject repeat testing occurred.

1.10 Frequency of Repeat Baseline Testing

Whilst it is generally accepted that athletes participating in contact sport would benefit from completing a baseline neurocognitive test (Merritt et al., 2016; Lovell, 2009), there are no firm guidelines and minimal consistent agreement as to how often baseline testing should be updated or repeated in high school athletes (Guskiewicz et al., 2004; Schatz, 2010). In fact, the ideal interval for repeat baseline testing on computerised neurocognitive tests has not been established (Harmon et al., 2013). The Centers for Disease Control and Prevention (2015) has recommended that “most components of baseline testing be repeated annually to establish a valid test result for comparison” (para. 9), although it also states that baseline testing may be repeated as infrequently as every 2 years. On their website, ImPACT purveyors recommend completing baseline testing annually to ensure an up-to-date record of a patient’s normal functioning, although they do not specify the age level that this applies to (i.e., high school, collegiate, or professional athletes etc.)

(ImPACT online, 2019d).

It is important to repeat baseline testing at the optimal intervals during the high school period. From a clinical perspective, if baseline tests are not updated frequently enough, the ability of the baseline/post-injury model to accurately detect the subtle neurocognitive deficits of concussion may be compromised. This is because cognitive improvement or development that takes place between baseline and post-injury "may offset any injury related cognitive impairment in concussed children and adolescents" (McCrory et al. 2005, p. 516). This may result in the clinician interpreting the post-injury scores as normal, when in truth the magnitude of the impairments has been obscured by the increased cognitive maturation that took place between baseline and post-injury. As concussions are increasingly being recognised for their adverse effects on the developing brain, and adolescents are particularly vulnerable to sustaining concussive injuries as discussed in section 1.6, being able to accurately detect the subtle neurocognitive deficits of the injury in adolescent athletes is incredibly important (Institute of Medicine and National Research Council, 2013; Carman et al., 2015). From a practical perspective, baseline testing that is updated too frequently or unnecessary is wasteful in terms of finances, resources and time for the client, albeit commercially rewarding for the test developers.

Furthermore, evidence-based information on the optimal frequency and timing of repeat baseline testing for high school athletes may aid test developers formulate firm 'repeat baseline testing' guidelines for this age level. Such information would be beneficial to education administrators and other stakeholders using the ImPACT test in the clinical setting. Finally, evidence-based information would also improve the overall administration of the ImPACT test in high school athletes. Despite the widespread use of CNTs, existing studies among high school athletes at risk of concussion indicate that in this population there is a need to improve the overall administration of computerised testing (Lichtenstein, Moser, & Schatz, 2013; Resch et al., 2013b).

1.10.1 Literature-based Recommendations for the Frequency of Repeat Baseline Testing

For college-aged athletes, the recommendation of baseline testing every two years is supported by the Schatz (2010) test-retest reliability study. This study found that cognitive performance of collegiate athletes at baseline remained considerable stable over a two-year period, indicating annual repeat baseline testing may not be necessary. Another study by Maerlender and Molfese (2015) analysed repeat baseline testing scores on the ImPACT test between 13 and 40 months apart in college-aged athletes to determine if repeat administration was necessary. Using t-test analyses, no significant differences in any test composite score were identified. Considering these findings, the study does not support the recommendation that baseline testing should be repeated in college athletes but noted that "this study has no bearing on the recommendation for retesting younger athletes every two years" (Maerlender & Molfese, 2015, p. 72).

Research has identified the need for more frequent updating of baseline testing for younger athletes (e.g. high school and earlier) compared with older athletes (Buzzini & Guskiewicz, 2006; Elbin et al., 2011; Schatz, 2010), as the neurocognitive baseline scores of young athletes may change often due to natural cognitive maturation and development that occurs rapidly during childhood and adolescence (Casey et al., 2000; Luria, 1980; Toledo et al., 2012; Yurgelun-Todd, 2007). As mentioned previously, developmental change has been observed on the CogSport neurocognitive test between the ages of 8 and 18, with the largest improvement between being between the ages of 8 and 15 (McCrory et al., 2004). These changes due to cognitive development are so large to be of "comparable magnitude to post-concussive impairments observed on computerised cognitive post-injury assessment in adults" (McCrory et al., 2004, p. 516). Due to the rapid cognitive maturation observed between age 8 and 15, baseline testing twice-annually was suggested to "enable accurate comparison for serial testing" (McCrory et al., 2004, p. 517). It was however noted that this would be too resource intensive for implementation, with it being suggested that children and adolescents (especially those playing contact sport where there is a greater risk of injury) should receive annual repeated baseline testing.

To the author's knowledge, most high schools (including those in South Africa) that provide computerised baseline testing do so on an annual pre-season basis. However, there are some high schools that conduct baseline assessments less frequently, presumably due to financial, personnel and/or resource limitations, as well as the general perception that serial assessments may not be necessary (Rogers, Smith, Stephenson, & Everhart, 2017). Clinicians have questioned the need to obtain annual baseline neurocognitive tests in high school athletes (Hunt & Ferrara, 2009). As mentioned previously, a between-subject cross-sectional study by Hunt and Ferrara (2009) examined differences at baseline testing on pencil-and-paper neurocognitive tests among grade levels in high school athletes, and subsequently observed that baseline neurocognitive test scores improved as a function of age, with significant differences between Grade 9 and Grade 11 and Grade 9 and Grade 12 on tests of information processing, attention, and motor speed. As the differences were driven by Grade 9 test scores, it was recommended that high school athletes should be baseline tested at least twice; once during their initial entrance into high school athletics (Grade 9) and again upon entering Grade 10. “Retesting upon entrance into 10th grade provides the opportunity to update test scores during cognitive growth” (Hunt & Ferrara, 2009, p. 408), which can be identified by a trend of improved performance. Whilst test scores appeared to remain stable after Grade 10, the study recommended annual baseline testing until additional research could confirm the results.

In the Elbin et al. (2011) test-retest reliability study it was found that the online ImPACT baseline composite scores were stable across a one-year time period for high school athletes (age range 13-18 years). Although the author’s interpretation of the intraclass correlation coefficient (ICC) data has been criticised (Brett et al., 2016), other analyses [Reliable Change Indexes (RCI) and Regression-based Methods (RBM)] provided support for test-retest stability. However, the authors did note that although baseline data from high school reflected stability across a one-year period, these scores were “not without degradation or change” (Elbin et al., 2011, p. 2322). The authors recommended that high school athletes receive updated baselines every two years, but also stated that “given the lack of data over a two-year period, when resources permit, high school athletes may benefit from updated baselines every year” (Elbin et al., 2011, p. 2322). The stability results in the Elbin et al. (2011) study are contrary to a study by MacDonald and Duerson (2015) using a different computerised neurocognitive test (the AxonSport), which found low to marginal test-retest reliability over a one year test-retest interval in high-school athletes.

A subsequent study by Tsushima et al. (2016) on the ImPACT test assessed test-retest reliability over two years in non-concussed athletes tested in Grade 9 (mean age 15.3 ± 0.5 years) and again in Grade 11 (mean age 17.3 ± 0.5 years). The study found not only evidence of stability but also evidence of developmental change. Using RBM analyses, results were stable after a two-year interval (i.e., showing no reliable change between initial baseline scores and follow-up baseline scores). These results indicate that repeat baseline testing of high school athletes every two years may not be necessary. Conversely, ICC and RCI analyses of the data found low test-retest reliability correlation coefficients and significant changes in RCIs. These results reflect developmental changes in adolescents that take place over the two-year period between Grade 9 and Grade 11 (Tsushima et al., 2016, p. 109) and offers some “support to the cautious practice of updating baseline testing every two years to account for any changes in young athletes” (Tsushima et al., 2016, p. 109). These mixed findings caused the authors to recommended future research into the consistency of the ImPACT test and other computerised neurocognitive tests over two years with high school participants.

A study by Brett et al. (2016) assessed test-retest reliability on the ImPACT test at mean time intervals of approximately one-, two-, and three- years using ICCs, reliable change indices and regression-based analysis. All athletes tested at either one-, two- or three- years between baseline assessments were only tested twice (Time 1 and Time 2). The results indicated that ImPACT composites remained considerably stable at one-, two-, and - year test-retest intervals in high-school athletes (age range 13-18 years), suggesting it does not matter if you wait one-, two-, or even three- years to update a baseline as there was minimal variation in reliability across all years (Brett et al., 2016). However, the authors noted that there is variance at the individual level and the analysis did not capture. It must also be noted that the test-retest interval ranges in this study were very broad. Those in Group 1 (*N* = 250) completed baseline anywhere between 6 and 18 months (mean = 11.6); Group 2 (*N* = 1146) anywhere between 19 and 30 months (mean = 23.21); and Group 3 (*N*= 114) anywhere between 31 and 42 months (mean = 33.83). These broad test-retest interval ranges could result in individuals with rather different time intervals between assessments (i.e., 6 months and 18 months) being collectively grouped together, potentially resulting in less accurate test-retest reliability results.

In response to the inconsistent results produced by test-retest studies (Elbin et al., 2011; Tsushima et al., 2016; Brett et al., 2016), a recent study by Rogers et al. (2017) assessed the pattern of change in neurocognitive test performance of high school athletes across each year of the four years of US high school (ages 15 to 18 years). One part of the study was to longitudinally assess athletes for changes in neurocognitive test performance across test intervals over time, in order to determine the frequency and timing of repeat baseline testing on the CNS Vital Signs computerised neurocognitive test in high school athletes. The longitudinal analysis involved a repeated measure mixed model analyses to examine participants with two or more assessment intervals. The study found moderate-to-large improvements between Freshman and Senior years (i.e. year 1 and year 4), with findings suggesting that “cognitive performance increases incrementally each year of high school, with substantial increases in executive functioning and processing speed” (Rogers et al., 2017, p. 1894), and large annual changes in cognitive flexibility and reaction time. Based on the change that occurs each year of high school in these domains, the authors recommended annual baseline neurocognitive testing for high school athletes. The large magnitude of change in executive functioning and processing speed indicates that frontal lobe-dependent tests (i.e. processing speed) may be particularly important to reassess each year. This study importantly noted that any change observed over time is plausibly “a combination or interaction of developmental differences and practice effects from having repeat test administrations” (Rogers et al., 2017, p. 1898). The authors went on to note that regardless of the etiology, the findings of the study “support the notion that such changes do occur and should be considered when making decisions about the frequency of baseline testing of high-school athletes” (Rogers et al., 2017, p. 1898).

It is worth noting methodological problems in the exisiting research that have provided recommendations for the frequency of repeat baseline testing. Two of the three existing test-retest reliability studies using the ImPACT test for adolescent athletes grouped all adolescent ages together at test and retest (age range 13-18) (Elbin et al., 2011; Brett et al., 2016). Neurocognitive performance in between-subject cross-sectional studies has been shown to differ at different adolescent ages, with older adolescents usually performing better on neurocognitive tests than younger adolescents (Hunt & Ferrara, 2009; Maruff et al., 2004). Furthermore, whilst cognitive development occurs throughout adolescence on some cognitive domains, the rate of cognitive development varies at different ages. For example, research indicates that more cognitive development occurs during late childhood and early adolescence, than during late adolescence (Womble et al., 2016; Casey et al., 2000). In situations where all adolescent ages are grouped together, the relative score stability expected during late adolescence may obscure or offset the more dramatic changes expected during early adolescence, potentially resulting in stability findings between test and retest that are not accurate. For example, it could be the case that a 17 year old athlete experiences little neurocognitive score change over one year, whereas a 13 year old experiences dramatic change over the same time period. Unfortunately, this vital distinction is obscured in most of existing test-retest reliability studies. With regard to Tsushima et al. (2016), whilst this study did group athletes of the same age together at test and retest (i.e., athletes were initially tested at Grade 9 and then again at Grade 11), the findings from different statistical analyses used in this study were mixed and conflicting, resulting in more confusion than clarity. The Rogers et al. (2017) study using the CNS Vital Signs neurocognitive test was the first and only study to date to assess the pattern of neurocognitive change using a longitudinal approach over the four academic years of US high school. High school athletes were assessed each year of high school for ages 15 to 18 years, thus allowing for the change in neurocognitive score performance across grades to be observed. This longitudinal design considers both the frequency (how often during the high school period) and timing (at which grades during the high school period) that baseline testing should be repeated. However, even this study is flawed in that not all participants completed four annual consecutive baseline assessments. Participants with as little as *two* assessment time points were included in the longitudinal analysis of change, which may have affected the results.

1.11 Aim and Rationale of Objective 1: Repeat Baseline Testing over Five Annual Test Intervals

On the basis of the literature review under section 1.10 above, it is evident that neurocognitive baseline testing is recommended (Lovell, 2009; Merritt et al., 2016), but that the frequency with which this should be applied throughout the adolescent years is not clear (Centre for Disease Control and Prevention, 2015; Guskiewicz et al., 2004; Harmon et al., 2013; Schatz, 2010). Furthermore, amongst existing computerised neurocognitive test studies, there is a general lack of consensus and conflicting recommendations for the frequency of repeat baseline testing of adolescent high school athletes (Brett et al., 2016; Elbin et al., 2011; Hunt & Ferrara, 2009; MacDonald & Duerson, 2015; McCrory et al., 2004; Rogers et al., 2017; Tsushima et al., 2016). On the ImPACT test, which is used in multiple countries around the world, evidence-based information on the optimal frequency and timing of repeat baseline testing for high school athletes may aid test developers formulate firm ‘repeat baseline testing’ guidelines for this age level. Such information would also be beneficial to clinicians, education administrators and other stakeholders using the ImPACT test in various settings, and would improve the overall administration of the ImPACT test for high school athletes.

Therefore, the primary objective of this study (Objective 1) was to assess the pattern of change in neurocognitive performance for adolescent athletes on baseline measures of the ImPACT test, over five consecutive years with test-taking intervals approximately one year apart. Participants were non-clinical predominantly South African high school athletes from Grades 8 to 12, in the overall age range 13 to 18 years, divided into five groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30). An enhanced understanding of the year-by-year neurocognitive score changes experienced by high school athletes across the 13 to 18 year old adolescent age range would allow for the development and implementation of more effective recommendations related to sports-concussion management. In particular, an indication of the optimal interval for the repeat baseline testing of high school athletes.

Currently there is no within-subject longitudinal study worldwide on any computerised neurocognitive test for sports concussion management that has assessed changes in neurocognitive scores on the *same participants every year over the five-year high-school period*, spanning from early adolescence to late adolescence. Such a longitudinal within-subject approach reflects the current clinical situation, where cognitively developing South African high school athletes are usually baseline tested annually throughout high school (every year over a five-year period). It is relevant to note that *any observed changes in neurocognitive scores from year-to-year in the present study could plausibly be a combination or interaction of naturally occurring cognitive development as high school athletes increase in age, and practice effects from repeat testing. Methodologically, on the basis of this study it will not be possible to differentiate between these two effects as the independent cause of any neurocognitive score change as they will be happening simultaneously. Therefore, regardless of specific causation between these two possible causes, the present study will seek to assess the actual neurocognitive score changes that occur in the current clinical situation, where cognitively developing South African high school athletes are usually baseline tested annually throughout high school (every year over a five-year period) and are plausibly vulnerable to both causes of change.*

It was hypothesised that neurocognitive score changes would occur on the ImPACT test each year from Grades 8 to 12 (overall age range 13 to 18 years) in non-clinical adolescent athletes during their high school careers. As mentioned, the expected changes would likely be due to the combination or interaction of a) cognitive development and maturation that occurs during adolescence (Casey et al., 2000; Luria, 1980; Toledo et al., 2012; Yurgelun-Todd, 2007) and b) practice effects from repeat testing (Rogers et al., 2017). Cognitive development is expected during adolescence, with reported improved performance on neurocognitive tests with increased age. In other words, older adolescents tend to perform better on neurocognitive tests than younger adolescents (Hunt & Ferrara, 2009; Maruff et al., 2004). In addition to cognitive development, practice effects are also expected in this study. Randomly-generated test items and alternate forms on computerised tests are used to minimise practice effects that occur at the retest interval. However, alternate forms only minimise practice effects caused by recall of content (i.e. remembering the answers to certain items) and do not control for practice effects caused by procedural learning or familiarity with the testing procedure (i.e. remembering how to do the test) (Hinton-Bayre & Geffen, 2005). In fact, most variance in retesting is believed to be the result of familiarity with the procedure as opposed to the specific content (Lezak et al., 2012). Moreover, of particular relevance to this study is that baseline testing is done repeatedly with the same baseline version of the ImPACT test (Baseline Form 1), as only the follow-up tests have multiple versions. Therefore, practice effects caused by recall of specific content and/or familiarity with the testing procedure are not minimised in repeated baseline testing. As such, it was expected that there would be an unknown amount of practice effects over the five repeated baseline tests caused by recall of content and/or familiarity with the testing procedure in the present study.

1.12 Aim and Rationale of Objective 2: Normative Data for Composite Scores, Subtest Scores, and Additional Memory Subcomponent Scores over Five Annual Test Intervals

On the basis of the literature review under section 1.7.3 above, it is evident that normative data are invaluable in managing a concussive injury when individualised baseline testing is unavailable, inappropriate, or invalid (Hinton-Bayre, 2015; Louey et al., 2014; Merritt et al., 2016). When comparing post-injury results to normative data, the use of normative data that closely match the demographics of the athlete provides the best comparison method (Merritt et al., 2016), and improves the accuracy of the test results and interpretation of the data. Normative profiles tailored to the high school athlete’s particular grade and age stage and specific number of prior baseline tests would accurately match the demographics of high school athletes in the current clinical situation, and be of clinical relevance. On the basis of the literature review under section 1.8.4 above, the available ImPACT normative data provided by existing studies (Henry & Sandel, 2015; Iverson et al., 2003a; Shuttleworth-Edwards et al., 2009) may not be the most accurate comparative indexes to use in the repeat testing situation, or the current clinical situation where athletes are exposed to two, three or four consecutive test administrations on the same test. Rather, group norms generated from a longitudinal within-subject design where *all* participants completed the *same* number of prior baseline tests at each grade or test interval, and thus *all* participants have the *same* exposure rate to practice effects may be a more accurate comparative index to use in the repeat testing situation. This normative information may aid clinicians in making more accurate comparisons and clinical diagnoses.

Therefore, the secondary objective was to generate normative tables for the ImPACT tests' five composite scores for the five participant groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30). This was done with the view to establishing normative platforms on which to base interpretations of concussion follow-up testing depending on the athlete's particular grade and age stage and specific number of prior baseline tests. The five normative participant groups are as follows, stratified by Grade as the primary variable, and age mean (SD), age range, and number of baseline tests previously completed as qualifiers: Grade 8 [mean age 13.69 (0.29) years, age range 12-13 years, 0 prior baseline tests]; Grade 9 [mean age 14.62 (0.32) years, 13-14 years, 1 prior baseline test]; Grade 10 [mean age 15.58 (0.31) years, age range 14-15 years, 2 prior baseline tests]; Grade 11 [mean age 16.62 (0.31) years, age range 16-17 years, 3 prior baseline tests]; Grade 12 [mean age 17.62 (0.30) years, 17-18 years, 4 prior baseline tests]. On this basis it was considered that normative data would be more focused and accurate, thereby providing a relevant basis for clinicians to refer to when assessing athletes of a particular grade and age stage, who have completed a specific number of prior baseline tests, thereby aiding clinicians in making more accurate comparisons and clinical diagnoses.

Currently there are no normative data for the ImPACT test worldwide that have been established prospectively over a five-year period, with a view to delineating normative indications for each year, from year one through to year five, so that a clinician can refer to a specific norm base depending on the concussed examinees particular grade and age stage, and specific number of baseline tests previously completed. In addition to this, providing normative data that is more age-specific and represents a wide spectrum of age groups in the adolescent period [i.e., Grade 8 (13-14 years); Grade 9 (14-16 years); Grade 10 (15-16 years); Grade 11 (16-17 years); Grade 12 (17-18 years)], than the broader adolescent age-group norms currently provided by ImPACT Applications [i.e., 13-15 years; 16-18 years] (ImPACT Applications, 2011), would allow clinicians to refer to more definitive year-by-year grade/age-group norms when comparing how an athlete is performing. This is particularly relevant considering high school adolescent athletes are going through a period of cognitive development and maturation (Casey et al., 2000; Luria, 1980; Toledo et al., 2012; Yurgelun-Todd, 2007), likely resulting in performance differences at different grades/ages. Such norms will likely result in better representation for performance comparison among 13 to 18 year olds.

In addition to generating normative tables for the ImPACT tests' five composite scores, it is also clinically useful to produce user-friendly normative data for the twelve subtest scores that make up the five ImPACT composite scores, as well as four additional subcomponent scores that make up the ImPACT Word Memory and Design Memory subtests. Such normative data is clinically useful as it allows for better comprehension of the ImPACT test that extends beyond the basic interpretation of composite scores and may be "helpful in the interpretation of the ImPACT clinical report and further delineation of areas of neurocognitive dysfunction" (Henry & Sandel, 2015, p. 266). It would also enrich the clinical interpretation of ImPACT test performance and provide additional interpretative information that may aid clinical judgement. Normative data on the immediate and delayed recall subcomponents of the Word Memory and Design Memory subtests would be especially relevant comparative indexes for clinicians managing concussive injuries, as delayed recall is particularly sensitive to brain impairment or dysfunction (Lezak et al., 2012). A falloff in delayed recall scores in the cognitive domains of verbal and/or visual memory is an indication of brain injury which may not be apparent on immediate recall. Even if immediate recall scores for an individual are "normal", there may be a deficit on delayed recall scores indicating the individual is cognitively impaired.

Considering the above, a subsidiary aspect of Objective 2 was to generate normative tables for the twelve subtest scores that make up the five ImPACT composite scores, as well as four additional memory subcomponent scores which were of clinical relevance (i.e., Word Memory immediate recall; Word Memory delayed recall; Design Memory immediate recall; Design Memory delayed recall). As Objective 2 is merely aimed at providing a platform for future normative comparisons, only broad observations on these sixteen additional normative data sets will be made on the basis of this research. A detailed neurocognitive analysis of how the subtest scores and additional memory subcomponent scores might be differentially affected in the event of a concussion is beyond the scope of the present research on a nonclinical population.

Overall, the secondary objective of this study (Objective 2) was to generate normative tables for the ImPACT tests' five composite scores, twelve subtest scores that make up the composite scores, as well as four additional memory subcomponent scores which were of clinical relevance. Descriptive data for the five participant groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30), will be tabled at each test interval for each of these five composite scores, twelve subtest scores, and four additional memory subcomponent scores. The results provide a clinical and heuristic normative platform for future use with brain injured individuals, which could be used to facilitate clinical interpretations of postconcussion assessment.

CHAPTER 2: METHODOLOGY

2.1 Research Design

A retrospective quasi-experimental longitudinal within-subject repeated-measures design was used in this study. In this design, grade level was treated as the primary within-subject variable, as each participant provides observations derived at each of the five consecutive grades tested in the study (McBride, 2013). This allowed for the dependent variables to be compared at different grades for the single group of participants over time. This study made use of pre-existing de-identified archival data obtained from the South African ImPACT test database. Participants were baseline tested at approximately one-year intervals for five consecutive years from Grade 8 to Grade 12 starting from either 2007, 2008, 2009, 2010 or 2011. The dependent variables were the five composite scores on the ImPACT test (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, Impulse Control). The independent variable was time (Interval 1, Interval 2, Interval 3, Interval 4, Interval 5).

2.2 Participants

2.2.1 Selection and Exclusion Criteria

A longitudinal view of an adolescent population was of interest as it could provide insight into neurocognitive score change currently lacking on the ImPACT test (Maelender & Molfese, 2015). In South Africa, the ImPACT programme has been established in eight educationally advantaged English medium well-resourced government funded model C and privately funded private schools for over a decade. In the apartheid era schools were racially separated for white children only, and Department of Education and Training (DET) poorly resourced schools in township and rural areas that were attended by black, coloured, Asian and Indian children only[1]. Since the dismantling of apartheid in 1994, schools are no longer formally reserved for exclusive attendance by any particular race, and the former categories have been replaced by a new system that defines schools

---

[1] "During the period between 1948 and 1994, all South Africans were classified according to the Population Registration Act (Act No. 30 of 1950) into four main divisions including 'white', 'black, 'coloured' and 'Asian/Indian'. These terms are thus utilised here on the understanding that they are socially constructed racial signifiers, which have had particular pertinence and continue to be salient in a complex relationship with each other within the South African society, and therefore have salience within the local research context (Van Ommen, 2013)" (Naidoo, Shuttleworth-Edwards, Botha, & Pienaar, 2019, p. 65).

in Quintiles from 1 to 5 denoting the extent to which they are fee paying or not, in accordance with the National Norms and Standards for School Funding of 1999 (NNSSF) (Department of Basic Education, 2017). Quintiles 4 and 5 equate to the former Model C/private schools that were traditionally reserved for white children, and Quintiles 1 and 2 equate to the former DET schools that were poorly resourced and reserved for the rest of the population. The participants in this study came from schools that were all in the Quintile 4 and 5 fee-paying categories, and are multi-racial. It has not been possible to establish extensive use of this programme within the non-fee-paying, poorly resourced Quintile 1 and 2 schools, traditionally reserved for the impoverished black, coloured and Indian/Asian populations in the Apartheid era, nor even in the better resourced government funded Quintile 4 and 5 fee-paying schools. The funding for this service has simply not been made available, and the facility has not been seen as a priority in these schools and incentivised. Therefore, it is in the sector of South African English medium private schools only where there exists longitudinal data on adolescent high school populations over a five-year period. Some schools in the country chose to use a different programme, viz., the Cogstate Sport programme. Even amongst the private schools in the country generally, the extensive use of computerized testing to facilitate concussion management has not been supported. Consequently, there were eight school populations using the ImPACT test on a regular basis over a five year period that were available for the purposes of the research, five of which were male-only schools, and three co-educational (i.e., the term used in South Africa for schools including both male and female learners).

A non-probability purposive sampling strategy was used to carefully select participants that were typical of the population being studied (adolescents) and met the longitudinal criteria established by the research question (i.e. baseline tested at approximately one-year intervals for five consecutive years from Grade 8 to 12) (Terre Blanche, Durrheim, & Painter, 2006).

Archival data of high school athletes from eight targeted South African schools were analysed. Of these eight schools, three schools had mandatory baseline testing as part of their schools' concussion management programme, whilst the remaining five schools had voluntary baseline testing. From the period 2007 to 2015, 5295 individuals (5178 male; 117 female) from the targeted schools were tested on the ImPACT test on at least one occasion. The following exclusion process occurred:

1. From the initial sample of 5295 (5178 male; 117 female), 2214 individuals first tested in 2012, 2013, 2014, and 2015 were immediately excluded as they began testing too late to accumulate five annual consecutive baseline tests by 2015 (when formatting of the data began), leaving 3081 individuals (2047 male; 34 female).

2. From the sample of 3081 (3047 male; 34 female), a further 2860 individuals were then excluded as they had at least one of the following criteria:
   a) Less than five consecutive annual baseline tests recorded, despite having begun testing before 2012.
   b) More than five consecutive annual baseline tests recorded, resulting from beginning testing in primary school (1 school had both a primary and a high school).
   c) More than one baseline test undertaken in any one year.
   d) Any post-injury tests recorded during the testing period.
   e) Any invalid baseline tests recorded as determined by ImPACT software (“Baseline ++”).

   What remained was a sample of 221 individuals (221 male; 0 female) who had baseline tests at approximately one-year intervals for five consecutive years from Grade 8 to 12 starting from either 2007, 2008, 2009, 2010 or 2011. Unfortunately, all the 117 females in the initial sample who attended any of the three co-education schools in the targeted sample were excluded as they qualified for one or more of the exclusion criteria. It is important to note that having a male-only sample was determined solely by the constraints of the data available and was not a predetermined characteristic of the sample.

3. From the sample of 221 (221 male; 0 female), a further 32 individuals were excluded (leaving 189 individuals, 189 male; 0 female), if they had Version 1 of the ImPACT test at any one baseline instead of Version 2/Version 2.1 consistently throughout the five-year testing period. Version 1 and 2 differed in some content and the manner in which some composite scores were conceptualised. However, the difference between Version 2 and Version 2.1 relates purely to internal database changes (i.e. security updates), and there are no differences in the test content.

4. From the sample of 189 individuals (189 male; 0 female), a further 2 individuals were excluded (leaving 187 individuals, 187 male; 0 female) if they completed one or more of the five baseline tests in an exam language other than English.

5. From the sample of 187 individuals (187 male; 0 female), a further 7 individuals were excluded (leaving 180 individuals, 180 male; 0 female) if they began testing in Grade 8 at the age of 12 instead of 13 years of age. This was done to ensure the sample contained very narrow age ranges in each grade group (i.e., at each test interval) [i.e., 13-14 years old in Grade 8 (Interval 1); 14-15 years old in Grade 9 (Interval 2)], in order to provide more specific normative data for the adolescent years.

6. From the sample of 180 individuals (180 male; 0 female), a further 52 individuals were excluded (leaving 128 individuals, 128 male; 0 female) if they self-reported any of the following exclusion criteria at any one test interval during the five-year testing period. These criteria are known to adversely affect test performance on neurocognitive assessment (McCrory et al., 2013). The criteria were monitored on the demographics section of the ImPACT programme, which is a standard part of the test and involved no additional testing. It must be noted that several individuals self-reported more than one of the following exclusion criteria.
    a) Diagnosis of learning disability (incl. dyslexia)
    b) Diagnosis of ADHD/ADD
    c) Treatment for neurological disorder (i.e. seizure and/or epilepsy, previous history of brain surgery)
    d) Treatment for substance/alcohol abuse
    e) Treatment for psychiatric/psychological disorder (i.e. depression and/or anxiety)
    f) Presence of ADD/ADHD-related/psychiatric/neurological medication (e.g. Ritalin, Concerta, Strattera, anti-depressants, anti-psychotics, epileptic/seizure medication)
    g) Concussive injury in the six months prior to Interval 1 testing
    h) History of three or more concussions prior to Interval 1 testing

7) From the remaining sample of 128 individuals (128 male; 0 female), a further 20 individuals were excluded (leaving 108 individuals, 108 male; 0 female) if they had outlying scores of three or more standard deviations from the sample mean on one or more composite scores at any of the five test intervals. This was done to ensure test administration validity and was in keeping with previous research on the ImPACT test (Bruce et al., 2014; Echemendia et al., 2012; Henry & Sandel, 2015).

2.2.2 Final Participants

After the 5187 exclusions (5070 male; 117 female) outlined above were removed, the final sample consisted of 108 non-clinical, male, predominantly South African adolescent high school athletes from Grades 8 to 12, in the overall age range 13 to 18 years, divided into five groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30). These participants were baseline tested on the ImPACT test at approximately one-year test intervals for five consecutive years, starting from either 2007, 2008, 2009, 2010 or 2011, and attended South African educationally advantaged English medium private schools. This sample is assumed to be non-clinical as participants complied with all the clinical exclusion criteria listed in point 6 above. As all the schools targeted in the study subscribe to the ImPACT programme as part of their sports concussion management programme, all participants were athletes and playing sport in each year of testing, by virtue of being baseline tested on the programme. At the targeted schools where baseline testing was mandatory (i.e. three out of the eight targeted schools), sports participation is likely also mandatory for every learner.

Table 2.1

Exclusion Criteria of Individuals with Baseline Tests at approximately One-year Intervals for Five Consecutive Years from Grades 8 to 12

Exclusion Criteria			(*N* = 5187)
1.	Began testing too late to accumulate five consecutive annual baseline tests		2214
2.	Less than five consecutive annual baseline tests		
	More than five consecutive annual baseline tests		
	More than one baseline in any one year		
	Any post-injury tests recorded during testing period		
	Any invalid baseline tests ("Baseline ++")		2860
3.	Incorrect software version of test		32
4.	Exam language other than English		2
5.	Began testing at 12 years-old		7
6.	Learning disability (incl. dyslexia)	11	52
	ADHD/ADD	32	
	Neurological disorder	1	
	Substance/alcohol abuse	4	
	Psychiatric/psychological disorder	18	
	Medication	29	
	Concussive injury in the six months prior to Interval 1 testing	2	
	History of three or more concussion prior to Interval 1 testing	3	
7.	Outliers of three of more SD from mean	20	

Note. The figures in the parenthesis do not total 52 as some participant had more than one of these exclusion criteria.

