RELIABILITY OF A COMPUTERIZED NEUROPSYCHOLOGICAL TEST

by

JACOB EARL RESCH

(Under the Direction of Michael S. Ferrara)

ABSTRACT

Since 2005, computerized platforms have been scrutinized for lack of psychometric data to support their use in the clinical setting. Poor to moderate reliability coefficients have been reported for commercially available computerized platforms. The purpose of this study was to replicate and extend the research addressing the reliability of computerized neuropsychological tests while controlling for effort, time of day, and mood state. One-hundred and fifty two (N=152) healthy, college aged students were recruited to participate in this study. Each participant completed the Profile of Mood States - Brief Form (POMS-B), Green's Word Memory Test (WMT), and the ImPACT computerized concussion assessment test at three clinically relevant time points: baseline, day 45 and day 50. ImPACT calculates six composite scores utilized to determine cognitive decline post-concussion. The POMS-B is a measure of six mood states and a composite total mood disturbance score and was utilized to measure mood state at the time of each session. Green's WMT was utilized to assess effort at each time point. Any baseline data which was considered invalid by ImPACT or was incomplete was removed from subsequent data analysis. Forty-five (n=45) and one hundred and eight (n=108) participants were included in studies one and two, respectively. For study one, intraclass correlation (ICC) values were calculated for each ImPACT composite score. The ICC values for each composite score were slightly higher than previously reported in the literature at the same clinically relevant time points but still fell below what is considered clinically acceptable. For the second study,

although one or more mood states were significantly correlated with one or more ImPACT composite scores, correlations coefficients were weak to moderate and did not appear to influence test performance. Values for the Green's WMT exceeded suggested criteria for good effort and improved over time. The test-retest reliability coefficients reported in this study are slightly higher than those previously reported. Differences may be due to the delivery of one computerized test at each time point. The added evidence that mood state does not influence test performance leads the authors to believe that the poor to moderate reliability coefficients are due to systematic rather than random error.

INDEX WORDS: concussion, test-retest reliability, ImPACT, mood state, random error

RELIABILITY OF A COMPUTERIZED NEUROPSYCHOLOGICAL TEST

by

JACOB E. RESCH

B.S., South Dakota State University, 2003

B.S., South Dakota State University, 2004

M.S., South Dakota State University, 2006

A dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2010

RELIABILITY OF A COMPUTERIZED NEUROPSYCHOLOGICAL TEST

by

JACOB EARL RESCH

Major Professor:	Michael S. Ferrara
Committee:	Ted Baumgartner Cathy Brown Stephen Macciocchi Kimberly Walpert

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August, 2010

DEDICATION

This dissertation is first and foremost dedicated to Celestial Leigh Resch who throughout this process has taught me no matter how busy you seem to be there is always time to live, laugh, love and to be loved.

I would also like to dedicate this dissertation to my family for their support throughout this process. This dissertation is also dedicated to my friends (extended family) namely Aaron and Chrystal, Chris and Leslie, Kevin and Karen, Leland, and last but not least Kaitlin who throughout the past four years have been there to provide unyielding support and who were always up for a little cussing and discussing.

ACKNOWLEDGEMENTS

First and foremost I would like to acknowledge Dr. Michael Ferrara for taking a chance on me and providing me with an opportunity. Your mentorship and friendship over the past years has allowed me to achieve what I have and has prepared me to accomplish so much more both in academia and more importantly life.

Dr. Ted Baumgartner, thank you for your guidance over the past four years, especially for always having time to help and to share a story or two.

Dr. Cathy Brown, thank you for helping me stay grounded throughout the past four years. You have helped me to realize that it's ok to look at the big picture as long as you don't forget to pay attention to the small details.

My committee members, Dr. Kimberly Walpert and Dr. Stephen Macciocchi for their support, patience, and guidance throughout the entirety of this process.

Dr. Steve Olejnik, thank you for your continued guidance and friendship, even after entering retirement. Your commitment to education (of one of the most difficult subjects) is only to be admired and I hope to carry on your tradition of excellence as I progress in my career.

Last but not least, a very special thank you to Dr. Bernadette Olson. Thank you Berni for seeing something in me that no one else did. Your mentorship and friendship during the past ten years has guided me to where I am today, thank you

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION

Background

Concussion has been defined as a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces. [1] [2] This definition may be used interchangeably with mild traumatic brain injury and mild head injury. The mechanism of injury for concussion most commonly includes motor vehicle accidents, falls, domestic abuse, assault, and sport. [3] [1] [4]

Approximately 1.6 traumatic brain injuries occur annually in the United States. Of these traumatic brain injuries 1.2 million, 290,000, and 51,000 resulted in emergency room visits, hospitalizations, and deaths, respectively. [4] Astonishingly, an estimated 1.6 to 3.8 million sports related concussion are thought to occur taking into account those injuries that do not receive medical care. [5]

The sport most frequently associated with concussion is American football. During a three year study, football accounted for over half of the reported mild traumatic brain injuries that occurred amongst 11 different sports. [6] Other sports associated with concussion are wrestling, soccer, basketball, softball, baseball, field hockey and volleyball. [6] Early research addressing high school football suggested 24% of all injuries were sport related concussion. [7] This estimate has decreased significantly due to improved equipment, rule changes, more focused definitions of concussion and education of health professionals, coaching staffs, and most importantly the athlete. [8] More current estimates suggest that approximately 5.2 percent of football athletes sustain a concussion throughout a season. It is important to note that occurrence of concussion ranges upon level of competition with the highest rates reported at the high school

and division III collegiate levels at 5.6 and 5.5 percent, respectively. These rates are slightly less at the division II and division I levels at 4.5 and 4.4 percent, respectively. Although research design and methodologies have become more specific to sport-related concussion, under-reporting of condition may lead to underestimates of the true prevalence. [9] Despite the measures employed to reduce the occurrence of sport-related concussion, it remains inherent with contact-sports.

Due to the inevitability of sport-related concussion the focus of the sports medicine team has been the immediate recognition and management of sport-related concussion leading to return to play decisions. Throughout the last three decades various modalities have been utilized to examine the effects and track the changes associated with concussion. These tools include self-reported symptom inventories, neuropsychological testing, posturography, physical examination, and diagnostic imaging. Numerous symptoms have been associated with mild head injuries. These symptoms include but are not limited to headache, nausea, loss of coordination, sensitivity to light and noise, etc… [1] Though clinically meaningful, athletes may under-report symptoms due to lack of education regarding the signs and symptoms of concussion and its implications. [9]

Neuropsychological testing has gained immense popularity in the management of mild traumatic brain injury. During the past two decades computerized neuropsychological testing has gradually replaced traditional paper and pencil tests. Computerized neuropsychological testing has purported advantages such as infinite randomized forms, ease of administration, ability to baseline test a large number of athletes in a short period of time, standardized self administration, rapid testing, internet based delivery, and centralized data storage, analysis, and reporting. [10, 11] Computerized neuropsychological testing has been recognized as the cornerstone of

concussion evaluation. [1] [12] Despite its advantages, computerized neuropsychological testing has its shortcomings including but not limited to a dearth in research regarding test reliability, validity, sensitivity and specificity, required training and qualifications, the need for licensed psychologists for clinical interpretation of tests, hard and software issues, and accessibility. [10, 11] More recently it was emphasized that computerized neuropsychological testing should not be the sole basis of concussion management decisions, rather a tool utilized in conjunction with clinical and supplemental tests. [2, 13]

Balance tests including the balance error scoring system (BESS) or the sensory organization test (SOT) have been recognized as a component of the concussion test battery. [1, 2, 11-14] Both forms of balance testing have been recognized as a reliable and valid means of measuring the motor domain of neurological functioning but should not be the sole indicator of concussion diagnosis or management. [2, 11, 15] A deficit in postural stability while concussed is thought to be a disturbance in interaction between visual, somatosensory, and vestibular inputs or sensations.[16] It has been evidenced that disruption of one or more of these inputs deficits in balance may persist up to 3 days post injury.[16-18] In regards to accessibility, computerized posturography may be too costly for most clinicians to employ in their concussion management protocol; given that; the BESS is a more economical means of integrating posturography.

The medical examination is considered one of the cornerstones of concussion diagnosis and management.[1, 2, 11-13] The physical examination includes a comprehensive history, detailed neurological examination, clinical status of patient in regards to time of injury, and the determination of emergent neuroimaging in order to exclude potentially life-threatening conditions.[2] The physician will also utilize all other aforementioned tests in order to make clinical decisions if available. During the past decade neuroimaging has played an increased role

in concussion research. Due to the subtleties of the injury, structural changes are rarely depicted using magnetic resonance imaging (MRI) or commuted tomography (CT) scans. In recent years, new techniques including diffusion tensor imaging (DTI) and functional MRI (fMRI) have provided new insight into the pathophysiology of concussion. [19-21] fMRI and DTI techniques are based on the paramagnetic properties of blood and water which provide sensitive images of the brain. Due to the complexity, sample sizes, cost, and time required to carry out research regarding these neuroimaging techniques studies are limited making current findings premature for clinical recommendations.[2]

Although numerous tests exist to aid clinicians in diagnosing concussions, many are misdiagnosed, especially in the emergency room setting. Reasoning for the misdiagnosis of concussion includes inaccurate definitions of concussion and lack of cognitive testing within the emergency room setting and a general lack of emergency room doctor knowledge.[22] The misdiagnosis of concussion in emergency room patients may lead to catastrophic consequences. Accurately depicting individuals as concussed in these settings may lead to information regarding follow-up, return to sport, and future activities.[22]

A relatively new area of concussion research is biomarkers associated with brain injury. Currently of special interest are apolipoprotein (APOE ε4, APOE promoter gene,) tau polymerase, and other genetic markers. APOE ε4 and tau polymerase have been of interest in concussion research in regards to predisposal for concussion and the development of Alzheimer's disease.[23] Blood serum markers including S-100β, neuron specific enolase, myelin basic protein, insulin-like growth factor-1 (IGF-1), fibroblast growth factor, glial fibrillary acidic protein, and others have also been studied in relation to sport-related concussion but results at this time are equivocal.[2, 24] A biomarker or series of markers that are valid and

reliable would make an invaluable contribution to current management concussion as well as provide insight into the long-term sequelae of single or multiple injuries.

One condition which currently employs tests for deficiencies in biomarkers is post-traumatic hypopituitarism (PTH). Hypopituitarism is defined as a documented biochemical deficiency in at least one endocrine axis with associated pathology either in the pituitary or the hypothalamus.[25] The etiology of hypopituitarism includes but is not limited to pituitary or hypothalamic tumors, necrosis of pituitary tissue via radiation therapy, hemorrhage, infiltrative, vascular, pituitary or cranial radiation and trauma.[25] Until a review published by Benvenga et al., the incidence of PTH was thought to be rare. In 1986, only 53 known cases of PTH had been published. Benvenga et al., found an additional 314 cases of PTH.[26] A systematic review by Schneider et al., in 2008 reported 809 cases of PTH. This same study provided the only known estimate of PTH among traumatic brain injury patients at 30 per 100,000 patients per year.[27] A more generalized estimate of PTH incidence is 20 to 80 percent of TBI patients will develop some form of hypopituitarism.[28] The increasing incidence of PTH may be due to increased clinical recognition post-trauma. Benvenga notes early texts in the 1970s briefly mentioned trauma as a cause of hypopituitarism, despite 40% of cases of hypopituitarism being termed cryptogenic or idiopathic hypopituitarism.[29] Agha et al., champion this statement by suggesting PTH symptoms mimic those of post-concussion syndrome and are overlooked.[30]

The most common causes of PTH include motor vehicle accidents, falls, abuse and assault.[26] Male patients under the age of 35 are most likely to sustain a TBI, suggesting male patients are most likely to endure PTH.[30] PTH patients may go unrecognized simply due to not seeking medical help post-insult or simply forgetting hospitalization which has been documented for the general public and in two cases, physicians.[26] The patients' neglect of seeking medical

attention coupled with a lack of endocrine testing post TBI suggests the incidence of PTH may be grossly underestimated.

Treatment for PTH patients is replacement of deficient hormone(s). Hormones most commonly deficient in PTH are growth hormone (GH) luteinizing hormone (LH) and follicle stimulating hormone (FSH), adrenocorticotrophic hormone (ACTH), thyroid stimulating hormone (TSH), and the peripheral hormones insulin-like growth factor (IGF-1), testosterone (T), tri-iodothyronine (T_3) and thyroxine (T_4).[3, 27, 28, 30-32] Hormone replacement is suggested to reverse the symptomology of the underlying hormone deficiency. [3, 27, 28] Damage to the pituitary may be subtle and only minor deficiencies and other underlying pathologies that occur make it unclear whether patients would benefit from hormone replacement.[3, 27, 28] It has also been suggested PTH may only be short-term which may not require long-term treatment.[33] To complicate matters, hormonal deficiencies may have cognitive, physical, and social sequelae decreasing patients' quality of life.[3, 33] To date, no conclusive recommendation exists regarding hormone replacement and short-term or panhypopituitarism.

As mentioned previously, the most common causes of PTH are motor vehicle accidents, falls and assaults.[26] Individuals participating in sport experience similar forces when victimized by sport-related concussion. Research suggests that professional football athletes experience head accelerations ranging from 16 to 25 miles per hour.[34] The average speed of these reconstructed collisions was 9.3 ± 1.9 m/s (20.8 ± 4.2 miles per hour).[35] Conflicting reporting of accelerometer data suggest collegiate and professional athletes sustain hits during practice and game settings between 20 and 103 g.[34-40] Biomechanical forces presented in this body of research are similar to those sustained in a motor vehicle accident.[41, 42] Replicating the impact

velocities associated professional football collision for injury and non-injury scenarios utilizing finite element analysis reveals a majority of significant strain rate was concentrated in the central core region of the brain, the midbrain, the upper brain stem and diencephalon.[43, 44] This body of literature suggests that a relationship between sport-related concussion and PTH is tenable.

Long-term hypopituitarism may lead to lack of energy, reduced lean body mass, muscle fatigue, decreased exercise capacity, and reduced bone mineral density, and neuropsychological issues in adults along with arrested puberty, secondary amenorrhea, and reduced libido in adolescent and pediatric populations.[45, 46] The recognition of anterior pituitary hormone deficiency may also lead to a series of biomarkers which may assist in the recognition, management, and treatment for sport-related concussion or mild traumatic brain injury.

Purpose

The purpose of this study is to investigate the reliability of a commonly used computerized neuropsychological test. ImPACT is a commonly utilized neuropsychological platform that has been employed in numerous studies and across all levels of sport.[11, 47, 48] Current literature has addressed the reliability of ImPACT in both concussed and healthy individuals.[49-51] The authors of these studies reported inter-class correlations ranging from .15 to .39 and .43 to .74 after 45 days and 2 years, respectively.[49, 51] The discrepancies in the reported values are of concern to the clinician since data derived from these tests are utilized to assist in making return-to-play decisions. This study will further investigate the reliability of ImPACT replicating methodologies employed by Broglio et al.,.[49]

The second aspect of this study will address mood states which may contribute to the variance associated with neuropsychological testing. Various mood states including fatigue-inertia, depression-dejection, and anger-hostility have been shown to impact neuropsychological

test performance.[52-54] This study will include a measure of mood state to further account for any variability associated with repeat administrations of ImPACT. Results from this component may provide greater insight in the extraneous variables associated with neuropsychological testing.

Statement of the Problem

Concussion has been well-documented as an emerging concern globally in health care. Specifically, sport-related concussion has drawn increasing attention during the past three decades. At this time, researchers and clinicians alike are investigating new modalities to accurately diagnose, manage and make return to play decisions in regards to sport-related concussion. The long term effect of an acute or multiple concussive events is currently unknown. Throughout the past three decades the body of concussion literature has proposed novel modalities to diagnose and manage concussion, the establishment of psychometric properties of those tests, investigating the changes of rules and equipment and its effect on the incidence of sport-related concussion, and the practice of clinicians confronted daily by this increasingly unsilent epidemic.[1, 2, 11-13, 16, 21, 34, 55-59] Still, when a battery of the aforementioned tests is delivered, sensitivity is still below 100 percent.[13]

Current statements regarding sport-related concussion have identified new modalities which show promise in detecting the pathophysiological subtleties which occur with this metabolic dysfunction.[1, 2, 12] These modalities include fMRI, DTI, and genetic and biomarker testing. The tools will help provide insight into the long-term effects of sport-related concussion. Specifically, current biomarkers of interest include APOE ε4, tau genotypes, neuron specific enolases, serum 100β and hormonal markers.[23, 60-62] The presence of these markers may predispose individuals to concussion, be related to the cumulative effects of concussion and

various neurocognitive disorders such as Alzheimer's disease, and or identify new implications of the injury. Despite the development of tools for the management of concussion, numerous questions exist regarding tools which are currently utilized in the management of concussion, specifically computerized neuropsychological testing.

In order to determine the validity of a test three criteria need to be met which are objectivity, reliability, and validity. Baumgartner et al., have defined objectivity in terms of the agreement of competent judges about the value of a measurement.[63] An advantage of utilizing computerized neuropsychological tests is standardized administration and scoring by the computer ensuring objectivity.[1, 2, 11, 12]

In terms of reliability, specifically test-retest reliability, conflicting results have been reported. Broglio et al., studied three computerized neuropsychological platforms in regards to test-retest reliability in a healthy collegiate sample. The authors administered each exam at three time points; baseline, forty-five days, and fifty days. Results of this study suggested weak to moderate reliability (.15 to .66) for the three computerized platforms tested.[49] Schatz performed a follow-up study which lengthened the time between baseline and post-testing to two years and reported a moderate reliability (.43 to .74) for all measures of the computerized neuropsychological platform ImPACT.[51] Variance associated with the data which impacts "*r*" may be random or systematic in nature. Random error affects an individual's performance purely by chance happening. Examples of random error may be the testing situation and or fluctuation in the individual examinee's state. Systematic error is defined as a particular characteristic of the person or the test that has nothing to do with the construct being measured.[64] One aspect of ImPACT which has not been addressed has been the equivalence of the alternate forms of the exam. Another advantage to computerized neuropsychological testing is random and or alternate

forms.[11] Although these forms are currently utilized in administration of the exam, no known research has addressed the coefficients of equivalence between the five alternate forms which may be a direct source of systematic error therefore decreasing test-retest reliability. This dearth of research is a cause of concern for clinicians since a lack of test-retest reliability means an absence of validity. That is, the interpretation of the results of computerized neuropsychological test may not reflect the construct in question which the presence of neurocognitive deficits.

Specific Aims and Null Hypotheses

Specific Aim 1: To determine test-retest reliability coefficients of ImPACT utilizing clinically relevant administration times.

Null Hypothesis 1: Intraclass correlation coefficients calculated between and among the three administrations of ImPACT, in a healthy sample, will be large enough for clinical interpretation (r = .75) or greater[49].

Specific Aim 2: To determine whether fatigue-inertia, vigor-activity, tension-anxiety, depression-dejections, anger-hostility, and or confusion bewilderment are sources of random error in relation to ImPACT sub-scores.

Null Hypothesis 2: Weak to moderate Pearson *r's* (.10 - .50) will exist for fatigue-inertia, vigor-activity, tension-anxiety, depression-dejections, and anger-hostility in relationship to ImPACT subscores.

The following chapters will outline the literature related to concussion, and two original research manuscripts. The first manuscript will address the test-retest reliability of impact utilizing clinically relevant time point and will be submitted to the Journal of Athletic Training. The second manuscript will be submitted to the British Journal of Sports Medicine. This

manuscript will address mood states and neuropsychological test performance. General conclusions, summary, and references complete the dissertation.

Limitations and Delimitations

Limitations:

1) Exercise prior to testing will not be controlled for
2) Consumption of food will not be controlled for
3) Sleep will not be controlled for

***Delimitations*:**

1) All participants will be drawn from the University of Georgia
2) Only participants between the ages of 18 and 24 will be included in this study
3) Each participant will have a minimum of a high school degree
4) Time of day of test administration and completion will be controlled for
5) The type and amount of caffeinated beverages will be accounted for
6) Participants will refrain from alcohol or drug use twenty-fours prior to testing

CHAPTER 2

REVIEW OF LITERATURE

Definition of Concussion

Concussion is derived from the Latin word *concussus,* which mean to shake violently.[65] Throughout the past three decades the definition of concussion has evolved to "a clinical syndrome characterized by immediate and transient posttraumatic impairment of neural function, such as alteration of consciousness and disturbance of vision or equilibrium due to brain-stem involvement."[6] In 1997, the American Academy of Neurologists defined the injury as "a trauma-induced alteration in mental status that may or may not involve loss of consciousness."[66] During the Vienna meeting in 2001, concussion was defined as "a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces.[1] The wording of the following reflects the most recent update to the 2001 and 2004 concussion in sport meetings.

1. Concussion 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]
2. Concussion typically results in the rapid onset of short-lived impairment of neurologic function that resolves spontaneously.[2]
3. Concussion may result in neuropathological changes but the acute clinical symptoms largely reflect a functional disturbance rather than a structural injury.[2]
4. Concussion results in a graded set 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 a small percentage of cases however, post-concussive symptoms may be prolonged.

5. No abnormality on standard structural neuroimaging studies is seen in concussion.[2]

Despite the latter portion of this definition remaining similar in wording, in 2004 an additional component involved defining simple versus complex concussions. Simple concussions were suggested to be those where an athlete progressively resolves without complication over 7 to 10 days. Complex concussions are when athletes suffer persistent symptoms, specific sequelae, prolonged loss of consciousness, or prolonged cognitive impairment after the injury.[12] During the 2009 meeting, this addition was removed.[2] Another change during the 2009 meeting was cessation of the interchangeability of concussion with mild traumatic brain injury.[2] The group felt that these two injuries were separate in nature and were not to be used interchangeably. What was once defined as "getting your bell rung," has evolved into the preceding definition(s). Current and future studies will dictate the future definition of this injury.

Pathophysiology of Concussion

As defined, concussion is partly described as a complex pathophysiological process.[2] Although the exact mechanism of concussion is currently unknown possible hypotheses include; shear strains generated by rotation; relative displacement due to impact from different locations, relative displacement due to impact from different directions; concussion is a result of shear stress, distortion, or mass movement in the brain stem resulting from pressure gradients due to impact loading. Linear acceleration is the most important mechanism; pressure waves traveling through the brain; strains affecting the brain in a centripetal sequence of disruptive effects on function and structure. The effects of this sequence always begin superficially and move central with increasing trauma severity; and impact pulses containing frequencies that are close to the nodal frequencies of thee skull/brain complex may cause concussion.[42]

To date, the most accurate depiction of the metabolic effects of concussion are best described by Giza and Hovda.[67] Upon biomechanical insult, neuronal disruption occurs leading to a significant extracellular K^+. This condition is exacerbated by an indiscriminate release of excitatory amino acids. An overabundance of extracellular K^+ is normally utilized by surrounding glial cells, this does not occur upon concussion. The non-relenting release of K^+ and subsequent release of excitatory amino acids leads to neuronal suppression affecting diffuse areas of the brain leading to the common signs and symptoms of concussion. Oxidative metabolism is the only means of restoring homeostasis.[67] During this time oxidative metabolism is supplemented by glycolysis. There are two types of glycolysis, aerobic and anaerobic. The former takes longer but produces limited or no lactic acid. The latter takes place in the limited presence or absence of oxygen and is quicker but forms lactic acid.[68] Given cerebral blood flow may be reduced to 50 percent anaerobic glycolysis is utilized for energy production. Lactic acid is suggested to lead to neuronal dysfunction. Given anaerobic glycolysis is the less preferred mechanism of energy, both oxidative and "fast" glycolysis are considered impaired which leads to an "energy crisis," and may lead to increased vulnerability to secondary ischemic injury.[67]

Also impairing oxidative metabolism is an excess of intracellular calcium leading to decreased adenosine tri-phosphate. An efflux of calcium is a consequence of the release of excitatory amino acids which activate N-methyl-D-aspartate receptors. The stretch of neuronal axons have also been shown to lead to an increase in intracellular calcium.[67] The overabundance of intracellular calcium may lead to morphologic neuronal damage or the initiation of apopotic signals if levels are still elevated past 96 hours.[67]

During the initial four days post-injury, intracellular magnesium is reduced which has been correlated with post-injury neurologic deficits. Reasoning for impaired neurologic function with decreased magnesium may be due to its relationship to impaired metabolism, impaired protein synthesis and its role in the efflux of overabundant intracellular calcium.[67]

The impaired metabolism has been noted to resolve in animals at approximately five to 10 ten days post injury. In humans positron emission tomography has detected impaired glucose metabolism up to two to four weeks.[67] The metabolic cascade proposed by Giza and Hovda correlates well to studies published regarding post-concussive symptomology, computerized neuropsychological testing, and posturography.[16, 17, 69, 70]

Epidemiology of Concussion

Sport-related concussion has been recognized as a major concern amongst athletes and their families and the sports-medicine community. An initial study 1983 utilizing high school football players revealed 19 out of 100 players sustained a possible concussion. The authors suggested 24 percent of all injuries were possible concussion. The authors also reported six permanent disabilities and an annual reporting of one high school athlete with extensive brain dysfunction and or quadriplegia or death as a result of head and or neck trauma in high school football.[7] A landmark study by Powell and Barber-Foss significantly reduced the estimate suggesting 5.5% of all injuries in sport were sport related concussion. The authors utilized a refined definition and studied football along with wrestling, girls' and boys' soccer, girls' and boys' basketball, softball, baseball, field hockey, and volleyball.[6] The sports listed are in order of incidence of concussion, highest to lowest. Football was still reported as having the highest incidence of concussion accounting for 63.4 percent of all reported concussions.[6] During the 1995 to 1997 football seasons Guskiewicz et al., furthered epidemiological research by incorporating high

school and collegiate football programs. The authors reported an average incidence of 5.1 percent among all levels of competition, high school, division III, division II, and division I reported incidence rates of 5.6, 5.5, 4.5 and 4.4 percent, respectively.[8] A systematic review done by Tommasone and McLeod reviewed all epidemiological studies between 1985 and 2000 (23 reviewed articles) and reported that team high school boys' hockey had the highest incidence of 3.6 concussions per 1000 athlete exposures. At the professional level, hockey and rugby had the highest incidence at 6.5 and 9.05 per 1000 player-games, respectively. In individual sports boxing both professional and amateur had the highest incidences at 0.8 per 10 rounds, and 7.9 per 1000 man-minutes.[71]

The dramatic decrease in early estimates of concussion prevalence in contact sport has been contributed to 1) rule changes that have outlawed spearing and butt blocking, 2) player education about the rule changes and the effects of multiple concussions, 3) availability of alternative assessment techniques, 4) implementation of equipment standards, 5) marked reduction of physical contact time in practice sessions, and 6) heightened awareness among clinicians of the dangers involved with returning an athlete to competition while still symptomatic.[8, 69] Despite these imperative changes in regulation, equipment and behavior, this significant decrease in incidence may be due to underreporting. McCrea et al., reported 52.7 percent of athletes who had possible concussions did not report them to their football staff or clinicians. The authors of this study suggest the actual rate of sport-related concussion in high school football is approximately 15 percent. Reasons for athletes not reporting their injury were, not thinking it was serious enough; not wanting to leave the game; being unaware of what a concussion was; not wanting to let teammates down, and other reasons.[9] The aforementioned research suggests despite an

under-reporting of possible concussions, the overall incidence has decreased throughout the past three decades.

Current Management of Concussion

One of the greatest difficulties confronting clinicians today is the recognition of a concussive event. Past literature suggested the hallmark signs of concussion included loss of consciousness and post-traumatic amnesia.[65] Although these are indeed signs of concussion, symptomology including headache, dizziness and nausea have been deemed the most common symptoms and may signal a concussive event has occurred.[2, 56, 72, 73] Other factors such as an athlete's inability to recognize that he or she may just sustained concussion, the injury's significance, its long term implications, a feeling that the athlete is letting their teammates down, external pressures the athlete experiences and the late on set of symptomology may limit the clinician's ability to diagnose the condition.[9] This section will provide an overview of what clinicians are currently utilizing to recognize concussion in the field and clinic.

Concussion Grading Scales

Numerous grading scales exist to classify concussion. Two which are widely accepted are scales proposed by the Colorado Medical Society, Cantu evidence based and American Academy of Neurologists (AAN).[65, 66, 74] These scales are similar in that length of symptomology, post-traumatic amnesia, and loss of consciousness dictate one of three levels of severity. The three levels of severity proposed by the AAN are:

Grade 1: Transient confusion, no loss of consciousness, and symptomology or mental status abnormalities on examination resolve in less than 15 minutes.

Grade 2: Transient confusion, no loss of consciousness, and symptomology or mental status abnormalities on examination last more than 15 minutes.

Grade 3: Any loss of consciousness, either brief (seconds) or prolonged (minutes).[66]

The grading scale proposed by Cantu is as follows:

Grade 1: No loss of consciousness; posttraumatic amnesia or post-concussion signs or symptoms lasting less than 30 minutes.

Grade 2: Loss of consciousness lasting less than one minute; post-traumatic amnesia or post-concussion signs or symptoms lasting longer than 30 minutes but less than 24 hours.

Grade 3: Loss of consciousness lasting more than one minute or post-traumatic amnesia lasting longer than 24 hours; post-concussion signs or symptoms lasting longer than seven days.[65]

The Colorado Medical Society Grading scales is as follows:

Grade 1: Confusion without amnesia; no loss of consciousness

Grade 2: Confusion with amnesia; no loss of consciousness

Grade 3: Loss of consciousness.[75]

Currently, clinicians utilize three approaches for labeling a concussion a particular grade, 1) grading the concussion a the time of injury 2) deferring final grading until all symptoms have resolved, or 3) not using a grading scale but rather focusing attention on the athlete's recovery via symptoms, neurocognitive testing, and postural-stability testing.[11] It is imperative no matter which approach is taken that the clinicians involved agree on which protocol to employ.

Concussion Field Tests

Another field test utilized to assess potentially concussed athletes is the standard assessment of concussion (SAC). The SAC was developed by McCrea et al., and is a brief sideline assessment and consists of four components: orientation, immediate memory, concentration, and

delayed recall. Auxiliary tests include neurological screening and exertional maneuvers.[76] The orientation section requires athletes to answer questions regarding date, time, and place. Immediate memory is examined by three trials of recalling five assigned words. Concentration is tested by recalling a series of numbers consisting of varying lengths and naming months of the year in reverse order. Delayed recall consists of recalling the five words previously assigned in the immediate memory task. Each component is allotted five to 15 points for a total score of 30 points.[76] The SAC has been recognized has been recognized as a tool for immediate recognition of concussion but does not replace a full physical and neuropsychological assessment.[2, 77, 78]

An enhanced form of the SAC is the sport concussion assessment tool (SCAT). The SCAT was developed during the Prague conference on concussion and sport. The purpose of the SCAT was to incorporate patient education into the clinician's examination. The SCAT incorporated the following: 1) SAC; 2) Management of concussion sports palm card; American Academy of Neurology and the Brain Injury Association; 3) Sideline evaluation of concussion; 4) Sideline concussion check; UPMC, Thinksafe, Sports Medicine of New Zealand Inc and the Brain Injury Association; 5) McGill abbreviated concussion evaluation; 6) National Hockey League physician evaluation form; 7) The UK Jockey Club assessment of concussion; and 8) Maddocks questions. Currently, normative data is being developed as the SCAT becomes established.[79]

Neuropsychological Testing

Paper-and-pencil tests

Since in the 1980's, neuropsychological testing in regards to sport-related concussion utilized paper and pencil testing.[80] Traditional paper and pencils tests included Hopkins Verbal

Learning Test, Symbol Digit Modalities Test, Controlled Oral Word Association Test, Trail Making Test, and Stroop Color Word Test.[57]

Hopkins Verbal Learning Test (The Johns Hopkins University, Baltimore, MD):

The Hopkins Verbal Learning Test (HVLT) is designed to measure a subject's verbal memory.[18]Twelve words are read aloud by an examiner three times. After the three learning trials, subjects are presented target words which may or may not have been on the original list of twelve words. Six alternate forms of the HVLT exist. The revised version of the HLVT includes a list of twelve words read aloud by an examiner followed by three recall trials followed by a twenty-five minute delay and a fourth recall trial.[81]

Symbol Digit Modalities Test (Western Psychological Service, Los Angeles, CA) :

The Symbol Digit Modalities Test (SDMT) is a measure of complex attention, information processing speed, and psychometric functioning. In this ninety-second test, subjects fill in as many numbers as possible that correspond to a given symbol according to a key at the top of the page. Scoring is based off the difference between correct and incorrect responses. The SDMT utilizes visual tracking and incidental learning by having a subject write the number which corresponds with a symbol on the page.

Controlled Oral World Association Test (Multilingual Aphasia Examination):

The COWAT was created by Benton and as a brief aphasia test.[82] The COWAT is designed to evaluate a subject's ability to make verbal associations to specific letters and detect word association fluency. Subjects are presented three letters. For instance in Version A, subjects are presented the letters C, F, and L. Subjects have one minute per letter to list as many words, excluding proper names, as they can that begin with the suggested letter. The final score is the sum of the three trials.[82]

Trail Making Test: Subjects are asked to sequentially trace a list of 25 numbers on a piece of paper as fast as possible using a pen. This task measures orientation, concentration, visuospatial capacity, and problem-solving abilities. The time required for completion is recorded. One second is added for each sequential error committed.[18]

Stroop Color Word Test (Stoelting Co., Wood Dale IL): There is no standardized version of the Stroop Color Word Test. [83] The original Stroop Test is a measure of mental flexibility and attention span by examining a subject's ability to separate word and color-naming stimuli through three separate subtests. Each subtest contains 100 items presented in five columns of 20 items. Subjects have 45 seconds to complete each subtest, with a total score calculated from the sum of each subject. The three sub-tests progress from reading the name of a color off of a card to identifying the colors red, green, and blue printed in "XXXX", and finally identifying a color of the print which does not match the word printed on the card.[18]

Digit Span Forward and Backward Test (Psychological Corporation, San Antonio, TX) : Subjects are presented a series of numbers and are asked to repeat the digits in either the same order (Digits Forward) or in the reverse order (Digits Backward). The number of successful trials for each part is recorded as the total score.[18]

All aforementioned tests measure aspects of memory, cognitive processing speed, working memory, or executive functions.[57] In 1998, Lovell and Collins published an article addressing a battery of paper and pencil exams utilized with the Pittsburgh Steelers. This was a modified battery which would later be considered the framework of their computerized neuropsychological test.

Despite paper and pencil tests ability to depict neurological deficits in concussed athletes it had shortcomings. Paper and pencil test batteries could be completed in approximately 30

minutes, but have to be administered by a trained examiner such as a licensed, preferably Board certified in clinical neuropsychological or with experience with sport-related concussion.[11, 57] Practice effects are also thought to occur with paper and pencil test delivery without the use of alternate forms.[84]

Computerized Neuropsychological Tests

In 2000, Maroon et al., reviewed the evidence of neuropsychological testing as an aid in the management of concussion. This article outlined Immediate Measurement of Performance and Cognitive Testing (ImPACT).[85] ImPACT represents one of many computerized neuropsychological platforms utilized in the assessment of sport-related concussion. Other platforms utilized by clinicians include Automated Neuropsychological Assessment Metrics (ANAM), CogSport and the Concussion Resolution Index. As with their predecessors computerized neuropsychological tests measure attention, memory, processing speed and reaction time in computerized forms. Advocates of computerized neuropsychological testing suggest it allows clinicians to measure reaction times within .001 second, administer neuropsychological tests without the assistance of a trained administrator, to test multiple subjects simultaneously, deliver an infinite amount of randomized forms, deliver a standardized battery of tests, and to maintain centralized data storage.[11, 57, 85, 86] These benefits, along with empirical data suggesting computerized neuropsychological tests and their ability to depict subtle neurological deficits have lead experts to suggest computerized testing is the cornerstone of concussion management.[1, 2, 10-13, 50, 69, 87-90]

Although computerized neuropsychological testing has gained acceptance into the sports medicine community, critiques exist. Randolph suggests a dearth of validity, reliability and sensitivity data limits their suggestion for full clinical application and if utilized should be

interpreted conservatively.[57] Broglio et al., investigated the sensitivity of two common computerized platforms and reported sensitivities of 78.6 and 79.2 percent. The authors of this study suggest that computerized testing is only one component of the concussion battery of tests and should be interpreted as such.[13] This statement is supported by the Zurich consensus statement.[2]

Posturography

One of the earlier assessments of balance in response to a possible concussion is Romberg's test. The test is performed by having the involved athlete stand with their feet together with their eyes closed. Any disturbance including anterior or posterior and or medial lateral sway indicates a positive test.[91] The assessment of balance and its meaning has developed vastly over the past three decades.

Balance is maintained when an individual's center of gravity remains within their base of support. The maintenance of balance requires visual, somatosensory, and vestibular inputs. Input from these systems is transmitted to multiple regions of the brain and spinal cord. The basal ganglia receive the initial inputs based on limb positioning. This signal is integrated with planned actions developed in the pre-motor and supplementary motor cortex in the cerebellum. The descending pathway continues via alpha motor neurons which innervate skeletal muscle allowing for regulation of balance.[92-94] Typically, the visual and somatosensory senses provide a majority of information to maintain postural stability. A conflict between the somatosensory and visual inputs allows the vestibular system to maintain balance according to linear and angular accelerations of the head.[14, 16, 92] Any disruption of the interaction between these inputs is thought to disturb an individual's ability to balance.

The means of isolating each sensory input has been the focus of numerous articles. An initial attempt was by Shumway-Cook and Horak. Utilizing a blindfold, a visual conflict dome constructed from a Chinese lantern, and solid and form surface the authors were able to successfully isolate was sensory input. The purpose of this study was to create a means of detecting sway patterns for those with neurological deficits.[95] Guskiewicz et al., replicated this study with the aide of a force plate and dynamics platform to test concussed athletes compared to healthy controls. Participants were initially base-line tested and returned for post-concussion testing on days 1, 3, 5, 10, and when asymptomatic. The authors reported that concussed athletes experienced sensory integration issues which were evidenced by larger sway index values when compared to control subjects 3 to 5 days after injury and complete resolution by day 10.[17] Riemann employed the Neurocom Smart Balance Master Sensory Organization Test (SOT) to further the investigation of computerized posturography in concussed athletes. The SOT is designed to systematically the sensory selection process by altering information available to sensory inputs to maintain balance.[18] The SOT consists of 18, 20 second trials consisting of three visual conditions (eyes open, eyes closed, sway referenced) and two surface conditions (fixed, sway referenced).[18] The data derived from the SOT are a composite score, visual, vestibular, somatosensory, and visual conflict score. The composite score describes the overall ability to balance and inherently the interaction of the three sensory inputs while the remaining sub-scores provide more specific information regarding balance in the absence of visual, vestibular, and or somatosensory inputs. This initial case study indicated the cumulative effects of concussion in regards to postural sway.[18] Despite the useful information provided by force plate analysis and the Neurocom Smart Balance Master, cost is a limiting factor for their utility.