2.2.3 Demographic Data

Demographic information regarding personal history (e.g., age, level of education, sex, first language, native country, and concussions history) and sporting history (sport played, position) was obtained from each participant at each baseline test administration. Each participant self-reported the information on the demographic profile (Section A) of the ImPACT test. There are instances where the demographic information provided by a participant differs from one test to the next. This is especially prevalent with regards to sport played, as a participant may for example have played a certain sport during the first three years of testing, and a different sport in the remaining two years. In such instances, the information that appeared most frequently was recorded for the purposes of this study. In the event where the researcher was unable to determine the most frequent information (i.e. an individual reporting a different sport on all five baseline tests or the same sport at Interval 1 and 2, a different sport at both Interval 3 and 4 and a third sport at Interval 5), such participants were recorded as ‘undetermined’ and can be viewed in Table 2.3 below.

2.2.3.1 Age and Education

Age of participants at each test interval was calculated based on the date of birth and date of testing. There was very little age difference between participants at each test interval, with the majority of participants being the same age, or within a year of each other. The age range of participants at each test interval (Grade) overlaps with the age range at the next test interval (Grade) by one year. This in the norm in the South African Education system, where children in all grades are either the same age or within a year of each other, having begun schooling at either 7-years-old or 8-years-old depending on the time of year they were born. The mean age, standard deviation, and age range of the participants at each test interval are documented in years in Table 2.2. below.

Quality and level of education were broadly controlled in this study by virtue of the sample being derived from an exclusively private school population. All participants attended South African educationally advantaged English medium private schools, thus there were no differences within the sample in terms of *quality* of education. The *level* of education at each of the five test intervals was the same for all participants and was recorded as the grade currently being completed at testing.

Table 2.2

Age and Education Level at Each Test Interval in years

Test Interval	1	2	3	4	5
Age					
Mean	13.69	14.62	15.58	16.62	17.62
SD	0.29	0.32	0.31	0.31	0.30
Range	13-14	14-15	15-16	16-17	17-18
Education Level					
Grade currently completing	8	9	10	11	12

2.2.3.2 Language and Race

The majority of participants (86.11%; *N*= 93) were white reporting English as their first language. 1.85% (*N*= 2) were white reporting Afrikaans as their first language. 12.04% (*N*= 13) were black reporting English as their first language[2]. Unfortunately, the ImPACT test does not have a category to identify race, which is a limitation of the demographic section of the test. In this study race was approximated based on the surnames of participants, which is a limitation of analysing retrospective data where some information is not available. Regardless of racial origins, individuals of apparent African indigenous origin based on their surnames all considered English to be their first language, most likely a result of being schooled in English. South African research by Shuttleworth-Edwards et al. (2004b) found that differences in neurocognitive test performance between white English first language and black African first language individuals were minimal when equivalently exposed to a relatively advantaged quality of education. In other words, *quality* of education rather than first language or race accounted for differing neurocognitive test performance. Considering these findings, it is unlikely that the marginal difference in racial composition of the sample would be of any significance as all participants had the same high quality of education, noted by their attendance at educationally advantaged English medium private schools. By virtue of this same attendance, the Afrikaans first language participants were considered to have a high English proficiency, resulting in first language differences in all probability being of no significance.

2.2.3.3 Nationality

The nationality of the sample was predominantly South African (93.52%; *N*= 101), with 0.93% (*N*= 1) self-reported Lesotho as his native country, 0.93% (*N*= 1) self-reported the United Kingdom, and 1.84% (*N*= 2) self-reported Canada. The nationality of 2.78% (*N*= 3) of the sample was recorded as 'undetermined'. As the non-South African participants were educated in the same

---

[2] "During the period between 1948 and 1994, all South Africans were classified according to the Population Registration Act (Act No. 30 of 1950) into four main divisions including 'white', 'black, 'coloured' and 'Asian/Indian'. These terms are thus utilised here on the understanding that they are socially constructed racial signifiers, which have had particular pertinence and continue to be salient in a complex relationship with each other within the South African society, and therefore have salience within the local research context (Van Ommen, 2013)" (Naidoo et al., 2019, p. 65).

schools as the South African participants in the sample and resided in South Africa during the testing period, it is unlikely this marginal difference in nationality would be of any significance.

2.2.3.4 Concussion History

Concussion history was recorded as the number of concussions sustained by each participant prior to Interval 1 testing. Individuals who sustained three or more concussions prior to Interval 1 testing were excluded, whilst individuals who sustained either one or two concussions prior to Interval 1 testing remained included in this study. These individuals remained included as research suggests that the effects of a concussive injury are largely undetectable on neurocognitive measures for athletes with a history of only one or two prior injuries (Macciocchi et al., 2001). Only 13.89% (*N*= 15) self-reported sustaining either one or two concussions (which would have occurred more than six months prior to Interval 1 testing as concussive injuries within six months of Interval 1 testing were already excluded), whilst the majority of the sample (86.11%; *N*= 93) self-reported no concussions prior to Interval 1 testing.

Table 2.3

Distribution of Sample for Sex, Country of Origin, Race, First Language, Sports Participation and Number of Prior Concussions

		N	*%*
Sex	Male	108	100.00
	Female	0	0
Country of Origin	South Africa	108	93.52
	Lesotho	1	0.93
	United Kingdom	1	0.93
	Canada	2	1.84
	Undetermined	3	2.78
Race	White	95	87.96
	Black	13	12.04
	Other	0	0
First Language	English	106	98.15
	Afrikaans	2	1.85
Primary Sport			
Contact	Rugby	78	72.22
Non-Contact	Field Hockey	24	22.22
	Waterpolo	1	0.93
	Basketball	1	0.93
	Undetermined	4	3.70
Number of concussions prior to Interval 1 testing	0	93	86.11
	1 or 2	15	13.89

2.2.3.5 Sport Participation

Participants played contact (rugby, soccer) and/or non-contact (field hockey, cricket, waterpolo, swimming, tennis, squash, basketball, baseball, and athletics) sports. The majority of the sample (53.70%; *N*= 58) played only contact sport each year of the five-year testing period. 19.45% (*N*= 21) played only non-contact sport, and the remaining 26.85% (*N*= 29) played a combination of both contact and non-contact sport across the five-year testing period. These details are summarised in Table 2.4. The sport that was reported most frequently on each participant's demographic profile across the five-year testing period was considered his ‘primary sport’. The majority of participants (72.22%; *N*= 78) played rugby as their primary sport across the five-year period, followed by field hockey (22.22%; *N*= 24) (see Table 2.3).

Table 2.4

Sports Participation Details of Participants across the Five-year Testing Period

	N		%
Contact Pure (played contact sport ONLY)	58	108	53.70
Non-contact Pure (played non-contact sport ONLY)	21		19.45
Contact/Non-contact Mixed (played BOTH contact and non-contact sports)	29		26.85
Rugby Pure (played Rugby ONLY)	58		53.70
Rugby Mixed (played Rugby on at least one baseline AND other sports)	29		26.85
Soccer Pure (played Soccer ONLY)	0		0.00
Soccer Mixed (played Soccer on at least one baseline AND other sports)	3		2.78
Field Hockey Pure (played Field Hockey ONLY)	14		12.96
Field Hockey Mixed (played Field Hockey on at least one baseline AND other sports)	23		21.30

In summary, the majority of participants sustained no previous concussion, played rugby as their primary sport, were white, and spoke English as their first language.

2.3 Measure

2.3.1 Immediate Post-Concussion Assessment and Cognitive Test (ImPACT)

As discussed previously in section 1.8, ImPACT is a standardised computer administrated neurocognitive test battery used to assess neurocognitive function and concussion symptoms (Elbin et al., 2011; Lovell, 2006). The test evaluates multiple aspects of cognitive functioning and is relatively brief, taking 20-25 minutes to administer. It is automatically scored and produces a computer generated clinical report. This report provides composite scores, subtest scores and additional subcomponent scores of the individual’s test performance, graphical representations of each composite score, as well as the correct percentage for the individual’s performance within the associated age-and gender-referenced percentile scores on the four neurocognitive composite scores (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time). Percentile scores for the Impulse Control composite, which acts as a validity indicator, were not provided as this composite has not been normed by ImPACT (ImPACT Applications, 2007; Van Kampen et al., 2006). The ImPACT online Version 2.0 was used for this study and consists of three parts: (i) demographics, background and concussion history questionnaire (ii) symptom inventory and (iii) neurocognitive test battery (which was the focus of this study).

Demographic/history questionnaire

This section of the ImPACT requires the individual to input basic demographic and descriptive information through a series of easy to follow instructional screens. The individual answers questions regarding height, weight, sport played, sporting position, prior concussion history, and history of learning disabilities as well as other important descriptive information (ImPACT Applications, 2004).

Symptom scale

This section consists of a scale that evaluates 22 common post-concussive symptoms (i.e., headache, nausea, dizziness, sensitivity to light, trouble sleeping), which individual are required to rate on a Likert scale from 0 (asymptomatic) to 6 (symptomatic) according to his/her condition at the time of testing. A total symptom score is produced by adding the individual items. This scale is administered at both baseline and post-injury testing sessions in order to document and compare pre-existing and post-injury symptoms (ImPACT Applications, 2004). The outcome on the Symptom Scale was not included for analysis. This was with a view to keeping the research focused on the neurocognitive aspect.

Neurocognitive test battery

The neurocognitive test battery of the ImPACT test evaluates multiple cognitive domains, and consisits of six test modules (i.e., Word Memory, Design Memory, X’s and O’s, Symbol Match, Colour Match, Three Letters Memory). From the six test modules, twelve subtest scores are automatically produced. These twelve subtest scores then contribute to the five composite scores (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control, which acts as a validity indicator) (ImPACT Applications, 2011). The five composite scores, twelve subtest scores and four additional subcomponent scores of the two memory subtests (i.e. Word Memory immediate recall, Word Memory delayed recall, Design Memory immediate recall, and Design Memory delayed recall) were used for separate analysis in this study, and are displayed in Table 2.5 below.

Table 2.5

ImPACT Composite Scores, Subtest Scores, and Additional Memory Subcomponent Scores

---

Five Composite Scores

1. Verbal Memory
2. Visual Memory
3. Visual Motor Speed
4. Reaction Time
5. Impulse Control

Twelve Subtest Scores

1. Word Memory: Total percent correct (immediate + delayed)
2. Symbol Match: Total correct (hidden)
3. Three Letters: Percent of total letters correct
4. Design Memory: Total percent correct (immediate + delayed)
5. X's and O's: Total correct (memory)
6. X's and O's: Total correct (interference)
7. Three Letters: Average counted correctly
8. X's and O's: Average correct reaction time (interference)
9. Symbol Match: Average correct reaction time (visible)
10. Colour Match: Average correct reaction time
11. X's and O's: Total incorrect (interference)
12. Colour Match: Total commissions

Four Subcomponent Scores of Memory Subtests

1. Word Memory: Learning percent correct (immediate recall)
2. Word Memory: Delayed memory percent correct (delayed recall)
3. Design Memory: Learning percent correct (immediate recall)
4. Design Memory: Delayed memory percent correct (delayed recall)

2.4 Procedure

Ethical approval to use archival data for secondary data analysis was obtained by the Rhodes University Research Projects and Ethics Review Committee (RPERC) (PSY2014/17) (Appendix E). ImPACT Applications-SA manages the sports concussion programme for several South

African schools. The programme involves baseline testing each year and post-injury assessment and management of a sports concussion should it occur. Parents of all participants at the targeted schools that implemented mandatory or voluntary ImPACT testing gave informed consent for their children to undergo testing. At the three schools that implemented mandatory testing, parents received a letter informing them of the ImPACT programme (Appendix B) as well as a consent form granting permission for their child to participate in the programme (Appendix C). At the remaining five schools where testing was voluntary, parents received a letter informing them of the ImPACT programme (Appendix D). If they wished to enrol their child on the programme they were required to notify the school via email and agree to be charged the programme fee thereby providing informed consent.

Whilst participants took part in ImPACT testing solely for the benefits of the programme, permission was granted by the schools to use the data for secondary analysis in the research context. In the contracts between ImPACT Applications-SA and the schools on the programme, consent was given for the data to be used in research studies conducted by the National Sports Concussion Initiative, Psychology Department, Rhodes University, provided that the data be de-identified in order to guarantee the anonymity of the research participants (Appendix A). This contract was signed annually by the Headmaster of each school, the medical doctor taking medical responsibility for concussed players, the accredited ImPACT psychologist taking clinical responsibility for the interpretation of the data, as well as the psychologist/technician taking responsibility for conducting baseline testing.

For this study, pre-existing archival data from South African schools dating from 2007 to 2015 was retrieved from the ImPACT online database. Access to this database was monitored by means of strict security protocol, which the researcher had access to. The data for each of the final participants included in the study was de-identified and manually captured on an Excel spreadsheet. This blinded data was then checked by a research assistant in order to minimise systematic/researcher error. The Excel spreadsheet was password protected and only the research and her two supervisors had access to the spreadsheet.

Participants completed the online Version 2.0 of the ImPACT test individually in the standardised automated manner on a desktop computer at all five test administrations. The exact same form of the ImPACT test (Baseline Form 1) was administered at the five different test intervals. The time between Test Intervals 1 to 2, 2 to 3, 3 to 4 and 4 to 5 (recorded in months) are presented in the following table.

Table 2.6

Time between Test Intervals in months

Time between Test Intervals	(1 to 2)	(2 to 3)	(3 to 4)	(4 to 5)
Mean	11.08	11.51	12.57	11.92
SD	1.06	0.97	1.39	1.45
Range	9-14	9-14	10-17	7-15

All testing of participants by ImPACT Applications-SA was done in accordance with the ImPACT guidelines for testing (ImPACT Applications, 2004). The ImPACT test was administered in the high school computer laboratories in groups of no more than 25 participants. Test administration was supervised by a trained ImPACT administrator or a board-registered psychologist. A keyboard and external mouse were always used to help avoid variability in some aspects of the data profile (ImPACT Applications, 2004). The desktop computers used did not have any additional applications running during the assessment so as to not influence the timing accuracy of the programme.

Before administering the test, the administrator would introduce the purpose of the assessment to participants, describe the importance of reporting concussions, emphasise the confidentiality of all test results, answer questions, and provide instruction if needed. Administration is however relatively self-explanatory as participants are able to read the on-screen instructions for each test module. The administrator ensured that the test environment was relatively free of noise and distraction. Although the same computer laboratories at each school may not have been used for all of five test intervals, nor was the same laboratory used across schools, the environment is relatively generic, and the testing conditions were likely to be highly similar at all of the five test intervals at each school and across schools. This was monitored by the trained neuropsychologists overseeing the concussion management programme in all schools from ImPACT Applications-SA.

The administrator also served to discourage horseplay amongst participants during testing by properly supervising the participants and not placing them too close together in the computer laboratories. Horseplay has been identified as a common cause of test invalidity at baseline testing, along with failure to read directions properly due to a reading disability or carelessness; presence of ADHD or hyperactivity; excessive fatigue; left-right confusion (usually the result of the reversal of left and right on the X's and O's distractor task); 'sandbagging' or intentional poor performance in order to set a low baseline standard to potentially be considered normal at post-injury testing when actually concussed (ImPACT Applications, 2007). The ImPACT test has a built-in validity indicator, designed to determine whether an examinee provided a good effort leading to a valid baseline. A baseline test would be flagged as invalid with the term "Baseline ++" (ImPACT Applications, 2007). Furthermore, a score of three or more standard deviations from the mean on any composite at any test interval was an indication of apparent performance. These scores were identified as outliers and participants with these scores were removed. Considering participants with "Baseline ++" scores as well as outlier scores were removed from the sample, it is likely that all final participants displayed normal effort, and adequately represent a non-clinical 'normal' sample.

As the data used is pre-existing and archival, the researcher did not personally collect the data. However, to familiarise herself with the content and administration of the ImPACT test, she worked as an assistant in the administration of both baseline and post-injury ImPACT tests at a local private school contracted into a concussion management programme. These administrations post-dated the data used for the present study.

2.5 Data Analysis

2.5.1 Objective 1: Repeat Baseline Testing over Five Annual Test Intervals

All data were initially subjected to descriptive analyses on the basis of which the means, standard deviations, and range values of each composite score at each test interval were determined. Five separate one-way repeated-measures ANOVAs examined differences in performance across the five test intervals for each of the five composites of the ImPACT test. The five dependent variables were the composite scores on the ImPACT test (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, Impulse Control) and the independent variable was time (Interval 1, Interval 2, Interval 3, Interval 4, Interval 5). Greenhouse-Geisser corrections were implemented when violations of sphericity were observed. A Bonferroni adjustment was made for multiple pairwise comparison tests used during *post hoc* analyses. To characterise the magnitude of cognitive change, data were also presented in standardised effect sizes using Cohen's *d*, which were calculated as $d = (M_1 - M_2) / SD_{pooled}$ (the differences between two groups' means divided by the average of their standard deviation). $d < .20$ denotes a trivial or "no effect" size (Bartels et al., 2010), $d = .20\text{-}.49$ denotes a small effect size, $d = .50\text{-}.79$ a moderate effect size, and $d \geq .80$ a large effect size (Cohen, 1988). All analyses were performed with IBM SPSS (Version 25), with statistical significance set at alpha ≤ .05.

The five composite scores *only* were analysed for Objective 1, and not the subtest scores that make up these composites, or the subcomponent scores of the two memory subtests. As discussed in section 1.11 previously, the aim of Objective 1 was to provide an indication of the optimal interval for repeat baseline testing of non-clinical predominantly South Africa high school athletes. It was considered that including the subtests and subcomponents of the two memory subtests, in addition to the main composites themselves, would be too cumbersome and not necessary to achieve the aim of Objective 1. In routine concussion management it is the composite scores that form the core data for assessment purposes, with subtest data analysis being a refinement for more in depth investigation only, used "especially in those cases that present unique challenges to basic test interpretation" (Henry & Sandel, 2015, p. 272). Accordingly it was considered that the main composite scores were the relevant variables required to evaluate the optimal frequency and timing of repeat baseline testing for high school athletes.

2.5.1.1 Rationale for Chosen Statistical Analyses

In this research multiple repeated-measures analysis of variance (ANOVAs) as opposed to one multivariate analysis of variance (MANOVA) were used as the former suited the nature of the research and the research question, which was univariate (as it pertained to individual outcome variables i.e., individual composite scores) and not multivariate. According to Huberty and Morris (1989), the respective multivariate and univariate methods address different research questions, and the "choice to conduct a strictly multivariate analysis or *multiple* univariate analyses is based on the purpose or purposes of the research effort" (p. 303). As the current research was not concerned with identifying or interpreting any underlying construct/s, nor any intercorrelation or interrelationships between the outcome/dependent variables, a MANOVA (which investigates these issues) was not used. Rather, the use of multiple ANOVAs was considered more appropriate as the outcome/dependent variables in this research were viewed individually, and considered 'conceptually independent'. It is appropriate to use multiple ANOVAs in a situation of ‘conceptually independent’ outcome/dependent variables, as the researcher would be interested in how **each** of the outcome/dependent variables (i.e. each composite score in the current research) is affected by the treatment/independent variable (i.e. ‘time’ across the five test intervals) (Huberty & Morris, 1989). As the current research is concerned with the performance of **each** composite score across the five test intervals, each composite score is considered ‘conceptually independent’.

Carrying out multiple statistical tests on the same experimental data increases the probability of committing a Type I error (i.e. rejecting the null hypothesis when it is true, thus finding significance that is erroneous) (Field, 2009). This may be a particular issue when conducting multiple ANOVAs. The more dependent variables that have been measured, the more ANOVAs would need to be conducted and the greater the chance of committing a Type I error (Field, 2009). Some research in the field of psychology and the behavioural sciences has attempted to control for Type I error probability by conducting a MANOVA as a preliminary to multiple ANOVAs. The rationale typically used, based on the notion of the protected (multivariate) *F* test (Bock, 1975, p. 422), is that if the MANOVA produces significance, then it is acceptable for the researcher to carry out multiple ANOVAs (with data interpretation being based on the results of the ANOVAs). However “the idea that one completely controls for Type I error probability by first conducting an overall MANOVA is open to question (Bird & Hadzi-Pavlovic, 1983; Bray & Maxwell, 1982), because the alpha value for each ANOVA would be less than or equal to the alpha employed for the MANOVA only when the MANOVA null hypothesis is true” (Huberty & Morris, 1989, p. 303). This notion lacks convincing empirical support in a MANOVA-ANOVA context, and according to Huberty and Morris (1989), “is seldom, if ever, appropriate” (p. 302) (please see Huberty & Morris (1989) for more detailed information). Considering this criticism, the current research did not precede multiple ANOVAs with a MANOVA in order to control for Type I error from carrying out multiple statistical tests on the same data.

The decision to use multiple repeated-measures ANOVAs as opposed to one single MANOVA or the MANOVA-ANOVA approach was also in keeping with existing research in the field of neurocognitive testing for sports concussion management. A study by Resch et al. (2018) that assessed the test-retest reliability of the ANAM, CNS Vital Signs and ImPACT neurocognitive concussion tests across three short-term test intervals (day 1, day 45, day 50) in healthy college-age participants, used multiple repeated-measures ANOVAs and *post hoc* multiple pairwise comparison tests (among other statistics). To examine differences in performance over time, each composite score on each neurocognitive test battery was analysed using a separate ANOVA test (e.g. the ImPACT test had five neurocognitive composites so five separate ANOVAs were conducted). This also occurred in a study by Littleton et al. (2015) that assessed the test-retest reliability of the CNS Vital Signs neurocognitive concussion test across three short-term test intervals one week apart, and used a series of repeated measures ANOVAs and *post hoc* tests for each composite score on the CNS Vital Signs battery. Another study by Miller et al. (2007) that assessed the test stability on the ImPACT test in non-concussed collegiate football players over a season (comparing preseason, midseason and postseason neurocognitive scores) also used a series of repeated measures ANOVAs and *post hoc* tests, with each composite score of the ImPACT test being analysed using a separate ANOVA. Commensurate with the statistical method employed for the present study, these prior studies appeared to view each composite score (i.e., each dependent variable) as 'conceptually independent' and thus carried out multiple repeated-measures ANOVAs as opposed to one MANOVA statistic.

2.5.1.2 The Bonferroni's Adjustment

As mentioned previously, Type I error probability increases when testing multiple hypotheses on the same experimental data (i.e., carrying out multiple statistical tests). This is known as the "familywise" or "experimentwise" error rate (Field, 2009). Type I error probability increases when carrying out multiple statistical tests on different dependent variables (Field, 2009) (i.e., multiple separate ANOVA tests), as well as when carrying out multiple statistical tests on the same dependent variable (Statistical Solutions, 2019) (i.e., *post hoc* multiple pairwise comparison tests or *post hoc* t-tests). To protect against Type I error caused by carrying out multiple statistical tests on the same experimental data, a Bonferroni's adjustment can be applied. An overly-stringent Bonferroni's adjustment to control for Type I error could increase the probability of committing a Type II error (i.e. failing to reject the null hypothesis when it is untrue, thus failing to find true significance when it exists) (Brandt, 2007; Hsu, 1996; Perneger, 1998). As such, it is necessary to find the correct balance of statistical power by neither under or over-correcting, as this would result in committing a Type I or Type II error respectively. For the purposes of this study, a decision was taken not to apply the Bonferroni's adjustment to protect against Type I error from carrying out five separate ANOVA tests as the Bonferroni's adjustment to control for Type I error from carrying out multiple pairwise comparison tests during *post hoc* analysis on each ANOVA was already applied (Bonferroni's adjustment applied during SPSS analysis). It was decided that any further

Bonferroni's adjustment to control for Type I error from carrying out multiple ANOVA statistical tests would be an over-correction towards stringency and risk a Type II error.

Whilst Type II errors are inherently no less problematic than Type I errors, in neuropsychological research (such as the current study) a Type II error is deemed to be particularly serious (Brandt, 2007). Over-correcting the Bonferroni's adjustment, especially in neuropsychological research where there is the expectation of relatively subtle findings, could result in an inappropriate loss of statistical sensitivity and a failure to identify clinically significant information when it exists. This could lead to insights of prime clinical relevance being overlooked, which could place individuals at risk. The current tendency to over-use the Bonferroni's adjustment in neuropsychological research has been criticised, with Brandt (2007) going so far as listing such erroneous use of the procedure as a neuropsychological "crime" or "misdemeanour" (p. 533). In the current research, failure to detect clinically significant relationships between test intervals when they exist due to a Bonferroni's over-correction (Type II error) could result in the failure to recommend repeat baseline testing at the necessary intervals or grades, which might result in failing to identify genuine injury related cognitive impairment in concussed adolescents (McCrory et al., 2004).

Further support for not over-correcting the Bonferroni's adjustment towards stringency is that the current neuropsychological research is exploratory in that it is the only study to date that assesses neurocognitive score changes on the same adolescent participants over such an extended length of time (i.e. every year over the five-year high school period). In exploratory neuropsychological research, over-correcting the Bonferroni's adjustment may result in missing clinically relevant indications in the results that serve a heuristic purpose.

Finally, support for this methodological stance is gained from being commensurate with that adopted in prior research that examined differences in neurocognitive score performance over time (i.e., Littleton et al., 2015; Miller et al., 2007; Resch et al., 2018) using the same statistical analysis as the current study (i.e., multiple repeated-measures ANOVAs). These researchers applied the Bonferroni's adjustment for Type I error from conducting *post hoc* multiple pairwise comparison tests on each separate ANOVA (statistical significance set at alpha ≤ .05) and did not further adjust the Bonferroni's adjustment for Type I error from carrying out multiple ANOVA tests.

2.5.1.3 The Assumption of Normality

Data was checked for normality using the Shapiro-Wilk's test of normality, and skewness and kurtosis values converted to z-scores for each composite at each test interval. The results of these normality tests can be found in Appendix F Table 1. Unfortunately, only the Visual Motor Speed composite reached normality on the Shapiro-Wilk's test at all test intervals (except Interval 5) as well as on the z-scores for skewness and kurtosis at all five test intervals. There was a lack of normality on the remaining composite scores.

On the Verbal memory composite, the Shapiro-Wilk's test indicated that scores were not normally distributed at all test intervals. Z-scores for skewness generally supported these findings, with most of the test intervals (Interval 1, 3, 4, and 5) indicating a non-normal distribution. At Interval 2, whilst the z-score for skewness indicated a normal distribution, the z-score for kurtosis indicated a non-normal distribution. On the Visual Memory composite, the Shapiro-Wilk's test indicated that scores were not normally distributed at Intervals 2, 3, 4, and 5. However, at Interval 1 the scores were normally distributed. Z-scores for skewness supported these findings. On the Reaction Time composite, the Shapiro-Wilk's test indicated that scores were not normally distributed at Intervals 2, 3, 4, and 5. However at Interval 1 the scores were normally distributed. Z-scores for skewness generally supported these findings, with scores at Intervals 2, 4, and 5 indicating a non-normal distribution and scores at Interval 1 and 3 indicating a normal distribution. On the Impulse Control composite, the Shapiro-Wilk's test indicated that scores were not normally distributed at all time intervals. Z-scores for skewness supported these findings. The lack of normality on these composite scores at most of the five test intervals is to be expected considering the sample is by nature quite variable, consisting of adolescents who are going through a period of rapid cognitive development (Casey et al., 2000; Luria, 1980; Toledo et al., 2012; Yurgelun-Todd, 2007) and display variance at the individual level (Brett et al., 2016).

Considering the lack of normality on four of the composite scores at most of the five test intervals, non-parametric tests that are distribution-free tests (i.e. make no assumptions of normality) were carried out (Field, 2009), including the non-parametric equivalent of the one-way repeated-measures ANOVA (i.e., the Friedman's ANOVA) and the non-parametric equivalent of *post hoc* paired t-tests (i.e., the Wilcoxon signed-rank tests). For the purposes of this research it was considered that if the results for significance on the non-parametric statistics (where there is no assumption of normality) were the same as the results for significance on the parametric one-way repeated-measures ANOVAs and *post hoc* multiple pairwise comparisons tests, it was acceptable to use the parametric statistics whilst reporting the non-parametric statistics in the Appendix.

Five separate Friedman's ANOVAs were carried out for the five dependent variables of the ImPACT test (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control composite scores). A Bonferroni's adjustment was made for multiple pairwise Wilcoxon signed-rank tests during *post hoc* analyses. The resulting *p*-values were adjusted by multiplying by the number of pairwise comparisons ($n = 10$). All analyses were performed with IBM SPSS (Version 25), with statistical significance set at alpha ≤ .05. There was a near-perfect alignment of results between the non-parametric and parametric tests and hence the use of separate one-way repeated measures ANOVAs and *post hoc* multiple pairwise comparison tests was considered acceptable for application in the current research. Whilst the non-parametric Friedman's ANOVAs produced the same results for significance as the parametric one-way repeated-measures ANOVAs on all composites, *post hoc* analysis revealed one pairwise exception on only one composite score. On the Verbal Memory composite, there was no significant difference between Interval 2 and 4 ($p = .069$) on the non-parametric Wilcoxon signed-rank test but a significant difference between Interval 2 and 4 ($p = .040$) on the parametric *post hoc* pairwise comparison test. These p-values reveal that this particular pairwise comparison is a borderline case that is *just* outside the significance boundary on the non-parametric test and just within on the parametric test, and is not a serious misalignment between the non-parametric and parametric results. Barring this one very minor misalignment, the results for significance between test intervals on the non-parametric and parametric tests were identical on all composite scores. As such, the results of the non-parametric Friedman's ANOVAs and *post hoc* Wilcoxon signed-rank tests were presented in the Appendix (please see Appendix F Tables 2-6). The results of the parametric one-way repeated-measures ANOVAs and *post hoc* multiple pairwise comparisons tests were presented and discussed in the following chapter.

2.5.2 Objective 2: Normative Data for Composite Scores, Subtests Scores, and Additional Memory Subcomponent Scores over Five Annual Test Intervals

Descriptive statistics (means, standard deviations and range values) for the ImPACT tests' five composite scores, twelve subtest scores, and four additional memory subcomponent scores were calculated at each test interval. The five normative participant groups are as follows, stratified by

Grade as the primary variable, and age mean (SD), age range, and number of baseline tests previously completed as qualifiers: Grade 8 [mean age 13.69 (0.29) years, age range 12-13 years, 0 prior baseline tests]; Grade 9 [mean age 14.62 (0.32) years, age range 13-14 years, 1 prior baseline test]; Grade 10 [mean age 15.58 (0.31) years, age range 14-15 years, 2 prior baseline tests]; Grade 11 [mean age 16.62 (0.31) years, age range 16-17 years, 3 prior baseline tests]; Grade 12 [mean age 17.62 (0.30) years, age range 17-18 years, 4 prior baseline tests]. All analysis was conducted using IBM SPSS (Version 25). Unlike Objective 1 where only the five main composite scores were analysed, Objective 2 analysed the five composite scores, in addition to the twelve subtests that make up these composite scores, and four additional memory subcomponent scores (i.e., Word Memory immediate recall, Word Memory delayed recall, Design Memory immediate recall, Design Memory delayed recall), with the aim of generating normative tables for the five participant groups according to the grade currently being completed (Grades 8 to 12) at each test interval.

CHAPTER 3: RESULTS

3.1 Objective 1: Repeat Baseline Testing over Five Annual Test Intervals

In response to the lack of normally distributed data on most of the composite scores at most of the test intervals, the non-parametric equivalent of the one-way repeated-measures ANOVA (i.e., Friedman’s ANOVA) and *post hoc* paired t-tests (i.e., Wilcoxon signed-rank tests) were carried out (Field, 2009). In addition to the non-parametric tests, separate parametric repeated-measures ANOVAs and *post hoc* multiple pairwise comparison tests were carried out for the five independent variables of the ImPACT test (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time and Impulse Control composite scores). As the non-parametric Friedman’s ANOVAs and *post hoc* Wilcoxon signed-rank tests (where there is no assumption of normality) produced the same results of significance as the parametric repeated-measures ANOVAs and *post hoc* pairwise comparisons, it was acceptable to use the parametric statistics whilst reporting the non-parametric tests in Appendix F.

Table 3.1

Descriptive Statistics for ImPACT Composite Scores at Each Test Interval (Grade, Mean Age and Age Range in years) and Results of One-way Repeated-measures ANOVAs

Test Interval	1 (*N* = 108)	2 (*N* = 108)	3 (*N* = 108)	4 (*N* = 108)	5 (*N* =108)	*F* Value	*P* Value
Grade	8	9	10	11	12		
Age Mean (SD)	13.69 (0.29)	14.62 (0.32)	15.58 (0.31)	16.62 (0.31)	17.62 (0.30)		
Age Range	13-14	14-15	15-16	16-17	17-18		
	Mean (SD) *Range (min; max)*	Mean (SD) *Range (min; max)*	Mean (SD) *Range (min; max)*	Mean (SD) *Range (min; max)*	Mean (SD) *Range (min; max)*		
Verbal Memory	84.73 (8.93) *38 (62; 100)*	84.39 (10.36) *36 (64; 100)*	87.78 (9.74)$^{b\ e}$ *38 (62; 100)*	87.24 (9.87) *40 (60; 100)*	88.50 (9.23)$^{d\ g}$ *36 (64; 100)*	*F*(3.505, 375.046) = 6. 378	< . 0001
Visual Memory	74.84 (12.20) *53 (47; 100)*	80.68 (11.56)^{a} *52 (48; 100)*	82.11 (11.05)^{b} *51 (49; 100)*	84.21 (10.94)$^{c\ f}$ *51 (49; 100)*	84.56 (9.09)$^{d\ g}$ *39 (61; 100)*	*F*(3.625, 387.898) = 22.635	< . 0001
Visual Motor Speed	33.12 (4.98) *21.78 (23.30; 44.08)*	36.16 (5.93)^{a} *30.97 (19.93; 50.90)*	39.73 (6.01)$^{b\ e}$ *26.70 (26.00; 52.70)*	40.73 (5.50)$^{c\ f}$ *26.05 (26.78: 52.83)*	42.46 (5.99)$^{d\ g\ i\ j}$ *24.88 (28.25; 53.13)*	*F*(4, 428) = 153.842	< . 0001
Reaction Time	0.62 (0.06) *0.34 (0.46; 0.80)*	0.58 (0.07)^{a} *0.30 (0.46; 0.76)*	0.57 (0.06)^{b} *0.29 (0.45; 0.74)*	0.58 (0.08)^{c} *0.38 (0.44; 0.82)*	0.59 (0.08)^{d} *0.38 (0.42; 0.80)*	*F*(3.377, 361.315) = 12.906	< . 0001
Impulse Control	7.31 (4.18) *23 (0; 23)*	7.32 (3.74) *19 (1; 20)*	7.07 (3.76) *20 (0; 20)*	6.69 (3.94) *17 (1; 18)*	6.71 (4.38) *20 (0; 20)*	*F*(4, 428) = 0.929	.447

Note. [a] Significant difference between Interval 1 and Interval 2 ($p \leq .05$).
[b] Significant difference between Interval 1 and Interval 3 ($p \leq .05$).
[c] Significant difference between Interval 1 and Interval 4 ($p \leq .05$).
[d] Significant difference between Interval 1 and Interval 5 ($p \leq .05$).
[e] Significant difference between Interval 2 and Interval 3 ($p \leq .05$).
[f] Significant difference between Interval 2 and Interval 4 ($p \leq .05$).
[g] Significant difference between Interval 2 and Interval 5 ($p \leq .05$).
[h] Significant difference between Interval 3 and Interval 4 ($p \leq .05$).
[i] Significant difference between Interval 3 and Interval 5 ($p \leq .05$).
[j] Significant difference between Interval 4 and Interval 5 ($p \leq .05$).

3.1.1 Verbal Memory

A one-way repeated-measures ANOVA with a Greenhouse-Geisser correction determined that mean Verbal Memory results differed significantly between test intervals (i.e., over time), [$F(3.505, 375.046) = 6.278$, $p < .001$]. *Post hoc* pairwise comparisons using the Bonferroni's adjustment revealed significant differences in mean values for Verbal Memory, all in the direction of improved performance: (i) between Interval 1 and Interval 3 (84.73 ± 8.93 vs 87.78 ± 9.74, $\bar{x}_d = 3.05$, $p = .024$); (ii) between Interval 1 and Interval 5 (84.73 ± 8.93 vs 88.50 ± 9.23, $\bar{x}_d = 3.77$, $p = .003$); (iii) between Interval 2 and Interval 3 (84.39 ± 10.36 vs 87.78 ± 9.74, $\bar{x}_d = 3.39$, $p = .013$); (iv) and between Interval 2 and Interval 5 (84.39 ± 10.36 vs 88.50 ± 9.23, $\bar{x}_d = 4.11$, $p = .001$). A comprehensive breakdown of *post hoc* pairwise comparison results for Verbal Memory is provided in Appendix G Table 7.

Effect sizes (Cohen's *d*) for the differences over time that reached significance are as follows. There were small effect size improvements in Verbal Memory between Interval 1 and 3 ($d = .33$), Interval 1 and 5 ($d = .42$), Interval 2 and 3 ($d = .34$), and Interval 2 and 5 ($d = .42$). The largest improvement in mean Verbal Memory values over time was between Interval 2 and 5.

A higher score for Verbal Memory indicates better performance. The pattern of change on the Verbal Memory composite across the five test intervals was staggered, with there being a very slight non-significant decrease in mean values between Interval 1 and 2 (84.73 ± 8.93 vs 84.39 ± 10.36, $\bar{x}_d = -0.34$), a *significant* albeit small increase between Interval 2 and 3 (84.39 ± 10.36 vs 87.78 ± 9.74, $\bar{x}_d = 3.39$, $p = .024$, $d = .34$), a very slight non-significant decrease between Interval 3 and 4 (87.78 ± 9.74 vs 87.24 ± 9.87, $\bar{x}_d = -0.54$), and a slight non-significant increase between Interval 4 and 5 (87.24 ± 9.87 vs 88.50 ± 9.23, $\bar{x}_d = 1.26$). Overall however, the mean values displayed a predominantly increasing trend, indicating improvement in performance on the Verbal Memory composite over time.

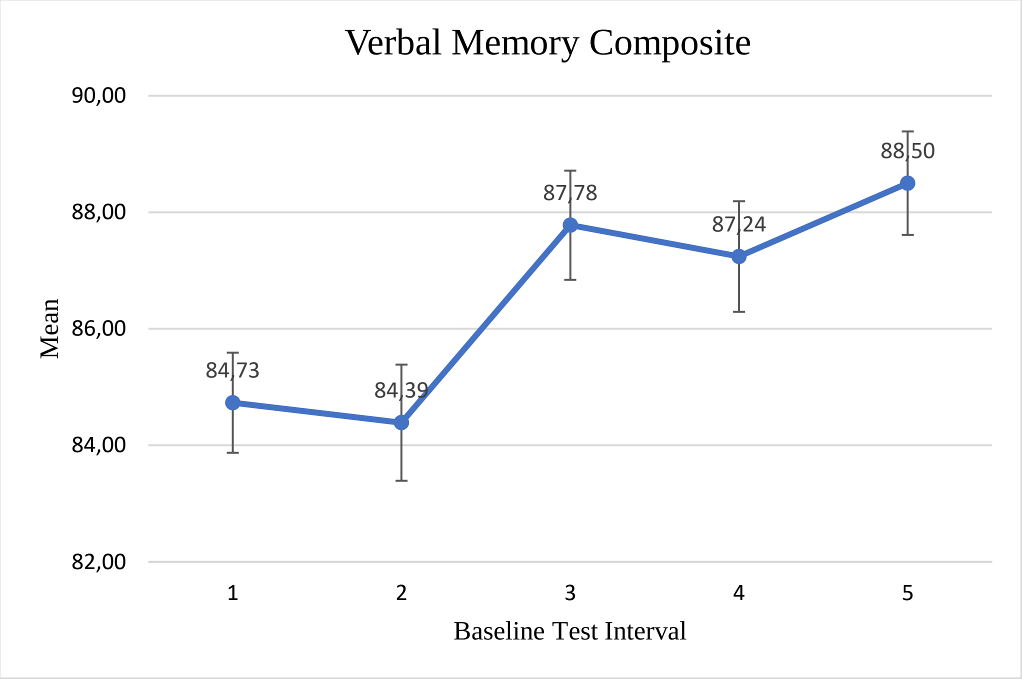

Figure 3.1. Mean baseline Verbal Memory composite score on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A higher score indicates better performance.

3.1.2 Visual Memory

A one-way repeated-measures ANOVA with a Greenhouse-Geisser correction determined that mean Visual Memory results differed significantly between test intervals (i.e., over time), [$F(3.625, 387.898) = 22.635$, $p < .001$]. *Post hoc* pairwise comparisons using the Bonferroni's adjustment revealed significant differences in mean values for Visual Memory, all in the direction of improved performance: (i) between Interval 1 and Interval 2 (74.84 ± 12.20 vs 80.68 ± 11.56, $\bar{x}_d = 5.84$, $p < .001$); (ii) between Interval 1 and Interval 3 (74.84 ± 12.20 vs 82.11 ± 11.05, $\bar{x}_d = 7.269$, $p < .001$); (iii) between Interval 1 and Interval 4 (74.84 ± 12.20 vs 84.21 ± 10.94, $\bar{x}_d = 9.37$, $p < .001$); (iv) between Interval 1 and Interval 5 (74.84 ± 12.20 vs 84.56 ± 9.09, $\bar{x}_d = 9.72$, $p = < .001$); (v) between Interval 2 and Interval 4 (80.68 ± 11.56 vs 84.21 ± 10.94, $\bar{x}_d = 3.53$, $p = .040$) and between Interval 2 and Interval 5 (80.68 ± 11.56 vs 84.56 ± 9.09, $\bar{x}_d = 3.88$, $p = .014$). A comprehensive breakdown of *post hoc* pairwise comparison results for Visual Memory is provided in Appendix G Table 8.

Effect sizes (Cohen's d) for the differences over time that reached significance are as follows. There were small effect size improvements in Visual Memory between Interval 1 and 2 ($d = .49$), Interval 2 and 4 ($d = .31$), and Interval 2 and 5 ($d = .38$). There was a moderate effect size improvement between Interval 1 and 3 ($d = .63$) and large effect size improvements between Interval 1 and 4 ($d = .81$) and Interval 1 and 5 ($d = .91$). The largest improvement in mean Visual Memory values over time was between Interval 1 and 5.

A higher score for Visual Memory indicates better performance. The pattern of change on the Visual Memory composite across the five test intervals showed a year-on-year increase in mean values from Interval 1 to 5, indicating improvement in performance on the Visual Memory composite over time. Whilst there was improvement year-on-year, it was generally at a decelerating rate (i.e., the rate of improvement decreased with increasing grade/age and number of baseline test administrations) (See Figure 3.2). There was a *significant* albeit small increase in mean values between Interval 1 and 2 (74.84 ± 12.20 vs 80.68 ± 11.56, $\bar{x}_d = 5.84$, $p < .001$, $d = .49$), which is during early adolescence. After Interval 2, neurocognitive score changes (all in the direction of improved performance) were small and non-significant between subsequent intervals [i.e., between Interval 2 and 3 (80.68 ± 11.56 vs 82.11 ± 11.05, $\bar{x}_d = 1.43$); between Interval 3 and 4 (82.11 ± 11.05 vs 84.21 ± 10.94, $\bar{x}_d = 2.10$); and between Interval 4 and 5 (84.21 ± 10.94 vs 84.56 ± 9.09, $\bar{x}_d = 0.35$)]. However, when these intervals were combined (i.e., between Interval 2 and 5)

(80.68 ± 11.56 vs 84.56 ± 9.09, $\bar{x}_{d} =$ 3.88, p = .014) there was a *significant* albeit small (d = .38) increase in mean values, indicating improvement in performance. Overall, the mean values depicted an increasing trend throughout the testing period, indicating improvement in performance on the Visual Memory composite over time.

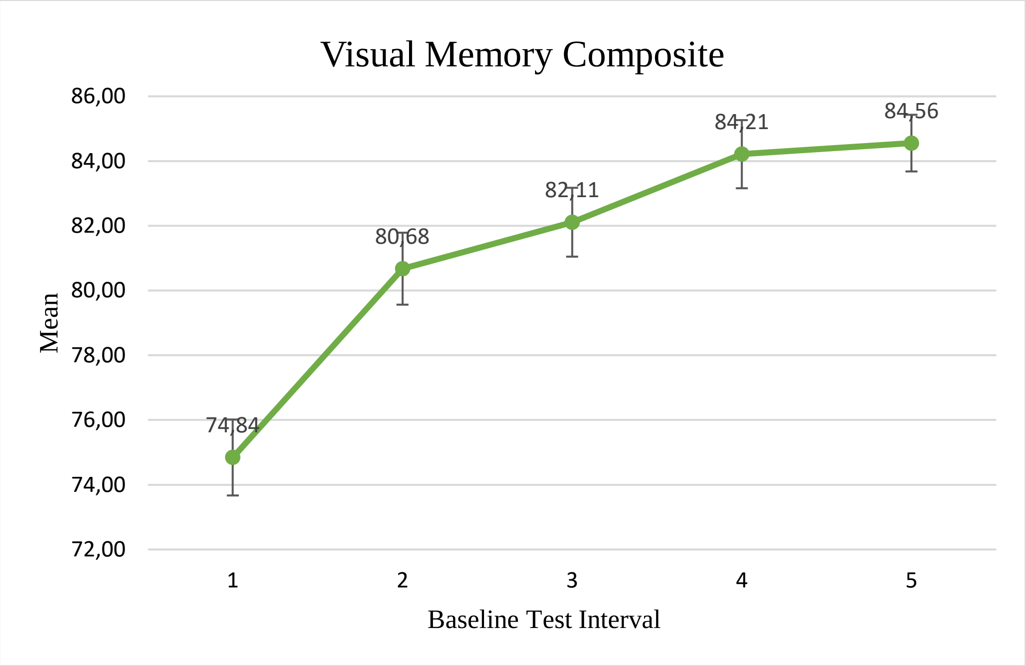

Figure 3.2. Mean baseline Visual Memory composite score on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A higher score indicates better performance.

3.1.3 Visual Motor Speed

A one-way repeated-measures ANOVA determined that mean Visual Motor Speed results differed significantly between test intervals (i.e., over time), [$F(4, 428) = 153.842$, $p < .001$]. *Post hoc* pairwise comparisons using the Bonferroni's adjustment revealed significant differences in mean values for Visual Motor Speed, between all test intervals except between Interval 3 and Interval 4. The estimated change in performance was in the direction of improved performance over time, with significant differences between the following test intervals;

(i) Interval 1 and Interval 2 (33.12 ± 4.98 vs 36.16 ± 5.93, $\bar{x}_d = 3.04$, $p < .001$);
(ii) Interval 1 and Interval 3 (33.12 ± 4.98 vs 39.73 ± 6.01, $\bar{x}_d = 6.61$, $p < .001$);
(iii) Interval 1 and Interval 4 (33.12 ± 4.98 vs 40.73 ± 5.50, $\bar{x}_d = 7.61$, $p < .001$);
(iv) Interval 1 and Interval 5 (33.12 ± 4.98 vs 42.46 ± 5.99, $\bar{x}_d = 9.34$, $p < .001$);
(v) Interval 2 and Interval 3 (36.16 ± 5.93 vs 39.73 ± 6.01, $\bar{x}_d = 3.57$, $p < .001$);
(vi) Interval 2 and Interval 4 (36.16 ± 5.93 vs 40.73 ± 5.50, $\bar{x}_d = 4.57$, $p < .001$);
(vii) Interval 2 and Interval 5 (36.16 ± 5.93 vs 42.46 ± 5.99, $\bar{x}_d = 6.30$, $p < .001$);
(viii) Interval 3 and Interval 5 (39.73 ± 6.01 vs 42.46 ± 5.99, $\bar{x}_d = 2.73$, $p < .001$);
(ix) Interval 4 and Interval 5 (40.73 ± 5.50 vs 42.46 ± 5.99, $\bar{x}_d = 1.73$, $p < .001$).

A comprehensive breakdown of *post hoc* pairwise comparison results for Visual Motor Speed is provided in Appendix G Table 9.

Effect sizes (Cohen's *d*) for the differences over time that reached significance are as follows. There were moderate effect size improvements in Visual Motor Speed between Interval 1 and 2 ($d = .56$) and Interval 2 and 3 ($d = .60$), and a small effect size improvement between Interval 4 and 5 ($d = .30$). There were very large effect size improvements between Interval 1 and 3 ($d = 1.20$), Interval 1 and 4 ($d = 1.45$), and Interval 1 and 5 ($d = 1.70$), and a large effect size improvement between Interval 2 and 4 ($d = .80$). The largest improvement in mean Visual Motor Speed values over time was between Interval 1 and 5.

A higher score for Visual Motor Speed indicates better performance. The pattern of change on the Visual Motor Speed composite across the five test intervals showed a year-on-year increase in mean values from Interval 1 to 5, indicating improvement in performance on this composite over time. There was *a significant* moderate increase in mean values between Interval 1 and 2 (33.12 ± 4.98 vs 36.16 ± 5.93, $\bar{x}_d = 3.04$, $p < .001$, $d = .56$) and between Interval 2 and 3 (36.16 ± 5.93 vs 39.73 ± 6.01, $\bar{x}_d = 3.57$, $p < .001$, $d = .60$), a non-significant slight increase between Interval

3 and 4 (39.73 ± 6.01 vs 40.73 ± 5.50, $\bar{x}_d$ = .1.00), and finally, a *significant* albeit small increase between Interval 4 and 5 (40.73 ± 5.50 vs 42.46 ± 5.99, $\bar{x}_d$ = 1.73, $p < .001$, $d = .30$). Overall, the mean values showed an increasing trend in the direction of improved performance, with significant change occurring throughout the testing period, and throughout adolescence (from early adolescence to late adolescence).

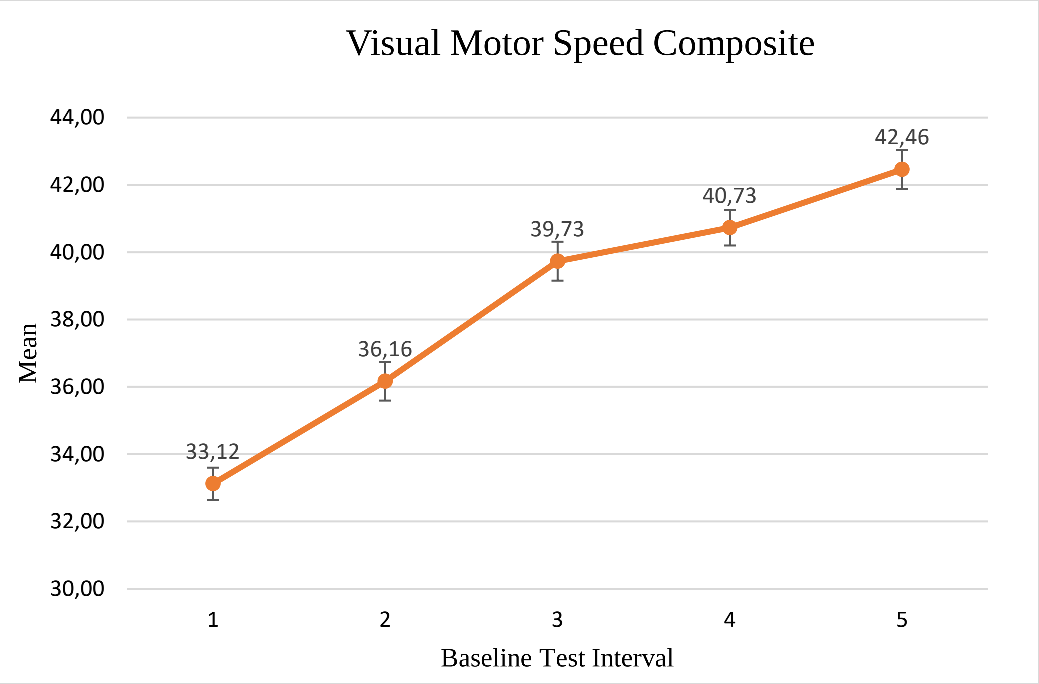

Figure 3.3. Mean baseline Visual Motor Speed composite score on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A higher score indicates better performance.

3.1.4 Reaction Time

A one-way repeated-measures ANOVA with a Greenhouse-Geisser correction determined that mean Reaction Time results differed significantly between test intervals (i.e., over time), [$F(3.377, 361.315) = 12.906$, $p < .001$]. *Post hoc* pairwise comparisons using the Bonferroni's adjustment revealed significant differences in mean values for Reaction Time, all in the direction of improved performance: (i) between Interval 1 and Interval 2 (0.62 ± 0.06 vs 0.58 ± 0.07, $\bar{x}_d = -0.04$, $p = < .001$); (ii) between Interval 1 and Interval 3 (0.62 ± 0.06 vs 0.57 ± 0.06, $\bar{x}_d = -0.05$, $p < .001$); (iii) between Interval 1 and Interval 4 (0.62 ± 0.06 vs 0.58 ± 0.08), $\bar{x}_d = -0.04$, $p < .001$) (iv) and between Interval 1 and Interval 5 (0.62 ± 0.06 vs 0.59 ± 0.08), $\bar{x}_d = -0.03$, $p = .006$). A comprehensive breakdown of *post hoc* pairwise comparison results for Reaction Time is provided in Appendix G Table 10.

As stated previously, a decrease in Reaction Time mean values indicates improved performance, whereas an increase in mean values indicates worsening performance. As such, improved Reaction Time performance results in negative effect sizes (Cohen's *d*), whilst worsening performance results in positive effect sizes. There were small effect size improvement in Reaction Time between Interval 1 and 2 ($d = -.49$), moderate effect size improvements between Interval 1 and 3 ($d = -.72$) and Interval 1 and 4 ($d = -.55$), and a small effect size improvement between Interval 1 and 5 ($d = -.43$). The largest improvement in mean Reaction Time values over time was between Interval 1 and 3.

The pattern of change on the Reaction Time composite across the five test intervals varied, with there being a *significant* decrease in mean values from Interval 1 to 3 indicating improved performance over time, and then a slight increase in mean values from Interval 3 to 5 indicating worsening performance over time. The neurocognitive score performance appeared to peak at Interval 3 (Grade 10). There was a *significant* albeit small improvement in performance between Interval 1 and 2 (0.62 ± 0.06 vs 0.58 ± 0.07, $\bar{x}_d = -0.04$, $p < .001$, $d = -.49$), a non-significant slight improvement in performance between Interval 2 and 3 (0.58 ± 0.07 vs 0.57 ± 0.06, $\bar{x}_d = -0.01$), and a non-significant slight worsening in performance between Interval 3 and 4 (0.57 ± 0.06 vs 0.58 ± 0.08, $\bar{x}_d = 0.01$) and Interval 4 and 5 (0.58 ± 0.08 vs 0.59 ± 0.08, $\bar{x}_d = 0.01$). Whilst there was a trend of improved performance followed by worsening performance, there was still an overall improvement in performance from Interval 1 to 5, indicating general improved performance on the Reaction Time composite over a five year period of time.

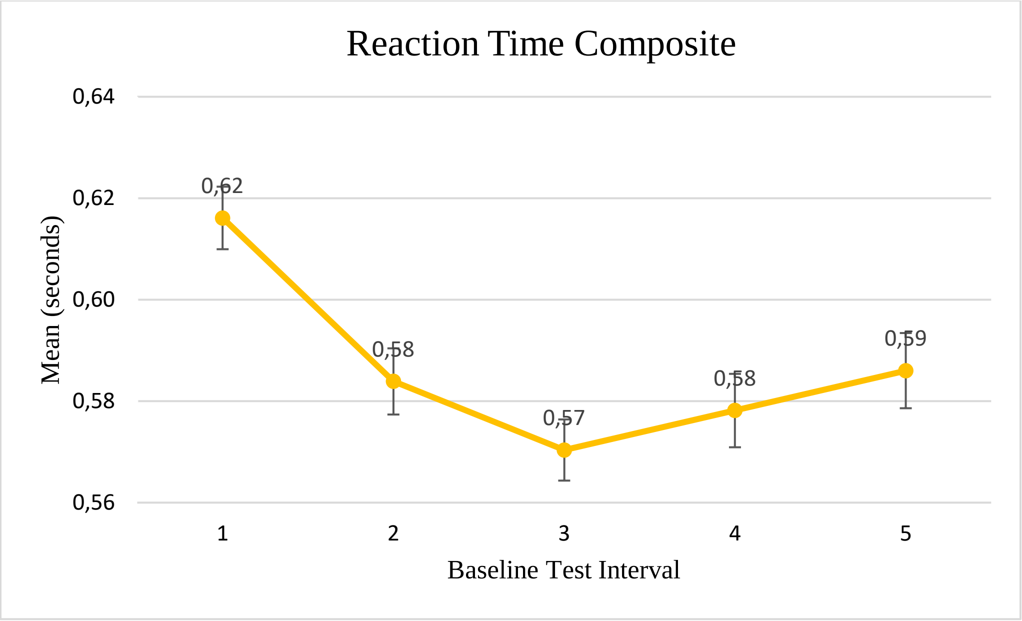

Figure 3.4. Mean baseline Reaction Time composite score on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A lower score indicates better performance.

3.2.5 Impulse Control

A one-way repeated-mesures ANOVA determined that mean Impulse Control results did not differ significantly between test intervals (i.e., over time), [$F(4, 428) = 0.929$, $p = .447$]. In other words, the Impulse Control composite was not significantly affected by time. Because of the non-significant results, *post hoc* tests were not warranted.

A lower score for Impulse Control indicates better performance. There was a slight increase in mean values between Interval 1 and 2 (7.31 ± 4.18 vs 7.32 ± 3.74, $\bar{x}_d = 0.01$) indicating worsening Impulse Control performance, a slight decrease in mean values between Interval 2 and 3 (7.32 ± 3.74 vs 7.07 ± 3.76, $\bar{x}_d = -0.25$) and between Interval 3 and 4 (7.07 ± 3.76 vs 6.69 ± 3.94, $\bar{x}_d = -0.38$) indicating improved performance, and finally a very slight increase in mean values between Interval 4 and 5 (6.69 ± 3.94 vs 6.71 ± 4.38, $\bar{x}_d = 0.02$) indicating worsening performance. Whilst there was fluctuating improved and worsening performance across the five test intervals, there was still an overall improvement in performance from Interval 1 to 5, indicating general improved performance on the Impulse Control composite over a five year period of time.

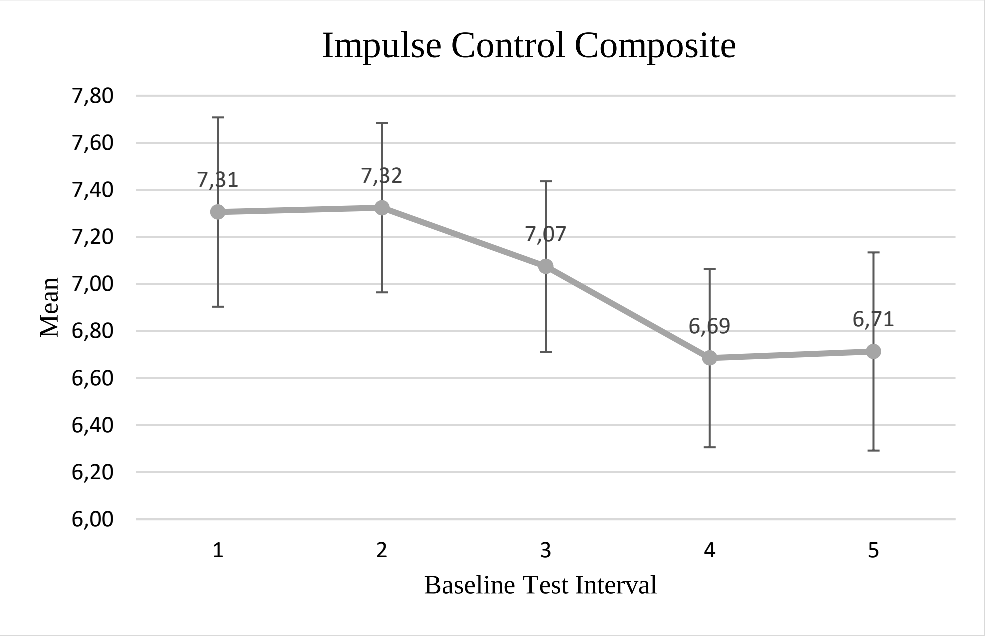

Figure 3.5. Mean baseline Impulse Control composite score on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A lower score indicates better performance.

3.2 Objective 2: Normative Data for Composite Scores, Subtest Scores, and Additional Memory Subcomponent Scores over Five Annual Test Intervals

Descriptive statistics (means, standard deviations and range values) for the ImPACT tests' five composite scores, twelve subtest scores, and four additional memory subcomponent scores at each test interval appears in Table 3.2. This table presents normative data for the five participant groups stratified by Grade as the primary variable, with the following respective mean ages (SD), age ranges, and number of baseline tests previously completed as qualifiers: Grade 8 [mean age 13.69 (0.29) years, age range 12-13 years, 0 prior baseline tests]; Grade 9 [mean age 14.62 (0.32) years, age range 13-14 years, 1 prior baseline test]; Grade 10 [mean age 15.58 (0.31) years, age range 14-15 years, 2 prior baseline tests]; Grade 11 [mean age 16.62 (0.31) years, age range 16-17 years, 3 prior baseline tests]; Grade 12 [mean age 17.62 (0.30) years, age range 17-18 years, 4 prior baseline tests].