In 1999, Riemann et al., introduced the Balance Error Scoring System (BESS) as an alternative means of measuring postural stability. The BESS consists of a 45 cm^2 x 13 cm thick, 60kg/m^3 medium density foam and three stance variations of Romberg's test. The varied stances consisted of double-leg, single-leg, and tandem stances which were performed on solid ground and once again on the foam. Participants are asked to place their hands on their iliac crests, and then to close their eyes. If participants broke from the assigned position it is considered an error. Errors include lifting hands off iliac crests, the opening of eyes, stepping, stumbling, or falling, remaining out of the testing position for more than 5 seconds, moving the hip into more than 30 degrees of flexion or abduction and or lifting the forefoot or feel, each resulting in an additional point.[18] The BESS was later validated by a subsequent study when compared to the SOT.[14, 70] Findings of this study and others also suggest that patients return to baseline performance three to five days post injury.[14, 70] The authors of these studies that posturography whether manual or computerized should only be part of a more thorough test battery including neuropsychological testing and symptomology. The statement is supported by the work of Broglio et al., which suggests that the gold standard, the SOT, possesses a sensitivity of 61.9 percent.[13]

Self-Reported Symptoms

Historically, concussion had been identified by a hallmark sign and symptom, loss of consciousness and post-traumatic amnesia.[7, 65] The former symptom, although, significant, is not as necessary to be diagnosed as concussed.[65, 72] Post traumatic amnesia is still considered a maker of concussion severity, regardless of type, retrograde or anterograde.[65]

Currently, it is know concussion is identified by a myriad of symptoms including headache, nausea, dizziness, vomiting, balance difficulty, fatigue, trouble falling asleep or sleeping more

than usual, drowsiness, sensitivity to light and noise, sadness, nervousness, numbness and or tingling, feeling slowed down, feeling as in a fog, and difficulty concentrating and remembering.[96] A construct validity study by Piland et al., suggested concussion is predominantly associated with three categories, somatic which consists of headache, nausea and difficulty balancing; neuropsychological symptoms consisting of fatigue, trouble falling asleep, and drowsiness; and cognitive symptoms consisting of feeling slowed down, feeling as in a fog, and difficulty concentrating.[96] The results of this study were substantiated with a larger sample with subsequent research.[56] The symptoms presented in Piland's work are supported by other studies addressing symptoms.[90]

Three post-concussion symptoms in particular, feeling as in a fog and headache have received special attention. Those who report fogginess reported an additional 32 points in a symptom inventory compared to those who did not. This sample also displayed slower reaction times, processing speed, and reduced memory performance compared to those who did not feel as in a fog.[97] Concussed athletes who report headache as a symptom report an average 18 points higher on a symptom inventory than those who did not as well as displayed slower reaction times and lower memory composite scores. Athletes who reported post-concussion headache were also sustained more severe concussions, and were four times more likely to possess three to four abnormal on-field injury makers and five times more likely to show sideline assessed mental status changes of five minutes or more.[72]

Neuroimaging

An initial step in the evaluation of those with closed head injuries is neuroimaging. Modalities consist of magnetic resonance imaging (MRI), commuted tomography (CT) and a more extensive study may be accomplished with positron emission tomography for investigate the

metabolic effects of concussion. Initially, neuroimaging was recognized as contributing little to the diagnosis and management of concussion, unless a more complicated injury was suspected.[1, 12] MRI and CT scans are more pertinent to those cases involving loss of consciousness, severe amnesia, abnormal or neurologic findings, or increasing intensified symptoms.[11] Recently, structural MRI modalities including gradient echo, perfusion and diffusion tensor imaging have been suggested to be more sensitive to structural abnormalities. Research utilizing these techniques is remains premature before any clinical suggestions can be made.[2, 12]

Making Return-to-Play Decisions

Return-to-play decisions will be dependent upon what framework clinicians utilize for concussion management. The concussion in sport group suggests a step-wise return to play protocol which is as follows:

1. No activity, complete rest; once asymptomatic, proceed to next step. This step is to aid in physiological recovery.[1, 12, 98]
2. Light aerobic exercise such as walking or stationary cycling. This step is to evoke a cardiovascular response.[1, 12, 98]
3. Sport-specific training which allows for additional movement.[1, 12, 98]
4. Noncontact training drills employing practice of coordination, exercise, and cognitive load.[1, 12, 98]
5. Full-contact training after medical clearance which allows for restoration of confidence and for assessment of athletic skills by the coach staff.
6. Game Play[1, 12, 98]

It is advised that with each progression, if an athlete becomes symptomatic, to drop back to the previous step and progress again after 24 hours. It has also been advised that the concussed athlete should not be taking any pharmacological agents/drugs that may affect or modify the symptoms of concussion. Anti-depressant medications, if prescribed, may be utilized upon return to sport under the supervision of a doctor. Higher levels of competition may expedite return-to-play given the advanced medical care and management.[2, 12] In addition to the aforementioned recommendations, the National Athletic Trainers' Association advises a minimum of seven days for the concussion to fully resolve and reduce the risk of secondary neuronal injury. The authors also advise a battery of tests including neuropsychological tests, posturography, symptoms, and physical examination.[1, 2, 11, 12]

The Future of Concussion Testing

Currently, concussion-related research is focused on assessing current instruments for sensitivity, reliability, and validity through varying protocols, investigating gender differences, utilizing new techniques in regards to MRI including fMRI and diffusion tensor imaging, rehabilitation protocols with some including virtual reality, and investigating biomarkers for predisposal, diagnosis, management, and the cumulative effects of sport-related concussion.[2] The section will elaborate on fMRI and biomarkers.

During the past two decades, fMRI has become increasingly utilized to investigate neurological disorders such as Alzheimer's and Parkison's disease. fMRI attempt to localize and map parts the brain which are correlated to specific mental processes.[99] fMRI utilizing the inherent oxygen carrying properties of erythrocytes, specifically hemoglobin (iron-sulfide based) to provide images of deoxygenated blood flow during cognitive processes. Generally, fMRI data details areas of increased blood flow and oxygen utilization due to increased activity. This, the

most common fMRI application, is called the blood oxygen level dependent (BOLD) technique. Researchers interested in the area of concussion utilize various neuropsychological tests such as the n-back or finger tapping tasks to detect differences in blood flow between concussed patients and either pre-concussed or post-concussed states and or normative data.

An initial study utilizing fMRI enrolled eight male subjects, each baseline tested, excluding one, with an fMRI neuropsychological protocol. Four of the eight suffered a concussion during their respective football season. Results showed increased BOLD signal in the bilateral inferior-superior parietal region and dorsal-lateral and frontal cortex which are associated with memory in a normal population.[21] The authors of this study reported that despite the absence of any behavioral issues related to their injury, fMRI was able to detect physiological differences suggesting greater sensitivity compared to other tests related to concussion.[21] Research by Lovell et al., employed fMRI and found relationships between Brodman's area and time to return-to-play, as well changes in activity in the dorsal attentional system and cognitive and somatic symptomology.[100] Overall the body of literature regarding fMRI reports changes in activation of the dorsal cingulate and ventrolateral prefrontal regions, Brodman's area, the bi-lateral inferior-superior parietal region, the dorsal-lateral and frontal cortex, cerebellar regions, and orbito-frontal cortices.[19, 21, 100, 101] These studies focused on the symptomology associated with sport-concussion as well as traditional areas of the brain associated with neuropsychological testing.

Although fMRI is relatively young in regards to concussion research, it shows promise to be a sensitive measure of concussion.[21] Difficulties arise when performing fMRI research due to its expensive and inaccessible nature, BOLD being an indirect measure of neural activity, and

difficulty in subject inclusion and selection.[102] Overall, fMRI will currently is considered a research tool with a dearth of literature supporting any clinical suggestions.[2, 102]

The Prague statement suggested research regarding biomarkers would provide valuable information for clinical management, return to play guidelines, and long-term outcomes in regards to sports-related concussion. Biomarkers such as APOE ε4, Serum 100β, neuron specific enolase (NSE), IGF-1, myelin basic protein (MBP), fibroblast growth factor, Cu-Zn superoxide dismutase and tau polymerase have been studied in relation to predisposal, immediate recognition, and long-term consequences of single and or multiple concussive events.[2, 12] Even though a body of research exists regarding these markers, it has been suggested it is too premature to make clinical recommendations.[2, 12] Currently, APOE ε4, tau polymerase and S 100β have been researched more extensively compared to the other aforementioned biomarkers and will be elaborated on in this section.

The APOE gene has 3 major isoforms and is responsible for lipid transport in the brain, maintaining neural structural integrity, and recovery after neurological injury[23] APOE ε4 inhibits neural outgrowth and is a known risk factor for Alzheimer's disease, chronic traumatic encephalopathy in boxer's, blood vessel integrity and coagulation, and outcome after TBI.[23, 96] Research involving intercollegiate athletes, both male and female, and each athlete's concussion history from the prior eight years was recorded. Genomic DNA results showed a statistically significant relationship between the APOE promoter genotype TT and a history of one or more concussions. Specifically this relationship suggested patients with a higher level of genotype TT had a three-fold increase in having one or more prior concussions and a four-fold increase in having lost consciousness during that event. This increased expression of APOE ε4 has been shown to lead to increased expression of amyloid β-protein deposits ultimately leading

to neurofibrillary tangles, diagnostic criteria for Alzheimer's disease.[23] APOE ε4 expression post head injury has also been documented to lead to increased aspartate and glutamate levels, multiple ionic fluxes, and hypoperfusion or hyperglycolysis.[23] Despite these reports, other articles suggest no relationship between the APOE gene and concussion.[60] Small sample sizes, the cost and limited expertise, and rarity of the injury make concrete results difficulty to ascertain.

The tau gene is located on chromosome 17 and is responsible for tau protein production and has 17 polymorphisms. Mutations in this gene and its polymorphisms have been linked to neurodegenerative diseases such as chromosome 17-frontotemporal dementia.[23] Various forms of tau have been tested in regards to concussion. Cleaved tau is the tau gene is proteolytically modified after axonal injury and is suggested to reflect central nervous system injury better than total tau.[103] Research performed by Ma et al., found no relationship between cleaved tau and concussion. These results are substantiated by Terrell et al.,. [23, 97] Although, tau protein has been associated with moderate and severe head trauma, it appears that it may not be associated with mild brain injuries.[23, 97] Once again, sample size, cost, and accessibility of such research may limit conclusive statements regarding the tau gene and its association with concussion.

Thus far, serum 100β has been the most extensively research biomarker in regards to concussion. S 100β is a member of calcium binding proteins that is found mainly in the cytosol of glial cells but may also be found in cartilage and skin.[24, 98] Some researchers believe that post-head injury concentrations of S 100β are elevated which provides as possible indicator of head injury and severity. An initial study by Raabe et al., found a strong association between S 100β and outcome in severe head injury patients which suggested its ability to detect more severe head injuries.[98] In regards to concussion, S 100β has been suggested to be both a

marker for diagnostic and injury resolution purposes. Ingebrigtsen et al., found elevated S 100β levels post concussion which returned to normative values for a majority of participants 12 hours post injury. This group also reported a relationship between post-concussion symptomology and elevated levels of S 100β.[104] This and other research suggests S 100β to be a possible marker of brain injury regardless of severity.[60, 98, 99]

Critiques of S 100β suggest small sample sizes, discrepancies in study methodology, failure to report various statistics, and physiological issues such as a short half-life of S 100β (97 minutes) make it difficult to interpret research findings.[105] A review regarding S 100β suggested there is no strong relationship between it and concussion. This review also suggests of the papers included, six of eleven did not show a relationship between S 100β and post concussion symptoms.[24] These and more recent studies suggest S 100β may be a poor predictor of concussion and long term outcome.[24, 96, 100-102]

Post-traumatic Hypopituitarism

Hypopituitarism is a heterogeneous disease with diverse underlying diagnosis and was first reported in 1918.[25] Despite the diverse nature of the disorder, definitions are similar defining hypopituitarism as a "documented biochemical deficiency in at least one endocrine axis with associated pathology either in the pituitary or the hypothalamus."[25]

Epidemiology of Post-traumatic Hypopituitarism

An average of 1.4 million TBIs occur each year, including 1.1 million emergency department visits, 235,000 hospitalizations, and 50,000 deaths which is thought to be an underestimate. This estimate does not include those treated in outpatient settings, military incidences, and those who do not seek care.[5] One issue often overlooked with survivors of severe, moderate, and mild head injuries is hypopituitarism. The first reported case of hypopituitarism was reported between

1914 and 1918.[27, 103] A review by Benvenga et al., reported the initial estimates of head trauma accounting for 4 out of 595 cases of TBI accounting for hypopituitarism was grossly underestimated.[26] Studies published between 1942 and 1998 reported only 53 known cases of PTH. Benvenga et al., reviewed literature from 1970 – 1998 and found 367 unreported cases of PTH which increased awareness of TBI and its relationship to hypopituitarism.[26] Recently, a systematic review by Schneider et al., reported the prevalence of PTH to be 27.5 percent of 809 reviewed cases of TBI.[27] The authors of this article report that the prevalence of PTH to be 15 to 68 percent of TBIs in all literature reviewed.[27] In Spain, the prevalence of PTH is suggested to be 4.21 per 100,000 cases.[25] No known study has addressed PTH within the United States although awareness of trauma induced hypopituitarism has increased dramatically during the past 10 years.

Anatomical Considerations of the Pituitary

The pituitary (hypophysis) consists of an anterior (adenohypophysis) and posterior (neurohypophysis) lobe that differ in origin, development and structure and sits in the sella turrica in the base of the skull.[103, 104] The pituitary is regulated by the hypothalamic neurohormones released from the median eminence as well as positive and negative feedback loops from peripheral endocrine tissues. The anterior lobe of the pituitary is innervated by parvocellular neurons and the posterior lobe by the magnocelluar neurons.[106] The anterior pituitary receives blood from two sources: the superior hypophyseal arteries and the hypophyseal portal veins which carry regulatory hormones from the hypothalamus. The posterior lobe of the pituitary receives its blood supply through the inferior hypophyseal artery which is not dependent on the pituitary stalk and its related blood supply making it less vulnerable to injury.[103]

In regards to hormone production, each lobe of the pituitary contains specific hormone producing cells. The adenohypophysis houses lactotrophs, corticotrophs, thyrotrophs, gonadotrophs, and somatotrophs. These cells are listed in order from deep to superficial location in the pituitary. Lactotrophs produce prolactin (Prl), corticotrophs produce proopiomelanocortin (POMC) which results in adrenocorticotropic hormone (ACTH), lipotropin, melanocyte-stimulating hormone (MSH) and β-endorphin, thyrotrophs produce and release thyrotropin-releasing hormone (TRH), gonadotrophs produce luteinizing hormone (LH) and follicle stimulating hormone (FSH) and somatotrophs produce and release growth hormone (GH).[106]

The posterior pituitary contains axon terminals of magnocellular neurons located in the supraoptic and paraventricular nuclei. Neuron generated action potentials result in the release of neuropeptides by magnocellular neurons along with oxytocin and arginine vasopressin (AVP) from the posterior pituitary.[106]

Each of the aforementioned hormones has a physiological effect. Prl is responsible for breast development and milk production. The cleavage of ACTH from POMC and subsequent release leads to release of cortisol, aldosterone and androgens which regulate fluid and electrolyte balance, the inflammation response, and metabolism. TSH stimulates the thyroid to release thyrotropin (T_4) and its conversion to triiodothyronine (T_3) which assist in metabolism , growth and differentiation. The gonadotropins, LH and FSH aid in the production of estrogen, progesterone and testosterone which allow for reproduction function and behavior. GH and its regulator insulin-like growth hormone (IGF-1) aid in growth and differentiation.[106] Each hormone's invaluable contribution aids in an individual daily functioning.

Mechanisms of Post-traumatic Hypopituitarism

Most commonly, the cause of acquired PTH is a pituitary tumor which either damages the anterior pituitary cells or interferes with their hypothalamic control.[25] Other common causes of hypopituitarism include tissue necrosis via radiation therapy, hemorrhage, hypothalamic tumor and cranial radiation.[25, 104] Until recently, trauma was less commonly noted to be a common cause of PTH. Benvenga suggests that during the 1970's little to no attention was given to trauma as a cause of PTH, yet 40 percent of hypopituitarism were idiopathic in nature.[29] In 2000, a review increased awareness of trauma as mechanism for PTH. This review identified motor vehicle accidents (which are suggested to account for 75 percent of PTH cases), falls, assaults, child abuse and cranial gunshot wounds as mechanisms for PTH.[25, 29] Since the article's release, the causes of PTH have been validated.[31, 105-107]

Thus far, a majority of research has dealt with moderate and severe head injury cases defined by the Glascow Coma Scale (GCS) score of 9 to 13, and 0 to 8, respecitively. Few articles include mild head injuries defined as Glascow Coma Scale scores of 14 and 15 into their analysis.[3, 28, 108] Research including individuals with concussions report provocative findings suggesting PTH may be more common in mild head injuries than previously thought.[45] Schneider et al., suggest that pituitary function should be addressed in all TBI patients regardless of GCS.[109] Agha suggests that no association between CT scans and GCS scores exist. These measures have been traditionally employed and positive findings increase suspicion of PTH which provides reasoning for a lack of testing with mild head injuries.[31]

Pathophysiology of Post-traumatic Hypopituitarism

Although the anatomical position of the pituitary provides protection from mechanical injury, its placement in the sella turrica, its infundibular hypothalamic structures and its vulnerable

vascular supply make it vulnerable to injury. Specifically, post-traumatic edema may compromise portal blood supply leading to necrosis, pituitary stalk transaction or rupture may lead to anterior lobe infarction due to loss of blood supply with or without the presence of a skull fracture.[3, 26, 103] It is noted that the necrotic pattern follow the path of the long hypophyseal portal veins which are fed by the superior hypophyseal arteries and other small branches of the circle of Willis. The long portal veins pass through the diaphragma sella, where they vulnerable to mechanical compression from both brain and pituitary gland swelling and direct stalk injury.[3] Even mechanical occlusion of these vessels disrupts of the 70 to 90 percent of blood supply to the adenohypophysis.[3] The involvement of the pituitary without the presence of a basal skull fracture may be due to the involved strain rate inflicted by the mechanism as reported by Zhang et al.,.[44] As mentioned previously, the blood supply of the anterior lobe of the pituitary is independent of the posterior lobe making it more vulnerable to ischemic insults. Regardless of the mechanism of injury, ischemia affects the aforementioned anterior pituitary cells and the production of their respective hormones.

Anterior pituitary cells are suggested to be affected in order from most superficial to deep in regards to lack of blood supply.[3] Somatotrophs are located in the lateral wings of the anterior lobe and gonadotrophs are found in the pars distalis and pars tuberalis making them both vulnerable to ischemic insult due to their blood supply by the long hypophyseal portal system. Corticotrophs and thyrotrophs are located in the anteromedial portion of the gland making them less vulnerable to ischemic injury due to their vasculature derived from the short hypophyseal portal system.[3] This arrangement is suggested to account for gonadotroph and somatotroph deficiencies being the most common forms of PTH.[3]

Physiological and Psychological effects of Post-traumatic Hypopituitarism

As mentioned previously, each hormone has a physiological effect on the body, so it would make sense that a deficiency in any one or more of these hormones would have physiological consequences. GH deficiency (GHD) is the most common deficiency in the literature. GHD as reported in a systemic review accounted for 8 to 32.7 percent cases of TBI cases.[27] The effects of GHD include decreased muscle mass and strength, visceral obesity, fatigue, decreased quality of life, impairment of attention and memory, intelligence quotient, high anxiety, growth retardation in children, dyslipidemia and premature atherosclerosis.[3, 32, 108] Growth hormones regulator and second messenger IGF-1 also appears to play a key role in brain development and enhances early recovery in experimental brain injury.[3, 108] Of recent interest is GH's role in quality of life. Bavisetty et al., investigated GHD and quality of life and found those who were GHD reported greater depression, fatigue, energy, emotional well-being, and a trend towards worse general health compared to non-deficient controls.[108]

Second to deficiency in somatotroph dysfunction are gonadotroph deficiencies of LH and FSH which occur in 9 to 20 percent of TBIs.[27] Deficiencies in LH and FSH include oligoamenorrhea, loss of libido, dyspareunia, and infertility in women. In men, effects include loss of libido, impaired sexual function, mood impairment, loss of facial, scrotal, and trunk hair, decreased muscle mass, osteoporosis and anaemia. The aforementioned effects are also due to a secondary deficiency in estrogen or testosterone.[3, 32]

ACTH deficiency has been reported to occur in 5.7 to 18.8 percent of TBIs.[27] Cortiocotroph deficiency results in chronic fatigue, pallor, anorexia, weight loss, hypoglycemia, hypotension, anaemia, lymphocytosis, eosinophilia, hyponatremia. Acutely deficiency results in weakness,

dizziness, nausea, vomiting, circulatory collapse, fever and shock. In children deficiency may result in delayed puberty and failure to thrive.[32, 103]

TSH deficiency is the forth most common deficiency and is reported to account for approximately 1 to 10 percent of TBI cases.[27] The effects of TSH deficiency include tiredness, cold intolerance, constipation, hair loss, dry skin, hoarseness, cognitive slowing, weight gain, bradycardia, and hypotension.[32] TSH is a regulator of the conversion of T_4 to its active form of T_3 which is considered normal in most TBI patients.[32] Despite this report, T_3 has been shown to be deficient acutely after injury and return to normal levels one week post trauma in MTBI patients[52].