Table 3.2

Descriptive Statistics for ImPACT Composite Scores, Subtest Scores, and Additional Memory Subcomponent Scores at Each Test Interval (Grade, Mean Age and Age Range in years)

Test Interval	1 (*N* = 108)			2 (*N* =108)			3 (*N* = 108)			4 (*N* = 108)			5 (*N* = 108)		
Grade	8			9			10			11			12		
Age Mean (SD)	13.69 (0.29)			14.62 (0.32)			15.58 (0.31)			16.62 (0.31)			17.62 (0.30)		
Age Range	13-14			14-15			15-16			16-17			17-18		
No. of Prior Baseline Tests	0			1			2			2			3		
	Mean	(SD)	Range	Mean	(SD)	Range	Mean	(SD)	Range	Mean	(SD)	Range	Mean	(SD)	Range
Composite Scores															
Verbal Memory Composite	84.73	(8.93)	62, 100	84.39	(10.36)	64, 100	87.78	(9.74)	62, 100	87.24	(9.87)	60, 100	88.50	(9.23)	64, 100
Visual Memory Composite	74.84	(12.20)	47, 100	80.68	(11.56)	48, 100	82.11	(11.05)	49, 100	84.21	(10.94)	49, 100	84.56	(9.09)	61, 100
Visual Motor Speed Composite	33.12	(4.98)	23.30, 44.08	36.16	(5.93)	19.93, 50.90	39.73	(6.01)	26.00, 52.70	40.73	(5.50)	26.78, 52.83	42.46	(5.99)	28.25, 53.13
Reaction Time Composite	0.62	(0.06)	0.46, 0.80	0.58	(0.07)	0.46, 0.76	0.57	(0.06)	0.45, 0.74	0.58	(0.08)	0.44, 0.82	0.59	(0.08)	0.42, 0.80
Impulse Control Composite	7.31	(4.18)	0, 23	7.32	(3.74)	1, 20	7.07	(3.76)	0, 20	6.69	(3.94)	1, 18	6.71	(4.38)	0, 20
Subtest Scores															
Word Memory: Total percent correct (immediate + delayed)	95.61	(3.92)	83.5, 100	96.11	(3.70)	83, 100	95.98	(4.57)	75, 100	95.53	(4.15)	79.5, 100	94.75	(5.02)	77, 100
Symbol Match: Total correct (hidden)	6.30	(1.93)	1, 9	6.18	(2.20)	1, 9	6.69	(1.96)	1, 9	6.73	(2.13)	0, 9	7.00	(2.02)	1, 9
Three Letters: Percent of total letters correct	88.64	(11.21)	60, 100	88.33	(12.60)	53.33, 100	93.15	(10.00)	53.33, 100	91.42	(10.35)	60, 100	93.02	(9.70)	60, 100
Design Memory: Total percent correct (immediate + delayed)	82.73	(11.08)	52, 100	88.56	(8.73)	64.5, 100	89.75	(9.58)	58.5, 100	88.64	(9.55)	58.5, 100	89.05	(9.12)	62.5, 100
X's and O's: Total correct (memory)	8.05	(2.32)	2, 12	8.75	(2.37)	2, 12	8.94	(2.13)	2, 12	9.57	(1.90)	3, 12	9.62	(1.70)	4, 12
X's and O's: Total correct (interference)	109.80	(7.01)	76, 123	113.38	(7.07)	91, 130	115.71	(6.29)	94, 130	115.54	(6.57)	94, 130	114.40	(6.55)	97, 127
Three Letters: Average counted correctly	12.93	(3.12)	7.20, 19.80	14.66	(3.74)	4.20, 23.60	16.85	(3.83)	8, 25	17.52	(3.55)	8.60, 24.80	18.77	(3.85)	9, 25
X's and O's: Average correct reaction time (interference)	0.53	(0.06)	0.39, 0.73	0.49	(0.05)	0.37, 0.69	0.47	(0.05)	0.39, 0.63	0.48	(0.05)	0.38, 0.62	0.48	(0.05)	0.38, 0.65
Symbol Match: Average correct reaction time (visible)	1.46	(0.20)	1.08, 2.00	1.45	(0.28)	1.06, 2.93	1.43	(0.28)	0.99, 2.27	1.50	(0.33)	0.96, 2.75	1.57	(0.36)	0.95, 2.90
Colour Match: Average correct reaction time	0.83	(0.13)	0.61, 1.27	0.78	(0.12)	0.54, 1.11	0.76	(0.13)	0.52, 1.19	0.75	(0.15)	0.00, 1.37	0.75	(0.13)	0.52, 1.17
X's and O's: Total incorrect (interference)	6.52	(3.99)	0, 22	6.79	(3.56)	1, 18	6.54	(3.46)	0, 19	6.31	(3.75)	1, 17	6.40	(4.23)	0, 20
Colour Match: Total commissions	0.79	(0.84)	0, 3	0.54	(1.30)	0, 12	0.54	(0.87)	0, 5	0.38	(0.56)	0, 2	0.31	(0.51)	0, 2
Additional Memory Subcomponent Scores															
Word Memory: Learning percent correct (immediate recall)	98.12	(3.02)	88, 100	98.21	(3.06)	83, 100	98.06	(3.74)	75, 100	97.55	(3.62)	83, 100	97.46	(3.86)	75, 100
Word Memory: Delayed memory percent correct (delayed recall)	93.11	(6.52)	75, 100	94.01	(5.67)	75, 100	93.91	(7.20)	71, 100	93.52	(6.23)	67, 100	92.05	(7.58)	67, 100
Design Memory: Learning percent correct (immediate recall)	84.44	(11.61)	50, 100	89.99	(8.75)	67, 100	90.77	(9.10)	63, 100	89.59	(9.51)	63, 100	90.19	(8.68)	63, 100
Design Memory: Delayed memory percent correct (delayed recall)	81.02	(12.82)	50, 100	87.12	(11.20)	46, 100	88.73	(11.33)	50, 100	87.69	(10.91)	50, 100	87.92	(11.25)	54, 100

For the five composite scores, statistical differences in scores across the five test intervals have been established and reported as part of Objective 1 (section 3.1). Objective 2 was to compile a set of descriptive normative data (means, standard deviations and range values) for the ImPACT tests' five composite scores, twelve subtest scores that make up the composite scores, and four additional subcomponent scores of the two memory subtests, for the purposes of clinical application when interpreting an athlete's postconcussion asssessment. A descriptive review follows with broad commentary on the overall trends of the mean values of ImPACT composite scores in comparison to the subtest scores and additional memory subcomponent scores over the five test intervals, without any indication of whether or not there are statistically significant changes between the test intervals:

On the Verbal Memory composite (Table 3.2, first section) there was a staggered pattern of change with slight decreases in mean values between some test intervals (Interval 1 and 2; Interval 3 and 4) and increases between other intervals (Interval 2 and 3; Interval 4 and 5). Overall, the mean values increased from Interval 1 to 5 (84.73 ± 8.93 vs 88.50 ± 9.23), indicating general improvement in performance on the Verbal Memory composite with increased grade/age and number of baseline test administrations. The Verbal Memory composite is made up of three subtests: Word Memory total percent correct (immediate + delayed); Symbol Match total correct (hidden); and Three Letters percent of total letters correct. Two additional subcomponents of the Word Memory total percent correct subtest, the Word Memory immediate recall and the Word Memory delayed recall, were included for their clinical relevance discussed previously in section 1.8.1.

The Word Memory subtest and its two subcomponents (Word Memory immediate recall and Word Memory delayed recall) all followed a similar trend, with mean values staying much the same across the five test intervals. Whilst there was a slight increase in mean values from Interval 1 to 2 followed by slight decreases across the remaining test intervals, there was an overall worsening in performance from Interval 1 to 5 on the Word Memory subtest and its two subcomponents, which were different in trend to the overall improvement pattern of the Verbal Memory composite. In contrast, the mean values on the Symbol Match and Three Letters subtests appear to fluctuate across the five test intervals. Despite fluctuating worsening and improved performance across the test intervals, there was an overall improvement in performance from Interval 1 to 5 on the Symbol Match (6.30 ± 1.93 vs 7.00 ± 2.02) and Three Letters (88.64 ± 11.21 vs 93.02 ± 9.70) subtests, similar to the overall improvement pattern of the Verbal Memory composite. Please see Figure 3.6 below for the pattern of neurocognitive score performance on the Verbal Memory composite, relevant subtests, and additional memory subcomponents over time.

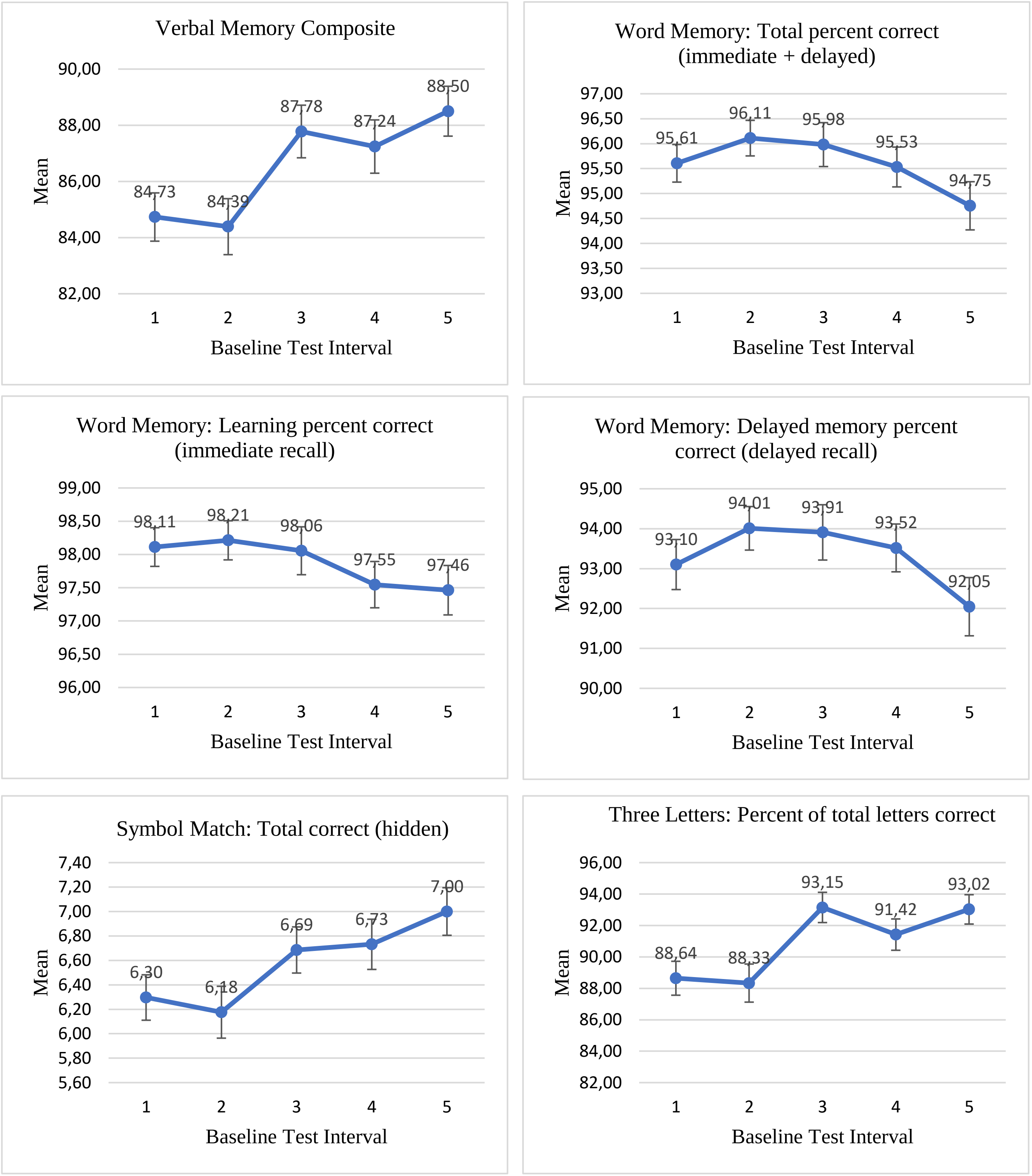

Figure 3.6. Mean baseline Verbal Memory composite score, relevant subtest scores, and additional memory subcomponent scores on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A higher score indicates better performance.

On the Visual Memory composite there was a year-on-year increase in mean values from Interval 1 to 5 (74.84 ± 12.20 vs 84.56 ± 9.09), indicating an increase in performance with increased grade/age and number of baseline test administrations. The Visual Memory composite is made up of two subtests: Design Memory total percent correct (immediate + delayed) and X's and O's total correct (memory). Two additional subcomponents of the Design Memory total percent correct subtest, the Design Memory immediate recall and Design Memory delayed recall, were included for their clinical relevance discussed previously in section 1.8.1.

The Design Memory subtest and its two subcomponents (Design Memory immediate recall and Design Memory delayed recall) all follow a similar trend, with there being an initial large increase in mean values from Interval 1 to 2, followed by stable mean values across the remaining test intervals. The Design Memory subtest and its two subcomponents all showed an overall improvement in performance from Interval 1 to 5. A similar trend was also observed on the X's and O's subtest, which showed an overall increase in mean values from Interval 1 to 5 (8.05 ± 2.32 vs 9.62 ± 1.70). The overall improvement in performance from Interval 1 to 5 on the Design Memory subtest, its two subcomponents, and the X's and O's subtest was similar to the overall improvement pattern of the Visual Memory composite. Please see Figure 3.7 below for the pattern of neurocognitive score performance on the Visual Memory composite, relevant subtests, and additional memory subcomponents over time.

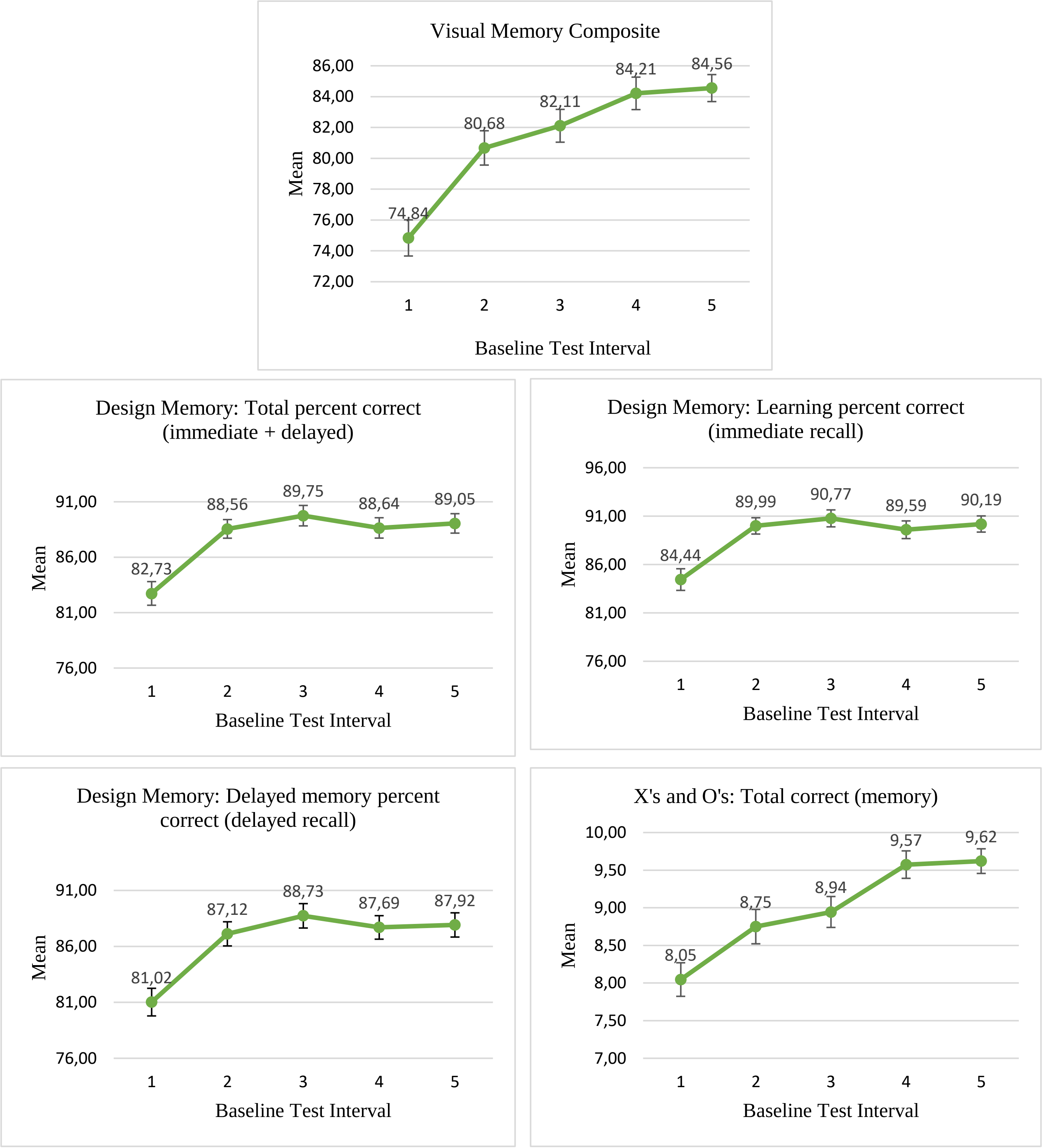

Figure 3.7. Mean baseline Visual Memory composite score, relevant subtest scores, and additional memory subcomponent scores on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A higher score indicates better performance.

On the Visual Motor Speed composite there was a year-on-year increase in mean values from Interval 1 to 5 (33.12 ± 4.98 vs 42.46 ± 5.99), indicating an improvement in performance with increased grade/age and number of baseline test administrations. The Visual Motor Speed composite is made up of two subtests: X’s and O’s total correct (interference) and Three Letters average counted correctly.

On the X’s and O’s subtest, the mean values increased from Interval 1 to 3, followed by a slight decrease from Interval 3 to 5. Despite fluctuating improved and worsening performance across the test intervals on this subtest, there was an overall improvement in mean performance from Interval 1 to 5 (109.80 ± 7.01 vs 114.40 ± 6.55), similar to the overall improvement pattern of the Visual Motor Speed composite. On the Three Letters subtest, there was a year-on-year improvement in mean performance from Interval 1 to 5 (12.93 ± 3.12 vs 18.77 ± 3.85), also similar to the overall improvement pattern of the Visual Motor Speed composite. Both the X’s and O’s and Three Letters subtests showed similar trends of overall improved performance from Interval 1 to 5. The pattern of neurocognitive score performance on the Visual Motor Speed composite and relevant subtests can be seen in Figure 3.8 below.

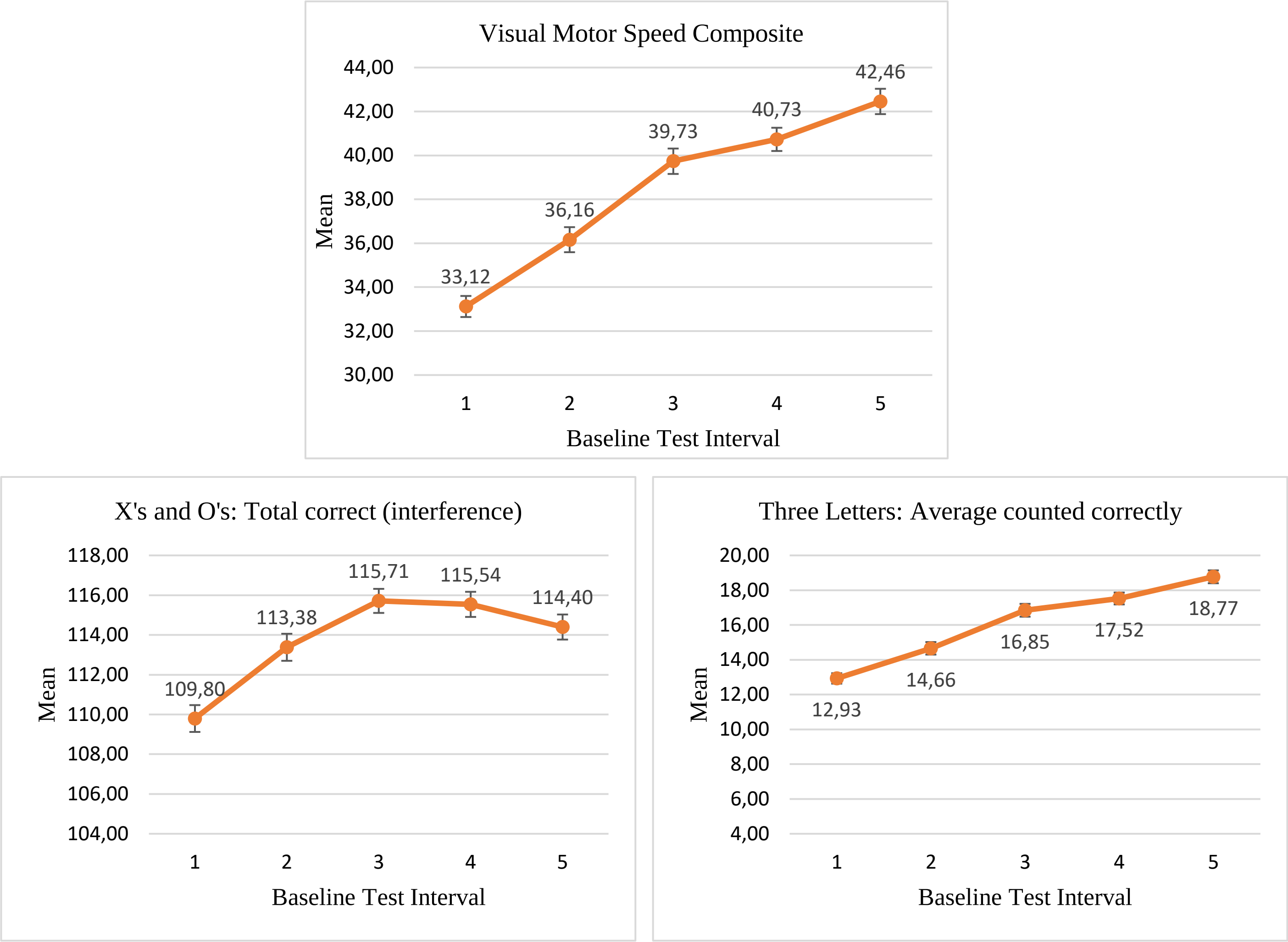

Figure 3.8. Mean baseline Visual Motor Speed composite score and relevant subtest scores on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A higher score indicates better performance.

On the Reaction Time composite, mean values decreased from Interval 1 to 3 indicating improved performance with increased grade/age and number of baseline test administrations, and then increased slightly from Interval 3 to 5 indicating worsening performance. However, despite fluctuating improved and worsening performance across the test intervals, there was an overall improvement in performance from Interval 1 to 5 (0.62 ± 0.06 vs 059 ± 0.08). The Reaction Time composite is made up of three subtests: X's and O's average correct reaction time (interference); Symbol Match average correct reaction time (visible); and Colour Match average correct reaction time. As a lower score on these subtests indicates better performance, a decreasing trend indicates improved performance over time, and vice versa.

On the X's and O's and the Symbol Match subtests the mean values decreased from Interval 1 to 3 indicating improved performance, and then increased from Interval 3 to 5 indicating worsening performance. Whilst both of these subtests followed a similar trend of improved and then worsening score performance over time, only the Symbol Match subtest showed an overall worsening in performance from Interval 1 to 5 (1.46 ± 0.20 vs 1.57 ± 0.36). In contrast, the X's and O's subtest showed an overall improvement in performance from Interval 1 to 5 (0.53 ± 0.06 vs 0.48 ± 0.05), which was similar to the overall improvement pattern of the Reaction Time composite. On the Colour Match subtest, the mean values decreased across the five test intervals, indicating an overall improvement in performance from Interval 1 to 5 (0.83 ± 0.13 vs 0.75 ± 0.13), which was also similar to the overall improvement pattern of the Reaction Time composite. The pattern of neurocognitive score performance on the Reaction Time composite and relevant subtests can be seen in Figure 3.9 below.

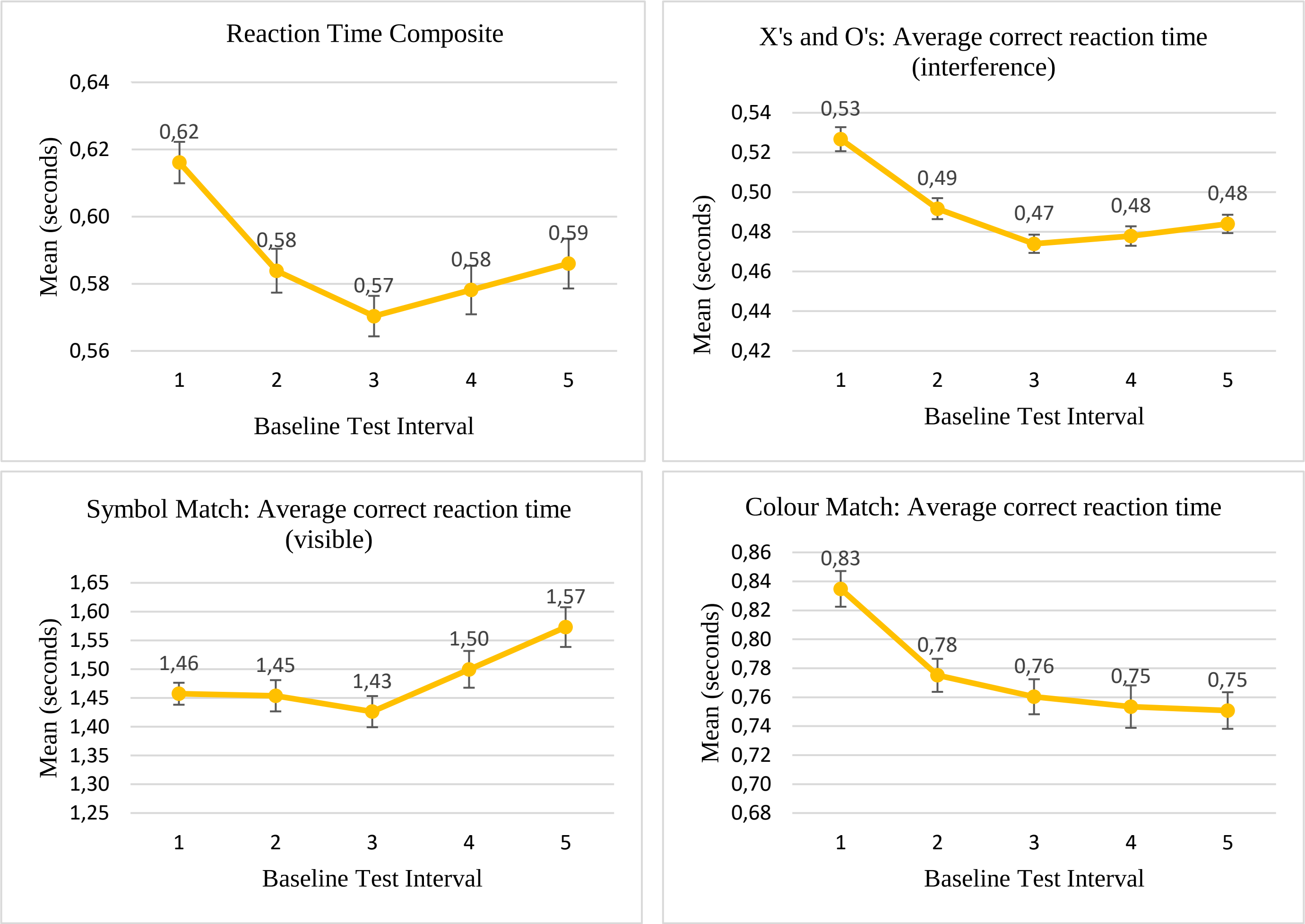

Figure 3.9. Mean baseline Reaction Time composite score and relevant subtest scores on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A lower score indicates better performance.

On the Impulse Control composite, mean values increased very slightly between some test intervals (Interval 1 and 2; Interval 4 and 5) indicting worsening performance, and decreased between other test intervals (Interval 2 and 3; Interval 3 and 4) indicating improved performance. The mean values decrease from Interval 1 to 5 (7.31 ± 4.18 vs 7.71 ± 4.38) indicating that, despite fluctuating improved and worsening performance across the five test intervals, there was an overall pattern of improved performance on the Impulse Control composite. The Impulse Control composite is made up of two subtests: X’s and O’s total incorrect (interference) and Colour Match total commissions. As a lower score on these subtests indicates better performance, a decreasing trend indicates improved performance over time, and vice versa.

On the X’s and O’s subtest, mean values stayed much the same across the five test intervals, with slight fluctuations of improved and worsening performance across the test intervals. There was a slight overall decrease in mean performance from Interval 1 to 5 (6.52 ± 3.99 vs 6.40 ± 4.23) indicating an overall improvement in performance across the test intervals, which was similar to the overall improvement pattern of the Impulse Control composite. On the Colour Match subtest, the mean values decreased across the five test intervals indicating an overall improvement in performance from Interval 1 to 5 (0.79 ± 0.84 vs 0.31 ± 0.51), which was similar to the overall improvement pattern of the Impulse Control composite. Both the X’s and O’s and Colour Match subtests showed similar trends of overall improved performance from Interval 1 to 5. The pattern of neurocognitive score performance on the Impulse Control composite and relevant subtests can be seen in Figure 3.10 below.

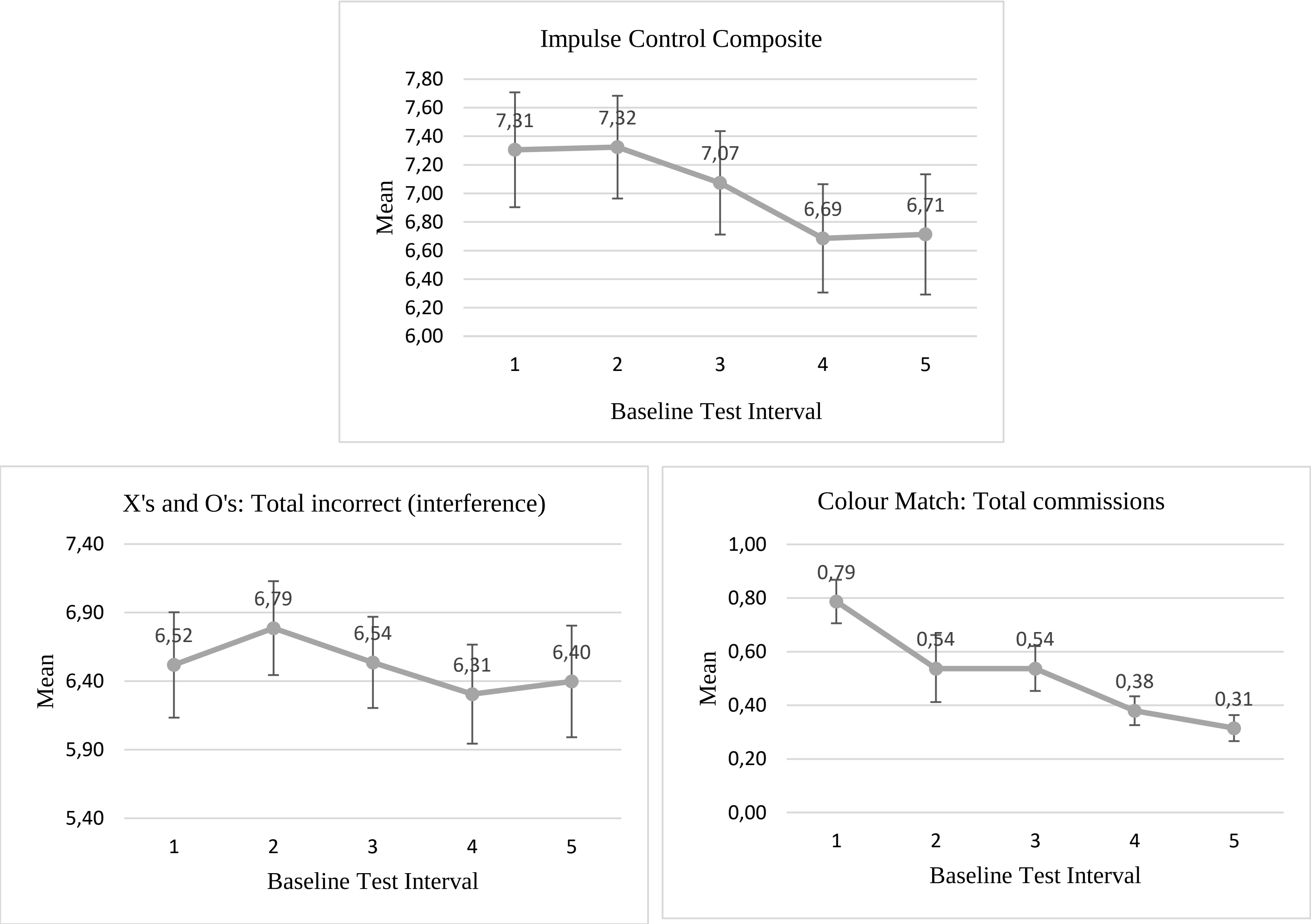

Figure 3.10. Mean baseline Impulse Control composite score and relevant subtest scores on the ImPACT test over five annual test intervals. Error bars represent standard error of mean. A lower score indicates better performance.

CHAPTER 4: DISCUSSION

The primary objective of the present study (Objective 1) was to assess the pattern of change in neurocognitive performance for adolescent athletes on baseline measures of the ImPACT test, over five consecutive years, with a view to providing an indication of the optimal interval for repeat baseline testing of high school athletes. Participants were non-clinical predominantly South African high school athletes from Grades 8 to 12, in the overall age range 13 to 18 years, divided into five groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30). Using a longitudinal within-subject repeated measures design, the same single sample of participants (*N* = 108) were followed over time and tested at five different test intervals. Five separate one-way repeated-measures ANOVAs examined differences in score performance across the five test intervals for each of the five composite scores of the ImPACT test. The five dependent variables were the composite scores on the ImPACT test (Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, Impulse Control) and the independent variable was time (Interval 1, Interval 2, Interval 3, Interval 4, Interval 5). It was hypothesised that neurocognitive score changes would occur on the ImPACT test each year from Grades 8 to 12 (overall age range 13 to 18 years) in non-clinical adolescent athletes during their high school careers.

In addition, a secondary objective (Objective 2) was to generate normative tables for the ImPACT tests' five composite scores, twelve subtest scores that make up the composite scores, as well as four additional memory subcomponent scores which were of clinical relevance (i.e., Word Memory immediate recall, Word Memory delayed recall, Design Memory immediate recall, Design Memory delayed recall). Descriptive data (means, standard deviations and range values) for the five participant groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30), were tabled at each test interval for each of these five composite scores, twelve subtest scores, and four additional memory subcomponent scores. The results provide a clinical and heuristic normative platform for future use with brain injured individuals, which could be used to facilitate clinical interpretations of postconcussion assessment. Discussion is restricted to a description of normative trends.

4.1 Objective 1: Repeat Baseline Testing over Five Annual Test Intervals

The results of repeated-measures ANOVA analyses confirm the hypothesis that neurocognitive score changes would occur on the ImPACT test each year from Grades 8 to 12 (overall age range 13 to 18 years) in non-clinical adolescent athletes during their high school careers. There were neurocognitive score changes over time (i.e., across the five test intervals) on all five ImPACT composite scores. On the Verbal Memory, Visual Memory, Visual Motor Speed and Reaction Time composites, there were significant neurocognitive score changes between several test intervals (although significant changes did not occur between all test intervals). Whilst there were neurocognitive score changes on the Impulse Control composite each year from Grades 8 to 12, these changes were not significant between any test intervals. First, the meaning of the observed patterns of change will be discussed in relation to the available literature, but in broad terms as the available literature does not examine changes year-on-year across the five test intervals. Following which the implications of the findings for the frequency of baseline testing will be examined, but in more specific terms (i.e., the pattern of change year-on-year across the five test intervals. In both cases, the pattern of change in baseline neurocognitive performance across the five test intervals will be discussed separately for each composite.

4.1.1 Meaning of Observed Patterns of Change in Baseline Neurocognitive Performance on ImPACT Composite Scores

The observed change (significant and non-significant) across the five test intervals on the Verbal Memory, Visual Memory, Reaction Time and Visual Motor Speed composites are plausibly due to a combination or interaction of both naturally occurring cognitive development and practice effects from repeat testing, with it not being possible to differentiate how much of the observed change is due to either factor in the current study. Nevertheless, it is interesting to note that existing cross-sectional between-subject research on age-related differences in cognitive performance (not influenced by practice effects) in the domains of verbal memory (Register-Mihalik et al., 2012), visual memory (Register-Mihalik et al., 2012), reaction time (McCrory et al., 2004; Register-Mihalik et al., 2012), and visual motor speed (Hunt and Ferrara, 2009; Register-Mihalik et al. 2012; Kumar Sharma et al. 2014), support the findings of this research.

For Verbal Memory, a study by Register-Mihalik et al. (2012) found no significant age-related differences between high school (age 16) and collegiate (age 20) athletes on the

Hopkins Verbal Learning Test and the ImPACT Verbal Memory composite. Similarly, for Visual Memory, Register-Mihalik et al. (2012) found no significant age-related differences between high school (age 16) and collegiate (age 20) athletes on the Brief Visual-Spatial Memory Test and the ImPACT Visual Memory composite. These findings indicate that cognitive maturity on the verbal memory and visual memory domains are likely obtained before the age of 16, having reached adult levels by this age or earlier, and align with the current findings on the ImPACT Verbal Memory and Visual Memory composites of there being significant change during early-to-mid adolescence (Interval 1 to 3; Grade 8 to 10), and non-significant minimal change during mid-to-late adolescence (Interval 3 to 5; Grade 10 to 12).

For Reaction Time, Register-Mihalik et al. (2012) found no significant age-related differences between high school (age 16) and collegiate (age 20) athletes on the ImPACT Reaction Time composite, indicating that cognitive maturity in the reaction time domain is likely obtained before the age of 16, having reached adult levels by this age or earlier. Another cross-sectional study, whereby participants between the ages of 8 and 18 were tested on the CogSport neurocognitive test, also indicates that reaction time may reach adult levels by the age of 16 or earlier (McCrory et al., 2004). According to McCrory et al. (2004), the largest performance improvements were found between the ages of 8 and 15 on the simple and choice reaction time tests, “with minimal changes after this age paralleling adult performance” (p. 516). The results of the Register-Mihalik et al. (2012) and McCrory et al., (2004) studies align with the current findings on the ImPACT Reaction Time composite of there being significant change during early-to-mid adolescence (Interval 1 to 3; Grade 8 to 10), and non-significant minimal change during mid-to-late adolescence (Interval 3 to 5; Grade 8 to 10).

It is also interesting to note that the theoretical view of the *power law of practice*, which describes the behaviour of practice effects when repeatedly administered, also supports the current findings on the Verbal Memory, Visual Memory and Reaction Time composites. According to the *power law of practice*, learning curves show diminishing gains over time, and after an initial phase of improvement eventually reaches a plateau (Scharfen et al., 2018). Practice effects are usually largest from the first to the second test administration, but decrease with the number of test administrations (Scharfen et al., 2018). The significant change during early-to-mid adolescence (Interval 1 to 3; Grade 8 to 10) followed by non-significant change during mid-to-late adolescence (Interval 3 to 5; Grade 10 to 12) on these composites could (in part) be due to practice effects decreasing with the number of test administrations.

For Visual Motor Speed, several cross-sectional between-subject studies found that visual motor (processing) speed develops and matures throughout adolescence and even into early adulthood, with improvement as age increases (Hunt and Ferrara, 2009; Register-Mihalik et al. 2012; Kumar Sharma et al. 2014). These improvements in processing speed with age further suggests that brain development continues throughout adolescence and into early adulthood (Kumar Sharma et al., 2004), and aligns with the current findings on the ImPACT Visual Motor Speed composite of significant change occurs throughout the testing period, and throughout adolescence (from early adolescence to late adolescence). Whilst there was significant change throughout the testing period, there was *more* significant change during early-to-mid adolescence (Interval 1 to 3; Grade 8 to 10) than during mid-to-late adolescence (Interval 3 to 5; Grade 10 to 12). These findings are consistent with previous research demonstrating maturational gains in processing speed during late childhood and early adolescence, which can be attributed to more efficient neuronal networks within the frontal and parietal lobes (Casey et al., 2000; Fry & Hale, 2000; Nagy, Westerberg, & Klingberg, 2004; Stevens, Skudlarski, Pearlson, & Calhoun, 2009).

The current findings on the Visual Motor Speed composite of *more* significant change occurring during early-to-mid adolescence (Interval 1 to 3; Grade 8 to 10) than during mid-to-late adolescence (Interval 3 to 5; Grade 10 to 12) also aligns with the theoretical view of the *power law of practice,* where the learning curve shows diminishing gains over time and practice effects decrease with the number of test administrations (Scharfen et al., 2018).

4.1.2 Implications for Frequency of Baseline Testing

For Verbal Memory, the significant neurocognitive score change that occurs on this composite between Interval 2 and 3 (Grade 9 and 10), and the non-significant change that occurs between Intervals 3 and 4 (Grade 10 and 11) and between Intervals 4 and 5 (Grade 11 and 12), indicates that annual baseline testing may be necessary during early-to-mid adolescence (Grade 9 and 10), and unnecessary during late asolescence (Grade 11 and 12). The non-significant slight decline in performance between Interval 1 and 2 (Grade 8 and 9) at first glance indicates that baseline testing may also be unnecessary at Grade 8. However, practice effects could be offsetting a greater decline in neurocognitive performance between Interval 1 and 2. Considering practice effects are usually largest between the first and second test interval and become smaller with each additional test administrations (Scharfen et al. 2018), practice effects could potentially be offsetting a significant decline between Interval 1 and 2. As such, it would be prudent to err on the side of caution and baseline test at Interval 1 (Grade 8), in addition to Interval 2 (Grade 9) and Interval 3 (Grade 10), to ensure a more accurate comparative index with which to compare post-injury test results.

For Visual Memory, the significant neurocognitive score change that occurs on this composite between Intervals 1 and 2 (Grade 8 and 9), and the non-significant change that occurs between Interval 2 and 3 (Grade 9 and 10), between Interval 3 and 4 (Grade 10 and 11), and between Interval 4 and 5 (Grade 11 and 12) initially indicates that annual baseline testing may be necessary during early adolescence (Grade 8 and Grade 9), and unnecessary during mid-to-late adolescence (Grade 10, Grade 11, Grade 12). However, the significant neurocognitive change between Interval 2 and 5 (Grade 9 and 12) indicates that too much change occurs between Grade 9 and 12 to forgo baseline testing entirely during this period. Whilst there was significant change between Interval 2 (Grade 9) and Interval 5 (Grade 12), it may be erroneous to recommend repeat baseline testing at Interval 5 as there is a chance the change between Interval 2 and 5 is only significant due to participants also being tested at Interval 3 and 4, and potentially benefitting from practice effects at these intervals. As it is unclear how much of the significant change between Interval 2 and 5 is due to practice effects from baseline testing at Interval 3 and 4, it is recommended that baseline testing be repeated at Interval 3 (Grade 10), instead of Interval 5 (Grade 12). In summary, it is recommended that baseline testing be repeated at Interval 1 (Grade 8), Interval 2 (Grade 9) and Interval 3 (Grade 10) on the Visual Memory composite.

For Visual Motor Speed, the significant neurocognitive score change that occurs on this composite between Intervals 1 and 2 (Grade 8 and 9), between Intervals 2 and 3 (Grade 9 and 10) and between Intervals 4 and 5 (Grade 11 and 12) indicates that annual baseline testing may be necessary throughout adolescence and the high school period (from Grades 8 to 12). Whilst the change between Interval 3 and 4 (Grade 10 and 11) was not significant, the significant change between Interval 2 and 3 and between Interval 4 and 5 indicates that baseline testing at Interval 4 (Grade 11) may be necessary. The significant change that occurs throughout adolescence (from early adolescence to late adolescence), indicates that annual baseline testing is necessary in order to prevent natural cognitive development that may occur between baseline and post-injury potentially offsetting any injury-related cognitive impairment in concussed adolescents. These results also reaffirms previous within-subject repeated-measures research by Rogers et al. (2017) which found substantial increases and large annual changes on the CNS Vital Signs Visual Motor (processing) Speed composite in US high school athletes over four years (ages 15 to 18 years), with these changes plausibly being a combination of both cognitive maturation and practice effects. The current findings also reaffirm discussion by previous authors (Hunt & Ferrara, 2009; Register-Mihalik et al., 2012; Rogers et al., 2017) that frontal lobe-dependent tests, such as visual motor speed (i.e. processing speed), may be particularly important to reassess each year.

For Reaction Time, the significant neurocognitive score change that occurs on this composite between Intervals 1 and 2 (Grade 8 and 9) and the non-significant change that occurs between Interval 2 and 3 (Grade 9 and 10), between Interval 3 and 4 (Grade 10 and 11) and between Interval 4 and 5 (Grade 11 and 12) indicates that annual baseline testing may be necessary during early adolescence (Grade 8 and 9), and unnecessary during mid-to-late adolescence (Grade 10, Grade 11, Grade 12).

For Impulse Control, there were no statistically significant differences on the Impulse Control composite over time between any of the test intervals. Whilst all the change between test intervals was non-significant, the pattern of change on this composite showed a general improvement in performance between Interval 1 and 5. As this composite is a validity indicator and not a neurocognitive composite, it is recommended that Impulse Control be repeated whenever the neurocognitive composites are repeated in order to ensure the validity of these composites. As the Visual Motor Speed composite should be repeated annually throughout high school (from Grades 8 to 12), so too should the Impulse Control composite.

In overview, interpretation of the observed pattern of change in baseline neurocognitive performance on the Visual Memory and Verbal Memory composites across the five test intervals indicates that annual baseline testing may be necessary during early-to-mid adolescence (i.e. Grade 8, Grade 9, Grade 10), and not during late adolescence (Grade 11, Grade 12) on these composites. Similarly, on the Reaction Time composite, annual baseline testing may be necessary during early adolescence (Grade 8, Grade 9), and not during mid-to-late adolescence (Grade 10, Grade 11, Grade 12). On the Visual Motor Speed composite, the pattern of significant change across the five test intervals indicated that annual baseline testing may be necessary throughout adolescence (Grade 8, Grade 9, Grade 10, Grade 11, Grade 12). The Impulse Control composite, being a validity indicator, should be repeated whenever the other neurocognitive composites are repeated in order to ensure the validity of the neurocognitive composites.

4.1.3 Additional Factors Supporting Annual Baseline Testing

Two additional factors provide support for repeat baseline testing on an annual basis on all composites, not only on the Visual Motor Speed composite where significant differences were found throughout the testing period. These factors are: 1) individual variability expected in a cognitively developing adolescent sample and 2) various practical considerations related to the design of the ImPACT test.

4.1.3.1 Individual Variability

Usually when assessing neurocognitive performance, "*measurement* takes place at the *individual* level, but analysis of the data takes place at the *group* level via the aggregation of individual responses and the derivation of a central tendencies (mean) score" (Jordan, 1997, p. 14). Consequently, in order to compare an element of cognitive performance across different groups, there is a comparison of *mean scores*, which are the scores considered to be most representative of the comparative groups. In the current research mean scores of groups (at each test interval) were compared to determine if there were significant differences between groups (Objective 1). Mean scores were also generated at each test interval for normative tables (Objective 2). A limitation of group analysis is that the results are based on group averages, and the full picture in terms of variance at the individual level is not considered. With regards to individual variability, it is necessary to distinguish *inter-individual variability* from *intra-individual variability*. Inter-individual variability is the extent to which raw scores vary about the mean score, or the variability amongst individuals *within* a group. Intra-individual variability is the extent to which scores on a task may vary for a particular individual at different periods of time (Jordan, 1997).

Inter-individual Variability

In the present study, broadly from a clinical perspective, there is evidence of a substantial degree of inter-individual variability on all composite scores (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, and Impulse Control) at all test intervals, as indicated by the wide possible ranges of scores that have occurred in this nonclinical population. The reason for this variability is explicable in terms of different levels of Brain Reserve Capacity (BRC). Formulated by Satz (1993), Brain Reserve Capacity (BRC) theory is a preliminary theoretical formulation of the threshold theory, which explains the presence of individual threshold differences in the onset of clinical symptoms or impaired neurocognitive test performance following a brain injury. Satz (1993) argues for individual differences in BRC, claiming that vulnerability factors such as education, intelligence quotient (IQ), age, the presence of a psychiatric or neurological disorder, a prior head injury, and/or a learning disability, account for variable instances of protection from or vulnerability to acquired brain injury. All participants in the current study had the same level (grade) and quality (educationally advantaged English medium private schools) of education, were approximately the same age at each test interval, had no psychiatric/neurological disorders, or learning disabilities. With regards to prior head injury, whilst 15 participants in the study (13.89%) had either one or two prior concussions, these were considered to have little or no effect as research suggests concussive injury effects are largely undetectable on neurocognitive measures for athletes with a history of only one or two prior injuries (Macciocchi et al., 2001). Estimated intelligence however was not assessed nor controlled for, and may be a contributing factor to the substantial degree of inter-individual variability observed in the current research. Differences in neurocognitive performance due to individual differences in IQ have been found not only on a single-point evaluation but also across two assessment points. According to Rapport, Brines, Axelrod, and Theisen (1997) educationally-matched normal adults with low IQ scores showed smaller score gains on repeat IQ testing than those with average and high IQ scores. Rapport et al. (1997) also found that the ‘rich get richer’ on memory tests.

The presence of inter-individual variability within a group implies that the group is not homogenous, and may comprise of individuals with pre-existing individual differences in cognitive reserve. Adolescents have been identified as a heterogenous group, with differences in the rate of adolescent brain development between individuals (Foulkes & Blakemore, 2018). According to Foulkes and Blakemore (2018) there are “striking individual differences in both behavioural and biological development” amongst adolescents, indicating that adolescents, and their brains, develop in meaningfully different ways (p. 2). Inspection of raw data in studies of brain development during adolescence (Tamnes et al., 2017; Tamnes et al., 2018; Vijayakumar et al., 2016; Wierenge, 2014) have revealed large variance in structural development trajectories in both cortical and subcortical regions (Foulkes & Blakemore, 2018). According to Foulkes and Blakemore (2018), it is likely that both the intercepts (overall level e.g. volume) and slope (i.e., trajectories) of brain development in these studies are subject to individual differences. Differences in the overall level of cognitive development in a group of same-aged adolescents indicates the presence of inter-individual variability.

The greater the inter-individual variability (i.e., variability *within* a group), the less reliable the group norm (i.e. the less likely the mean result is representative of the group). It is important to note that the presence of inter-individual variability does not invalidate the group norm, as there are always going to be individuals who do not fall exactly on the mean. In situations where there is significant inter-individual variability, group norms would function only as important *general guidelines* in the clinical interpretive process of a concussive injury, caution would be applied in the clinical interpretation of a concussed athlete, and interpretation made in conjunction with other clinical data. Group norms with inter-individual variability around the mean do however present a limitation in that, being only group averages, they do not consider individual differences in neurocognitive performance.

On the individual level, every baseline test completed helps to gauge where the individual lies in relation to the average for his/her age and gender, and therefore to gauge the extent of a fall-off in the event of a concussive injury. For example, if a non-injured individual consistently scores below average (+1 SD below the norm) on several baseline tests and then scores below average (+1 SD below the norm) after a concussive injury, it might be assumed he is showing no significant effects of a brain injury. However, if he falls +2 SD below the norm, it might be assumed he is showing significant effects of a brain injury. Yet if another non-injured individual consistently scores above average on several baseline tests and then scores average (on par with the norm) after a concussion, it might be assumed he is showing significant effects of a brain injury. The threshold for what is considered 'normal' and subsequently what is considered a significant brain injury effects is very different for these two individuals due to their individual neurocognitive differences. The group average (i.e. the group norm) is a less reliable comparative index for concussive injury when such individual differences in performance are present, as seen in the above example. In response to the problem of inter-individual variability, information on what is 'normal' neurocognitive performance for each individual should be gathered by collecting as much baseline information on each individual in a non-injured state as is reasonably possible. The more baseline tests collected for an individual, the more certain the clinician of what is the 'norm' for a particular individual (i.e., the more certain the clinician of where the individual falls regularly in relation to the mean), and the more valid the clinical interpretation.

Intra-individual Variability

As mentioned above, intra-individual variability is the extent to which scores on a task may vary for a particular individual at different periods of time (Jordan, 1997). Intra-individual variability was not determined in the current study as data was analysed as groups, and not individually (i.e., the neurocognitive performance of a particular individual across the five test intervals was not determined). Be this as it may, from a theoretical perspective, it is likely intra-individual variability exists in any developing adolescent sample (including the current sample) due to individual differences in the rate of adolescent cognitive development, as discussed previously (Foulkes & Blakemore, 2018), which includes individual differences in cognitive developmental surges. Whilst many studies have shown cognitive development to demonstrate some stage-like properties and some consistency across domains, at the same time studies have also shown that "environmental and organismic factors have powerful effects on [developmental] levels, and individual differences in development seem to be common"

(Fischer & Silvern, 1985, p. 643). Different individuals seem to experience different cognitive development patterns or trajectories (and developmental surges) as a result of both environmental and organismic factors (Fischer & Silvern, 1985).

Intra-individual variability in the developing adolescent adds to the instability of the individual's 'norm'. Individual differences in the rate of cognitive development (Foulkes & Blakemore, 2018) and in developmental surges, would result in the clinician being uncertain as to what is the neurocognitive 'norm' for a particular individual during the volatile period of adolescence. To ensure the clinical certainty of an individual's 'norm', which is used during baseline-post-injury comparisons, it is important that baseline testing occurs on a frequent basis. Frequent baseline testing during this period increases the likelihood of recording accurate and up-to-date cognitive information as close as possible in time to the occurrence of a concussive event, which increases the reliability of any baseline-post-injury comparison. For example, an athlete may score below average (i.e. below the norm) on certain composites on several baseline tests (i.e., baseline 1 and 2) and then subsequently experience a surge of cognitive development recorded on baseline 3. The athlete's improved cognitive performance on baseline 3 (his most recent and best prior performance) should be used for comparison with his post-injury results to ensure the most accurate and reliable baseline-post-injury comparison.

4.1.3.2 Practical Considerations

It is prudent from a practical perspective to baseline test annually *throughout* high school, from Grades 8 to 12, on all the composites, and not only on the Visual Motor Speed composite where significant differences were found between all test intervals. This is because on the ImPACT test it is currently not possible to administer certain composites, such as Visual Motor Speed, at the exclusion of other composites due to the design of the test (ImPACT Applications, 2011). Whilst the test developers could adjust the design to ensure the separate administration of composites, administering a full comprehensive assessment battery on an annual basis as opposed to separate composites has fewer challenges and appears more clinically useful. For example, the difference in the time to set up a computer and administer one composite of the ImPACT test, compared with a full battery, may be seen as negligible (Rogers et al., 2017). Furthermore, if clinicians only have access to the baseline test results of one composite (e.g. Visual Motor Speed) at a particular grade, they will need to locate the most recent *full* battery to have access to baseline test results for the other composites (i.e., Verbal Memory, Visual Memory, Reaction Time, Impulse Control). This results in the clinician having to locate and compile two separate sets of baseline results (one full battery and one more recent Visual Motor

Speed composite), which is an administrative problem for clinicians. More importantly, separate sets of baseline results could complicate the interpretation of results and the comparison with post-injury results.

Furthermore, as discussed previously in section 1.8.5., measures of visual motor speed (information processing speed) and reaction time have demonstrated greater test-retest reliability than measures of verbal and visual memory on the ImPACT test, as well as other computerised neurocognitive tests (Bruce et al., 2014; Cole et al., 2013; Elbin et al., 2011; Littleton et al., 2015; Nakayama et al., 2014; Nelson et al., 2016; Randolph et al., 2005; Resch et al., 2013a; Resch et al., 2013b; Resch et al., 2018; Schatz, 2010; Schatz & Ferris, 2013; Womble et al., 2016). On the ImPACT test, the Visual Motor Speed composite is the most clinical reliable composite, having consistently recorded adequate (0.70 – 0.79) to high (< 0.80) test-retest reliability (Register-Mihalik et al., 2012; Resch et al., 2013a; Cole et al., 2013; Schatz & Ferris et al., 2013; Nakayama et al., 2014; Womble et al., 2016; Elbin et al., 2011; Schatz, 2010; Tsushima et al., 2016; Brett et al., 2016). This composite is also the only composite in the current research to record significant differences between test intervals *throughout* the high school period, resulting in the recommendation for repeat baseline testing annually throughout this period. In contrast, repeat baseline testing is recommended during early and early-to-mid adolescence *only* on the less reliable Reaction Time and Verbal and Visual Memory composites respectively. For clinical decision-making or formulating strong conclusions in research (i.e., recommendations for repeat baseline testing), it is advisable to follow the recommendations of the most reliable composite on the ImPACT test (i.e., the Visual Motor Speed composite), and apply these recommendations to the other less reliable composites.

4.1.4 Conclusion of Objective 1 of this Study

In conclusion, the observed pattern of change in baseline neurocognitive performance across the five test intervals indicates annual baseline testing may be necessary at Grade 8, 9, and 10 on the Visual Memory and Verbal Memory composites, at Grade 8 and 9 on the Reaction Time composites, at Grade 8, 9, 10, 11, and 12 on the Visual Motor Speed composite, and whenever the other neurocognitive composites are repeated on the Impulse Control composite. Forgoing baseline testing at Grade 11 and 12 on the Visual Memory and Verbal Memory composites, and at Grade 10, 11, and 12 on the Reaction Time composite is not recommended as expected individual variability in the adolescent sample and various practical considerations related to the design of the ImPACT test provide support for repeat baseline testing *annually* throughout high school (from Grades 8 to 12) on all composites of the ImPACT test (i.e., Verbal Memory, Visual Memory, Visual Motor Speed, Reaction Time, Impulse Control).

The ideal interval for repeat baseline testing on computerised neurocognitive tests has not been established (Harmon et al., 2013), and there are no firm guideline and minimal consistent agreement as to how often baseline testing should be updated or repeated in high school athletes (Guskiewicz et al., 2004; Schatz, 2010). Considering this, the results of the present study contribute to the growing body of evidence informing the optimal interval for the repeat baseline testing of high school athletes, and may help test developers formulate firm ‘repeat baseline testing’ guidelines for the high school population, thus improving the administration of the ImPACT test in this population. The results of the present study also provide support for the current practice, in South Africa and in most other countries, of administering pre-season baseline testing on the ImPACT test on an annual basis to high school athletes, and may also be beneficial to clinicians, education administrators, and other stakeholders who are using the ImPACT test in the current clinical situation. Whilst this study documented ongoing neurocognitive change on the ImPACT composite scores from one year to the next, the data did not allow for analysis of exact time periods when change might have occurred in intervals less than one year. It is acknowledged that significant change may occur between annual tests, however it is impractical to assume that baseline neurocognitive testing would be conducted more frequently in schools than on an annual basis as per other standardised assessments (Rogers et al., 2017).

4.2 Objective 2: Normative Data for Composite Scores, Subtest Scores, and Additional Memory Subcomponent Scores over Five Annual Test Intervals

The normative data for the ImPACT tests’ five composite scores, twelve subtest scores that make up the composite scores, as well as four additional memory subcomponent scores that make up the ImPACT Word Memory and Design Memory subtests were provided in Table 3.2. Descriptive data for the five participant groups according to the grade currently being completed, including Grades 8, 9, 10, 11 and 12, with mean ages respectively of 13.69 (SD = 0.29), 14.62 (SD = 0.32), 15.58 (SD = 0.31), 16.62 (SD = 0.31) and 17.62 (SD = 0.30), were tabled at each test interval for each of these five composite scores, twelve subtest scores, and four additional memory subcomponent scores.

4.2.1 Normative Data for Composite Scores

The grade-stratified normative data controlled for prior baseline test exposure, thus controlling for the level of exposure to practice effects by ensuring *all* participants were exposed to the *same* number of prior baseline tests at each grade or test interval. As such, clinicians can use the current normative tables to compare an athlete's post-injury test results with norms that accurately match his particular grade and age stage, and specific number of prior baseline tests, thus allowing for a more accurate interpretation of results and clinical diagnosis. Generating normative data using participants that have *all* had the *same* exposure to practice effects ensures practice effects are accounted for (that is, factored into the equation). Accounting for practice effects when generating normative data on the ImPACT test accurately reflects the current clinical situation, where cognitively developing predominantly South African high school athletes are usually baseline tested annually throughout high school (every year over a five-year period) and are thus exposed to practice effects from repeat testing. Accounting for these practice effects ensures the normative data used to compare post-injury test results in the absence of individualised baseline testing are accurate and appropriately normed to the relevant population. By providing more accurate and appropriate normative information reflective of the current clinical situation, the current research adds value to the utility of the ImPACT test.

It is relevant to note that whilst the *ideal* would have been to eliminate or at least significantly minimise practice effects in serial test administrations as they are a source of measurement error, this is not practically possible. Even if alternate forms were used in the serial administration of ImPACT baseline tests to minimise practice effects (which they are not as Baseline Form 1 was administered at all five test intervals), alternate forms would only control for practice effects caused by recall of content (i.e. remembering the answers to certain items). They unfortunately do not control for practice effects caused by procedural learning or familiarity with the testing procedure (i.e. remembering how to do the test) (Hinton-Bayre & Geffen, 2005). In fact, most variance in retesting is felt to be due to familiarity with the procedure as opposed to the specific content (Lezak et al., 2012). Whilst not possible to eliminate or significantly minimise practice effects when serially administering the Baseline Form of the ImPACT test, it is still possible to provide clinicians with accurate normative data by accounting for practice effects, (that is, factoring them into the equation). This has occurred in the current study, where practice effects have been 'controlled for' in the sense that *all* participants have completed the *same* number of prior baseline tests at each test interval.

The normative data (Table 3.2) is more age-specific, representing a wide spectrum of age groups in the adolescent period [i.e., Grade 8 (13-14 years); Grade 9 (14-15 years); Grade 10 (15-16 years); Grade 11 (16-17 years); Grade 12 (17-18 years)], than the broader adolescent age-group norms currently provided by ImPACT Applications [i.e., 13-15 years; 16-18 years] (ImPACT Applications, 2011). Such normative data will allows clinicians to refer to more definitive year-by-year grade/age-group norms when comparing how an athlete is performing, which is particularly relevant considering high school adolescent athletes are going through a period of cognitive development and maturation (Casey et al., 2000; Luria, 1980; Toledo et al., 2012; Yurgelun-Todd, 2007), likely resulting in performance differences at different grades/ages. Such norms will likely result in better representation for performance comparison among 13 to 18 year olds.

4.2.2 Normative Data for Subtest Scores and Additional Memory Subcomponent Scores

Normative data for the twelve subtest scores that make up the five ImPACT composite scores and the four additional subcomponent scores (provided in Table 3.2) are clinically useful in that the data increase the utility of the ImPACT test by providing clinicians with a "deeper level of interpretation that may be missed" if only the ImPACT composite scores were used. (Henry & Sandel, 2015, p. 272). Whilst ImPACT composite scores adequately assess neurocognitive functioning and chart progress in many cases, there are cases with a slow or complicated recovery course, requiring additional information that can be interpreted from the subtest scores that make up the composite scores (Henry & Sandel, 2015), and from the additional memory subcomponent scores. Such normative information could "provide more specific details of post-injury neurocognitive impairment and, in conjunction with a clinical interview, assist in guiding decisions regarding academic accommodations and appropriate activity restrictions during an athlete's recovery" (Henry & Sandel, 2015, p. 271). The subtest and additional memory subcomponent norms provided in the current study also increase the utility of the ImPACT test "in those cases that present unique challenges to basic test interpretation", such as complex and complicated cases involving comorbid disorders (Henry & Sandel, 2015, p. 272). Where the composite score may indicate a deficit, closer inspection of the subtests (and additional memory subcomponents) using the norms provided by the current study can provide significant insight into a particular athlete's performance (Henry & Sandel, 2015).

From a neurocognitive perspective, normative data on the immediate and delayed recall subcomponents of the Word Memory and Design Memory subtests would be especially relevant comparative indexes for clinicians managing concussive injuries. This is because delayed recall is particularly sensitive to brain impairment or dysfunction (Lezak et al., 2012), and a falloff in verbal and/or visual memory delayed recall scores is an indication of brain injury, which may not be apparent on immediate recall. Even if immediate recall scores for an individual are “normal”, there may be a deficit on delayed recall scores indicating an individual is cognitively impaired.

As Objective 2 was merely aimed at providing a platform for future normative comparisons, as provided in Table 3.2, detailed neurocognitive analysis of how the composite scores, subtest scores and four additional memory subcomponent scores might be differentially affected in the event of a concussion is beyond the scope of the present research on a non-clinical population. However, broad observations on the following will be made in this discussion: a) the normative trends of the composite scores, and b) similarities and/or differences in the trends of the subtest scores and additional memory subcomponent scores when compared to the trends of the overall composite scores. Such information may be useful to the clinician when interpreting the performance of a particular individual, and be a platform for further research on what might be typically affected by a concussive injury.

4.2.3 Normative Trend of the Composite Scores

The normative data revealed that the composites scores displayed an overall trend of improved performance over time, with increased grade/age and number of baseline test administrations. On the Visual Memory and Visual Motor Speed composites, there was an overall improvement in performance across the five test intervals. On the Verbal Memory and Impulse Control composites, despite fluctuating improved and worsening performance across the test intervals, there was also an overall improvement in performance across the five test intervals. On the Reaction Time composite, performance improved over the years from Interval 1 to 3 (early-to-mid adolescence) and then worsened from Interval 3 to 5 (mid-to-late adolescence). Despite this worsening in performance, the score performance at Interval 5 (Grade 12) was still better than the performance at Interval 1 (Grade 8), indicating older adolescents performed better on this composite than younger adolescents.

As mentioned previously, improved performance on neurocognitive tests with increased grade/age and number of test administrations could plausibly be due to a combination or interaction of naturally occurring cognitive development that occurs during adolescence, and practice effects from repeat testing. Whilst not influenced by practice effects, previous cross-sectional research supports the current findings, as older adolescents demonstrated higher baseline mean scores than younger adolescents on information processing, attention, and motor tests (Hunt & Ferrara, 2009). Longitudinal research on the annual neurocognitive performanceof US high school athletes over four years (ages 15 to 18 years), plausibly influenced by both cognitive development and practice effects, also found improvement with increased grade/age in the domains of executive functioning, cognitive flexibility, psychomotor speed and reaction time (Rogers et al., 2017). In further support, the results of a normative study on the ImPACT composites by Shuttleworth-Edwards et al. (2009) using the age groups 11-13 years; 14-16 years; and 17-21 years, showed improved performance on neurocognitive tests with each subsequent age-group on all composites in a South African sample. Whilst direct comparison of these norms to those of the current study is not possible due to the different age-group classifications and the likelihood that the level of exposure to practice effects was not controlled for, it still stands that there is improved performance on neurocognitive tests with increased grade/age.