Treatment for Post-traumatic Hypopituitarism

The effects of hormonal deficiencies may corrected with replacement therapy given the underlying cause is corrected. For instance, hypogonadal men benefit from returned bone and muscle mass, sexual function, and a normalized hematocrit after testosterone replacement begins. GH replacement has been shown to enhance body composition, lipid variables, and quality of life.[32, 108] Schneider et al., suggest that it is important to remember that hypopituitarism may be accompanied by various disorders dependent upon the deficiency. These events must also be treated in concert with hormone replacement to make a full recovery.[32] The initial step to correcting PTH recognition. After recognition, hormone replacement and or correcting the underlying cause may drastically change the outcome of a PTH patient.

Psychometric Properties of Tests

In order for a test or measure to be considered valid it must first demonstrate the psychometric properties of objectivity and reliability. An absence of the aforementioned properties inherently leads to a lack of validity.

Baumgartner et al define objectivity as the close agreement between the scores assigned to each person by two or more judges. If no differences are expected in regards to the scoring of the instrument, no significant differences would be observed. Objectivity is also known as rater reliability. Intraclass correlations (ICCs) are utilized as evidence of objectivity.[63, 64]

In regards to reliability two types exist, internal and stability. Internal reliability refers to a consistent rate of scoring by the individual being tested throughout a test or, when multiple trials are administered, from trial to trial. Evidence of internal reliability for non-physical measures such as personality or intelligence tests may be provided by the split-half method, the Kuder Richardson 20, Cronbach's alpha, and Hoyt's analysis of variance.[64]

Stability reliability is also known as test-retest reliability. The basis for test-retest reliability is that a person will perform equally on the same test on two or more days. More specifically, reliability is defined as the consistency of test scores.[107] Factors affecting test-retest reliability are variance in subject performance, the measurement instrument may operate or be applied differently, or the person administering the measurement may change.[63] Stability reliability is most commonly calculated utilizing an ICC. Calculation of an ICC is based of the results of an Analysis of Variance. The formula to calculate an ICC (R) is shown below:

$$R = \frac{MS_A - MS_W}{MS_A + MS_W}$$

The value of R will fall between 0 and 1 with a coefficient closer to 1 representing a more reliable measure. In the past, Pearson correlations have been used to estimate test-retest reliability. This use of the Pearson coefficient is inappropriate. A Pearson *r* statistic is a bivariate measure of the relationship between 2 independent variables. Pearson's *r* is also insensitive to random or systematic error. Pearson's *r* is also known to overestimate the correlation when

sample sizes are limited. The ICC is the appropriate choice to assess reliability due to its ability to estimate agreement between scores on the same test at 2 or more points in time.[49]

Another form of reliability is alternate or equivalent form reliability. Determining this form of reliability is appropriate when alternate forms of the same test are created to reduce practice effects, cheating, or achieve a better estimate of a person's true score on a given measurement.

The alternate form method requires the same participants to complete all equivalent forms of a given test. Forms must be administered within in a short time period, but not so brief that examinee fatigue occurs. The correlation of equivalence is determined by calculating an ICC. Although no minimum requirements exist in regards to determining test equivalence, correlations of .80 to .90 have been suggested.[64]

Conclusion

Currently, research in sport-related concussion include increasing the body of literature regarding psychometric tests of current and recently developed assessment tools, investigating gender effects on injury risk, severity, and outcome, novel approaches to rehabilitation, establishment and implementation of new imaging techniques, possible biomarkers utilized as predictors and identifiers of sport-related concussion.[2] The past several years have been spent identifying the injury and creating tools to identify, manage, and make return-to-play decisions for athletes. This body of literature and future studies addressing the aforementioned topics suggests clinicians are and will provide better care to those victimized by this injury.

CHAPTER 3

METHODS

The purpose of this dissertation is to expand on the literature regarding the psychometric properties of computerized neuropsychological testing. The first study will address poor to moderate test-retest reliability coefficients which have been reported in previously published literature.[49, 108] The second study will address sources of error potentially associated with computerized neuropsychological testing. Specially, the second study will address mood state, a source of random error, and its potential influence on test performance.

Study 1: Reliability of a computerized neuropsychological exam

The purpose of this study is to estimate reliability coefficients for ImPACT at three clinically relevant time points. In addition, participant effort will be measured during each session using Green's Word Memory Test.

Participants: A total of fifty eight (N = 58) participants will be recruited to satisfy the requirements of have a power of .80 and an effect size of .75 as suggested by Maxwell and Delaney.[109] Participants were recruited from the undergraduate student body of the University of Georgia.

Exclusion criteria for this study includes English not being the participant's primary language; if he or she has been diagnosed by a physician with a learning disability; attention deficit disorder; or if a participant has been diagnosed as concussed within 6 months before or during the study.

Testing Sessions

Session 1: Recruited participants reported to Room 110 of the Ramsey Center where they received a detailed explanation of the study. After addressing any questions, each participants

reviewed and signed a University of Georgia Institutional Review Board Consent Form (Appendix 1). Initial measures collected included a health questionnaire (Appendix 2) and the POMS-B inventory. Participant were then divided into six groups.

The POMS-B inventory is a brief measure of six mood states including fatigue-inertia, vigor-activity, tension-anxiety, depression-dejection, confusion-bewilderment, and anger-hostility. Completion of the POMS-B inventory will take approximately five minutes. The Profile of Mood States (POMS) standard form was originally designed to assess the effects of various drugs on six hypothesized mood states. The hypothesized mood states served as constructs formed by 55 adjectives on a four-point likert scale. As the POMS evolved it provided 55, 57, and 67 adjectives rated a five point likert scale and was presented to a variety of populations ranging in health status, gender, and age. The present POMS standard form consists of 65 adjectives four point likert scale to formulate the six mood states and the composite total mood disturbance score.[110] The POMS-B was developed in 1989 as a briefer version of the original form to be delivered to patients dealing with intense stress or pain. The shorter form was also more tolerable than the 65-item version. The 30-item POMS-B consists of the five adjectives which had the highest factor loading on each of the six mood states. Likert scores for each adjective corresponding to each of the six mood states are summed providing six composite scores. Total mood disturbance is calculated via summing raw scores for fatigue-inertia, tension-anxiety, depression-dejection, confusion-bewilderment, and anger-hostility and subtracting the raw score for vigor-activity.[110] Currently, only experimental research studies provide any form of normative data for the POMS-B.

ImPACT is a computerized neuropsychological test which tests attention, memory, reaction time, and information processing speed.[111] ImPACT consists of eight tests including

immediate and delayed word recall, immediate and delayed design recall, a symbol match test, a three letter recall, the X's and O's test, and color-match test. For both word and design recall, examinees are presented 12 individually presented stimuli followed by a pause, and one more presentation of the same words or designs.

Following the second delivery, participants are asked whether a word or design was or not shown during the prior delivery. For the symbol match test participants are presented with a symbol at the bottom of the monitor. On the top of the screen each of the delivered symbols sits on top of a corresponding number. Participants are to click on number of the corresponding symbol presented. After three deliveries of each symbol the symbols above the numbers disappear. Examinees are then to click on the number which corresponded to the previously shown matching symbol. The X's and O's test asks participants to remember three highlighted X's and or O's amongst a field of randomly placed X's and O's. After an interference test, participants are then to click on the previously highlighted stimuli. The distraction test consists of responding to two stimuli as quickly and accurately as possible by clicking on the right or left button of a serial mouse. The three letter recall test presents examinees with three random letters followed by an interference test. The interference test is composed of counted backwards from 25 to 1 by clicking on a field of randomly placed numbers. After completing the interference test participants then type in the three previously presented letters. The final two tests ask participants to discern between words and designs that may or may not have appeared during the earlier word and design tests.

Combinations of two or more of the aforementioned tests are utilized to calculate five sub-scores including visual and verbal memory, reaction time, visual motor speed, and impulse control. Composite verbal memory is composed of the total memory percent correct, the number

of hidden symbols correctly recalled during the symbol match tests, and the total number letters correctly recalled during the three-letter test. Composite visual memory is percentage of designs and total X's and O's correctly recalled. Visual processing speed is the sum of X's and O's recalled divided by four and the average of the correctly recalled three-letters multiplied by three. Composite reaction time is composed of the reaction time during interference task, the reaction time of the symbols which were correctly recalled divided by three, and reaction time for the correction responses of the color match test. Impulse control is the composite of incorrect responses for the interference task for the X's and O's test and the number of commissions during the Color Match test.[111, 112]

In order to determine whether or not examinees provided a good effort leading to a valid test, ImPACT provides and utilizes invalidity criteria. These criteria can only be employed during baseline examinations given subsequent deliveries would occur post-concussion. In order for a baseline examination to be determined invalid a participant must meet one of the following criteria, score greater than 30 for the number of X's and O's incorrect; have a score greater than 30 for composite Impulse Control; correctly respond to less than 69% word memory; score less than 50% correct for design memory; and correctly recall less than eight letters for the Three Letter test.[112] Common causes of invalid baseline tests as reported by the manufacturer include failure to adhere or read directions, presence of Attention Deficit Disorder and or hyperactivity, excessive fatigue, "horseplay," while testing, left-right confusion during the interference task corresponding to the X's and O's test, and purposeful sabotage of one's baseline to set a low standard, potentially to be considered normal when actually concussed.[112] The ImPACT test battery has been shown to possess high sensitivity (79.2

percent) and specificity (89.4 percent) to sport-related concussion.[13] Participants will receive forms 1, 2 and 3 of ImPACT. [13, 54]

The Green's Word Memory Test (WMT) is a measure of participant effort. The Green's WMT presents 20 pairs of words to each participant. Each participant is scored based on five of six subtests in order to determine effort. Score calculated include immediate and delayed recall, consistency of responses which is defined as the percent agreement between immediate and delayed recall sections of the test. The last three scores are multiple choice, paired associates, and free recall. Green's WMT is divided into two components, divided by a 30 minute delay. Effort is based on scores of immediate recall, delayed recall, consistency of responses, and paired associates. Completion of the test takes approximately 40 minutes including the 30 minute delay. ImPACT and Green's WMT were administered via desktop computer and external mouse. Total testing time for each of the three sessions was approximately forty-five minutes.

Testing Environment: Each session took place in the St. Mary's Athletic Training Laboratory (Room 110B) of the Ramsey Center on the campus of the University of Georgia. Participants were comfortably seated in front of a desktop computer where they worked at their own pace. During the first session, the principal investigator assisted each participant with entry of demographic information and left during neuropsychological test delivery. The investigator did not actively participate in test administration and did not actively monitor participants during testing with the exception of the set-up of the computerized tests. At the completion of the first session the principal investigator scheduled the second testing session at approximately the same time of day as the baseline session.

Sessions 2 and 3: Forty-five days after the first session, participants returned to Room 110B of the Ramsey Center and completed the second administration of the POMS-B, ImPACT and

the Green's WMT. The forty-five day interval has been suggested to represent the estimated time between baseline testing and the first concussion assessment for a division 1 collegiate athlete.[49] The final assessment period occurred approximately five days after session two. This time interval represented the mean amount of time between the initial post-concussion evaluation and when an athlete reports asymptomatic.[49] The protocol on day 50 replicated both baseline and day 45 with the exception of an alternate form of ImPACT was delivered.

Study 2: The effect of mood states on computerized neuropsychological test performance.

The purpose of this study was to evaluate the influence of various mood states on the results of three independent administrations of a computerized neuropsychological test, specifically ImPACT. Mood states were assessed by the POMS-B inventory. Additionally, participant effort was measured at teach session.

Participants: Data from the same 156 participants was utilized to examine the influence of mood state on ImPACT composite scores. All participants were recruited from undergraduate courses held on the campus of the University of Georgia. Sample size was estimated based to satisfy the desired power of .80 and effect size of .75.[109]

Exclusion criteria included English not being the participant's primary language, if he or she had been diagnosed by a physician with a learning disability, attention deficit disorder, or as concussed within 6 months before or throughout the study.

Testing Sessions

Session 1: After recruitment, participants reported to Room 110 of the Ramsey Center where they received a detailed explanation of the study. After addressing any questions, reviewed and signed a University of Georgia Institutional Review Board Consent Form (Appendix 1). Once informed consent was obtained, participants completed a health questionnaire (Appendix 2) and

the POMS-B inventory. Participants then completed the initial portion of Green's WMT followed by ImPACT. The session concluded with the administration of the delayed component of the Green's WMT.

The testing protocol was identical to study 1 and employed ImPACT, the POMS-B, and the Green's WMT. Each session took place in the St. Mary's Athletic Training Laboratory (Room 110B) of the Ramsey Center on the campus of the University of Georgia. Participants were comfortably seated in front of a desktop computer where they worked at their own pace. During the first session, the principal investigator assisted each participant with entry of demographic information and then left during neuropsychological test delivery. The investigator did not actively participate in test administration and did not actively monitor participants during test delivery. At the completion of the first session the principal investigator scheduled each participant's second session at approximately the same time of day.

Sessions 2 and 3: Forty-five days after the first session, participants returned to Room 110B of the Ramsey Center and completed the POMS-B, an alternate form of ImPACT and the Green's WMT. The forty-five day interval has been suggested to replicate the estimated time between baseline testing and the first concussion assessment for a division 1 collegiate athlete.[49] The final assessment period occurred five days after session two. This time interval represents the mean amount of time between the initial post-concussion evaluation and when an athlete is re-tested after reporting asymptomatic.[49] The protocol on day 50 replicated the protocol from the previous two sessions with the exception of an alternate form of ImPACT will was delivered.

Statistical Analysis:

Specific Aim 1: Intraclass correlation coefficients (ICC) will be calculated for each composite score of ImPACT. Three ICC calculations were mad for baseline to day 45, baseline to day 50, and day 45 to day 50. Repeated measures analyses of variance were utilized to detect significant changes in effort across time.

Specific Aim 2: Pearson correlation coefficients were calculated for each mood state and each ImPACT composite score, at each time point. Identification of a moderate to strong correlation between any mood state and neuropsychological test performance may explain variability found in test performance. Repeated measures analyses of variance were utilized to detect significant changes in effort and individual test performance across time.

All data analyses were conducted using SPSS version 17.0 (SPSS, Chicago, IL) and statistical significance was set *a priori* $\alpha = .05$.

CHAPTER 4

Reliability of a Computerized Neuropsychological Test

[1] Resch JE, Ferrara MS, Brown CN, Macciocchi SN, Baumgartner TA, Walpert KM. To be submitted to the *Journal of Athletic Training.*

Abstract

Context Computerized neuropsychological testing is commonly utilized to assess sport-related concussion at all levels of competition. Reliability of such testing must be established prior to establishing validity and clinical utility as a measure of neurocognitive function pre- and post-concussion.

Objective The purpose of this study was provide test-retest reliability evidence for a commercially available computerized neuropsychological assessment over clinically relevant time points while concurrently assessing effort.

Design, Setting, and Patients Forty-five (N = 45) college aged healthy participants were enrolled. Each participant completed forms one, two and three of ImPACT at three separate time points: baseline, day 45 to day 50 after baseline, respectively. Participants also completed the Green's Word Memory Test pre- and post- completion of ImPACT as a measure of effort.

Main Outcome Measure Intraclass correlation coefficients were calculated for each composite score of ImPACT between baseline and day 45 and day 50, and between day 45 and day 50. Repeated measures analysis of variance (ANOVA) evaluated change in effort over time.

Results A majority of ICC values (.374 - .756) for each of the four ImPACT composite scores fell below what are considered acceptable ($x \geq .75$) for clinical interpretation with most interpreted as poor to moderate reliability. ImPACT incorrectly classified 22.2 and 28.9% of healthy participants at day 45 and day 50, respectively. Analysis of the Green's WMT suggests all participants demonstrated high effort at each time point. Significantly improved effort was indicated for delayed recall ($F_{(1,44)} = 10.133$, $p = .003$), consistency ($F_{(1,44)} = 7.367$, $p = .003$), multiple choice ($F_{(1,44)} = 3.791$, $p = ..031$), paired associates ($F_{(1,44)} = 9.246$, $p = .001$), and free recall ($F_{(1,44)} = 20.432$, $p = .000$) compared to baseline.

Conclusion The results of our study support previous research suggesting poor to moderate reliability exists for computerized neuropsychological testing at clinically relevant time points. Participants exhibited high effort throughout the study which may suggest that the poor to moderate ICC values may be due to systematic rather than random error.

Key Words intraclass correlation coefficient, ImPACT, reliability, concussion, Green's Word Memory Test

Introduction

During the past three decades sport-related concussion has emerged as a major concern amongst athletes and the sports medicine community. Although the physiology of concussion from time of insult to resolution remains largely ambiguous, various tests have been developed to manage this injury. One facet of the suggested battery of tests to assess sport-related concussion is neurocognitive testing.[1] [2] Neurocognitive testing has proven beneficial to depict subtle neurocognitive deficits associated with sport-related concussion and has been suggested to be the most sensitive measure of the concussed state when compared to self-reported symptoms and computerized posturography.[3, 4]

Since the 1980's, neuropsychological testing in the form of paper and pencil testing has been utilized to manage sport-related concussion.[5] During the late nineties, computerized neuropsychological testing became commercially available. Computerized neuropsychological tests are suggested to possess numerous benefits including standardized and rapid delivery, a centralized means of data storage, and multiple forms reducing the potential for practice effects while potentially measuring the same neurocognitive constructs as paper and pencil tests.[2] Despite the benefits and empirical evidence supporting the use of computerized testing, questions regarding the psychometric properties and ultimately the clinical utility of these tests have been raised.

In order for a measurement tool to be considered valid it must first possess objectivity and reliability. Baumgartner defines objectivity as the degree which multiple scorers agree on the values of the collected measures/scores. Reliability is defined as the degree which a measure is consistent and unchanged over a short period of time. Only after these criteria have been met may evidence of validity be gathered.[6] Randolph et al., reviewed the psychometric properties

of both paper and pencil and computerized neuropsychological testing platforms. This review suggested a dearth in the literature regarding the reliability of commercially available computerized platforms. Authors addressing the reliability of these programs report coefficients ranging from .32 to .82 for any one measure.[7] Broglio et al., conducted a reliability study employing multiple computerized platforms across clinically relevant time points (baseline, day 45, and day 50) in a healthy college aged sample in order to address the lack of reliability evidence.[8] Intraclass correlation values between baseline to day forty-five ranged from .23 to .66 across programs. The authors of this study reported several indices across platforms met or exceeded the .60 level which is considered acceptable by some for clinical decision making, but fell short of .90 which is suggested by others.[7-10] In a similar study, Schatz conducted a reliability study utilizing ImPACT (Immediate Post-Concussion Assessment and Cognitive Testing). Participants were tested at two time points (baseline and 2 years later) and they reported larger ICC coefficients (.47 to .75) across indices compared the findings of Broglio et al.,, but remained below .90. [7, 8, 11]

Although a multi-facet approach to concussion management is advised, various levels of sport have limited access to the recommended battery of tests due to time or financial restraints. [1, 2] Clinicians employed by these institutions may rely more heavily on computerized neuropsychological testing to determine the concussed state of an athlete and when making return to play decisions. The emphasis placed on computerized neuropsychological testing due to such restraints warrants increased focus on the psychometric properties, ultimately determining test validity.

The purpose of this study was to replicate the work performed by Broglio et al., utilizing solely ImPACT and Green's Word Memory Test as a measure of effort during clinically relevant

time points. We hypothesized that ImPACT indices would achieve acceptable ICC coefficients (> .75) inferring acceptable reliability for clinical decision making.[8]

Methods

Participants consisted of 51 (N = 51) healthy, college aged students from the general university population. Sample size was based on prior research and guidelines provided by Baumgartner and Chung for reliability studies using a one-way analysis of variance model to estimate the ICC for a single score.[8, 12] Exclusion criteria for participants included English not being their primary language, history of learning disability/attention deficit disorder, and/or participants has a history of a concussive injury within six months prior to or during the study.