4.2.4 Normative Trend of the Subtest Scores and Additional Memory Subcomponent Scores Compared to the Composite Scores

For Verbal Memory, the Word Memory subtest (Word Memory total percent correct), and its two subcomponents (Word Memory immediate recall and Word Memory delayed recall) stayed much the same across the test intervals with a slight overall worsening in performance, which were different in trend to the overall improvement pattern of the Verbal Memory composite. This “dissociation” in trend between the Verbal Memory composite on the one hand, and the Word Memory subtest and its immediate and delayed recall subcomponents (which goes to make up the composite) on the other hand, is very important for interpretive comparisons. Whilst the clinician can expect an individual’s scores on the Verbal Memory composite to increase across the five test intervals, the Word Memory subtest and its immediate and delayed recall subcomponents can be expected to stay much the same. Therefore, stable scores across test intervals on the Word Memory subtest and its immediate and delayed recall subcomponents are not to be seen as a fall-off in function if there is an an intervening concussion. In contrast, the Symbol Match and Three Letters subtests displayed an overall improvement in performance across the test intervals, which was similar in trend to the overall improvement pattern of the Verbal Memory composite.

For Visual Memory, the Design Memory subtest (Design Memory total percent correct), its two subcomponents (Design Memory immediate recall and Design Memory delayed recall), and the X's and O's subtest all displayed an overall improvement in performance across the test intervals, which were similar in trend to the overall improvement pattern of the Visual Memory composite.

For Visual Motor Speed, the X's and O's and the Three Letters subtests both displayed an overall improvement in performance across the test intervals, which were similar in trend to the overall improvement pattern of the Visual Motor Speed composite.

For Reaction Time, the Symbol Match subtest displayed an overall worsening in performance across the test intervals, which was different in trend to the overall improvement pattern of the Reaction Time composite. This "dissociation" in trend between the Reaction Time composite and the Symbol Match subtest (which goes to make up the composite) may provide useful information for in-depth clinical evaluations. Whilst the trend differences are subtle, they may assist the clinician in conducting an in-depth investigation of a post-concussion test profile. In contrast, the X's and O's subtest and the Colour Match subtest both displayed an overall improvement in performance across the test intervals, which were similar in trend to the overall improvement pattern of the Reaction Time composite.

For Impulse Control, the X's and O's subtest stayed much the same across the test intervals, with a slight overall improvement in performance. The Colour Match subtest also displayed an overall improvement in performance across the test intervals. Both the X's and O's and the Colour Match subtests displayed an overall improvement in performance across the test intervals which were similar in trend to the overall improvement pattern of the Impulse Control composite.

4.2.5 Conclusion of Objective 2 of this Study

Concussive injuries can result in long-term neurological difficulties and chronic sequelae, which for some can be permanent and disabling (Lovell & Collins, 1998; McKee et al., 2013). Youth athletes in particular are vulnerable to detrimental concussion effects as their brains are still developing. Relative to adults, adolescent athletes have a longer post-injury recovery period, thus increasing their changes of returning to play before being fully recovered (Field et al., 2003; McCrory et al., 2013). Considering these factors, it is very important to ensure neurocognitive tests used to measure post-injury cognitive impairment and inform concussion management are appropriately administered and interpreted, especially when used to manage youth concussion. Accurate and appropriate normative data on the ImPACT composite scores, subtest scores, and additional memory subcomponent scores will allow for better comprehension of the ImPACT test, which will aid in the assessment and management of concussive injuries.

The current normative data are the first on the ImPACT test worldwide to be generated for such a wide spectrum of grade and associated age stages that have been exposed to multiple test administrations of the same test, and adds to the limited literature on normative data on the ImPACT test. The norms generated in this study have great clinical utility in that when compared with the post-injury results of a concussed athlete, they allow the clinician to refer to a specific norm base depending on the concussed examinees particular grade and age stage, and specific number of prior baseline tests. By closely reflecting the current clinical situation where adolescent athletes are exposed to two, three, or four consecutive test administrations on the same test during the high school period, the current norms are an accurate and appropriate record of ‘normal’ neurocognitive performance for adolescent high school athletes, particularly in the South African context. These normative results provide a clinical and heuristic normative platform for future use with brain injured individuals, which could be used to facilitate clinical interpretations of postconcussion assessment.

The normative data on the ImPACT subtest scores and additional memory subcomponent scores are also clinically useful as they allow for better comprehension of the ImPACT test that extends beyond the basic interpretation of composite scores. The deeper level of interpretation offered by these norms may aid clinicians in interpreting the ImPACT clinical report and help further delineate areas of neurocognitive dysfunction, as well as generally enriching the clinical interpretation of ImPACT test performance. In particular, the clinician would likely find the normative information on the immediate and delayed recall subcomponents of the Word Memory and Design Memory subtests the most useful comparative indexes for managing concussive injuries, as delayed recall is particularly sensitive to brain impairment or dysfunction (Lezak et al., 2012). Importantly, the current research found “dissociations” in the pattern or trend of performance between 1) the Verbal Memory composite and its contributing Word Memory subtest and immediate and delayed recall subcomponents and 2) the Reaction Time composite and its contributing Symbol Match subtest. Information on such “dissociations” may assist the clinician in the in-depth clinical evaluation of a post-concussion test profile.

4.3 Evaluation of the Study

The methodological strengths of this study include:

1) This study provides a descriptive indication of year-on-year neurocognitive score change on each of the five ImPACT composites using a longitudinal approach. No other study on the ImPACT test provides such year-on-year descriptive information of the neurocognitive change that occurs *throughout* adolescence on various cognitive domains. Only one other study by Rogers et al. (2017) assess the pattern of change that occurs from ages 15 to 18 years, but on a different neurocognitive test (the CNS Vital Signs). The Rogers et al. (2017) study also only provides results for the overall change that occured throughout this period, not for the change year-on-year. As such, the current research adds to the limited literature on neurocognitive change that occurs throughout adolescence by providing descriptive year-on-year information, that may be useful to future researchers.
2) This study is the only within-subject longitudinal study on any computerised neurocognitive test, including the Rogers et al. (2017) study, that assesses changes in neurocognitive scores on the *same participants every year over the five-year high school period* across the 13 to 18 year adolescent age range. This research design most accurately reflects the current clinical situation, where cognitively developing South African high school athletes are usually baseline tested annually throughout high school (every year over a five-year period). In this situation any change in neurocognitive scores observed over time can plausibly be due to a combination or interaction of naturally occurring cognitive maturation as the participant's age and practice effects from repeat annual testing. Regardless of the etiology of any observed change, the fact that change does occur, and the location and magnitude of this change, is valuable information that should be considered when making decisions about the frequency and timing of baseline testing in high school athletes. As no similar study exists, the current research provides clinicians and researcher with valuable preliminary information on 'normal' neurocognitive score change for cognitively developing adolescent athletes as they increase in grade/age and are exposed to consecutive test administrations on the ImPACT test.
3) Information on medical or neurodevelopmental conditions that may influence participant performance and thus the overall pattern of baseline test performance were obtained, and participants with these conditions were excluded from the final sample.

These included a learning disability, ADHD/ADD, a neurological disorder, substance/alcohol abuse, a psychiatric/psychological disorder, the presence of certain medication, a concussive injury six months prior to Interval 1 testing, and a history of three or more concussions prior to Interval 1 testing. As these clinical variables known to influence neurocognitive performance were excluded, the dataset can be considered relatively clean and consisting of a non-clinical sample.

The methodological limitations of this study include:

1) This study controlled for significant demographic variables such as age and education level by ensuring all participants were approximately the same age and exactly the same grade at each testing interval. However, IQ is another potentially influential variable that was neither assessed nor controlled for. Differences in IQ amongst participants may have influenced neurocognitive performance and may have contributed to the inter-individual variability observed at most of the test intervals. As all participants attended educationally advantaged English medium private schools and had not repeated any grades (indicating they could perform to at least an acceptable academic level), it is likely that there were no participants included who were intellectually compromised. However, it would still be preferable in a future study to control for an estimated level of IQ if possible, so as to formally exclude any outlying individuals on this parameter that might confound overall scores.
2) This study consisted of participants who played only contact sport (53.70%), only non-contact sport (19.45%), and a mixture of both (26.85%) over the five-year testing period. Athletes participating in contact sport have an increased risk of sustaining neurocognitive injuries known as sub-concussive blows, compared to non-contact athletes. Further, participation in contact sports with a high exposure to sub-concussive blows negatively affects neurocognitive functioning in young athletes (Tsushima et al. 2016). Whilst the current study controlled for a history of concussive injuries, it did not control for exposure to sub-concussive blows. As sub-concussive blows do not cause observable symptoms or meet the criteria of a clinically diagnosed concussion (Belanger, Vanderploeg, & McAllister, 2016; Concussion Legacy Foundation, 2018), they often go unnoticed/unrecorded and are difficult to control for. It is likely athletes participating in contact sport or a mixture of both contact and non-contact sport would exhibit a deficit in neurocognitive functioning following multiple seasons of play compared to athletes participating in non-contact sport, with athletes participating in

only contact sport having a greater deficit than athletes participating in a mixture of both contact and non-contact sport. Considering this, the results in this study could have been biased by differences in sports participant (i.e. contact sport; non-contact sport; and a mixture of both), potentially resulting in differential levels of exposure to sub-concussive blows and thus differential levels of neurocognitive performance among participants. A future study might analyse data such as these by doing a comparative analysis of those involved in contact sports versus those who are not.

3) There is a lack of generalisability of the findings of this study to other populations. The present study was carried out on a highly specific cohort, where the same participants in the same context were used to minimise the influence of extraneous variables. As such, the findings are limited to generalising elite, privileged schools, where all the male athletes speak English and the academic standards are high. The same pattern of results may not occur in less advantaged and/or disadvantaged schools where academic and cognitive functioning may be negatively impacted by a lower standard of education and poorer resources. Certainly, replication with a less advantaged sample would be highly desirable. Furthermore, as the sample was of high school adolescent athletes, the age range may limit inference to college-aged and older adults, as well as children. Cognitive maturation and development occur more during adolescence than during adulthood, and more during childhood than during adolescence (Luria, 1980; Casey et al., 2000), resulting in limited inference of the current study results across these broad age groups. Finally, as the sample was of all male participants, the findings are not generalisable to female same-aged athletes.
4) The sample size of the present study of $N = 108$ is not small, but also not particularly large. Curtailing the sample size significantly was the need to have participants who had been assessed on an annual basis over a five-year period, which is a very lengthy assessment period, together with compliance to strict exclusion criteria that were demanded by the study. Therefore, the sample size is considered adequate under the sampling circumstances, and preferable to a larger sample without suitable control for potentially confounding factors. However, it is possible that some of the small changes observed between intervals on the composites may have been statistically significant with a larger sample size.
5) The data used in this study was collected retrospectively from archival data on the ImPACT database. The researcher of this study did not have control over test administration and test conditions, and therefore, administration may not have been

standardised. However, considering that ImPACT is a standalone computerised test battery requiring limited interaction from the test administrator, this potential confound was minimised. Also, as all test administrations were supervisored by a trained ImPACT administrator or a board-registered psychologist, it is likely many factors that could bias results were minimised. Furthermore, any effects that non-standard administration may have on performance were also minimised by eliminating very extreme scores. Nonetheless, it is possible that athletes' performance on the ImPACT test could have been biased by unforeseen factors such as environmental distration or misreported demographic information (Henry & Sandel, 2015).

4.4 Recommendations for Future Research

1) As the current research is the only within-subject longitudinal study on any computerised neurocognitive test that assesses changes in neurocognitive scores on the *same participants every year over the five-year high school period* across the 13 to 18 year adolescent age range, the study needs to be replicated, preferably with a larger sample, and controlled for IQ, in order to validate and corroborate the current findings. Further, replication with other populations would be highly desirable (i.e., different age ranges; female participants; less educationally advantaged populations).
2) Future research should control for type of sports participation to ensure all participants played the same type of sport. This would ensure results are more conclusive and generalisable to athletes that play either contact, non-contact, or a mixture of both contact and non-contact sports. Future research can also go one step further and control for the exposure rate to sub-concussive blows. There is likely a diversity in exposure to sub-concussive blows within a contact sport, which may cause individual differences in neurocognitive performance within the same sports participation group. Whilst a challenging measure to obtain, the rate of exposure can be estimated by obtaining information on the participant's playing position (i.e. forward vs backline players) and the number of minutes played during games and practices (Broglio, Martini, Kasper, Eckner, & Kutcher, 2013). Some playing positions involve more contact whilst other involve less, such as forward as opposed to backline playing positions in rugby. This results in some players being exposed to more sub-concussive blows than others. Similarly, some individual may only play a small portion of a game or practice, whereas others in the same sport may play considerably more, and thus have a much greater

exposure to sub-concussive blows. Ensuring all participants in a study have the same exposure rate to sub-concussive blows would control for a confounding variable that may influence participant neurocognitive performance, and thus the overall pattern of baseline test performance.

3) The same database used in this study can be further analysed by comparing results for those involved in different sports, or contact versus non-contact sport, or adding in a comparative group of those excluded from the study for various reasons (i.e., participants with ADHD, learning disability, neurological disease, those with more than two prior concussions), or using them as a combined neurologically compromised group.
4) Finally, prospective research would be preferable to retrospective analysis using archival data, allowing for better control of possible confounding variables and test-taking conditions than has not been possible in a study such as this. However, the possibility of retrospective research of this type using a bank of available data, has allowed for important clinical indications in the here and now, that might take many years to achieve with the forward planning demanded of a worthwhile study going forward.

REFERENCES

Albers, F., & Hoeft, S. (2009). Do it again and again and again – practice effects in the computer-based test of spatial ability, *Diagnostica, 55*(2), 71-83.

Alexander, D. G., Shuttleworth-Edwards, A. B., Kidd, M., & Malcolm, C. M. (2015). Mild traumatic brain injuries in early adolescent rugby players: Long-term neurocognitive and academic outcomes. *Brain Injury, 29*(9), 1113-1125. doi: 10.3109/02699052.2015.1031699.

Arnett, P., Rabinowitz, A., Vargas, G., Ukueberuwa, D., Merritt, V., & Meyer, J. (2014). Neuropsychological testing in sports concussion management: An evidence-based model when baseline is unavailable. In S. Slobounov, & W. Sebastianelli (Eds.), *Concussion in athletics: Current understanding from basic brain science to clinical research* (pp. 35-48). New York, NY: Springer.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., ... & Schamasch, P. (2002). Summary and Agreement Statement of the First International Conference on Concussion in Sport, Vienna 2001. *British Journal of Sports Medicine, 36*(1), 6-7.

Bailes, E. J. (2009). Sports-related concussion: What do we know in 2009 – A Neurosurgeon's perspective. *Journal of the International Neuropsychological Society, 15*, 509-511.

Bakhos, L. L., Lockhart, G. R., Myers, R., & Linakis, J. G. (2010). Emergency department visits for concussion in young child athletes. *Pediatrics, 126*(3), e550-e556.

Barr, W. B. (2003). Neuropsychological testing for assessment of treatment effects: Methodologic issues. *CNS Spectrums, 7*(04), 300-306.

Bartels, C., Wegrzyn, M., Wiedl, A., Ackermann, V., & Ehrenreich, H. (2010). Practice effects in healthy adults: A longitudinal study on frequency repetitive cognitive testing. *Neuroscience, 11*, 118-129.

Barth, J. T., Alves, W. M., Ryan, T. V., Macciocchi, S. N., Rimel, R. W., Jane, J. A., & Nelson, W. E. (1989). Mild head injury in sports: Neuropsychological sequelae and recovery of function. In H. S. Levin, H. M. Eisenberg, & A. L. Benton (Eds.), *Mild head injury* (pp. 257-275). New York, NY: Oxford University Press.

Barth, J. T., Freeman, J. R., Broshek, D. K., & Varney, R. N. (2001). Acceleration-deceleration sport-related concussion: the gravity of it all. *Journal of Athletic Training, 36*(3), 253-256.

Bauer, R., & Fritz, H. (2004). Pathophysiology of traumatic injury in the developing brain: An introduction and short update. *Experimental and Toxicologic Pathology*, *56*(1), 65-73.

Belanger, H. G., & Vanderploeg, R. D. (2005). The neuropsychological impact of sports-related concussion: A meta-analysis. *Journal of the International Neuropsychological Society*, *11*(4), 345-357.

Belanger, H. G., Curtiss, G., Demery, J.A., Lebowitz, B. K., & Vanderploeg, R. D. (2005). Factors moderating neuropsychological outcome following mild traumatic brain injury: A meta-analysis. *Journal of the International Neuropsychological Society, 11,* 215-227.

Belanger, H. G., Vanderploeg, R. D., & McAllister, T. (2016). Subconcussive blows to the head: A formative review of short-term clinical outcomes. *Journal of Head Trauma Rehabilitation, 31*(3), 159-166. doi: 10.1097/HTR0000000000000138

Benes, F. (2001). The development of the frontal cortex: the maturation of neurotransmitter systems and their interactions. In C. Nelson & M. Luciana (Eds.), *Handbook of Developmental Cognitive Neuroscience* (pp. 79-82). Cambridge, MA: MIT Press.

Bird, K. D., & Hadzi-Pavlovic, D. (1983). Simultaneous test procedures and the choice of a test statistic in MANOVA. *Psychological Bulletin, 93*, 167-178.

Bissell, T. (2018). Study suggests blows to the head that don’t cause concussion can still cause CTE. *SB Nation Bloody Elbow*. Retrieved October 4, 2018 from https://www.bloodyelbow.com/2018/1/22/16909086/blows-head-not-concussion-cause-cte-mma-ufc-nfl-sports-tbi-brain-trauma-science-health-news.

Bleiberg, J., Cernich, A. N., Cameron, K., Sun, W., Peck, K., Ecklund, L. P. J., ... & Warden, D. L. (2004). Duration of cognitive impairment after sports concussion. *Neurosurgery, 54*(5), 1073-1080.

Bleiberg, J., Halpern, E., Reeves, D., & Daniels, J. C. (1998). Future directions for the neuropsychological assessment of sports-related concussion. *Journal of Head Trauma Rehabilitation, 13*(2), 36-44.

Bless, C., Higson-Smith, C., & Kagee, A. (2006). *Fundamentals of social research methods: An African perspective*. Cape Town, South Africa: Juta and Company Ltd.

Bock, R. D. (1975*). Multivariate statistical methods in behavioural research*. New York: McGraw-Hill.

Boelema S, Harakeh Z, Ormel J, Hartman C, Vollebergh W, van Zandvoort M. (2014). Executive functioning shows differential maturation from early to late adolescence: longitudinal findings from a TRAILS study. *Neuropsychology, 28*(2), 177–87.

Bornstein, R. A. (1985). Normative data on selected neuropsychological measures from a nonclinical sample. *Journal of Clinical Psychology, 41*(5), 651-659.

Brain Injury Research Institute. (2019). What is CTE?. *Brain Injury Research Institute: Protect the Brain*. Retrieved January 2, 2019, from http://www.protectthebrain.org/Brain-Injury-Research/What-is-CTE-.aspx

Brandt, J. (2007). 2005 INS presidential address: neuropsychological crimes and misdemeanors. *The Clinical Neuropsychologist, 21*, 553-568.

Bray, J. H., & Maxwell, S. E. (1982). Analysing and interpreting significant MANOVAs. *Review of Educational Research, 52*, 340-367.

Brett, B. L., Smyk, N., Solomon, G., Baughman, B. C., & Schatz, P. (2016). Long-term Stability and Reliability of Baseline Cognitive Assessments in High School Athletes Using ImPACT at 1-, 2-, and 3-year Test–Retest Intervals. *Archives of Clinical Neuropsychology, 31*(8), 904-914.

Broglio, S. P., Ferrara, M. S., Macciocchi, S. N., Baumgartner, T. A., & Elliott, R. (2007). Test-retest reliability of computerized concussion assessment programs. *Journal of Athletic Training, 42*(4), 509-514.

Broglio, S. P., & Puetz, T. W. (2008). The effect of sport concussion on neurocognitive function, self-report symptoms and postural control. *Sports Medicine, 38*(1), 53-67.

Broglio, S. P., Martini, D., Kasper, L., Eckner, J. T., & Kutcher, J. S. (2013). Estimation of head impact exposure in high school football: implications for regulating contact practices. *American Journal of Sports Medicine, 41*(12), 2877-2884. Doi: 10. 1177/0363546513502458.

Bruce, J., Echemendia, R., Meeuwisse, W., Comper, P., & Sisco, A. (2014). 1 year test–retest reliability of ImPACT in professional ice hockey players. *The Clinical Neuropsychologist, 28*(1), 14-25.

Buzzini, S. R. R., & Guskiewicz, K. M. (2006). Sport-related concussion in the young athlete. *Current Opinion in Pediatrics, 18*(4), 376-382.

Calamia, M., Markon, K., & Tranel, D. (2012). Scoring higher the second time around: Meta-analyses of practice effects in neuropsychological assessment. *The Clinical Neuropsychologist, 26*(4), 543-570.

Carman, A. J., Ferguson, R., Cantu, R., Comstock, R.D., Dacks, P. A., DeKosky, S. T., … & Fittit, H.M. (2015). Expert consensus document: mind the gaps-a advancing research into short-term and long-term neuropsychological outcomes of youth sport related concussions. *Nature Reviews Neurology, 11*, 230-244.

Cantu, R. C. (1998). Second-impact syndrome. *Clinics in Sports Medicine, 17*(1), 37-44.

Cantu, R. C. (2001). Posttraumatic retrograde and anterograde amnesia: pathophysiology and implications in grading and safe return to play. *Journal of Athletic Training, 36*(3), 244-248.

Cantu, R. C., & Hyman, M. (2012). *Concussions and Our Kids. America's Leading Expert on How to Protect Young Athletes and Keep Sports Safe.* New York, NY: Houghton Mifflin Harcourt Publishing Co.

Casey, B., Giedd, J. N., & Thomas, K. M. (2000). Structural and functional brain development and its relation to cognitive development. *Biological Psychology, 54*(1), 241-257.

Casey, B.J., Jones, R. M., & Hare, T.A. (2008). The Adolescent Brain. *Annals of the New York Academy of Sciences, 1124,* 111-126.

Center for Disease Control and Prevention. (2015). FAQs about baseline testing among young athletes. *Centers for Disease Control and Prevention*. Retrieved December 10, 2018, from https://www.cdc.gov/headsup/basics/baseline_testing.html

Cohen, J (1988). *Statistical power for the behavioural sciences*. Hillsdale, NJ: Lawrence Erlnaum.

Cohen, R. A., Malloy, P. F., Jenkins, M. A., & Paul, R. H. (2006). Disorders of attention. In P. J. Snyder (Ed.), *Clinical neuropsychology: A pocket handbook for assessment* (2nd ed., pp. 572–606). Washington, DC: American Psychological Association.

Cole, W. R., Arrieux, J. P., Schwab, K., Ivins, B. J., Qashu, F. M., & Lewis, S. C. (2013). Test-retest reliability of four computerized neurocognitive assessment tools in an active duty military population. *Archives of Clinical Neuropsychology, 28*(7), 732-742. doi:10.1093/arclin/act040.

Collie, A., Darby, D., & Maruff, P. (2001). Computerised cognitive assessment of athletes with sports related head injury. *British Journal of Sports Medicine*, *35*(5), 297-302.

Collins, M. W., Lovell, M. R., Iverson, G. L., Cantu, R. C., Maroon, J. C., & Field, M. (2002). Cumulative effects of concussion in high school athletes. *Neurosurgery, 51*(5), 1175-1181.

Collins, M., Lovell, M. R., Iverson, G. L., Ide, T., & Maroon, J. (2006). Examining concussion rates and return to play in high school football players wearing newer helmet technology: a three-year prospective cohort study. *Neurosurgery*, *58*(2), 275-286.

Concussion Legacy Foundation. (2018). CTE Resources: Subconcussive Impacts. *Concussion Legacy Foundation.* Retrieved October 4, 2018 from https://concussionfoundation.org/CTE-resources/subconcussive-impacts

Colvin, A.C., Mullen, J., Lovell, M. R., West, R. V., Collins, M.W., & Groh, M. (2009). The role of concussion history and gender in recovery from soccer-related concussion. *The American Journal of Sports Medicine, 37*, 1699-1704.

Congress of Neurosurgeons. (1968). Clinical neurosurgery. Proceedings of the congress of neurological surgeons, San Juan, Puerto Rico, 1966. *British Journal of Surgery*, *55*(4), 319-320. doi: 10.1002/bjs.1800550431.

Cook, R. S., L. Schweer, K. F. Shebesta, K. Hartjes, and R. A. Falcone. (2006). Mild traumatic brain injury in children: Just another bump on the head? *Journal of Trauma Nursing, 13*(2), 58-65.

Coolidge, F. L., Middleton, P. A., Griego, J. A., & Schmidt, M. M. (1996). The effects of interference on verbal learning in multiple sclerosis. *Archives of Clinical Neuropsychology, 11*(7), 605-611.

Covassin, T., Elbin, R., Kontos, A., & Larson, E. (2010). Investigating baseline neurocognitive performance between male and female athletes with a history of multiple concussions. *Journal of Neurology, Neurosurgery & Psychiatry, 81*, 597-601.

Covassin, T., Elbin, R., Harris, W., Parker, T., Kontos, A. (2012). The role of age and sex in symptoms, neurocognitive performance, and postural stability in athletes after concussion. *The American Journal of Sports Medicine, 40*, 1303-1312.

Cromer, J.A., Schembri, A. J., Harel, B. T., Maruff, P. (2015). The nature and rate of cognitive maturation from late childhood to adulthood. *Frontiers in Psychology*, *6*, 1-12. doi: 10.3389/fpsyg.2015.00704.

Department of Basic Education. (2017). Select Committee on Education and Recreation: Poverty Ranking of Schools (Quintiles). *PMG Parlimentary Monitoring Group*. Retrieved March 3, 2019 from http://pmg.org.za/files/171129Quintiles.pptx

Duff, M. C. (2009). Management of sports-related concussion in children and adolescents. *The ASHA Leader*, *14*(9), 10-13. doi: 10.1044/leader.FTR1.14092009.10.

Duff, K. (2012). Evidence-based indicators of neuropsychological change in the individual patient: Relevant concepts and methods. *Archives of Clinical Neuropsychology*, 27, 248-261.

Dunlop, P. D., Morrison, D. L., & Cordery, J. L. (2011). Investigating retesting effects in a personnel selection context. *International Journal of Selection and Assessment*, *19*(2), 217-221.

Echemendia, R. J., Bruce, J. M., Bailey, C. M., Sanders, J. F., Arnett, P., & Vargas, G. (2012). The utility of post-concussion neuropsychological data in identifying cognitive change following sports-related MTBI in the absence of baseline data. *The Clinical Neuropsychologist, 26,* 1077-1091.

Echemendia, R. J., Iverson, G. L., McCrea, M., Macciocchi, S. N., Gioia, G. A., Putukian, M., & Comper, P. (2013). Advances in neuropsychological assessment of sport-related concussion. *British Journal of Sports Medicine,47*(5), 294-298.

Elbin, R. J., Schatz, P., & Covassin, T. (2011). One-year test-retest reliability of the online version of ImPACT in high school athletes. *The American Journal of Sports Medicine*, *39*(11), 2319-2324. doi:10.1177/0363546511417173

Elbin, R. J., Kontos, A. P., Kegel, N., Johnson, E., Burkhart, S., & Schatz, P. (2013). Individual and combined effects of LD and ADHD on computerised neurocognitive concussion test performance: Evidence for separate norms. *Archives of Clinical Neuropsychology, 28*, 476-484.

Ellemberg, D., Henry, L. C., Macciocchi, S. N., Guskiewicz, K. M., & Broglio, S. P. (2009). Advances in sport concussion assessment: from behavioral to brain imaging measures. *Journal of Neurotrauma*, *26*(12), 2365-2382.

Erlanger, D. M., Kutner, K. C., Barth, J. T., & Barnes, R. (1999). Forum neuropsychology of sports-related head injury: Dementia pugilistica to post-concussion syndrome. *The Clinical Neuropsychologist*, *13*(2), 193-209.

Erlanger, D., Saliba, E., Barth, J., Almquist, J., Webright, W. & Freeman, J. (2001). Monitoring resolution of postconcussion symptoms in athletes: Preliminary results of web-based neuropsychological test protocol. *Journal of Athletic Training, 36*(3), 280-287.

Fazio, V. C., Lovell, M. R., Pardini, J. E., & Collins, M. W. (2007). The relation between post-concussion symptoms and neurocognitive performance in concussed athletes. *Neuro Rehabilitation*, *22*(3), 207-216.

Field, M., Collins, M. W., Lovell, M. R., & Maroon, J. (2003). Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. *The Journal of Paediatrics*, *142*(5), 546-553.

Field, A. (2009). *Discovering Statistics using SPSS* (3rd ed.). London, United Kingdom: Sage Publications Ltd.

Fischer, K.W., & Silvern, L. (1985). Stages and Individual Differences in Cognitive Development. *Annual Reviews in Psychology, 36*, 613- 648.

Fisher, J. M., & Vaca, F. E. (2004). Sport-related Concussions in the Emergency Department. *Advanced Emergency Nursing Journal*, *26*(3), 260-266.

Flin, R. H. (1985). Development of visual memory: An early adolescent regression. *The Journal of Early Adolescence*, *5*(2), 259-266.

Foulkes, L., & Blakemore, S. J. (2018). Studying individual differences in human adolescent brain development. *Nature Neuroscience, 21*(3), 315-323. doi: 10.1038/s41593-018-0078-4.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology*,*54*(1), 1-34.

Gardner, A., Shores, E. A., & Batchelor, J. (2010). Reduced processing speed in rugby union players reporting three or more previous concussions. *Archives of Clinical Neuropsychology, 25*(3), 174-181. Doi: 10.1093/arclin/acq007

Gessel, L. M., Field, S.K., Collins, C.L., Dick, R. W., & Comstock, R. (2007). Concussions among United States high school and collegiate athletes. *Journal of Athletic Training*, *42*(4), 495.

Giedd J, N., Blumenthal, J., Jeffries, N. O., Rajapakse, J. C., Vaituzis, A. C., Liu, H., … & Castellanos, F.X. (1999). Development of the human corpus callosum during childhood and adolescence: a longitudinal MRI study. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 23(4), 571–88.

Gioia, G. A., Schneider, J. C., Vaughan, C. G., & Isquith, P. K. (2009). Which symptom assessments and approaches are uniquely appropriate for paediatric concussion?. *British Journal of Sports Medicine*, *43*(Suppl 1), i13-i22.

Giza, C. C., & Hovda, D. A. (2001). The neurometabolic cascade of concussion. *Journal of Athletic Training*, *36*(3), 228-235.

Guskiewicz, K. M., McCrea, M., Marshall, S. W., Cantu, R. C., Randolph, C., Barr, W., ... & Kelly, J. P. (2003). Cumulative effects associated with recurrent concussion in collegiate football players: the NCAA Concussion Study. *Journal of the American Medical Association*, *290*(19), 2549-2555.

Guskiewicz, K. M., Bruce, S. L., Cantu, R. C., & Ferrara, M. S., Kelly, J.P., McCrea, M., … & McLeod, T.C. (2004). National Athletic Trainers' Association position statement: management of sport-related concussion. *Journal of Athletic Training*, *39*(3), 280-297.

Harmon, K. G., Drezner, J. A., Gammons, M., Guskiewicz, K. M., Halstead, M., Herring, S. A., ... & Roberts, W. O. (2013). American Medical Society for Sports Medicine position statement: concussion in sport. *British Journal of Sports Medicine*, *47*(1), 15-26.

Health Professions Council of South Africa. (2006). List of tests classified as being psychological tests. Professional Board of Psychology Form 207, 1-9. Ref Type: Bill/Resolution. *Health Professions Council of South Africa.* Retrieved October 2, 2018 from https://www.hpcsa.co.za/Uploads/editor/UserFiles/downloads/psych/psychom_form_207.pdf

Henry, L. C., & Sandel, N. (2015). Adolescent subtest norms for the ImPACT neurocognitive battery. *Applied Neuropsychology: Child*, *4*(4), 266-276.

Hillis, W. S., McIntyre, P. D., Maclean, J., Goodwin, J. F., & McKenna, W. J. (1994). ABC of sports medicine: Sudden death in sport. *British Medical Journal*, *309*(6955), 657-660. doi: 10.1136/bmj.309.6955.657

Hinton-Bayre, A. & Geffen, G. (2005). Comparability, Reliability, and Practice Effects on Alternate Forms of the Digit Symbol Substitution and Symbol Digit Modalities Test. *Psychological Assessment*, *17*(2), 237-241.

Hinton-Bayre, A. D. (2015). Normative Versus Baseline Paradigms for Detecting Neuropsychological Impairment Following Sports-Related Concussion. *Brain Impairment*, *16*(2), 80-89.

Holtzhausen, L., Schwellnus, M., Jakoet, I., Pretorius, A. (2002). Pre-season assessment of South African players in the 1999 rugby Super 12 competitions. *South African Journal of Sports Medicine, 9*, 15-21.

Horsman, M. (2010). *Concussion in Contact Sport: Investigating the Neurocognitive Profile of Afrikaans Adolescent Rugby Players.* (Unpublished Masters dissertation.) Rhodes University, Grahamstown, South Africa.

Hsu, L. M. (1996). Regression towards the mean associated with measurement error and the identification of improvement and deterioration in psychotherapy. *Journal of Consulting and Clinical Psychology*, *63*(1), 141- 144.

Huber, B.R., Stein, T. D., Alosco, M. L., & McKee, A.C. (2016). Potential Long-Term Consequences of Concussive and Subconcussive Injury. *Physical Medicine & Rehabilitation Clinics North America, 27*(2), 503-511. doi: 10.1016/j/pmr. 2015.12.007

Huberty, C. J., & Morris, J.D. (1989). Multivariate Analysis Versus Multiple Univariate Analysis. *Psychological Bulletin, 105*(2), 302-308.

Hunt, T. N., & Ferrara, M. S. (2009). Age-related differences in neuropsychological testing among high school athletes. *Journal of Athletic Training*, *44*(4), 405-409.

Ikeda, Y., Okuzumi, H., & Kokubun, M. (2013). *Stroop/reverse-Stroop interference in typical development and its relation to symptoms of ADHD. Research in Developmental Disabilities, 34*, 2391–2398. doi: 10.1016/j.ridd.2013.04.019

ImPACT Applications. (2004). ImPACT version 2.0 clinical user's manual. *ImPACT Applications, Inc.* Retrieved February 18, 2016 from https://www.impacttest.com/dload/ImPACT21_usermanual.pdf

ImPACT Applications. (2007). ImPACT 2007 (6.0) Clinical Interpretation Manual: Fall 2007. *ImPACT Applications, Inc.* Retrieved February 10, 2016 from https://www.sjschools.org/images/Athletics_HS/ImPACT_Clinical_Interpretation_Manual.pdf

ImPACT Applications. (2011). Technical manual- online ImPACT 2007-2012. *ImPACT Applications, Inc.* Retrieved February 10, 2016 from http://impacttest.com/pdf/ImPACT-TechnicalManual.pdf.

ImPACT online. (2019a). ImPACT Applications: A Story of Excellence. *ImPACT Applications, Inc.* Retrieved January 4, 2019 from https://concussioncareresources.com/our-story/

ImPACT online. (2019b). Buyer's Guide. *ImPACT Applications, Inc*. Retrieved January 4, 2019 from http://impacttest.com/why-choose-impact-applications-tools-for-concussion-management.

ImPACT online. (2019c). ImPACT Pediatric. *ImPACT Applications, Inc.* Retrieved January 4, 2019 from https://pediatric.impacttest.com/.

ImPACT online. (2019d). Concussion Protocol Basics: What is baseline cognitive testing. *ImPACT Applications, Inc*. Retrieved January 4, 2019 from https://impacttest.com/concussion-protocol-101-guide/#what_is_baseline_cognitive_testing

Institute of Medicine and National Research Council. (2013). *Sports-related concussions in youth: improving the science, changing the culture.* Washington, DC: The National Academies Press.

Isbell, E., Fukuda, K., Neville, H. J., & Vogel, E. K. (2015). Visual working memory continues to develop through adolescence. *Frontiers in Psychology*, *6*(696), 1-10, doi: 10.3389/fpsyg.2015.00696

Iverson, G. L., Lovell, M. R., & Collins, M. W. (2003). Interpreting change on ImPACT following sport concussion. *The Clinical Neuropsychologist, 17*(4), 460-467.

Iverson, G. L., Gaetz, M., Lovell, M. R., & Collins, M. W. (2004). Cumulative effects of concussion in amateur athletes. *Brain Injury, 18*(5), 433-443.

Iverson, G. L., Lovell, M. R., & Collins, M. W. (2005). Validity of ImPACT for measuring processing speed following sports-related concussion. *Journal of Clinical and Experimental Neuropsychology*, *27*(6), 683-689.

Iverson, G. L., Brooks, B. L., Collins, M.W., Lovell, M.R. (2006). Tracking neuropsychological recovery following concussion in sport. *Brain Injury, 20*, 245-252.

Iverson, G.L. (2011). Sport-Related Concussion. In M.R. Schoenberg and J.G. Scott (Eds.), *The Little Black Book of Neuropsychology: A Syndrome-Based Approach* (pp. 721-744). New York, NY: Springer Publications.

Iverson, G. L., & Schatz, P. (2015). Advanced topics in neuropsychological assessment following sport-related concussion. *Brain Injury*, *29*(2), 263-275.

Johnson, S. B., Blum, R.W., Giedd, J. N. (2009). Adolescent Maturity and the Brain: The Promise and Pitfulls of Neuroscience Research in Adolescent Health Policy. *Journal of Adolescent Health, 45*(3), 216-221. Doi: 10.1016/j.jadohealth.2009.05.016

Jordan, A. B. (1997). *The Shuttle Effect: The development of a model for the prediction of variability in cognitive test performance across the adult life span.* (Unpublished PhD dissertation.) Rhodes University, Grahamstown, South Africa.

Jotwani, V. & Harmon, K. G. (2010). Postconcussion syndrome in athletes. *Current Sports Medicine Reports, 9*(1), 21-26. Doi: 10.1249/JSR.0b013e3181ccbsse

Kelly, J. P. (2001). Loss of consciousness: pathophysiology and implications in grading and safe return to play. *Journal of Athletic Training, 36*(3), 249-252.

Kemp, S.P., Hudson, Z., Brooks, J. H., Fuller, C. W. (2008). The epidemiology of head injuries in English professional rugby union. *Clinical Journal of Sports Medicine, 18* (3), 277-234. doi: 10.1097/JSM. 0b013e31816a1c9a

Kieslich, M., A. Fielder, C. Heller, W. Kreuz, and G. Jacobi. (2002). Minor head injury as cause and co-factor in the aetiology of stroke in children: A report of eight cases. *Journal of Neurology, Neurosurgery, and Psychiatry, 73*, 13-16.

Kinnaman, K. A., Mannix, R. C., Comstock, R. D., & Meehan, W. P. (2014). Management of paediatric patients with concussion by emergency medicine physicians. *Pediatric Emergency Care, 30*(7), 458-461.

Kirkwood, M., Randolph, C., & Yeates, K. (2009). Returning pediatric athletes to play after concussion: the evidence (or lack thereof) behind baseline neuropsychological testing. *Acta Paediatrica, 98*(9), 1409-1411. doi: 10.1111/j.1651-2227.2009.01448.x

Konrad K, Firk C, Uhlhaas P. (2013). Brain development during adolescence: neuroscientific insights into this developmental period. *Deutsches Ärzteblatt International, 110*(25), 425–31.

Kumar Sharma, V., Kumar Subramanian, S., Vinayathan, A., Sarah, R., Balasubramaniam, S. R., & Velkumary, S. (2014). Study of Effect of Age and Gender Related Differences on Common Paper and Pencil Neurocognitive Tests in Adolescents. *Journal of Clinical and Diagnostic Research, 8*(11), 5-10.

Lang, D. A., Teasdale, G. M., Macpherson, P., & Lawrence, A. (1994). Diffuse brain swelling after head injury: more often malignant in adults than children?. *Journal of Neurosurgery, 80*(4), 675-680.

Leclerc, S., Lassonde, M., Delaney, J. S., Lacroix, V. J., & Johnston, K. M. (2001). Recommendations for grading of concussion in athletes. *Sports Medicine, 31*(8), 629-636.

Lezak, M. D., Howieson, D. B., Bugler, E. D., & Tranel, D. (2012). *Neuropsychological assessment*. (5th ed.). New York, NY: Oxford University Press.

Lichtenstein, J., Moser, R., Schatz, P. (2013). Age and test setting affect the prevalence of invalid baseline scores on neurocognitive tests. *The American Journal of Sports Medicine, 42*(2), 479-484.

Littleton, A. C., Register-Mihalik, J. K., & Guskiewicz, K. M. (2015). Test-Retest Reliability of a Computerized Concussion Test: CNS Vital Signs. *Sports Health*, 7(5), 443-447. doi: 10.1177/19417381155869971941738115586997.

Louey, A. G., Cromer, J. A., Schembri, A. J., Darby, D. G., Maruff, P., Makdissi, M., … & McCrory, P. (2014). Detecting cognitive impairment after concussion: Sensitivity of change from baseline and normative data methods using the Cogsport/Axon cognitive test battery. *Archives of Clinical Neuropsychology, 29*, 432-441.

Lovell, M. R., & Collins, M. W. (1998). Neuropsychological assessment of the college football player. *The Journal of Head Trauma Rehabilitation, 13*(2), 9-26.

Lovell, M. R. (1999). Evaluation of the professional athlete. In J.E. Bailes, M. Lovell, & J. Maroon (Eds.), *Sports-Related Concussion* (pp. 200-214). New York, NY: CRC Press.

Lovell, M. R., Iverson, G. L., Collins, M. W., McKeag, D., & Maroon, J. C. (1999). Does Loss of Consciousness Predict Neuropsychological Decrements After Concussion?. *Clinical Journal of Sport Medicine*, *9*(4), 193-198.

Lovell, M.R., & Burke, C., J. (2000). Concussion management in professional hockey. In R. Cantu (Ed.), *Neurological athletic head and spine injury* (pp. 109-116). Philadelphia, PA: WB Saunders.

Lovell, M. R., Collins, M. W., Iverson, G. L., Field, M., Maroon, J. C., Cantu, R., ... & Fu, F. H. (2003). Recovery from mild concussion in high school athletes. *Journal of Neurosurgery, 98*(2), 296-301.

Lovell, M. R., Collins, M. W., Iverson, G. L., Johnston, K. M., & Bradley, J. P. (2004). Grade 1 or “ding” concussions in high school athletes. *The American Journal of Sports Medicine*, *32*(1), 47-54.

Lovell, M. R. (2006). The ImPACT Neuropsychological Test Battery. In R. J. Echemendia (Ed.), *Sport Neuropsychology: Assessment and Management of Traumatic Brain Injury* (pp. 193-215). New York, NY: Guilford Press.

Lovell, M. (2009). The management of sports-related concussion: current status and future trends. *Clinics in Sports Medicine*, *28*(1), 95-111.

Lovell, M. R., & Pardini, J. E. (2010). Neuropsychological assessment and sports-related mild traumatic brain injury (concussion). In T.D. Marcotte & I Grant (Eds.), *Neuropsychology of everyday functioning* (pp. 331-356). New York, NY: Guilford Press.

Lovell, M. R., & Solomon, G. S. (2011). Psychometric data for the NFL neuropsychological test battery. *Applied Neuropsychology*, *18*(3), 197-209.

Luria, A.R. (1980). *Higher Cortical Functions in Man*. New York, NY: Basic Books.

Macciocchi, S. N., Barth, J. T., Alves, W., Rimel, R. W., & Jane, J. A. (1996). Neuropsychological functioning and recovery after mild head injury in collegiate athletes. *Neurosurgery*, *39*(3), 510-514.

Macciocchi, S. N., Barth, J. T., Littlefield, L., & Cantu, R. C. (2001). Multiple concussions and neuropsychological functioning in collegiate football players. *Journal of Athletic Training, 36*(3), 303-306.

MacDonald, J., & Duerson, D. (2015). Reliability of a computerized neurocognitive test in baseline concussion testing of high school athletes. *Clinical Journal of Sport Medicine, 25*(4), 367-372.

Macleod, D. (1993). Risks & injuries in rugby football. In G.R. McLatchie, & C.M.E. Lennox (Eds.), *The soft tissues: trauma & sports injuries* (p. 371-381). London, UK: Butterworth Heinemann Ltd.

Maerlender, A., & Molfese, D. L. (2015). Repeat baseline assessment in college-age athletes. *Developmental Neuropsychology, 40*(2), 69-73. doi: 10.1080/87565641.2015.1014089

Maerlender, A. C., Masterson, C. J., James, T. D., Beckwith, J., Brolinson, P. G., … McAllister, T.W. (2016). Test-retest, retest: Growth curve models of repeat testing with Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT). *Journal of Clinical and Experimental Neuropsychology, 38*(8), 869-874, doi: 10.1080.13803395.2016.1168781.

Makdissi, M., Collie, A., Maruff, P., Darby, D. G., Bush, A., McCrory, P., & Bennell, K. (2001). Computerised cognitive assessment of concussed Australian Rules footballers. *British Journal of Sports Medicine, 35*(5), 354-360.

Marar, M., McIlvain, N. M., Fields, S. K., & Comstock, R. D. (2012). Epidemiology of concussions among United States high school athletes in 20 sports. *The American Journal of Sports Medicine*, *40*(4), 747-755.

Maruff, P., Collie, A., Anderson, V., Mollica, C., McStephen, M., & McCrory, P. (2004). Cognitive development in children: Implications for concussion management. *British Journal of Sports Medicine, 38*, 654-655.

Matser, J.T., Kessels, A.G., Lezak, M., Jordan, B.D., & Troost, J. (1999). Neuropsychological impairment in amateur soccer players. *Journal of the American Medical Association, 282(10)*, 971-973.

McBride, D.M. (2013). *The process of research in psychology (2nd ed.)* Thousand Oaks, CA: Sage Publications, Inc.

McCaffrey, R. J., Duff, K., & Westervelt, H. J. (Eds.). (2000). *Practitioner's guide to evaluating change with neuropsychological assessment instruments*. New York, NY: Springer Science & Business Media.

McClincy, M. P., Lovell, M. R., Pardini, J., Collins, M. W., & Spore, M. K. (2006). Recovery from sports concussion in high school and collegiate athletes. *Brain Injury*, *20*(1), 33-39.

McCrea, M., Hammeke, T., Olsen, G., Leo, P., & Guskiewicz, K. (2004). Unreported concussion in high school football players: implications for prevention. *Clinical Journal of Sport Medicine*, *14*(1), 13-17.

McCrea, M., Guskiewicz, K., Marshall, S., Barr, W., Randolph, C., Cantu, R. C., & Kelly, J. P. (2003). Acute effects and recovery time following concussion in collegiate football players: the NCAA concussion study. *Journal of the American Medical Association, 290*(19), 2556-2563. doi: 10.1001/jama.290.19.2556

McCrory, P. R., Bladin, P. F., & Berkovic, S. F. (1997). Retrospective study of concussive convulsions in elite Australian rules and rugby league footballers: phenomenology, aetiology, and outcome. *British Medical Journal*, *314*(7075), 171-174.

McCrory, P. (2001). Does second impact syndrome exist?. *Clinical Journal of Sport Medicine*, *11*(3), 144-149.

McCrory, P., Collie, A., Anderson, V., & Davis, G. (2004). Can we manage sport related concussion in children the same as in adults? *British Journal of Sports Medicine*, *38*(5), 516-519.

McCrory, P., Johnston, K., Meeuwisse, W., Aubry, M., Cantu, R., Dvorak, J., ... & Schamasch, P. (2005). Summary and Agreement Statement of the 2nd International Conference on Concussion in Sport, Prague 2004. *British Journal of Sports Medicine*, *39*(4), 196-204.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., & Cantu, R. (2009). Consensus Statement on Concussion in Sport: The 3rd International Conference on Concussion in Sport held in Zurich, November 2008. *Journal of Athletic Training, 44*(4), 434-448. doi: 10.4085/1062-6050-44.4.434

McCrory, P., Meeuwisse, W., Aubry, M., Cantu, B., Dvorak, J., Echemendia, R. J., & Turner, M. (2013). Consensus Statement on Concussion in Sport: The 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, *47*(5), 250-258. doi: 10.1136/bjsports-2013-092313

McCrory, P., Meeuwisse, W. H., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., … & Vos, P. (2017). Consensus Statement on Concussion in Sport: The 5th International Conference on Concussion in Sport held in Berlin, October 2016. *British Journal of Sports Medicine, 51*(11), 838-847. doi. 10.1136/bjsports-2017-097699

McKee, A. C., Cantu, R. C., Nowinski, C. J., Hedley-Whyte, E. T., Gavett, B. E., Budson, A. E., ... & Stern, R. A. (2009). Chronic traumatic encephalopathy in athletes: progressive tauopathy after repetitive head injury. *Journal of Neuropathology & Experimental Neurology*, *68*(7), 709-735.

McKee, A. C., Stern, R. A., Nowinski, C. J., Stein, T. D., Alvarez, V. E., Daneshvar, D. H., … & Cantu, R. C. (2013). The spectrum of disease in chronic traumatic encephalopathy. *Brain, 136*, 43–64. doi: 10.1093/brain/aws307

McKinlay, A. (2010) Controversies and outcomes associated with mild traumatic brain injury in childhood and adolescences. *Child: Care, Health and Development, 36*(1), 3-21.

Meehan, W. P., & Bachur, R. G. (2009). Sport-related concussion. *Pediatrics, 123*(1), 114-123.

Meehan, W., d'Hemecourt, P., Collins, C., Taylor, A., & Comstock, R. (2011). Computerised neurocognitive testing for the management of sport-related concussions. *Pediatrics, 124*(1), 38-44.

Merritt, V.C., Meyer, J.E., Cadden, M.H., Roman, C.A.F., Ukueberuwa, D.M., Shapiro, M.D., Arnett, P.A. (2016). Normative Data for a Comprehensive Neuropsychological Test Battery used in the Assessment of Sports-Related Concussion. *Archives of Clinical Neuropsychology, 32*(2), 168-183. doi: 10.1093/arclin/acw090

Mez, J., Daneshvar, D. H., Kiernan., Abdolmohannadi, B., Alvarez, V. E., Huber, B.R., …& McKee, A. (2017). Clinicopathological Evaluation of Chronic Traumatic Encephalopathy in Players of American Football. *Journal of the American Medical Association*, *318*(4), 360-370. doi: 10.1001/jama.2017.8334.

Mihalik, J. P. (2012). Biomechanics of sports concussion. In M.W. Kirkwood & K. O. Yeates (Eds.), *Mild Traumatic Brain Injury in Children and Adolescents: From Basic Science to Clinical Management* (pp. 38-52). New York, NY: The Guildford Press.

Miller, J. R., Adamson, G. J., Pink, M. M., Sweet, J. C. (2007). Comparison of Preseason, Midseason, and Postseason Neurocognitive Scores in Uninjured Collegiate Football Players. *The American Journal of Sports Medicine, 35*(8), 1284-1288. doi: 10.1177/0363546507300261.

Moser, R. S., Schatz, P., & Jordan, B. D. (2005). Prolonged effects of concussion in high school athletes. *Neurosurgery, 57*(2), 300-306.

Moser, R. S. (2007). The growing public health concern of sports concussion: The new psychology practice frontier. *Professional Psychology: Research and Practice, 38*(6), 699-704.

Moser, R. S., Iverson, G. L., Echemendia, R. J., Lovell, M. R., Schatz, P., Webbe, F. M., ... & Bush, S. S. (2007). Neuropsychological evaluation in the diagnosis and management of sports-related concussion. *Archives of Clinical Neuropsychology, 22*(8), 909-916. doi: 10.1016/j.acn.2007.09.004

Moser, R. S., Schatz, P., & Lichtenstein, J. D. (2015). The importance of proper administration and interpretation of neuropsychological baseline and postconcussion computerized testing. *Applied Neuropsychology: Child, 4*(1), 41-48.

Nagy, Z., Westerberg, H., & Klingberg, T. (2004). Maturation of white matter is associated with the development of cognitive functions during childhood. *Journal of Cognitive Neuroscience, 16*(7), 1227-1233.

Naidoo, R., Shuttleworth-Edwards, A.B., Botha, R., & Pienaar, I. (2019). The Tower of London-DX 2nd Edition test: Preliminary norms for educationally disadvantaged Xhosa-speaking individuals. *Journal of Psychology in Africa, 29*(1), 60-66, doi: 10.1080/14330237.2019.1568074

Nakayama, Y., Covassin, T., Schatz, P., Nogle, S., & Kovan, J. (2014). Examination of the test-retest reliability of a computerized neurocognitive test battery. *The American Journal of Sports Medicine, 42*(8), 2000-2005, doi:10.1177/0363546514535901

Nathan, M., Goedeke, R., & Noakes, T. D. (1983). The incidence and nature of rugby injuries experienced at one school during the 1982 rugby season. *South African Medical Journal, 64*(4), 132-137.

Nelson, L., La Roche, A., Pfaller, A., Lerner, E., Hammeke, T., Randolph, C., Barr, W., Guskiewicz, K., McCrea, M. (2016). Prospective, head-to-head study of three computerised neurocognitive tools (CNTs): reliability and validity for the assessment of sport concussion. *Journal of the International Neuropsychological Society, 22*, 33-37.

Nell, V., & Brown, D. S. (1991). Epidemiology of traumatic brain injury in Johannesburg--II. Morbidity, mortality and etiology. *Social Science Medicine, 33*(3), 289-296.

Newell, A., & Rosenbloom, P. S. (1981). Mechanism of skill acquisition and the law of practice. In J. R. Anderson (Ed.). *Cognitive skills and their acquisition* (pp.1-55). Hillsdale, NJ: Erlbaum.

Olvey, S. (2002, July). *Concussion in motorsports*. Paper presented at the New Developments in Sports-related Concussion Conference, Pittsburgh, Pennsylvania.

Omalu, B., Bailes, J., Hamilton, R. L., Kamboh, M. L., Hammers, J., Case, M., Fitzsimmons, R. (2011). Emerging Histomorphologic Phenotypes of Chronic Traumatic Encephalopathy in American Athletes. *Neurosurgery*, *69*(1), 173-183. doi: 10.1227/NEU.0b013e218212bc7b

Ommaya, A. K., Goldsmith, W., & Thibault, L. and L. (2002) Biomechanics and neuropathology of adult and paediatric head injury. *British Journal of Neurosurgery, 16*(3), 220-242.

Patel, D. R., Shivdasani, V., & Baker, R. J. (2005). Management of sport-related concussion in young athletes. *Sports Medicine, 35*(8), 671-684.

Patoine, B (2010). Sports Concussions & The Immature Brain (2): Young Athletes May Be More Vulnerable to Mild Brain Injury. *The DANA Foundation.* Retrieved October 10, 2018 from http://www.dana.org/News/Details.aspx?id=43489

Patricios, J.S., Kohler, R.M.N., & Collins, R. (2010). Boksmart 2010: Concussion in Rugby Literature Review. *Boksmart*. Retrieved October 10, 2018 from http://www.boksmart.co.za/pdf/BokSmart%202010-concussion%20in%20Rugby%20Literature%20Review.pdf

Perneger, T.V. (1998). What's wrong with Bonferroni adjustments? *British Medical Journal, 31*(6), 1236-1238.

Powell, J. W., & Barber-Foss, K. D. (1999). Traumatic brain injury in high school athletes. *Journal of the American Medical Association, 282*(10), 958-963.

Proctor, M. R., & Cantu, R. C. (2000). Head and neck injuries in young athletes. *Clinics in Sports Medicine*, *19*(4), 693-715.

Puddey, I. B., Mercer, A., Andrich, D., & Styles, I. (2014). Practice effects in medical school entrance testing with the undergraduate medicine and health sciences admission test (UMAT). *Medical Education, 14*, 48-62.

Randolph, C., McCrea, M., Barr, W. B., & Macciocchi, S. N. (2005). Is Neuropsychological Testing Useful in the Management of Sport-Related Concussion? *Journal of Athletic Training*, *40*(3), 139-154.

Randolph, C. (2011). Baseline neuropsychological testing in managing sport-related concussion: does it modify risk?. *Current Sports Medicine Reports*, *10*(1), 21-26.

Rapport, L. J., Brines, D. B., Axelrod, B. N., & Theisen, M. E. (1997). Full scale IQ as a mediator of practice effects: The rich get richer. *The Clinical Neuropsychologist, 11* (4), 375–380.

Register-Mihalik, J. K., Kontos, D. L., Guskiewicz, K. M., Mihalik, J. P., Conder, R., & Shields, E.W. (2012). Age-related differences and reliability on computerized and paper-and-pencil neurocognitive assessment batteries. *Journal of Athletic Training, 47*(3), 297-305.

Reitan, R. M., & Wolfson, D. (1999). The two faces of mild head injury. *Archives of Clinical Neuropsychology*, *14*(2), 191-202.

Resch, J., Driscoll, A., McCaffrey, N., Brown, C., Ferrara, M. S., Macciocchi, S., …& Walpert, K. (2013a). ImPACT test-retest reliability: Reliably unreliable? *Journal of Athletic Training, 48*(4), 506-511.

Resch, J. E., McCrea, M. A., & Cullum, C. M. (2013b). Computerized neurocognitive testing in the management of sport-related concussion: An update. *Neuropsychology Review*, *23*(4), 335-349.

Resch, J.E., Schneider, M. W., Cullum, C. M. (2018). The test-retest reliability of three computerised neurocognitive tests used in the assessment of sport concussion. International *Journal of Psychophysiology, 132*, 31-38. doi.org/10.1016/j.ijpsycho.2017.09.011.

Reynolds, E., Fazio, V. C., Sandel, N., Schatz, P., & Henry, L. C. (2016). Cognitive development and the immediate postconcussion assessment and cognitive testing: a case for separate norms in preadolescents. *Applied Neuropsychology: Child*, *0,* 1-11.

Rogers, S., Smith, P. J., Stephenson, A., Everhart, D. (2017). A Retrospective Cross-Sectional and Longitudinal Study of the Effects of Age on CNS Vital Signs Scores in High-School Athletes, *Sports Medicine*, *47*, 1893-1899, doi:10.1007/s40279-017-0686-2

Roux, C.E., Goedeke, R., Visser, G. R., van Zyl, W.A., & Noakes, T.D. (1987). The epidemiology of schoolboy rugby injuries. *South African Medical Journal, 71*(5), 307-313.

Salmon-Godlo, N.C. (2006). *The Establishment of Normative Data on Xhosa-Speaking High School Learners using the ImPACT 3.0 Programme*. (Unpublished Masters dissertation.) Rhodes University, Grahamstown, South Africa.

Salthouse, T.A., Schroeder, D. H., & Ferrer, E. (2004). Establishing retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. *Developmental Psychology*, *40*(5), 813-822.

Salthouse, T. A., & Tucker-Drob, E. M. (2008). Implications of short-term retest effects for . the interpretation of longitudinal change. *Neuropsychology*, *22*(6), 800-811.

Salthouse, T. A. (2010). Influence of age on practice effects in longitudinal neurocognitive change. *Neuropsychology*, *24*(5), 563-572.

Satz, P. (1993). Brain reserve capacity on symptom onset after brain injury: a formulation and review of evidence for threshold theory. *Neuropsychology, 7*(3), 273-295.

Sim, A., Terryberry-Spohr, L., Wilson, K.R. (2008). Prolonged recovery of memory functioning after mild traumatic brain injury in adolescent athletes. *Journal of Neurosurgery, 108*, 511-516.

Scharfen, J., Peters, J. M., Holling, H. (2018). Retest effects in cognitive ability tests: A meta-analysis. *Intelligence, 67*, 44-66, doi: 10.1016/j.intell.2018.01.003.

Schatz, P., & Browndyke, J. (2002). Applications of computer-based neuropsychological assessment. *The Journal of Head Trauma Rehabilitation*,*17*(5), 395-410.

Schatz, P., & Zillmer, E. A. (2003). Computer-based assessment of sports-related concussion. *Applied Neuropsychology*, *10*(1), 42-47.

Schatz, P., Pardini, J. E., Lovell, M. R., Collins, M. W., & Podell, K. (2006). Sensitivity and specificity of the ImPACT Test Battery for concussion in athletes. *Archives of Clinical Neuropsychology*, *21*(1), 91-99.

Schatz, P. (2010). Long-term test-retest reliability of baseline cognitive assessments using ImPACT. *The American Journal of Sports Medicine*, *38*(1), 47-53. doi:10.1177/0363546509343805

Schatz, P., & Moser, R. S. (2011). Current issues in pediatric sports concussion. *The Clinical Neuropsychologist*, *25*(6), 1042-1057.

Schatz, P., Moser, R. S., Covassin, T., & Karpf, R. (2011). Early indicators of enduring symptoms in high school athletes with multiple previous concussions. *Neurosurgery, 68*(6), 1562-1567.

Schatz, P., & Ferris, C. S. (2013). One-month test-retest reliability of the ImPACT test battery. *Archives of Clinical Neuropsychology: The Official Journal of the National Academy of Neuropsychologists, 28*(5), 499-504. doi:10.1093/arclin/act034 [doi]

Schmidt, J. D., Register-Mihalik, J. K., Mihalik, J. P., Kerr, Z. Y., & Guskiewicz, K. M. (2012). Identifying Impairments after Concussion: Normative Data versus Individualised Baselines. *Journal of the American College of Sports Medicine, 44*(9), 1621-1628. doi: 10.1249/MMS.0b013e318258a9fb.

Shuttleworth-Edwards, A. B., Border, M., Reid, I., & Radloff, S. (2004a). South African Rugby Union. In M.R. Lovell, R.E. Echemendia, J.T. Barth, & M.W. Collins (Eds.), *Traumatic Brain Injury in Sports: An International Neuropsychological Perspective* (pp. 149-168). The Netherlands: Swets & Zeitlinger.

Shuttleworth-Edwards, A. B., Kemp, R. D., Rust, A. L., Muirhead, J. G., Hartman, N. P., & Radloff, S. E. (2004b). Cross-cultural effects on IQ test performance: A review and preliminary normative indications on WAIS-III test performance. *Journal of Clinical and Experimental Neuropsychology, 26*(7), 903-920.

Shuttleworth-Edwards, A. B., & Whitefield, V. J. (2007). Ethically we can no longer sit on the fence-a neuropsychological perspective on the cerebrally hazardous contact sports. *South African Journal of Sports Medicine*, *19*(2), 32-38.

Shuttleworth-Edwards, A. B., & Radloff, S. E. (2008). Compromised visuomotor processing speed in players of Rugby Union from school through to the national adult level. *Archives of Clinical Neuropsychology*, *23*(5), 511-520.

Shuttleworth-Edwards, A. B., Noakes, T. D., Radloff, S. E., Whitefield, V. J., Clark, S. B., Roberts, C. O., ... & Smith, I. P. (2008a). The comparative incidence of reported concussions presenting for follow-up management in South African Rugby Union. *Clinical Journal of Sport Medicine*, *18*(5), 403-409.

Shuttleworth-Edwards, A.B., Smith, I., & Radloff, S. (2008b). Neurocognitive vulnerability amongst university rugby players versus non-contact sport controls. *Journal of Clinical and Experimental Neuropsychology*, *30(8)*, 870-884.

Shuttleworth-Edwards, A. B., Whitefield-Alexander, V. J., Radloff, S. E., Taylor, A. M., & Lovell, M. R. (2009). Computerized neuropsychological profiles of South African versus US athletes: A basis for commentary on cross-cultural norming issues in the sports concussion arena. *The Physician and Sports Medicine*, *37*(4), 45-52.

Shuttleworth-Edwards, A. B. (2011). Response to the Article on Baseline Neuropsychological Testing: Throwing Away Clinical Gold with the Statistical Bathwater. *Current Sports Medicine Reports*, *10*(6), 391-392.

Shuttleworth-Edwards, A. B., Whitefield-Alexander, V. J., Radloff, S.E. (2013). The ImPACT neurocognitive screening test: A Survey of South African research including current and projected applications. In: S. Laher & K. Cockcroft (Eds.), *Psychological Assessment in South Africa: Research and Application* (pp. 443-460). Johannesburg, South Africa: Wits University Press.

Smith, A. (1991). *Symbol Digit Modalities Test*. Los Angeles, CA: Western Psychological Services.

Sowell, E. R., Thompson, P. M., Holmes, C. J., Jernigan, T. L., & Toga, A. W. (1999). In vivo evidence for post-adolescent brain maturation in frontal and striatal regions. *Nature Neuroscience*, *2*(10), 859-861.