ImPACT (ImPACT Applications, Pittsburgh, PA) is a computerized neuropsychological test which tests attention, memory, reaction time, and information processing speed.[111] ImPACT consists of eight tests including immediate and delayed word recall, immediate and delayed design recall, a symbol matching, three letter recall, X's and O's, and color-match tests. Combinations of two or more of the aforementioned tests are utilized to calculate five sub-scores including visual and verbal memory, reaction time, visual motor speed, and impulse control. In order to determine whether or not examinees provided a good effort leading to a valid test, ImPACT provides and utilizes invalidity criteria.[13] Common causes of invalid baselines as reported by the manufacturer include failure to adhere or read directions, presence of Attention deficit disorder and or hyperactivity, excessive fatigue, "horseplay," while testing, left-right confusion during the interference task corresponding to the X's and O's test, and purposeful sabotage of one's baseline to set a low standard, potentially to be considered normal when actually concussed.[13] The ImPACT test battery has been shown to possess high sensitivity

(79.2 percent) and specificity (89.4 percent) to sport-related concussion. Participants will receive forms 1, 2 and 3 of ImPACT. [13, 54]

The Green's Word Memory Test Test (WMT) (Edmonton, Canada) is measure of participant effort. The Green's WMT presents 20 pairs of words to each participant. Each participant is scored based on four subtests based on tests of immediate and delayed recall, which occur after a 30 minute delay. Effort is based on scores of immediate recall, delayed recall, consistency of responses, multiple choice, paired associates, and free recall of the word pairs. Completion of the test takes approximately 40 minutes including the 30 minute delay. ImPACT and Green's WMT were administered via desktop computer and external mouse. Total test time to complete both tests was approximately 45 minutes.

Testing Protocol

Baseline: The first session consisted of participants reading and signing an institutional review board approved informed consent form. Participants then completed a health questionnaire consisting of demographic information, concussion history, and current health status. At this time it was determined if participants would continue with the study. Participants then completed two computerized tests, ImPACT 6.7.723 and Green's WMT (Edmonton, Canada). Participants completed the initial portion of the Green's WMT followed by ImPACT, then the delayed portion of the Green's WMT.

Days 45 and 50: Participants returned to the research laboratory and completed the same questionnaire delivered at baseline testing at approximately the same time as their baseline session. This was done to account for a greater portion of variability in test performance. Completion of the health questionnaire at day 45 and fifty ensured participants were still eligible to participate in the study (i.e. a participants may have endured a concussion between baseline

and day 45.) Once the health questionnaire was completed and reviewed, the Green's WMT was delivered as previously described. Participants then completed forms two and three of ImPACT on days 45 and 50, respectively followed by the latter portion of Green's WMT.

Statistical procedures: To determine stability reliability of ImPACT, ICCs were calculated for verbal and visual memory, visual motor reaction time, reaction time, impulse control, and symptomology. Reliability coefficients were calculated each variable for baseline and day 45, baseline and day 50, and day 45 and day 50 utilizing the one-way random model (1). The one-way model is calculated as $[MS_A – MS_W] / [MS_A + (k –1) MS_W]$ where MS_A is the mean squares among participants and MS_W is equal to the mean squares within participants and *k* equals the number of observations.[14, 15] Intraclass correlation coefficients range from zero to one. Larger coefficients suggest higher reliability.[16-18] Acceptable reliability is difficult to define. Pedhazur suggests that in determination of what acceptable reliability is depends on what types of decisions are made on the basis of the scores and the possible consequences of the decisions.[17] It has been suggested coefficients greater than .70 are acceptable while others suggest .90 is needed for clinical decision making.[7, 16] For this study, an ICC of .75 will be considered acceptable for clinical decision making.[16]

Effort was determined utilizing Green's WMT according to the manufacturer instructions. A repeated measures analysis of variance (ANOVA) was utilized to detect differences in effort across the three time points. Greenhouse-Geisser corrections were implemented when sphericity violations occurred. A Bonferroni adjustment was made for multiple pairwise comparisons used during post-hoc analysis. Pearson correlations were utilized to determine any relationship between academic achievement exemplified by the Scholastic Aptitude Test (SAT) and cumulative grade point average and each ImPACT composite score. All data analyses were

performed utilizing SPSS version 17.0 (SPSS, Chicago, IL) and statistical significance was set at $\alpha \leq .05$.

Results

If a participant did not complete all three time points or if his or her baseline assessment was determined to be invalid by the criteria suggested by ImPACT[13] then their data were removed from subsequent analyses. Removal of incomplete and or invalid data allowed for optimal ICC calculations. There were a total of 51 participants (N= 51) and six did not meet the criteria.

The final analysis consisted of 45 participants (17 males and 28 females) aged 20.94 $\pm$ 1.72 years, height 171.02 $\pm$ 10.37 cm, weight 67.13 $\pm$ 13.91 kg, and self-reported an average cumulative grade point average of 3.409 $\pm$ .39, and Scholastic Aptitude Test score of 1212.78 $\pm$ 122.84, verbal and math only. Due to inconclusive evidence to support differences between genders on computerized neuropsychological test performance, the final sample is representative of both males and females. Correlational analysis between SAT and ImPACT composite scores revealed one significant positive correlation between composite visual memory at day 45 and SAT (r = .400, p = .05). Analysis between cumulative grade point average and each composite score of ImPACT at each time point resulted in a positive correlation for GPA and composite visual memory (r = .418, p = .004) at day 45, and significant negative correlations for GPA and impulse control (r = -.365, p = .014) and (r = -.343, p = .021) for baseline and day 50, respectively.

Six participants (13.3%) reported previous history of concussion as diagnosed by a medical doctor or certified athletic trainer. No more than two prior concussions were reported by these six participants. No participants were excluded from this study due to sustaining a concussion six months prior to or during the study. Participants were tested at 47.27 $\pm$ 2.74 days after baseline

within 44.89 ± 67.87 minutes of their baseline time. Day 50 occurred approximately 6.90 ± 1.10 days after day 45 within approximately 52.22 ± 61.45 minutes of the cohort's baseline time. Descriptive data for this sample are presented in table 4.1.

Neuropsychological ICC Results

Mean scores and standard deviations for each ImPACT composite score by time point are presented in table 4.2. Our results were similar to those previously reported in the literature [8] Calculated ICC values for each ImPACT sub-score for baseline to day 45, baseline to day 50 and day 45 to day 50 are presented in table 4.3. The majority of ICC values between baseline and day 50 were higher than baseline to day 45 and day 45 to day 50. The highest ICC values calculated were for composite visual motor speed and reaction time scores. The ICC values for composite verbal memory were considerably lower than the other composite scores. ICC values for composite verbal and visual memory and visual motor speed only differed by .024 to .094 between time points suggesting relatively consistent ICC values. These results suggest consistently poor to moderate reliability. The consistent poor to moderate values, all ICC values for each measure fell below what is considered acceptable for clinical utility.

We also assessed for practice effects across testing sessions. Repeated measures ANOVA indicated violations of sphericity for reaction time ($W = .835$, $p = .020$). Post Greenhouse-Geisser corrections were employed to account for this violation. Significant differences were noted across time for visual motor speed ($F_{(2,88)} = 4.078$, $p = .020$). Post hoc paired *t* tests revealed significant decreases between baseline and day 50 ($t_{(44)} = -2.122$, $p = .039$) and day 45 and day 50 ($t_{(44)} = -2.521$, $p = .015$). An increased score for visual motor speed indicates an improvement in performance. Since there was an increase of approximately 2 points compared to day 45, this finding suggests potential practice effects exist for this measure when delivered over

a short period of time. All other composite scores, although not significantly different across time, provided evidence of potential practice effects from baseline to day 45 and day 45 to day 50.

Test-retest reliability coefficients remained consistent with previously reported values.[11] Although it is difficult to ascertain a standard ICC value for acceptable reliability, a majority of ImPACT composite scores fell below .75 which is suggested to be acceptable for clinical decision making.

ImPACT is designed to determine a significant change in performance compared to baseline measures. This study utilized this feature to compare day 45 and 50 to baseline. No participant sustained a concussion throughout the duration of the study nor reported a concussive event six months prior to enrollment. This criteria deems any significant change a false positive for cognitive impairment. On day forty-five, 22.2% (n = 10) of participants achieved a score significantly different on one or more composite scores compared to baseline. On day fifty, 28.9% (n = 13) of participants were considered significantly different than baseline. Detailed results of this analysis are presented in table 4.4.

Analysis of Green's WMT resulted in violations of sphericity for all scores excluding multiple choice which resulted in a significant difference across days ($F_{(2,88)} = 3.791$, $p = .026$. Specifically violations of sphericity were noted for immediate recall ($W = .256$, $p = .000$), delayed recall ($W = .843$, $p = .025$), consistency ($W = .739$, $p = .001$), paired associates ($W = .805$, $p = .009$), and free recall ($W = .771$, $p = .004$). Post Greenhouse-Geisser correction for these violations; significant differences across days existed for delayed recall ($F_{(1.729,76.061)} = 6.000$, $p = .006$), consistency ($F_{(1.586,69.778)} = 7.367$, $p = .003$), paired associates ($F_{(1.673, 73.618)} = 9.246$, $p = .001$), and free recall ($F_{(1.627, 71.579)} = 20.432$, $p = .000$). Scores for immediate recall,

delayed recall, and consistency variables exceeded 85% indicating participants provided good effort at each time point.[19] Means and standard deviations for each variable at each time point can be found in table 4.5. A review of each participant's scores revealed no instance of poor effort at any time point.

Discussion

The purpose of this study was to replicate and extend the literature addressing reliability of computerized neuropsychological tests utilized to assess sport-related concussion.[8, 11] Our hypothesis was each composite score of ImPACT would achieve an ICC value equal to or exceeding what is suggested for clinical interpretation.[7] Our design employed clinically relevant time points and a measure of participant effort and revealed a range of ICC values similar to those previously reported in the literature.[8] Specifically, ICC values for composite scores of ImPACT ranged from .374 to .756 falling short of what is acceptable for clinical decision making.[7] Despite three ICC values exceeding what is considered acceptable reliability for clinical decision making, the majority of values fell into the range of acceptable reliability for research purposes.[17] Pedhazuar suggests that ICC values of .50 to .60 are acceptable for the early stages of research testing predictor tests or hypothesized measures. [17]

Intraclass correlations are a measure of reliability or reliability. Higher ICC values suggest a more consistent measure of a particular construct, in this case, cognitive functioning. In classical test theory the formula $X = T + E$ defines what is accounted for in an individual's observed score where X represents an individual's observed score, T represents an individual's true ability and E represents error.[16] Error can be dissected into random and systematic components. Random error affects an individual's score based on chance and may be introduced via testing time and location, administration errors, and fluctuation in an examinee's state.[16] Our study controlled

for random error by testing individuals at approximately the same time of day, testing in the same environment, following standardized testing procedure for both the study and ImPACT as suggested by manufacturer's instructions[20], and by controlling for effort. Participants who were identified as having one or more invalid baseline measures were also removed from analysis. Systematic error is defined as some particular characteristic of the examinee or test that has nothing to do with the construct of interest.[16] In the context of this study, sources of systematic error could have been a learning disability, being diagnosed with a concussion during testing, or unclear instructions. Systematic error was controlled for via exclusion criteria, implementing a standardized protocol as well as abiding by the manufacturers' instructions.[20] Despite our efforts to control for both random and systematic error performance across the ImPACT indices remained relatively inconsistent. Further studies are needed to discern between the influence of random and systematic error and an individual's performance on computerized neuropsychological tests.

In 2005, Randolph questioned the use of neuropsychological testing in the management of sport-related concussion detailing lack of psychometric evidence supporting the reliability and validity of paper and pencil and computerized neuropsychological platforms.[7] Broglio et al., addressed the reliability of three commercially available platforms and reported ICC values ranging from .15 to .66 utilizing three clinically relevant time points. The authors reported the ICC values of ImPACT to range from .23 to .61. [8] Schatz et al., reported the test-retest reliability of ImPACT to range between .46 and .74 over a longer two year period. Our study differs from that of Broglio et al., due to using one computerized test to assess cognitive status while controlling for effort and time of day. Comparisons of our ICC values to the aforementioned authors' initial time points reveal our values are more similar to those of Schatz

et al.,.[8, 11] That said, although our ICC values are slightly higher than Broglio et al.,, a majority of the values fall below what is suggested for clinical interpretation.[7, 8, 16, 17]

Computerized neuropsychological testing, specifically ImPACT, has been shown to exhibit a sensitivity of approximately 79 to 93 percent when classifying individuals as concussed.[3, 21, 22] Individuals scores are "bolded," if they scores significantly different compared to baseline performance utilizing reliable change methodology. Significant differences are thought to occur due to cerebral insult resulting in cognitive deficits. Considering our sample was composed of healthy college aged adults who provided good effort throughout testing, ImPACT identified 20 to 29 percent of the sample as impaired on one or more index at one or both time points following baseline testing.

As a cohort, with the exclusion of visual motor speed, all ImPACT indices improved over baseline performance at day 45 which remained consistent with group performance at day 50. These findings support suggestions that two administrations of computerized neuropsychological tests should occur to achieve a closer estimate of an individuals' true score.[7, 8, 23] The delivery of two tests, preferably within the same testing session would increase an individual's familiarity with test and reduce practice effects which may ultimately increase ICC values allowing for more certainty when interpreting results at the expense of temporal and financial cost.[7, 8, 23]

Green's WMT was employed to determine individuals who provided less than optimal effort at each time point. Results of the Green's WMT reveal individuals as well as the sample provided a good effort at each time point. Good effort is defined by scoring greater than 85% on immediate recall, delayed recall, and consistency portions of the WMT test.[8, 19] All measures of the WMT increased significantly, excluding immediate recall, when compared to baseline.

Scores for day 50 when compared to day 45 for delayed recall, multiple choice, paired associates, and free recall were significantly higher. Due to these results we feel that results of ImPACT testing were not a result of poor effort. Good effort of the sample is also evident due to only 1 of 51 participants (1.9%) was invalidated by ImPACT at baseline.

In regards to intellectual ability, our sample scored higher than the national average in regards to the SAT.[24] Our correlational analysis revealed that SAT performance explained approximately 16% of variance associated with composite visual memory score at day 45. No other significant relationships were found suggesting SAT performance does not influence performance on ImPACT. The same analysis was conducted utilizing a measure of academic achievement, cumulative grade point average and each composite score of ImPACT at each time point. A significant positive correlation between GPA and composite visual memory score at day 45 accounted for approximately 13% of variance associated with performance. Significant negative correlations for GPA and impulse control at baseline and day 45.The latter negative correlations suggest a higher GPA was associated with a lower composite impulse control score suggesting better effort and or concentration. Analysis of the results of Green's WMT, SAT scores and the negative correlations between high GPA and composite impulse control suggest suboptimal effort was an unlikely source of low ICC values.

The purpose of this study was to replicate the work performed by Broglio et al., utilizing solely ImPACT and Green's Word Memory Test as a measure of effort during clinically relevant time points. This study is not without limitations including unequal numbers of males and females. Further investigation of the reliability of computerized neuropsychological testing needs to address differing time points, various sources of random error such as mood, sleep

deprivation, etc…, as well as performance by examinees of differing academic backgrounds and age groups.

Conclusion

Our study addressed the test-retest reliability of a single commercially available computerized neuropsychological test. Our results showed slightly better ICC scores than reported by Broglio et al., and our data support the more recent findings by Schatz regarding the stability-reliability of ImPACT delivered in a healthy sample.[8, 11] Despite this slight improvement no ICC value met the suggested .75 which is suggested for clinical utility, with all values ranging from poor to moderate reliability. Broglio et al., hypothesized that a longer time period between baseline and follow up testing would further decrease ICC values; this hypothesis which was contradicted by Schatz et al.,. [8, 11] Despite our effort to control for random and systematic error associated with participants' true scores, performance remained inconsistent across time.

During the past five years a considerable amount of literature has addressed a multi-facet approach to sport-related concussion management.[1-3] Despite these recommendations, which suggest baseline and post-injury assessment of self-reported symptoms, neuropsychological testing, postural stability and a physical examination time and financial restraints may limit a clinicians access to one or more of the above. This scenario may prove problematic if clinical diagnosis and evaluation are based solely on one or two measures. The findings of this study support caution when interpreting the results of computerized neuropsychological tests and encourage a multi-facet approach including a stepwise approach to exertional testing.[1] The establishment of a protocol incorporating the aforementioned measures based on the consensus of the sports medicine team coupled with education of athletes and their parents, coaches and other medical personal will help ensure an appropriate return-to-play decision.

Sample Demographics									
Gender	Gender (n = 45)	Age (Baseline)	Age (Day 45)	Age (Day 50)	Height (cm)	Weight (kg)	Cumulative GPA	Years of Education	SAT
Males	(n = 17)	21.11 (1.716)	21.259 (1.710)	21.277 (1.723)	179.593 (8.742)	79.097 (14.482)	3.376 (.474)	13.706 (1.047)	1223.462 (153.669)
Females	(n = 28)	20.832 (1.739)	20.953 (1.751)	20.957 (1.751)	165.826 (7.472)	59.869 (6.808)	3.430 (.333)	13.536 (1.290)	1202.857 (90.335)
Total	(N=45)	20.938 (1.717)	21.069 (1.722)	21.078 (1.729)	67.333 (4.084)	150.378 (31.165)	3.409 (.388)	13.600 (1.195)	1212.78 (122.838)

Table 4.1: Means and (Standard Deviations) of Participant Demographics

ImPACT Composite scores Means (SD) (n = 45)					
	Memory Composite Verbal	Memory Composite Visual	Visual Motor Speed	Composite Reaction Time	Impulse Control
Baseline	.895 (.088)	.788 (.125)	42.003 (6.183)	.557 (.085)	7.356 (5.824)
Day 45	.924 (.082)	.819 (.097)	41.572 (5.870)	.545 (.070)	6.711 (4.966)
Day 50	.920 (.076)	.805 (.116)	43.436*† (6.745)	.539 (.068)	6.533 (4.595)

Table 4.2: Means and (Standard Deviations) for ImPACT Composite Scores by Time Point. * = significant difference from baseline $p \leq .05$ † = significant difference from day 45 $p \leq .05$

ImPACT Intraclass Correlation Values (Forms 1,2,3) *n = 45*					
	Memory Composite Verbal	Memory Composite Visual	Visual Motor Speed	Composite Reaction Time	Impulse Control
Baseline to Day 45	.454	.523	.756	.568	.819
Baseline to Day 50	.374	.523	.738	.487	.426
Day 45 to Day 50	.397	.547	.662	.705	.506

Table 4.3: ICC values for each ImPACT composite score

	ImPACT						
	Memory Composite Verbal	Memory Composite Visual	Visual Motor Speed	Composite Reaction Time	Impulse Control	Self Reported Symptoms	Impaired on any variable
Day	n = 4	n = 1	n = 7	n = 2	n = 0	n =1	n = 10
45	8.9%	2.2%	15.6%	4.4%	0%	2.2%	22.2%
Day	n = 5	n = 4	n = 3	n = 4	n = 0	n = 1	n = 13
50	11.1%	8.9%	6.7%	8.9%	0 %	2.2%	28.9%

Table 4.4: Number of participants, and percentage of total labeled as significantly impaired by ImPACT's automated feature indicating significant statistical change compared to baseline

Green's Word Memory Test Composite Scores						
	Immediate Recall	Delayed Recall	Consistency	Multiple Choice	Paired Associates	Free Recall
Baseline	96.837	98.667	97.722	96.889	95.667	70.611
(n = 45)	(14.506)	(2.105)	(3.007)	(4.168)	(8.893)	(12.760)
Day 45	98.844	99.244	98.756 *	97.00	98.000 *	78.156 *
(n = 45)	(4.539)	(1.468)	(1.845)	(4.931)	(6.431)	(16.106)
Day 50	99.611	99.744 * †	99.356 *	98.889 * †	99.444 * †	84.900 * †
(n = 45)	(1.06)	(.736)	(1.384)	(2.798)	(3.057)	(10.577)

Table 4.5: Means and (Standard Deviations) for each Composite Score of Green's Word Memory Test. * = $p \leq .05$† = $p \leq .01$

References

1. McCrory P, Meeuwisse W, Johnston K, et al.: Consensus statement on concussion in sport: the 3rd International Conference on Concussion in Sport held in Zurich, November 2008. J Athl Train 2009; 44(4): 434-48.

2. Guskiewicz KM, Bruce SL, Cantu RC, et al.: National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion. J Athl Train 2004; 39(3): 280-297.

3. Broglio SP, Macciocchi SN, Ferrara MS: Sensitivity of the concussion assessment battery. Neurosurgery 2007; 60(6): 1050-7; discussion 1057-8.

4. Collins MW, Grindel SH, Lovell MR, et al.: Relationship between concussion and neuropsychological performance in college football players. JAMA 1999; 282(10): 964-70.

5. Macciocchi SN, Barth JT, Littlefield L, Cantu RC: Multiple Concussions and Neuropsychological Functioning in Collegiate Football Players. J Athl Train 2001; 36(3): 303-306.

6. Baumgartner TA, Hensley, L.D.: Conducting and Reading Research in Health and Human Performance 4th ed. New York, NY: McGraw-Hill, 2006.

7. Randolph C, McCrea M, Barr WB: Is neuropsychological testing useful in the management of sport-related concussion? J Athl Train 2005; 40(3): 139-52.

8. Broglio SP, Ferrara MS, Macciocchi SN, Baumgartner TA, Elliott R: Test-retest reliability of computerized concussion assessment programs. J Athl Train 2007; 42(4): 509-14.

9. Linda Crocker JA: Introduction to Classical & Modern Test Theory. Mason, OH: Cengage Learning, 2008.

10. Baumgartner TA, Hensley, L.D.: Conducting and Reading Research in Health and Human Performance 4th ed. New York, NY: McGraw-Hill, 2006.

11. Schatz P: Long-term test-retest reliability of baseline cognitive assessments using ImPACT. Am J Sports Med; 38(1): 47-53.

12. Baumgartner TA, Chung, H: Confidence Limits for Intraclass Reliability Coefficients Measurement in Physical Education and Exercise Science 2001; 5(3): 179-188.

13. Lovell MR: ImPACT 2007 (6.0) Clinical Interpretation Manual. Pittsburgh, PA ImPACT, 2007.

14. McGraw KO, S.P. Wong: Forming Inferences About Some Intraclass Correlation Coefficients Psychological Methods 1996; 1(1): 30-46.

15. Shrout PE, Fleiss, J.L.: Intraclass Correlations: Uses in assessing Assessing Rater Reliability. Psychological Bulletin 1979; 86(2): 420-428.

16. Crocker L AJ: Introduction to Classical & Modern Test Theory. Mason, OH: Cengage Learning, 2008.

17. Pedhazur EJ, Pedhazur Schmelkin, L.: Measurement, Design, and Analysis: An Integrated Approach. Hillsdale, NJ: Lawrence Erlbaum Associates 1991.

18. Maxwell SE, Delaney, H.D.: Designing Experiments And Analyzing Data 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates 2004.

19. Green P: Green's Word Memory Test for Windows: User's Manual and Program. . Edmonton, Canada: Green's Publishing Inc. , 2005.

20. Lovell M: ImPACT 2007 (6.0) Software User's Manual Pittsburgh, PA ImPACT Applications Inc. , 2007.

21. Schatz P, Pardini JE, Lovell MR, Collins MW, Podell K: Sensitivity and specificity of the ImPACT Test Battery for concussion in athletes. Arch Clin Neuropsychol 2006; 21(1): 91-9.

22. Van Kampen DA, Lovell MR, Pardini JE, Collins MW, Fu FH: The "value added" of neurocognitive testing after sports-related concussion. Am J Sports Med 2006; 34(10): 1630-5.

23. Hinton-Bayre AD, Geffen GM, Geffen LB, McFarland KA, Friis P: Concussion in contact sports: reliable change indices of impairment and recovery. J Clin Exp Neuropsychol 1999; 21(1): 70-86.

24. 2009 College-Bound Seniors Total Group Profile Report Total Group. CollegeBoard Inspiring Minds, 2009; 1 - 18.

CHAPTER 5

The Influence of Mood State on a Computerized Neuropsychological Test

[1] Resch JE, Ferrara MS, Brown CN, Baumgartner TA, Macciocchi SN, Walpert KM. To be submitted to the *British Journal of Sports Medicine.*

Abstract

Context One component of the suggested battery of tests utilized to assess sport-related concussion is neuropsychological (CNP) testing. The reliability of CNP testing has demonstrated poor to moderate which limits clinical utility. Potential reasoning for low test-retest reliability is random error caused by extraneous variables.

Objective The purpose of this study was to potentially identify mood state as a source of random error which would help explain the reported low test-retest reliability of computerized neuropsychological tests.

Design, Setting, and Patients One hundred and eight (n = 108) college aged healthy participants were enrolled. Each participant completed the Profile of Mood States Brief Form (POMS-B) prior to completing a commercially available computerized neuropsychological test at three clinically relevant time points, baseline, day 45, and day 50. Participants also completed the Green's Word Memory Test (WMT) prior to and after the administration of the computerized neuropsychological test as a measure of effort.

Main Outcome Measure Pearson correlation coefficients were calculated for each of the seven POMS-B factors and total symptom and each composite score of ImPACT. Repeated measures analysis of variance (ANOVA) was utilized to determine significant differences across time in regards to ImPACT composite scores, POMS-B factors, and Green's (WMT)

Results One significant correlation existed between the ImPACT composite score impulse control and the POMS-B factor, fatigue-inertia which occurred at baseline. ImPACT's total symptom score was most consistently correlated to multiple POMS-B factors. Repeated measures ANOVA indicated an improvement of effort on the Green's WMT for delayed recall ($F_{(1.719,\ 183.907)} = 20.925$, $p < .001$), consistency ($F_{(1.750,185.476)} = 20.186$, $p < .001$), multiple choice

($F_{(1.481, 156.958)} = 9.007$, $p = .001$), paired associates ($F_{(1.298, 137.590)} = 29.811$, $p < .001$), and free recall ($F_{(1.827, 193.679)} = 97.671$, $p < .001$) over baseline.

Conclusion Our results suggest that although mood states as measured by the POMS-B were significantly correlated total symptom score of ImPACT, performance remained either consistent or improved. Future research is needed in a collegiate athletic sample to determine if potentially higher mood state values influence computerized neuropsychological test results.

Key Words mood state, ImPACT, random error, concussion, Green's Word Memory Test, correlation

Introduction

Current management of sport-related concussion includes a physical examination, self-reported symptoms, balance assessment, and neuropsychological testing. Each of the aforementioned tests is most beneficial when compared to a baseline examination.[1, 2] The baseline examination serves to establish the individual athlete's "normal," pre-injury performance taking into account age, education and extraneous variables such as learning disabilities and or attention deficit disorder.[1] In order for the results of this testing to be the most accurate it is important for the athlete to achieve a score as close to his or her true score as possible.

Computerized neuropsychological testing has been recognized as the most sensitive of the aforementioned tools utilized to assess sport-related concussion.[3] Despite this finding, psychometric issues including reliability and subsequent validity have been brought into question.[4, 5] Test-retest reliability coefficients have ranged from poor to acceptable for clinical interpretation dependent upon the computerized platform utilized, construct in question, and time intervals between tests.[4, 6-8] The inconsistency associated with reported reliability coefficients may be due to either systematic and or random error. Systematic error is defined as error which consistently affects an individual's score because of some particular characteristic of the person or test and has nothing to do with the construct being measured.[9] Examples of systematic error in this context are an athlete with attention deficit disorder or a learning disability or a poorly designed test. Random error is defined as error which affects an individual's score purely be chance happening. Sources of random error in this context are lack of effort, administration errors, and fluctuation's in the examinee's state.[9] Systematic errors do not result in an inconsistent instrument but may be inaccurate and limit clinical utility. Random error reduces

both the consistency and clinical utility of test results.[9] Each source of error impacts an individual's true score resulting in the observed score.

The purpose of this study was to determine whether mood state, a source of random error, was correlated to one or more composite scores of ImPACT. Identification of various mood states which influence an individual's neurocognitive performance may help clinicians obtain a more accurate neuropsychological baseline measurement.

Methods

This project was part of a larger study involving the analysis of ImPACT scores. Although methodologies are similar, results will differ due to sample composition. One hundred and fifty-two (N = 152) college aged participants were recruited from the general student body to participate. All participants reviewed and signed an Institutional Research Board consent form prior to beginning the study. Participants then completed a health history questionnaire. Exclusion criteria included English not being his or her primary language, the participant being diagnosed by a physician or certified athletic trainer as concussed within six months of the start of the study, and if they had been diagnosed by a physician with Attention Deficit Disorder or a learning disability. All exclusion criteria were included in the form of a health questionnaire item and upon review participants were either included or excluded for participation at that point and time.

ImPACT (ImPACT Applications, Pittsburgh, PA) is a computerized neuropsychological test which tests attention, memory, reaction time, and information processing speed.[111] ImPACT consists of eight tests including immediate and delayed word recall, immediate and delayed design recall, symbol match, three letter recall, X's and O's, and color-match tests. Combinations

of two or more of the aforementioned tests are utilized to calculate five sub-scores including visual and verbal memory, reaction time, visual motor speed, and impulse control.

ImPACT's composite symptom score consists of 22 self-reported symptoms ranked on a likert scale from 0 (mild) to 6 (severe). To simplify these results the authors' utilized Piland's 3-factor model based on 16 self-reported symptoms for the Head Injury Scale including somatic, neurobehavioral, and cognitive constructs. Piland's study focused on only 16 of the 22 self-reported symptoms measured by ImPACT due to the self-reported symptoms most commonly reported in the literature.[10, 11] For this correlational analysis only those sixteen symptoms were utilized. Somatic self-reported symptoms included headache, nausea, vomiting, balance problems, and numbness and tingling. Neurobehavioral self-reported symptoms included trouble falling asleep, sleeping more or less than usual, drowsiness, sadness, and nervousness. The cognitive construct consisted of feeling slowed down, feeling mentally foggy, difficulty concentrating, and difficulty remembering.

In order to determine whether or not examinees provided a good effort leading to a valid test, ImPACT provides and utilizes invalidity criteria.[10] Common causes of invalid baselines as reported by the manufacturer include failure to adhere or read directions, presence of Attention deficit disorder and or hyperactivity, excessive fatigue, "horseplay," while testing, left-right confusion during the interference task corresponding to the X's and O's test, and purposeful sabotage of one's baseline to set a low standard, potentially to be considered normal when actually concussed.[10]

The POMS-B inventory is a brief 30-item measure of six mood states including fatigue-inertia, vigor-activity, tension-anxiety, depression-dejection, confusion-bewilderment, and anger-hostility. Completion of the POMS-B inventory took approximately five minutes. The POMS-B

was developed in 1989 as a briefer version of the original form to be delivered to patients dealing with intense stress or pain. The shorter form was also more tolerable than the 65-item version. The 30-item POMS-B consists of the five adjectives which had the highest factor loading on each of the six mood states. Likert scores for each adjective corresponding to each of the six mood states are summed providing six composite scores. Total mood disturbance is calculated via summing raw scores for fatigue-inertia, tension-anxiety, depression-dejection, confusion-bewilderment, and anger-hostility and subtracting the raw score for vigor-activity.[12] Currently, only experimental research studies provide any form of normative data for the POMS-B.

The Green's Word Memory Test (WMT) is measure of participant effort. The Green's WMT presents 20 pairs of words to each participant. Each participant is scored based on five of six subtests in order to determine effort. Score calculated include immediate and delayed recall, consistency of responses which is defined as the percent agreement between immediate and delayed recall sections of the test. The last three scores are multiple choice, paired associates, and free recall. Green's WMT is divided into two components, divided by a 30 minute delay. Effort is based on scores of immediate recall, delayed recall, consistency of responses, and paired associates. Completion of the test takes approximately 40 minutes including the 30 minute delay.[113]

Participants were required to attend three testing sessions at baseline, 45 and day 50 at approximately the same time of day. During each session participants initially completed the same health questionnaire completed at enrollment to ensure health status had not changed. Participant's then completed the Profile of Mood States (POMS) (Multi-Health Systems Inc., North Tonawanda, NY) with the instructions to respond to each mood description, "as of right now." Participants then completed the initial component of the Green's Word Memory Test

(WMT) (Green's Publishing, Edmonton, Alberta, Canada) followed by delivery of the concussion management system ImPACT Version 6.723 (ImPACT Applications, Pittsburgh, PA). Following the completion of ImPACT participants then completed the delayed component of Green's WMT. Total testing time for each session was approximately 45 to 60 minutes. All testing was performed in a controlled laboratory setting. All tests were delivered according to the manufacturers' instructions.[12, 14, 15]

Statistical Analysis: Analysis of variance was utilized to determine group differences between POMS scores, ImPACT composite scores, and Green's WMT at each time point. Pearson correlation coefficients were calculated between each of the seven factors of the POMS and the total symptom score and composite scores of ImPACT. A repeated measures analysis of variance (ANOVA) was utilized to detect differences in effort over time. Post hoc testing was performed utilizing paired *t* tests. Significance testing was performed with $\alpha = .05$. Data analysis was performed SPSS Version 17.0 (Chicago, IL).

Results

Participants

Results from 108 of the recruited 152 participants were included in the final analysis. Thirty-four participants were excluded due to meeting exclusion criteria or did not complete one or more testing sessions. Three additional participants were excluded from data analysis due to an invalid baseline test as determined by the ImPACT manufacturer's instructions. An additional seven participants were chosen at random to be excluded from the final data analysis to ensure equal sample size among the six groups. Incomplete Green's WMT data for one participant for day 45 results in the loss of one degree of freedom and did not significantly impact the results of

the analysis. Due to this participant's complete and valid data for both ImPACT and the POMS, this participant's data was not excluded from data analysis.

The sample utilized for data analysis consisted of 108 participants (33 males, 75 females). Demographics for the sample were as follows: age 20.99 $\pm$ 1.43 years, weight 66.86 $\pm$ 11.92 kg, height 171.57 $\pm$ 9.71 cm, average Scholastic Aptitude Test score 1224.50 $\pm$ 117.66, years of education 13.81 $\pm$ 1.30 years, and average cumulative grade point average based off of a four point scale 3.45 $\pm$ .31. Demographic data for participants may be found in table 5.1. Participants were tested at approximately 46.10 $\pm$ 3.39 days after baseline and 7.09 $\pm$ 2.51 days after day 45. Participants were tested approximately 31 $\pm$ 51 minutes, and 63 $\pm$ 98 minutes compared to their baseline testing time for day 45 and day 50, respectively. Eighteen participants (16.7%) reported having been diagnosed with one or more concussions. No participant reported having a concussion during the testing process or six months prior to participation.

Analysis of ImPACT

The mean scores for ImPACT may be found in table 5.2. A review of the ImPACT composite scores revealed comparable performance to previously reported results with a similar sample.[4] A repeated measures ANOVA revealed significant differences across time for composite verbal memory ($F_{(2,214)} = 12.493$, $p < .001$) and visual motor speed ($F_{(2,214)} = 9.231$, $p < .001$) indicating practice effects . Post hoc analysis for composite verbal memory was performed via paired *t* tests and revealed significant increases between baseline and day 45 ($t_{(107)} = -4.241$, $p < .001$) and baseline and day 50 ($t_{(107)} = -3.968$, $p < .001$). An increase of four and three percent occurred between baseline and day 45 and day 50, respectively, for composite verbal memory. Significant increases were indicated for visual motor speed between baseline and day 45 ($t_{(107)} = -2.101$, $p = .038$), baseline and day 50 ($t_{(107)} = -4.086$, $p < .001$) and between

day 45 and 50 ($t_{(107)} = -2.310$, $p = .023$). An increase of one and two points occurred between baseline and day 45 and 50, respectively and an increase of one point between day 45 and day 50. An increased visual motor speed score indicates better performance on the X's and O's, and Three letters components of the exam.

ImPACT utilizes the reliable change indices in order to indicate significant change. It is important to note that multiple formulas for reliable change exist. ImPACT calculates reliable change in the form of standardized error of difference (S_{diff}).[16] The formula for $S_{diff} = \sqrt{SEM_1^2 + SEM_2^2}$, where SEM represents the standard error of measurement. Hinton-Bayre suggests the correct formula for the reliable change is $(\overline{X}_1 - \overline{X}_2)/S_{diff}$, where $\overline{X}$ represents the mean of the score in question.[17] Results of the two formulas if utilizing the same scores vary greatly. For this study significant change was determined for both day 45 and day 50 compared to baseline. No participant sustained a concussion throughout the study making any significant change a false-positive for cognitive impairment. For day 45 approximately 33 participants (31%) were considered false-positives on one or composite scores and 28 participants (26%) were considered to have a significant change in performance on Day 50. The number of false-positives for each composite score at day 45 and day 50 may be found in table 5.3.

Mean scores and standard deviations for the POMS are presented in table 5.4. Results of the repeated measures ANOVA indicated one violation of sphericity for the factor depression ($W = .846$, $p < .001$). A Greenhouse-Geisser correction was made prior to interpretation of the results for this factor. Significant differences were indicated for two factors of the POMS, vigor ($F_{2,214)} = 8.844$, $p < .001$) and fatigue ($F_{2,214)} = 6.061$, $p = .003$). Post hoc analyses revealed significant differences for vigor between baseline and day 50 ($t_{(107)} = 4.394$, $p < .001$) and day 45 and day 50 ($t_{(107)} = 2.714$, $p = .008$). Post hoc analysis of the factor fatigue revealed significant

differences between baseline and day 45 ($t_{(107)} = 3.262$, $p = .001$) and between day 45 and day 50 ($t_{(107)} = 2.840$, $p = .005$).

Correlational analysis between the composite scores of ImPACT and each factor of the POMS at baseline resulted in significant correlations between the composite score impulse control and fatigue ($r = -.232$, $p = .016$) Significant correlations were also indicated for self-reported symptom score and POMS tension-anxiety ($r = .258$, $p = .007$) depression-dejection ($r = .517$, $p \leq .001$), anger-hostility ($r = .354$, $p \leq .001$), fatigue-inertia ($r = .481$, $p \leq .001$), confusion-bewilderment ($r = .364$, $p \leq .001$), and total mood disturbance ($r = .481$, $p \leq .001$). At day 45 a myriad of significant correlations existed between composite symptom score and tension-anxiety ($r = .294$, $p = .002$), depression-dejection ($r = .514$, $p \leq .001$), anger-hostility ($r = .259$, $p = .007$), fatigue-inertia ($r = .432$, $p \leq .001$), confusion-bewilderment ($r = .307$, $p = .001$), and total mood disturbance ($r = .474$, $p \leq .001$). Day 50 resulted in significant correlations between ImPACT's total symptom score and tension-anxiety($r = .470$, $p < .001$), depression-dejection ($r = .405$, $p \leq .001$), anger-hostility ($r = .398$, $p \leq .001$), fatigue-inertia ($r = .520$, $p \leq 001$), confusion-bewilderment ($r = .396$, $p \leq .001$) and total mood disturbance ($r = .562$, $p < .001$).

Self-Reported Symptoms

Correlational analysis of each mood state and each of the 22 self-reported symptoms resulted in one or more significant correlations ($p \leq .05$) per symptom as measured by ImPACT with the exclusion of nausea, vomiting, balance problems, and dizziness. Correlational analysis of somatic, neurobehavioral, and cognitive constructs and each mood state resulted in varying correlations at baseline, day 45, and day 50.At baseline, significant correlations were noted between the somatic construct and confusion-bewilderment ($r = .254$, $p = .008$). Significant correlations existed between the neurobehavioral construct and tension-anxiety ($r = .356$, $p <$

.001), depression-dejection ($r = .557$, $p < .001$), anger-hostility ($r = .271$, $p = .005$), fatigue-inertia ($r = .405$, $p < .001$), confusion-bewilderment ($r = .278$, $p = .004$), and total mood disturbance ($r = .405$, $p < .001$). The cognitive construct was significantly correlated with depression-dejection ($r = .393$, $p < .001$), anger-hostility ($r = .255$, $p = .008$), fatigue-inertia ($r = .391$, $p < .001$), confusion-bewilderment ($r = .351$, $p < .001$), vigor ($r = -.199$, $p = .039$), and total mood disturbance ($r = -.396$, $p < .001$).