Statistical Solutions. (2019). Bonferroni Correction. *Statistical Solutions*. Retrieved August 3, 2019 from https://www.statisticssolutions.com/bonferroni-correction/

Stevens, M. C., Skudlarski, P., Pearlson, G. D., & Calhoun, V. D. (2009). Age-related cognitive gains are mediated by the effects of white matter development on brain network integration. *Neuroimage*, *48*(4), 738-746.

Stroop, J. (1935). Studies of interference in several verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643-662. doi: 10.1037/h0054651

Strauss, E., Sherman, E. M., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd ed.). New York, NY: Oxford University Press.

Sukel, K. (2017). When is the Brain “Mature”? *The DANA Foundation.* Retrieved March 29, 2019 from https://www.dana.org/article/when-is-the-brain-mature/

Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., … & Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *Journal of Neuroscience, 37*, 3402-3412.

Tamnes, C. K., Bos, M.G.N., van de Kamp, F.C., Peters, S. & Crone, E.A. (2018). Longitudinal development of hippocampal subregions from childhood to adulthood. *Developmental Cogntive Neuroscience, 30*, 212-222.doi: 10.1106/j.dcn.2018.03.009

Terre Blanche, M., Durrheim, K & Painter, D. (2006). *Research in Practice: Applied Methods for the Social Sciences.* Cape Town, South Africa: UCT Press.

Toledo, E., Lebel, A., Becerra, L., Minster, A., Linnman, C., Maleki, N., …& Borsook, D. (2012). The young brain and concussion: Imaging as a biomarker for diagnosis and prognosis. *Neuroscience and Biobehavioural Review, 3*(6), 1510-1531. doi: 10.1016/j.neubiorev.2012.03.007

Tsushima, W. T., Siu, A. M., Pearce, A. M., Zhang, G., & Oshiro, R. S. (2016). Two-year Test–Retest Reliability of ImPACT in High School Athletes. *Archives of Clinical Neuropsychology*, *31*(1), 105-111.

Udall, T. (2011). Tackle the concussion epidemic. *CNN.com.* Retrieved November 7, 2018 from https://edition.cnn.com/2011/10/05/opinion/udall-football-concussions/

UPMC Sports Medicine, (2019). Concussion Facts and Statistics. *UPMC Sports Medicine.* Retrieved January 2, 2019, from http://www.upmc.com/services/sports-medicing/services/concussion/facts-statistics

Van Kampen, D. A., Lovell, M. R., Pardini, J. E., Collins, M. W., & Fu, F. H. (2006). The "value added" of neurocognitive testing after sports-related concussion. *The American Journal of Sports Medicine, 34*(10), 1630-1635. doi:0363546506288677

Van Ommen, C. (2013). Resisting the obfuscation of ‘race’: The importance of acknowledging the socio-political in normative data research. *South African Journal of Psychology*, *43*(2), 198-207.

Vijayakumar, N., Allen, N. B., Youssef, G., Dennison, M., Yucel, M., Simmons, J. G., … & Whittle, S. (2016). Brain development during adolescence: A mixed-longitudinal investigation of cortical thickness, surface area, and volume. *Human Brain Mapping, 37*(6), 2027-2038.

Wierenga, L. et al. (2014). Typical development of basal ganglia, hippocampus, amygdala and cerebellum from age 7 to 24. *NeuroImage, 96*, 67-72.

Wechsler, D. (2008). *WAIS-IV administration and scoring manual.* San Antonio, TX: Psychological Corporation.

Wekesa, M., Asembo, J. M., & Njororai, W. W. (1996). Injury surveillance in a rugby tournament. *British Journal of Sports Medicine*, *30*(1), 61-63.

Wills, S. M., & Leathem, J. M. (2001). An investigation of brain injury incurred in New Zealand club-grade rugby. *Journal of the International Neuropsychological Society, 7,* 405.

Womble, M.N., Reynolds, E., Schatz, P., Shah, K. M., Kontos, A. P. (2016). Test-retest Reliability of Computerized Neurocognitive Testing in Youth Ice Hockey Players. *Archive of Clinical Neuropsychology.* doi: 10.1093/arclin/acw011

Wurz, C. (2016). *An investigation into the utility of the ImPACT neurocognitive screening toll with patients diagnosed with Multiple Sclerosis.* (Unpublished Masters dissertation.) Rhodes University, Grahamstown, South Africa.

Yeudall, L.T., Reddon, J. R., Gill, D. M., Stefanyk, W.O. (1987). Normative data for the Halstead-Reitan neuropsychological tests stratified by age and sex. *Journal of Clinical Psychology, 43*(3), 346-367.

Yurgelun-Todd, D. (2007). Emotional and cognitive changes during adolescence. *Current Opinion in Neurobiology, 17*(2), 251-257.

Zemper, E. D. (2003). Two-year prospective study of relative risk of a second cerebral concussion. *American Journal of Physical Medicine & Rehabilitation, 82*(9), 653-659.

APPENDIXES

Appendix A

ImPACT Contract Agreement

ImPACT™
SOUTH AFRICA

The Best Approach to
Concussion
Management

CONTRACT AGREEMENT (Year) – ANNUAL SUBSCRIPTION FOR THE ImPACT© CONCUSSION MANAGEMENT PROGRAMME (ImPACT©) TO BE USED BY (Insert School Name)

*(Payment from the organization is due to ImPACT Applications-SA at the completion of baselining, by cheque or via direct deposit, at **First National Bank Branch code 210717; Account Number 62085086760**)*

In signing this document, **THE HEADMASTER** duly authorised, agrees on behalf of …………………….. to accept all conditions of ImPACT Applications- SA listed below. Violation of any portion of this agreement could result in immediate termination of this agreement and removal of …………………….. from the Programme.

1. As ImPACT Application's Software (ImPACT) is protected by copyright and is being provided only for use at …………………….., the school will not copy, alter or distribute the Software without written permission from ImPACT Applications-SA. This agreement allows the school to download the ImPACT software from the Website below, solely for the purposes

http://www.impacttestonline.com/southafrica

of concussion management at the school. Use of the Software other than for the purposes of concussion management and/or outside …………………….. is strictly prohibited.

2. ………………………… understands that all information gathered during evaluation using the Software needs to be strictly controlled for confidentiality, in accordance with normal ethical procedures for the protection of clinical and research material. …………………….. further understands and agrees that the information gathered from the Software may be utilized in studies currently being conducted in collaboration with the University of Pittsburgh Medical School, Sports Concussion Programme and the National Sports Concussion Initiative, Psychology Department, Rhodes University.

3. All ImPACT tests downloaded from the Website for baseline testing and/or follow-up testing and administered for the purposes of concussion management only at ………………………… will be supervised under strict testing conditions in accordance with appropriate testing procedures.

4. ……………………….. agrees to the following annual subscription, that is payable by ………………………… to ImPACT Applications – SA at the completion of baselining which will be completed no later than ……………………:

R………. per individual for a minimum of at least……..individuals enlisted in the programme.

5. ………………………… understands that the contract includes the price of the Software (including the baseline test and multiple follow-up tests), interpretation of the ImPACT data by an ImPACT accredited consultant and ImPACT technical support from www.impacttest.com for the duration of the year ……………………… has made payment. This includes all upgrades, updates and add-on at no cost during the year that ……………………….. has made payment.

6. We understand and accept that ImPACT Applications -SA or representatives of ImPACT Applications-SA make no claim as to the current diagnostic accuracy of the Software. Under the terms of this agreement, ImPACT is to be used as a clinical/research instrument in conjunction with standard medical care. The ImPACT web site including the ImPACT neurocognitive test available through the website http://www.impacttestonline.com/southafrica and the results generated by the ImPACT test are provided to designated officials at the school, designated medical professionals, and **if**

especially requested to parents and the athletes themselves, for informational purposes only. Should an athlete suffer from symptoms that may be related to a concussion, professional medical advice, diagnosis and/or treatment from a doctor or qualified medical personnel will be sought out by **……………………** who may use the data derived from the ImPACT test and its interpretation by a registered psychologist as additional information, but not as a replacement for standard medical management.

7. We understand that ImPACT Applications-SA do not take the professional responsibility for their representatives responsible for the baseline testing, the follow-up testing and its interpretation, and the overall medical management of the concussed athlete. The designated persons in this regard as delineated below, are contracted to take individual professional responsibility for these tasks in their private capacity, and any malpractice liabilities that may occur will be subject to their own reimbursement or that of their own professional indemnity insurance.

8. ImPACT Applications-SA disclaims any and all promises, representations, and warranties, except as expressly set forth in this Agreement, with respect to any data, information, the Software, or other material furnished to pursuant to this agreement, including their condition; conformity to any representation or description; the existence of any latent or patent defects; and, merchantability, or fitness for a particular purpose or use.

Signature: ________________________**(Headmaster) Date:** ____________________

Print Name: ____________________________ **Email:** _____________________

School: ____

Address: ____

Phone: __________________________ **Fax:** ______________________

Medical doctor taking medical responsibility for concussed players signed up for the programme at ……………………….

Name: ______________________________ **Telephone:** ____________________

Email Address: ____________________

Signature: ____________________________________ **Date:** ______________

Accredited ImPACT Psychologist who will be reading your ImPACT Reports, who will take clinical responsibility for the interpretation of the data and who will provide feedback of the data within 24 to 48 hours. The psychologist will also be responsible for providing appropriate supervision to the technician who does the baselining, and the registered persons (e.g. nurses, doctors, biokineticists etc) affiliated to ……………………… who are responsible for administering the follow-up assessments at ……………………

SOUTH AFRICA

Name: **Dr Victoria Alexander** Telephone: **072-3456770**

Email Address: **mastermindz@icon.co.za**

Signature: ____________________________________ Date: ______________

Psychologist / Technician taking responsibility for conducting the baseline testing at ………………………… for all athletes signed up for the Programme.

Name: Telephone:

Email Address:

Signature: ____________________________________ Date: ______________

ImPACT

SOUTH AFRICA

ImPACT representative responsible for the initial and on-going liaison with ……………………… This involves introducing the school to the Programme, setting up contact with the sports heads and medical personnel, organizing and monitoring the concussion management team at the school.

Name: Telephone:

Email Address:

Signature: ____________________________________ Date: ______________

Appendix B

Parent Letter for Mandatory Participation in the ImPACT Programme

Dear Parents,

(School name) has had the ImPACT assessment program in place for the last (number) years for our students. This program assists the medical staff in evaluating and treating head injuries (e.g. concussion). ImPACT, (Immediate Post Concussion Assessment and Cognitive Testing), is a computerized exam utilized across the country to successfully diagnose and manage concussions. If a student is believed to have suffered a head injury, ImPACT is used to help determine the severity of head injury and when the injury has fully healed.

ImPACT baseline assessments, compulsory for all (School name) students, will be implemented during the first week of the first term. This non-invasive, simple assessment is set up in "video-game" type format and takes about 15-20 minutes to complete. Essentially, the ImPACT assessment is a preseason physical of the brain. It tracks information such as memory, reaction time, speed, and concentration. It, however, is not an IQ test.

If a concussion is suspected, the students will be required to re-take the test once he is asymptomatic after his injury. Both preseason and post-injury test data are sent to our designated specialist neuropsychologists via email, following which their reports are sent to Doctors at (Name) and/or Doctor (Name) who evaluate the boys clinically. The information gathered can also be forwarded to your family doctor on request. The test data will enable these health professionals to determine when return-to-play is appropriate and safe for the injured students.

I wish to stress that the ImPACT assessment procedures are non-invasive, and they pose no risks to your son. The program costs R_____ inclusive per person. This is not claimable through your medical aid

If you would like to read more on the impact program visit www.impacttest.com.

Yours sincerely

(Name)

SENIOR SANATORIUM SISTER

Appendix C

Parent Consent Form for Mandatory Participation in the ImPACT Programme

Consent Form

For use of the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT)

I have read the attached information. I understand its contents. I have been given an opportunity to ask questions and all questions have been answered to my satisfaction. I agree to participate in the ImPACT Concussion Management Program.

Printed Name of Athlete ____

Sport ____

Signature of Athlete ____________________________

Date _________________

Signature of Parent _____________________________

Date _________________

Appendix D

Parent Letter for Voluntary Participation in the ImPACT Programme

Dear Parents

RE: CONCUSSION MANAGEMENT AND CONTACT SPORT

Boys who engage in contact sports are at risk of sustaining head injuries and concussions. In the past it was common practice for an athlete to be off-sport for a three week period following a concussion. There is, however, little evidence-based medical research to support the idea that all athletes take three weeks to recover from a concussion or that a three week period is sufficient for a full recovery following some concussions.

Best practice in the assessment and management of concussion indicates that neuropsychological testing should form part of the medical assessment and management of concussions. The school has thus established a partnership with ImPACT, an internationally renowned computerised concussion evaluation system. Dr Vicky Alexander, a psychologist with extensive expertise in this area, and sport physician Dr Jason Suter offer a comprehensive concussion management programme which combines best practice in the medical and neuropsychological management of concussion.

Athletes enrolled in this programme will undergo pre-season (base-line) neuropsychological testing (which takes approximately 40 mins) at the school. In the event of a head injury your son will have direct access to the sports physicians and will be re-assessed using the ImPACT protocol to help determine the severity of injury and to monitor recovery. This protocol will be used along with other medical data to determine when it is safe to return to play.

This concussion management system is currently utilized throughout professional sports and is considered the most effective method for the recognition and management of concussion in sport. You can read more about the ImPACT tests and concussion management on the following website: http://www.concussion.co.za

The cost of the concussion management programme will be **R_____** per individual for the academic year. This will include the cost of pre-season (baseline) testing and all neuropsychological follow-up concussion monitoring for the duration of the year. This fee does not include any medical examinations which may be indicated. It is worth pointing out that if your son is not on the programme, and suffers a concussion, and you want him to be assessed and monitored following the concussion, each visit will cost R_____, and he may well need two or three assessments.

It is necessary for boys to be re-tested at the start of every season. This is because adolescence is a time of significant neurological development and the neuropsychological tests we use are very sensitive (so the results we obtained 12 months ago will not be accurate enough for the clinical assessment required to manage a head injury).

Should you wish to enrol your son on this programme please email (Name) (email address) indicating that you would like your son to be part of the programme and that you give us permission to charge R_____ to your school account.

It is important that the baseline testing be done before the season gets underway. We have arranged to have the baseline assessments conducted in the Computer room (room #) during Open period and straight after school on the following dates:

- (Date)
- (Date)
- (Date)
- (Date)

There will be lists up in the______where boys will choose a day and time that suits them to have their baseline test. There are 8 different time slots for baseline testing — this will give ample opportunity for every boy who enrols in the programme to have his baseline test done. Boys must sign up for a specific session, and then ensure that they attend that session. **No individual baseline testing will be done after (Date).**

If you have any questions regarding this programme please feel free to contact us.

Yours sincerely

(Name)	(Name)	(Name)
Psychologist (Sport)	Psychologist	Deputy Headmaster
(Place of Employment)	(Place of Employment)	(Email)
(Email)	(Email)	

Appendix E

Ethical Clearance

RHODES UNIVERSITY
Grahamstown • 6140 • South Africa

DEPARTMENT OF PSYCHOLOGY
Tel: +27 (0)46 603 8500 • Fax: +27 (0)46 622 4032 • Website: http://www.rhodes.ac.za/academic/departments/psychology

RESEARCH PROJECTS AND ETHICS REVIEW COMMITTEE

15 October 2014

Marcelle Reichling
Department of Psychology
RHODES UNIVERSITY
6140

Dear Marcelle

ETHICAL CLEARANCE OF PROJECT PSY2014/17

This letter confirms your research proposal with tracking number PSY2014/17 and title, 'The utility of annual baseline testing on the ImPACT neuropsychological screening test: A five year prospective study on South African high school athletes', served at the Research Projects and Ethics Review Committee (RPERC) of the Psychology Department of Rhodes University on 15 October 2014. The project has been given ethics clearance.

Please ensure that the RPERC is notified should any substantive change(s) be made, for whatever reason, during the research process. This includes changes in investigators.

Yours sincerely

Professor Michael Guilfoyle
CHAIRPERSON OF THE RPERC

Appendix F

Non-Parametric Tests

Table 1

Normality Results for ImPACT Composite Scores at Each Test Interval

Test Interval	ImPACT Composite	Sample size (*n*)	Skewness	Standard Error	*z*-score for Skewness	Kurtosis	Standard Error	*z*-score for Kurtosis	Shapiro-Wilk		
									Statistic	df.	Sig.
1	Verbal Memory	108	-.557	.233	-2.39[a]	-.136	.461	-0.30	.961	108	.003*
2	Verbal Memory	108	-.344	.233	-1.48	-.997	.461	-2.16[a]	.946	108	.000*
3	Verbal Memory	108	-.656	.233	-2.82[a]	-.298	.461	-0.65	.930	108	.000*
4	Verbal Memory	108	-.651	.233	-2.79[a]	-.283	.461	-0.61	.940	108	.000*
5	Verbal Memory	108	-.637	.233	-2.73[a]	-.413	.461	-0.90	.932	108	.000*
1	Visual Memory	108	-.240	.233	-1.03	-.630	.461	-1.37	.981	108	.116
2	Visual Memory	108	-.673	.233	-2.89[a]	.148	.461	0.32	.950	108	.000*
3	Visual Memory	108	-.632	.233	-2.71[a]	.053	.461	0.11	.960	108	.003*
4	Visual Memory	108	-.951	.233	-4.08[a]	.876	.461	1.90	.938	108	.000*
5	Visual Memory	108	-.507	.233	-2.18[a]	-.105	.461	-0.23	.969	108	.014*
1	VMS	108	.262	.233	1.12	-.668	.461	-1.45	.978	108	.074
2	VMS	108	.066	.233	0.28	-.363	.461	-0.79	.990	108	.637
3	VMS	108	.084	.233	0.36	-.558	.461	-1.21	.987	108	.359
4	VMS	108	.046	.233	0.20	-.651	.461	-1.41	.986	108	.336
5	VMS	108	-.433	.233	-1.86	-.564	.461	-1.22	.968	108	.011*
1	Reaction Time	108	.312	.233	1.34	.032	.461	0.07	.981	108	.123
2	Reaction Time	108	.694	.233	2.98[a]	.118	.461	0.26	.955	108	.001*
3	Reaction Time	108	.443	.233	1.90	.226	.461	0.49	.976	108	.046*
4	Reaction Time	108	.922	.233	3.96[a]	.999	.461	2.17[a]	.940	108	.000*
5	Reaction Time	108	.696	.233	2.99[a]	.251	.461	0.54	.957	108	.002*
1	Impulse Control	108	.974	.233	4.18[a]	1.379	.461	2.99[a]	.943	108	.000*
2	Impulse Control	108	.576	.233	2.47[a]	.364	.461	0.79	.961	108	.003*
3	Impulse Control	108	.689	.233	2.96[a]	.533	.461	1.16	.964	108	.005*
4	Impulse Control	108	.975	.233	4.18[a]	.629	.461	1.36	.920	108	.000*
5	Impulse Control	108	1.047	.233	4.49[a]	.938	.461	2.03[a]	.920	108	.000*

Note. VMS= Visual Motor Speed Composite;

[a] An absolute z-score value greater than 1.96 (when the minus sign is ignored) is significant at $p \leq .05$, indicating scores are significantly different from a normal distribution.

* S-W test is significant at $p \leq .05$, indicating scores are significantly different from a normal distribution.

Table 2

Results of Friedman's ANOVAs for ImPACT Composite Scores

Test Interval	ImPACT Composite	Mean Rank	Chi-Square	df.	Asymp. Sig
1	Verbal Memory	2.61	19.562	4	.001*
2	Verbal Memory	2.69			
3	Verbal Memory	3.22			
4	Verbal Memory	3.13			
5	Verbal Memory	3.35			
1	Visual Memory	1.96	67.478	4	.000*
2	Visual Memory	2.96			
3	Visual Memory	3.14			
4	Visual Memory	3.48			
5	Visual Memory	3.46			
1	Visual Motor Speed	1.32	251.637	4	.000*
2	Visual Motor Speed	2.20			
3	Visual Motor Speed	3.46			
4	Visual Motor Speed	3.73			
5	Visual Motor Speed	4.28			
1	Reaction Time	3.91	54.539	4	.000*
2	Reaction Time	2.91			
3	Reaction Time	2.43			
4	Reaction Time	2.78			
5	Reaction Time	2.97			
1	Impulse Control	3.17	8.279	4	.082
2	Impulse Control	3.19			
3	Impulse Control	3.11			
4	Impulse Control	2.79			
5	Impulse Control	2.75			

Note. Asymptotic significances (2-sided tests) are displayed.
* Conditions are significantly different at $p \leq .05$.

Table 3

Results of post hoc Wilcoxon Signed Ranks Test for Verbal Memory

Comparison (Test Interval)	Z	Sig.	Adj. Sig. [a]
Interval 1-2	-.454	.650	1.000
Interval 1-3	-2.958	.003	.031*
Interval 1-4	-2.065	.039	.389
Interval 1-5	-3.509	.000	.004*
Interval 2-3	-3.033	.002	.024*
Interval 2-4	-2.430	.015	.151
Interval 2-5	3.589	.000	.003*
Interval 3-4	-.537	.591	1.000
Interval 3-5	-.761	.446	1.000
Interval 4-5	-1.128	.259	1.000

Note. Asymptotic significances (2-sided tests) are displayed.

[a] Adjustment for multiple comparisons: Bonferroni.

* Comparisons are significantly different at $p \leq .05$.

Table 4

Results of post hoc Wilcoxon Signed Ranks Test for Visual Memory

Comparison (Test Interval)	Z	Sig.	Adj. Sig. [a]
Interval 1-2	-4.495	.000	.000*
Interval 1-3	-5.305	.000	.000*
Interval 1-4	-6.020	.000	.000*
Interval 1-5	-6.414	.000	.000*
Interval 2-3	-1.108	.268	1.000
Interval 2-4	-2.704	.007	.069
Interval 2-5	-2.967	.003	.030*
Interval 3-4	-1.725	.085	.845
Interval 3-5	-2.215	.027	.268
Interval 4-5	-.028	.977	1.000

Note. Asymptotic significances (2-sided tests) are displayed.

[a] Adjustment for multiple comparisons: Bonferroni.

* Comparisons are significantly different at $p \leq .05$.

Table 5

Results of post hoc Wilcoxon Signed Ranks Test for Visual Motor Speed

Comparison (Test Interval)	Z	Sig.	Adj. Sig.[a]
Interval 1-2	-6.590	.000	.000*
Interval 1-3	-8.619	.000	.000*
Interval 1-4	-8.966	.000	.000*
Interval 1-5	-8.990	.000	.000*
Interval 2-3	-6.642	.000	.000*
Interval 2-4	-7.655	.000	.000*
Interval 2-5	-8.472	.000	.000*
Interval 3-4	-2.195	.028	.282
Interval 3-5	-5.438	.000	.000*
Interval 4-5	-4.334	.000	.000*

Note. Asymptotic significances (2-sided tests) are displayed.

[a] Adjustment for multiple comparisons: Bonferroni.

* Comparisons are significantly different at $p \leq .05$.

Table 6

Results of post hoc Wilcoxon Signed Ranks Test for Reaction Time

Comparison (Test Interval)	Z	Sig.	Adj. Sig.[a]
Interval 1-2	-4.797	.000	.000*
Interval 1-3	-6.186	.000	.000*
Interval 1-4	-4.939	.000	.000*
Interval 1-5	-3.896	.000	.001*
Interval 2-3	-2.182	.029	.291
Interval 2-4	-.960	.337	1.000
Interval 2-5	-.192	.848	1.000
Interval 3-4	-1.380	.168	1.000
Interval 3-5	-2.090	.037	.366
Interval 4-5	-.643	.520	1.000

Note. Asymptotic significances (2-sided tests) are displayed.

[a] Adjustment for multiple comparisons: Bonferroni.

* Comparisons are significantly different at $p \leq .05$.

Appendix G

Parametric Tests

Table 7

Results of Verbal Memory post hoc Pairwise Comparisons, and Effect Size

(I) Interval	(J) Interval	Mean Differences (I-J)(*)	Std. Error	Sig. [a]	95% Confidence Interval for Difference		Effect Size (Cohen's *d*)
					Lower Bound	Upper Bound	
1	2	.343	1.143	1.000	-2.935	3.620	.04
	3	-3.046*	.981	.024	-5.857	-.235	**-.33**
	4	-2.509	1.194	.380	-5.932	.913	-.27
	5	-3.769*	1.013	.003	-6.673	-.864	**-.42**
2	1	-.343	1.143	1.000	-3.620	2.935	-.04
	3	-3.389*	1.023	.013	-6.320	-.458	**-.34**
	4	-2.852	1.245	.239	-6.420	.717	-.28
	5	-4.111*	1.019	.001	-7.031	-1.192	**-.42**
3	1	3.046*	.981	.024	.235	5.857	**.33**
	2	3.389*	1.023	.013	.458	6.320	**.34**
	4	.537	.981	1.000	-2.276	3.350	.05
	5	-.722	.855	1.000	-3.174	1.729	.08
4	1	2.509	1.194	.380	-.913	5.932	.27
	2	2.852	1.245	.239	-.717	6.420	.28
	3	-.537	.981	1.000	-3.350	2.276	-.05
	5	-1.259	.860	1.000	-3.723	1.205	-.13
5	1	3.769*	1.013	.003	.864	6.673	**.42**
	2	4.111*	1.019	.001	1.192	7.031	**.42**
	3	.722	.855	1.000	-1.729	3.174	.08
	4	1.259	.860	1.000	-1.205	3.723	.13

Note. [a] Adjustment for multiple comparisons: Bonferroni.

* The mean difference is significant at $p \leq .05$.

Cohen’s $d = (M_1 – M_2) / SD_{\text{pooled}}$, with $d < .20$ denoting a trivial effect size, $d = .20\text{-}.49$ denoting a small effect size, $d = .50\text{-}.79$ a moderate effect size, and $d \geq .80$ a large effect size (Cohen, 1988). Effect size **bolded** only if respective differences over time are significant ($p \leq .05$).

Table 8

Results of Visual Memory post hoc Pairwise Comparisons, and Effect Size

(I) Interval	(J) Interval	Mean Differences (I-J)(*)	Std. Error	Sig. [a]	95% Confidence Interval for Difference		Effect Size (Cohen's *d)*
					Lower Bound	Upper Bound	
1	2	-5.833*	1.170	.000	-9.187	-2.479	**-.49**
	3	-7.269*	1.245	.000	-10.838	-3.699	**-.63**
	4	-9.370*	1.335	.000	-13.197	-5.544	**-.81**
	5	-9.713*	1.251	.000	-13.299	-6.127	**-.91**
2	1	5.833*	1.170	.000	2.479	9.187	**.49**
	3	-1.435	1.155	1.000	-4.747	1.877	-.13
	4	-3.537*	1.203	.040	-6.985	-.089	**-.31**
	5	-3.880*	1.183	.014	-7.269	-.490	**-.38**
3	1	7.269*	1.245	.000	3.699	10.838	**.63**
	2	1.435	1.155	1.000	-1.877	4.747	.13
	4	-2.102	1.150	.704	-5.398	1.194	-.19
	5	-2.444	.946	.111	-5.155	.266	-.24
4	1	9.370*	1.335	.000	5.544	13.197	.81
	2	3.537*	1.203	.040	.089	6.985	**.31**
	3	2.102	1.150	.704	-1.194	5.398	.19
	5	-.343	.997	1.000	-3.200	2.515	-.03
5	1	9.713*	1.251	.000	6.127	13.299	**.91**
	2	3.880*	1.183	.014	.490	7.269	**.38**
	3	2.444	.946	.111	-.266	5.155	.24
	4	.343	.997	1.000	-2.515	3.200	.03

Note. [a] Adjustment for multiple comparisons: Bonferroni.

* The mean difference is significant at $p \leq .05$.

Cohen's $d = (M_1 – M_2) / SD_{pooled,}$ with $d < .20$ denoting a trivial effect size, $d = .20$-.49 denoting a small effect size, $d = .50$-.79 a moderate effect size, and $d \geq .80$ a large effect size (Cohen, 1988). Effect size **bolded** only if respective differences over time are significant ($p \leq .05$).

Table 9

Results of Visual Motor Speed post hoc Pairwise Comparisons, and Effect Size

(I) Interval	(J) Interval	Mean Differences (I-J)(*)	Std. Error	Sig.[a]	95% Confidence Interval for Difference		Effect Size (Cohen's *d*)
					Lower Bound	Upper Bound	
1	2	-3.042*	.417	.000	-4.237	-1.848	**-.56**
	3	-6.614*	.450	.000	-7.905	-5.323	**-1.20**
	4	-7.609*	.395	.000	-8.743	-6.476	**-1.45**
	5	-9.337*	.433	.000	-10.579	-8.094	**-1.70**
2	1	3.042*	.417	.000	1.848	4.237	**.56**
	3	-3.572*	.439	.000	-4.829	-2.314	**-.60**
	4	-4.567*	.456	.000	-5.873	-3.261	**-.80**
	5	-6.294*	.470	.000	-7.642	-4.946	**-1.06**
3	1	6.614*	.450	.000	5.323	7.905	**1.20**
	2	3.572*	.439	.000	2.314	4.829	**.60**
	4	-.995	.403	.152	-2.151	.161	-.17
	5	-2.723*	.443	.000	-3.993	-1.453	**-.45**
4	1	7.609*	.395	.000	6.476	8.743	**1.45**
	2	4.567*	.456	.000	3.261	5.873	**.80**
	3	.995	.403	.152	-.161	2.151	.17
	5	-1.727*	.370	.000	-2.787	-.668	**-.30**
5	1	9.337*	.433	.000	8.094	10.579	**1.70**
	2	6.294*	.470	.000	4.946	7.642	**1.06**
	3	2.723*	.443	.000	1.453	3.993	**.45**
	4	1.727*	.370	.000	.668	2.787	**.30**

Note. [a] Adjustment for multiple comparisons: Bonferroni.

* The mean difference is significant at $p \leq .05$.

Cohen's $d = (M_1 – M_2) / SD_{pooled,}$ with $d < .20$ denoting a trivial effect size, $d = .20$-.49 denoting a small effect size, $d = .50$-.79 a moderate effect size, and $d \geq .80$ a large effect size (Cohen, 1988). Effect size **bolded** only if respective differences over time are significant ($p \leq .05$).

Table 10

Results of Reaction Time post hoc Pairwise Comparisons, and Effect Size

(I) Interval	(J) Interval	Mean Differences (I-J)(*)	Std. Error	Sig. ^{a}	95% Confidence Interval for Difference		Effect Size Cohen's *d*
					Lower Bound	Upper Bound	
1	2	.032*	.006	.000	.015	.050	**.49**
	3	.046*	.006	.000	.028	.064	**.72**
	4	.038*	.008	.000	.016	.060	**.55**
	5	.030*	.009	.006	.006	.055	**.43**
2	1	-.032*	.006	.000	-.050	-.015	**-.49**
	3	.014	.006	.182	-.003	.030	.21
	4	.006	.007	1.000	-.014	.026	.08
	5	-.002	.008	1.000	-.024	.020	-.03
3	1	-.046*	.006	.000	-.064	-.028	**-.72**
	2	-.014	.006	.182	-.030	.003	-.21
	4	-.008	.006	1.000	-.024	.009	-.11
	5	-.016	.006	.167	-.034	.003	-.22
4	1	-.038*	.008	.000	-.060	-.016	**-.55**
	2	-.006	.007	1.000	-.026	.014	-.08
	3	.008	.006	1.000	-.009	.024	.11
	5	-.008	.007	1.000	-.028	.013	-.10
5	1	-.030*	.009	.006	-.055	-.006	**-.43**
	2	.002	.008	1.000	-.020	.024	.03
	3	.016	.006	.167	-.003	.034	.22
	4	.008	.007	1.000	-.013	.028	.10

Note. ^{a} Adjustment for multiple comparisons: Bonferroni.

* The mean difference is significant at $p \leq .05$.

Cohen's $d = (M_1 - M_2) / SD_{pooled,}$ with $d < .20$ denoting a trivial effect size, $d = .20\text{-}.49$ denoting a small effect size, $d = .50\text{-}.79$ a moderate effect size, and $d \geq .80$ a large effect size (Cohen, 1988). Effect size **bolded** only if respective differences over time are significant ($p \leq .05$).