Correlational analysis at day 45 resulted in a significant correlations between the somatic construct and tension-anxiety($r = .214$, $p = .026$), depression-dejection ($r = .200$, $p = .038$), fatigue-inertia ($r = .316$, $p < .001$), and total mood disturbance ($r = .290$, $p = .002$. Correlational analysis resulted in significant correlations between the neurobehavioral construct and the POMS factors tension-anxiety ($r = .267$, $p = .005$), depression-dejection ($r = .509$, $p < .001$), anger-hostility ($r = .228$, $p = .018$), fatigue-inertia ($r = .399$, $p < .001$), confusion-bewilderment ($r = .247$, $p = .010$) and total mood disturbance ($r = .401$, $p < .001$). Significant correlations were indicated between the cognitive construct and depression-dejection ($r = .321$, $p = .001$), fatigue-inertia ($r = .230$, $p = .017$), confusion-bewilderment ($r = .311$, $p < .001$) and total mood disturbance ($r = .329$, $p < .001$).

At day 50, significant correlations were indicated for the somatic construct and tension-anxiety ($r = .212$, $p = .028$), depression-dejection ($r = .583$, $p < .001$), anger-hostility ($r = .225$, $p = .019$), fatigue-inertia ($r = .264$, $p = .006$), confusion-bewilderment ($r = .354$, $p < .001$), and total mood disturbance ($r = .289$, $p = .002$. Significant correlations existed for the neurobehavioral construct and the POMS factors tension-anxiety ($r = .507$, $p < .001$), depression-dejection ($r = .361$, $p = .002$), anger-hostility ($r = .304$, $p = .001$), fatigue-inertia ($r = .428$, $p < .001$), confusion-bewilderment ($r = .324$, $p = .001$) and total mood disturbance ($r =$

.487, p <.001). Significant correlations were indicated between the cognitive construct and depression-dejection (r = .293, p = .002), anger-hostility (r = .251, p = .009), fatigue-inertia (r = .345, p < .001), confusion-bewilderment (r = .355, p < .001) and total mood disturbance (r = .366, p < .001). Correlation values for each time point are located in table 5.5.

Means and standard deviations for Green's WMT are presented in table 5.8 Results of the repeated measures ANOVA indicated violations of sphericity for each composite score of Green's WMT. Violations included immediate recall (W = .108, p < .001), delayed recall (W = .836, p < .001), consistency (W = .857, p < .001), multiple choice (W = .649, p < .001), paired associates (W = .459, p < .001), and free recall (W = .905, p = .005). Post Greenhouse-Geisser corrections for these violations, significant differences across time were indicated for delayed recall ($F_{(1.719, 183.907)}$ = 20.925, p < .001), consistency ($F_{(1.750,185.476)}$ = 20.186, p < .001), multiple choice ($F_{(1.481, 156.958)}$ = 9.007, p = .001), paired associates ($F_{(1.298, 137.590)}$ = 29.811, p < .001), and free recall ($F_{(1.827, 193.679)}$ = 97.671, p < .001). Post hoc analyses for delayed recall indicated significant differences between baseline and day 45 ($t_{(107)}$ = -3.911, p < .001) baseline and day 50 ($t_{(107)}$ = -6.313, p < .001) and between day 45 and day 50 ($t_{(107)}$ = -2.087, p = .039). Significant differences existed for consistency between baseline and day 45 ($t_{(106)}$ = -4.062, p < .001) and baseline and day 50 ($t_{(107)}$ = -5.989, p < .001). Post-hoc analyses for multiple choice indicated significant differences between baseline and day 45 ($t_{(106)}$ = -3.125, p = .002) and baseline and day 50 ($t_{(107)}$ = -3.316, p = .001). Analysis of paired associates revealed significant differences between baseline and day 45 ($t_{(106)}$ = -4.528, p < .001), baseline and day 50 ($t_{(107)}$ = -6.317, p < .001), and between day 45 and day 50 ($t_{(106)}$ = -4.574, p < .001). Post hoc analysis of free recall indicated significant differences between baseline and day 45 ($t_{(106)}$ = -8.168, p < .001), baseline and day 50 ($t_{(107)}$ = -13.119, p < .001), and between day 45 and day 50 ($t_{(106)}$ = -5.860, p < .001).

Scores for all composites Green's WMT scores exceeded 85%, indicating the sample provided good effort at each time point.[14]

Discussion

Our study sought to account for sources of random error affecting performance on a computerized neuropsychological test resulting in poor to moderate test-retest reliability coefficients previously reported. [4-6, 8] Our hypothesis was that mood state would be significantly correlated to one or more composite scores calculated by ImPACT while concurrently controlling for effort and time of day. To the author's knowledge this was the first study to address mood state's influence in a healthy population on a computerized neuropsychological test.

Our results indicated varying significant correlations at each time point excluding those between one or more POMS-B factors and ImPACT's composite symptom score. Most notably at baseline, vigor and total mood disturbance accounted for 27% to 29% of variance associated with composite reaction time, respectively. At day 45, fatigue-inertia accounted for 26% of variance associated with composite visual memory score. Composite verbal score was strongly correlated with vigor and total mood disturbance which accounted for 53% and 25% of the variance associated with performance, respectively. At day 50, depression-dejection accounted for 36% of the variance associated with composite verbal score. More consistent were the correlations between one or more mood states and composite symptom score, accounting for 30% to 66% of variance. Composite symptom score is measured via likert scale ranging from 0 (minor) to 6 (severe) for 22 symptoms associated with concussion. When dividing 16 of the 22 symptoms into somatic, neurobehavioral, and cognitive constructs it appears that depression-dejection, fatigue-inertia, and total mood disturbance consistently accounted for the most

variance (.1% to 31%) associated with each construct across time. Despite these finding, ImPACT performance remained consistent or significantly improved over time suggesting that although significant correlations were indicated, performance was not affected by one or more mood states.

To the author's knowledge only normative values for the POMS-B exist for psychiatric outpatients. [12] Upon comparing our POMS-B factor scores to normative values for college students for the POMS standard form, our sample scored considerably lower for each factor.[18] This was also true when comparing our findings to international, club, and recreational sport groups. Although our samples' scores were closer to the values reported by Terry and Lane, they were still significantly less.[19] When comparing our results to those of Simpson and Newby, collegiate football players scored considerably higher on each component of the POMS standard form compared to our sample.[20] It is important to note participants completed the POMS standard form in regards to "the past week including today," opposed to "as of right now," which was employed in the current study.[19, 20] This difference in methodology may be responsible for our sample's lower factor scores. Another potential reason for the lower reported scores is participants completed the POMS-B is a controlled setting free of potential distractions. Given these small values it is difficult to discern whether athletes, who have been noted to have higher mood state values, would differ in results.[20] Further research utilizing a collegiate athlete sample is needed to further investigate the relationship between computerized neuropsychological test performance and mood state.

In regards to ImPACT, our samples' results for each composite score are similar to those previously reported.[4] One suggested advantage of utilizing computerized neuropsychological platforms is randomized forms utilized to minimize practice effects.[5] Our results show

significant improvement between baseline and day 45 and baseline to day 50 in regards to composite verbal memory score and visual motor speed. Although not significant, results for composite visual memory score also indicated improvements between baseline and day 45. The amount of change for composite verbal memory and visual processing speed scores, although significant, was not as large between days 45 and 50 as compared to baseline values. Although this evidence does not conclusively suggest practice effects occur with repeated testing, it does support the delivery of two forms of ImPACT at baseline to although an individual to score closer to his or her true score.[21]

ImPACT utilizes the reliable change indices to determine if an individual's performance any factor is significantly different than baseline test values.[15] Clinically, significant declines are indicative of the concussed state. Any differences in one or more ImPACT composite scores may be interpreted as cognitive decline and delay an athlete's return to play. No participant in this study sustained a concussion or any head trauma six months prior to- or throughout the study. ImPACT determined cognitive declines in approximately 26 to 31 percent of our sample. The results provide evidence supporting a lack of sensitivity associated with computerized neuropsychological testing. Our percentage of false positives is comparable to that previously reported.[3]

The Green's WMT was employed to ensure participants provided good effort at each testing session. The sample scoring above 85% on all measures excluding free recall indicated participants provided good effort. [14] Results of this study are comparable to those previously reported.[4] No cutoff value for interpretation of free recall was provided. [3, 5]

To the author's knowledge this is the first study to address mood state and computerized neuropsychological test performance in a healthy collegiate sample while controlling for time of

day and effort at clinically relevant time points. Current recommendations suggest all athletes at high risk for sport-related concussion are to be baseline tested in regards to the suggested battery of tests, which includes neuropsychological testing, prior to the season.[1, 2] For a majority of athletes this occurs during his or her freshman year after making a transition from high school to college. Despite of the results of this study showing limited influence of mood state on computerized neuropsychological test performance further research needs to address this area of research in a collegiate athlete sample. The clinician's ability to ascertain a close estimate of an athlete's true ability is vital to ensure an accurate baseline test.

Conclusion

This study addressed potential sources of random error which may affect performance on computerized neuropsychological tests. In order to assess mood state we employed the POMS-B prior to delivery of Green's WMT and ImPACT, a commercially available computerized neuropsychological platform. Despite our results suggesting significant correlations between various mood states and composite ImPACT scores, it appears that mood state did not influence performance on ImPACT. Related literature addressing normative values for collegiate aged students and student athletes suggests higher mood state scores than observed in this study. Future research is needed to assess the relationship between mood state and computerized neuropsychological test performance in collegiate aged athletes.

Computerized neuropsychological testing has been shown to exhibit the highest sensitivity of the measures of the suggested concussion battery.[1, 2] Regardless, our data showed a false-positive rate of 26 to 31% in healthy collegiate aged sample. Although computerized neuropsychological testing is an important component of the suggested concussion battery is should not be the sole measure of an athlete's health status post concussion.[2] The delivery of

computerized neuropsychological tests along with self-reported symptoms, physical/neurological examination along with an assessment of balance at baseline and post- concussion is advocated to manage sport-related concussion.[1, 2] When baseline testing, it is imperative that clinicians ensure each athlete understands the purpose of each measure and that good effort is provided throughout testing. Despite the high sensitivity of proposed concussion battery[3], the value of post-concussion testing is only as good as the baseline test it is compared to.

Gender	Gender (n = 45)	Age (Baseline)	Age (Day 45)	Age (Day 50)	Height (cm)	Weight (kg)	Cumulative GPA	Years of Education	SAT
Males	(n = 33)	21.452 (1.542)	21.600 (1.549)	21.615 (1.551)	181.033 (7.107)	76.556 (10.761)	3.394 (.356)	14.091 (1.284)	1182.292 (132.976)
Females	(n = 75)	20.788 (1.342)	20.932 (1.379)	20.952 (1.384)	167.403 (7.553)	62.595 (1.329)	3.484 (.293)	13.693 (1.294)	1245.604 (104.384)
Total	(N=108)	20.991 (1.432)	21.136 (1.459)	21.154 (1.462)	171.568 (9.713)	66.861 (11.915)	3.456 (.314)	13.815 (1.298)	1293.630 (228.918)

Table 5.1: Means and (Standard Deviations) of participant demographics.

	ImPACT Composite Scores (N = 108)					
	Composite Symptom Scores	Memory Composite Verbal	Memory Composite Visual	Visual Motor Speed	Composite Reaction Time	Impulse Control
Baseline	3.870 (7.249)	.907 (.073)	.810 (.112)	42.094 (6.258)	.541 (.062)	6.361 (5.105)
Day 45	3.796 (7.165)	.941* (.068)	.820 (.105)	43.007* (5.607)	.538 (.060)	6.426 (5.873)
Day 50	3.093 (5.791)	.937*† (.069)	.820 (.117)	44.020*† (6.284)	.543 (.067)	6.815 (5.358)

Table 5.2 Means and (Standard Deviations) for composite ImPACT scores. † represents a significant difference from baseline

$p < .05$ † = represents a significant difference from day 45 $p < .05$

	ImPACT False-Positives (N = 108)						
	Memory Composite Verbal	Memory Composite Visual	Visual Motor Speed	Composite Reaction Time	Impulse Control	Self Reported Symptoms	Impaired on any variable
Day 45	n = 11 10.2%	n = 7 6.5%	n = 9 8.3%	n = 8 7.4%	n = 0 0%	n =8 7.4%	n = 33 30.6%
Day 50	n = 6 5.6%	n = 9 8.3%	n = 5 4.6%	n = 12 11.11%	n = 0 0 %	n = 3 2.8%	n = 28 25.9%

Table 5.3: Number of participants, and percentage of total sample labeled as significantly impaired by ImPACT's automated feature indicating significant statistical change compared to baseline.

	Profile of Mood States Factors (N = 108)						
	Tension/ Anxiety	Depression/ Dejection	Anger/ Hostility	Fatigue/ Inertia	Confusion/ Bewilderment	Vigor/ Activity	Total Mood Disturbance
Baseline	1.602 (1.894)	.565 (1.474)	.370 (.112)	3.759 (3.090)	2.528 (1.430)	7.954 (3.992)	1.324 (8.356)
Day 45	1.333 (1.919)	.630 (1.655)	.435 (1.088)	3.556† (3.319)	2.796 (1.465)	7.352 (4.292)	1.185 (8.909)
Day 50	1.213 (2.162)	.546 (1.500)	.556 (1.130)	2.611† (3.223)	2.870 (1.231)	6.333*† (4.447)	1.648 (9.212)

Table 5.4: Means and (Standard Deviations) of each factor of the Profile of Mood States. * = significantly different than baseline $p \leq .05$ † = significantly different than day 45 $p \leq .05$

Self-Reported Symptom Constructs	Profile of Mood State Factors						
	Baseline ($N = 108$)						
	Tension/ Anxiety	Depression/ Dejections	Anger/ Hostility	Fatigue/ Inertia	Confusion/ Bewilderment	Vigor/ Activity	Total Mood Disturbance
Somatic	.027	.042	.163	.152	.254†	-.022	.127
Neurobehavioral	.356†	.557†	.271†	.405†	.278†	-.059	.405†
Cognitive	.067	.393†	.255†	.391†	.351†	-.199*	.396†
	Day 45 ($N = 108$)						
Somatic	.214*	.200*	.011	.316†	.066	-.155	.290†
Neurobehavioral	.267†	.509†	.228*	.399†	.247*	-.132	.401†
Cognitive	.125	.321†	.173	.230*	.311†	-.161	.329†
	Day 50 ($N = 108$)						
Somatic	.212*	.225*	.106	.264†	.354†	-.113	.289†
Neurobehavioral	.507†	.361†	.304†	.428†	.324†	-.187	.487†
Cognitive	.164	.293†	.251†	.345†	.355†	-.180	.366†

Table 5.5: Correlation values for each self-reported symptom construct and each Profile of Mood States factor. * = $p \leq .05$ † = $p \leq .01$

Green's Word Memory Test Composite Scores (N = 108)						
	Immediate Recall	Delayed Recall	Consistency	Multiple Choice	Pair Associates	Free Recall
Baseline	97.492	98.542	97.617	96.963	96.589	71.308
(n = 107)	(10.856)	(1.937)	(3.001)	(5.695)	(5.346)	(12.680)
Day 45	99.306	99.426*	98.953*	98.738*	98.738*	81.589*
(n = 106)	(1.958)	(1.429)	(2.170)	(2.578)	(2.578)	(12.182)
Day 50	99.630	99.755*†	99.379*	98.879*	99.720*†	87.023*†
(n = 107	(1.015)	(.735)	(1.305)	(2.506)	(1.344)	(9.904)

Table 5.6: Means and (Standard Deviations) for each composite score of the Green's Word Memory Test. * = significantly different than baseline $p \leq .05$ † = significantly different than day 45 $p \leq .01$

References

1. Guskiewicz KM, Bruce SL, Cantu RC, et al.: National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion. J Athl Train 2004; 39(3): 280-297.

2. McCrory P, Meeuwisse W, Johnston K, et al.: Consensus statement on concussion in sport: the 3rd International Conference on Concussion in Sport held in Zurich, November 2008. J Athl Train 2009; 44(4): 434-48.

3. Broglio SP, Macciocchi SN, Ferrara MS: Sensitivity of the concussion assessment battery. Neurosurgery 2007; 60(6): 1050-7; discussion 1057-8.

4. Broglio SP, Ferrara MS, Macciocchi SN, Baumgartner TA, Elliott R: Test-retest reliability of computerized concussion assessment programs. J Athl Train 2007; 42(4): 509-14.

5. Randolph C, McCrea M, Barr WB: Is neuropsychological testing useful in the management of sport-related concussion? J Athl Train 2005; 40(3): 139-52.

6. Schatz P: Long-term test-retest reliability of baseline cognitive assessments using ImPACT. Am J Sports Med; 38(1): 47-53.

7. Van Kampen DA, Lovell MR, Pardini JE, Collins MW, Fu FH: The "value added" of neurocognitive testing after sports-related concussion. Am J Sports Med 2006; 34(10): 1630-5.

8. Collie A, Maruff P, Makdissi M, McCrory P, McStephen M, Darby D: CogSport: reliability and correlation with conventional cognitive tests used in postconcussion medical evaluations. Clin J Sport Med 2003; 13(1): 28-32.

9. Crocker L AJ: Introduction to Classical & Modern Test Theory. Mason, OH: Cengage Learning, 2008.

10. Lovell MR: ImPACT 2007 (6.0) Clinical Interpretation Manual. Pittsburgh, PA ImPACT, 2007.

11. Piland SG, Motl RW, Ferrara MS, Peterson CL: Evidence for the Factorial and Construct Validity of a Self-Report Concussion Symptoms Scale. J Athl Train 2003; 38(2): 104-112.

12. McNair DM, Heuchert, JW.P. : Profile of Mood States Technical Update North Tonawanda, NY: Multi-Health Systems Inc. , 2005.

13. Maxwell SE, Delaney, H.D.: Designing Experiments And Analyzing Data 2nd ed. Mahwah, NJ: Lawrence Erlbaum Associates 2004.

14. Green P: Green's Word Memory Test for Windows: User's Manual and Program. . Edmonton, Canada: Green's Publishing Inc. , 2005.

15. Lovell M: ImPACT 2007 (6.0) Software User's Manual Pittsburgh, PA ImPACT Applications Inc. , 2007.

16. Iverson GL, Lovell MR, Collins MW: Interpreting change on ImPACT following sport concussion. Clin Neuropsychol 2003; 17(4): 460-7.

17. Hinton-Bayre AD, Geffen GM, Geffen LB, McFarland KA, Friis P: Concussion in contact sports: reliable change indices of impairment and recovery. J Clin Exp Neuropsychol 1999; 21(1): 70-86.

18. Nyenhuis DL, Yamamoto C, Luchetta T, Terrien A, Parmentier A: Adult and geriatric normative data and validation of the profile of mood states. J Clin Psychol 1999; 55(1): 79-86.

19. Terry PC, Lane, A.M.: Normative Values for the Profile of Mood States for Use with Athletic Samples. Journal of Applied Sport Psychology 2000; 12: 93-109.

20. Simpson S, Newby RW: Scores on profile of mood states of college football players from nonscholarship and scholarship programs. Percept Mot Skills 1994; 78(2): 635-40.

21. Hinton-Bayre AD, Geffen G, McFarland K: Mild head injury and speed of information processing: a prospective study of professional rugby league players. J Clin Exp Neuropsychol 1997; 19(2): 275-89.

CHAPTER 6

SUMMARY

This dissertation served to replicate and extend research regarding the test-retest reliability of computerized neuropsychological tests. For this study we focused on one commercially available platform, ImPACT. Utilizing clinically relevant time points we concurrently controlled for time of day and effort. Test-retest reliability coefficients in previous literature range from poor to moderate. To account for a greater amount of random error associated with this testing format we employed the POMS-B to investigate the influence of mood on test performance. The authors are unaware of any study which has addressed mood state and computerized neuropsychological test performance.

Each participant was tested at three time points (baseline, day 45, and day 50). Participants were administered a health questionnaire, the POMS-B, Green's WMT, followed by ImPACT. Test-retest reliability coefficients for ImPACT ranged from .374 to .756, which were slight better than previously reported when utilizing the same time points but still fall below what is suggested for clinical interpretation. It is also important to note that ImPACT incorrectly classified approximately 20 to 30% of participants as cognitively impaired when healthy. This finding adds to the empirical evidence supporting poor sensitivity with computerized neuropsychological testing. In regards to mood state, although multiple significant correlations between various mood states and ImPACT composite scores, specifically visual and verbal memory, reaction time, and composite symptom score, mood did not appear to influence test performance. Results of Green's WMT suggested the sample provided good effort at all three time points. Despite accounting for time of day, effort, and mood state, test-retest reliability coefficients

remained poor to moderate. This leads the authors to believe that the poor to moderate reliability coefficients are the result of poor test design rather than random error.

The results of this study add to the body of literature regarding psychometric properties of computerized neuropsychological testing. Despite its common use in sports medicine, computerized neuropsychological testing has yet to achieve clinically acceptable reliability and subsequent validity in regarding to measuring neurocognitive deficits. Until the appropriate evidence is produced to support the clinical use of computerized platforms for neurocognitive assessment, the sports medicine clinician is urged to utilize multiple measures when managing sport-related concussion.

The increased focus on sport-related concussion during the past ten years has lead to advancements in the management of this ambiguous injury. It is important for the sport-medicine clinician to critically evaluate each of these instruments prior to utilization in clinical practice. Currently the suggested battery of tests is composed of self-reported symptoms, neurocognitive testing, balance assessment, and the physical/neurological examination. Despite the sensitivity of this suggested battery being greater than 90%, the greatest tool in the management of sport-related concussion is the education of athletes, coaches, parents, and other sports medicine professionals about this ambiguous brain injury.

REFERENCES

1. Aubry M, Cantu R, Dvorak J, et al.: Summary and agreement statement of the 1st International Symposium on Concussion in Sport, Vienna 2001. Clin J Sport Med 2002; 12(1): 6-11.

2. McCrory P, Meeuwisse W, Johnston K, et al.: Consensus Statement on Concussion in Sport: the 3rd International Conference on Concussion in Sport held in Zurich, November 2008. Br J Sports Med 2009; 43 Suppl 1: i76-90.

3. Kelly DF, Gonzalo IT, Cohan P, Berman N, Swerdloff R, Wang C: Hypopituitarism following traumatic brain injury and aneurysmal subarachnoid hemorrhage: a preliminary report. J Neurosurg 2000; 93(5): 743-52.

4. Rutland-Brown W, Langlois JA, Thomas KE, Xi YL: Incidence of traumatic brain injury in the United States, 2003. J Head Trauma Rehabil 2006; 21(6): 544-8.

5. Langlois JA, Rutland-Brown W, Wald MM: The epidemiology and impact of traumatic brain injury: a brief overview. J Head Trauma Rehabil 2006; 21(5): 375-8.

6. Powell JW, Barber-Foss KD: Traumatic brain injury in high school athletes. JAMA 1999; 282(10): 958-63.

7. Gerberich SG, Priest JD, Boen JR, Straub CP, Maxwell RE: Concussion incidences and severity in secondary school varsity football players. Am J Public Health 1983; 73(12): 1370-5.

8. Guskiewicz KM, Weaver NL, Padua DA, Garrett WE, Jr.: Epidemiology of concussion in collegiate and high school football players. Am J Sports Med 2000; 28(5): 643-50.

9. McCrea M, Hammeke T, Olsen G, Leo P, Guskiewicz K: Unreported concussion in high school football players: implications for prevention. Clin J Sport Med 2004; 14(1): 13-7.

10. Collie A, Darby D, Maruff P: Computerised cognitive assessment of athletes with sports related head injury. Br J Sports Med 2001; 35(5): 297-302.

11. Guskiewicz KM, Bruce SL, Cantu RC, et al.: National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion. J Athl Train 2004; 39(3): 280-297.

12. McCrory P, Johnston K, Meeuwisse W, et al.: Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. Br J Sports Med 2005; 39(4): 196-204.

13. Broglio SP, Macciocchi SN, Ferrara MS: Sensitivity of the concussion assessment battery. Neurosurgery 2007; 60(6): 1050-7; discussion 1057-8.

14. Riemann BL, Guskiewicz KM: Effects of Mild Head Injury on Postural Stability as Measured Through Clinical Balance Testing. J Athl Train 2000; 35(1): 19-25.

15. Broglio SP, Ferrara MS, Sopiarz K, Kelly MS: Reliable change of the sensory organization test. Clin J Sport Med 2008; 18(2): 148-54.

16. Guskiewicz KM: Postural stability assessment following concussion: one piece of the puzzle. Clin J Sport Med 2001; 11(3): 182-9.

17. Guskiewicz KM, Perrin DH, Gansneder BM: Effect of Mild Head Injury on Postural Stability in Athletes. J Athl Train 1996; 31(4): 300-306.

18. Guskiewicz KM, Riemann BL, Perrin DH, Nashner LM: Alternative approaches to the assessment of mild head injury in athletes. Med Sci Sports Exerc 1997; 29(7 Suppl): S213-21.

19. Chen JK, Johnston KM, Collie A, McCrory P, Ptito A: A validation of the post concussion symptom scale in the assessment of complex concussion using cognitive testing and functional MRI. J Neurol Neurosurg Psychiatry 2007; 78(11): 1231-8.

20. Chen JK, Johnston KM, Frey S, Petrides M, Worsley K, Ptito A: Functional abnormalities in symptomatic concussed athletes: an fMRI study. Neuroimage 2004; 22(1): 68-82.

21. Jantzen KJ, Anderson B, Steinberg FL, Kelso JA: A prospective functional MR imaging study of mild traumatic brain injury in college football players. AJNR Am J Neuroradiol 2004; 25(5): 738-45.

22. Naunheim RS, Matero D, Fucetola R: Assessment of patients with mild concussion in the emergency department. J Head Trauma Rehabil 2008; 23(2): 116-22.

23. Terrell TR, Bostick RM, Abramson R, et al.: APOE, APOE promoter, and Tau genotypes and risk for concussion in college athletes. Clin J Sport Med 2008; 18(1): 10-7.

24. Begaz T, Kyriacou DN, Segal J, Bazarian JJ: Serum biochemical markers for post-concussion syndrome in patients with mild traumatic brain injury. J Neurotrauma 2006; 23(8): 1201-10.

25. Popovic V, Aimaretti G, Casanueva FF, Ghigo E: Hypopituitarism following traumatic brain injury. Growth Horm IGF Res 2005; 15(3): 177-84.

26. Benvenga S, Campenni A, Ruggeri RM, Trimarchi F: Clinical review 113: Hypopituitarism secondary to head trauma. J Clin Endocrinol Metab 2000; 85(4): 1353-61.

27. Schneider HJ, Kreitschmann-Andermahr I, Ghigo E, Stalla GK, Agha A: Hypothalamopituitary dysfunction following traumatic brain injury and aneurysmal subarachnoid hemorrhage: a systematic review. JAMA 2007; 298(12): 1429-38.

28. Aimaretti G, Ambrosio MR, Di Somma C, et al.: Residual pituitary function after brain injury-induced hypopituitarism: a prospective 12-month study. J Clin Endocrinol Metab 2005; 90(11): 6085-92.

29. Benvenga S: Brain injury and hypopituitarism: the historical background. Pituitary 2005; 8(3-4): 193-5.

30. Agha A, Phillips J, Thompson CJ: Hypopituitarism following traumatic brain injury (TBI). Br J Neurosurg 2007; 21(2): 210-6.

31. Agha A, Rogers B, Sherlock M, et al.: Anterior pituitary dysfunction in survivors of traumatic brain injury. J Clin Endocrinol Metab 2004; 89(10): 4929-36.

32. Schneider HJ, Aimaretti G, Kreitschmann-Andermahr I, Stalla GK, Ghigo E: Hypopituitarism. Lancet 2007; 369(9571): 1461-70.

33. Estes SM, Urban RJ: Hormonal replacement in patients with brain injury-induced hypopituitarism: who, when and how to treat? Pituitary 2005; 8(3-4): 267-70.

34. Pellman EJ, Viano DC, Withnall C, Shewchenko N, Bir CA, Halstead PD: Concussion in professional football: helmet testing to assess impact performance--part 11. Neurosurgery 2006; 58(1): 78-96; discussion 78-96.

35. Pellman EJ, Viano DC, Tucker AM, Casson IR, Waeckerle JF: Concussion in professional football: reconstruction of game impacts and injuries. Neurosurgery 2003; 53(4): 799-812; discussion 812-4.

36. Duma SM, Manoogian SJ, Bussone WR, et al.: Analysis of real-time head accelerations in collegiate football players. Clin J Sport Med 2005; 15(1): 3-8.

37. Guskiewicz KM, Mihalik JP, Shankar V, et al.: Measurement of head impacts in collegiate football players: relationship between head impact biomechanics and acute clinical outcome after concussion. Neurosurgery 2007; 61(6): 1244-52; discussion 1252-3.

38. McCaffrey MA, Mihalik JP, Crowell DH, Shields EW, Guskiewicz KM: Measurement of head impacts in collegiate football players: clinical measures of concussion after high- and low-magnitude impacts. Neurosurgery 2007; 61(6): 1236-43; discussion 1243.

39. Mihalik JP, Bell DR, Marshall SW, Guskiewicz KM: Measurement of head impacts in collegiate football players: an investigation of positional and event-type differences. Neurosurgery 2007; 61(6): 1229-35; discussion 1235.

40. Newman JA, Beusenberg MC, Shewchenko N, Withnall C, Fournier E: Verification of biomechanical methods employed in a comprehensive study of mild traumatic brain injury and the effectiveness of American football helmets. J Biomech 2005; 38(7): 1469-81.

41. Viano DC, Lovsund: Biomechanics of Brain and Spinal Cord Injury: Analysis of Neuropathologic and Neurophysiologic Experiments Journal of Crash Prevention and Injury Control 1999; 1(1): 35-43.

42. King AI, King, H., Zhang, L., Hardy, W., Viano, D.C. : Is head injury caused by linear or angular acceleration. International Research Conference on the Biomechanics of Impacts, Lisbon, Portugal, 2003.

43. Viano DC, Casson IR, Pellman EJ, Zhang L, King AI, Yang KH: Concussion in professional football: brain responses by finite element analysis: part 9. Neurosurgery 2005; 57(5): 891-916; discussion 891-916.

44. Zhang L, King, H.Y., King, A.I., Viano, D.C.: A new biomechanical predictor for mild traumatic brain injury - a preliminary finding. Summer Bioengineering Conference. Key Biscayne, Florida, 2003; 137 - 138.

45. Acerini CL, Tasker RC, Bellone S, Bona G, Thompson CJ, Savage MO: Hypopituitarism in childhood and adolescence following traumatic brain injury: the case for prospective endocrine investigation. Eur J Endocrinol 2006; 155(5): 663-9.

46. Popovic V, Pekic S, Pavlovic D, et al.: Hypopituitarism as a consequence of traumatic brain injury (TBI) and its possible relation with cognitive disabilities and mental distress. J Endocrinol Invest 2004; 27(11): 1048-54.

47. Collins MW, Grindel SH, Lovell MR, et al.: Relationship between concussion and neuropsychological performance in college football players. JAMA 1999; 282(10): 964-70.

48. Guskiewicz KM, McCrea M, Marshall SW, et al.: Cumulative effects associated with recurrent concussion in collegiate football players: the NCAA Concussion Study. JAMA 2003; 290(19): 2549-55.

49. Broglio SP, Ferrara MS, Macciocchi SN, Baumgartner TA, Elliott R: Test-retest reliability of computerized concussion assessment programs. J Athl Train 2007; 42(4): 509-14.

50. Broglio SP, Macciocchi SN, Ferrara MS: Neurocognitive performance of concussed athletes when symptom free. J Athl Train 2007; 42(4): 504-8.

51. Schatz P: Long-Term Test-Retest Reliability of Baseline Cognitive Assessments Using ImPACT. Am J Sports Med 2009.

52. Kizilbash AH, Vanderploeg RD, Curtiss G: The effects of depression and anxiety on memory performance. Arch Clin Neuropsychol 2002; 17(1): 57-67.

53. Mahmood O, Rapport LJ, Hanks RA, Fichtenberg NL: Neuropsychological performance and sleep disturbance following traumatic brain injury. J Head Trauma Rehabil 2004; 19(5): 378-90.

54. Stanford MS, Greve, K.W., Gerstle, J.E.: Neuropsychological Correlates of self-reported impulsive aggression in a college sample. Person. individ. Diff. 1997; 23(6): 961-965.

55. Ferrara MS, McCrea M, Peterson CL, Guskiewicz KM: A Survey of Practice Patterns in Concussion Assessment and Management. J Athl Train 2001; 36(2): 145-149.

56. Piland SG, Motl RW, Guskiewicz KM, McCrea M, Ferrara MS: Structural validity of a self-report concussion-related symptom scale. Med Sci Sports Exerc 2006; 38(1): 27-32.

57. Randolph C, McCrea M, Barr WB: Is neuropsychological testing useful in the management of sport-related concussion? J Athl Train 2005; 40(3): 139-52.

58. Schatz P, Pardini JE, Lovell MR, Collins MW, Podell K: Sensitivity and specificity of the ImPACT Test Battery for concussion in athletes. Arch Clin Neuropsychol 2006; 21(1): 91-9.

59. Van Kampen DA, Lovell MR, Pardini JE, Collins MW, Fu FH: The "value added" of neurocognitive testing after sports-related concussion. Am J Sports Med 2006; 34(10): 1630-5.

60. Kristman VL, Tator CH, Kreiger N, et al.: Does the apolipoprotein epsilon 4 allele predispose varsity athletes to concussion? A prospective cohort study. Clin J Sport Med 2008; 18(4): 322-8.

61. Raabe A, Grolms C, Sorge O, Zimmermann M, Seifert V: Serum S-100B protein in severe head injury. Neurosurgery 1999; 45(3): 477-83.

62. Schneider M, Schneider HJ, Yassouridis A, Saller B, von Rosen F, Stalla GK: Predictors of anterior pituitary insufficiency after traumatic brain injury. Clin Endocrinol (Oxf) 2008; 68(2): 206-12.

63. Baumgartner TA, Jackson, A.S., Mahar, M.T., Rowe, D.A.: Measurement for Evaluation in Physical Education and Exercise Science 8th ed. New York, New York: McGraw Hill, 2007.

64. Crocker L, Algina, J.: Introduction to Classical and Modern Test Theory. Mason, Ohio: Cengage Learning, 2008.

65. Cantu RC: Posttraumatic Retrograde and Anterograde Amnesia: Pathophysiology and Implications in Grading and Safe Return to Play. J Athl Train 2001; 36(3): 244-248.

66. Practice parameter: the management of concussion in sports (summary statement). Report of the Quality Standards Subcommittee. Neurology 1997; 48(3): 581-5.

67. Giza CC, Hovda DA: The Neurometabolic Cascade of Concussion. J Athl Train 2001; 36(3): 228-235.

68. Brooks GA, Fahey, T.D., White, T.P., Baldwin, K.M.: Exercise Physiology: Human Bioenergetics and Its Applications, 3rd ed. New York: McGraw-Hill, 2000.

69. Collins M, Lovell MR, Iverson GL, Ide T, Maroon J: Examining concussion rates and return to play in high school football players wearing newer helmet technology: a three-year prospective cohort study. Neurosurgery 2006; 58(2): 275-86; discussion 275-86.

70. Guskiewicz KM, Ross SE, Marshall SW: Postural Stability and Neuropsychological Deficits After Concussion in Collegiate Athletes. J Athl Train 2001; 36(3): 263-273.

71. Tommasone BA, Valovich McLeod TC: Contact sport concussion incidence. J Athl Train 2006; 41(4): 470-2.

72. Collins MW, Field M, Lovell MR, et al.: Relationship between postconcussion headache and neuropsychological test performance in high school athletes. Am J Sports Med 2003; 31(2): 168-73.

73. Piland SG, Motl RW, Ferrara MS, Peterson CL: Evidence for the Factorial and Construct Validity of a Self-Report Concussion Symptoms Scale. J Athl Train 2003; 38(2): 104-112.

74. Notebaert AJ, Guskiewicz KM: Current trends in athletic training practice for concussion assessment and management. J Athl Train 2005; 40(4): 320-5.

75. Report of the Sports Medicine Committee. Guidelines for the Management of Concussion in Sports. Denver, CO: Colorado Medical Society, 1990.

76. McCrea M, Kelly JP, Kluge J, Ackley B, Randolph C: Standardized assessment of concussion in football players. Neurology 1997; 48(3): 586-8.

77. Barr WB, McCrea M: Sensitivity and specificity of standardized neurocognitive testing immediately following sports concussion. J Int Neuropsychol Soc 2001; 7(6): 693-702.

78. McCrea M: Standardized Mental Status Testing on the Sideline After Sport-Related Concussion. J Athl Train 2001; 36(3): 274-279.

79. Shehata N, Wiley JP, Richea S, Benson BW, Duits L, Meeuwisse W: Sport Concussion Assessment Tool: Baseline Values For Varsity Collision Sport Athletes. Br J Sports Med 2009.

80. Macciocchi SN, Barth, J.T., Alves, W., Rimel, R.W., Jane, J.A.: Neurological functioning and recovery after mild head injury in collegiate athletes Neursurgery 1996; 39: 510-514.

81. Shapiro AM, Benedict RH, Schretlen D, Brandt J: Construct and concurrent validity of the Hopkins Verbal Learning Test-revised. Clin Neuropsychol 1999; 13(3): 348-58.

82. Ruff RM, Light RH, Parker SB, Levin HS: Benton Controlled Oral Word Association Test: reliability and updated norms. Arch Clin Neuropsychol 1996; 11(4): 329-38.

83. Jensen AR, Rohwer WD, Jr.: The Stroop color-word test: a review. Acta Psychol (Amst) 1966; 25(1): 36-93.

84. Macciocchi SN: "Practice makes perfect:" retest effects in college athletes. J Clin Psychol 1990; 46(5): 628-31.

85. Maroon JC, Lovell MR, Norwig J, Podell K, Powell JW, Hartl R: Cerebral concussion in athletes: evaluation and neuropsychological testing. Neurosurgery 2000; 47(3): 659-69; discussion 669-72.

86. Randolph C: Implementation of Neuropsychological Testing Models for the High School, Collegiate, and Professional Sport Settings. J Athl Train 2001; 36(3): 288-296.

87. Collins MW, Lovell MR, Iverson GL, Cantu RC, Maroon JC, Field M: Cumulative effects of concussion in high school athletes. Neurosurgery 2002; 51(5): 1175-9; discussion 1180-1.

88. Echemendia RJ, Julian LJ: Mild traumatic brain injury in sports: neuropsychology's contribution to a developing field. Neuropsychol Rev 2001; 11(2): 69-88.

89. Lovell MR, Collins MW: Neuropsychological assessment of the college football player. J Head Trauma Rehabil 1998; 13(2): 9-26.

90. Lovell MR, Iverson GL, Collins MW, et al.: Measurement of symptoms following sports-related concussion: reliability and normative data for the post-concussion scale. Appl Neuropsychol 2006; 13(3): 166-74.

91. Booher JM, Thibodeau, G.A.: Athletic Injury Assessment, 4th ed. New York: McGraw Hill, 2000.

92. Broglio SP, Tomporowski PD, Ferrara MS: Balance performance with a cognitive task: a dual-task testing paradigm. Med Sci Sports Exerc 2005; 37(4): 689-95.

93. Carlson NR: Physiology of Behavior, 9th ed. Amherst, MA: Allyn & Bacon, 2006.

94. Martin JH: Neuroanatomy: Text and Atlas, 2nd ed. Stamford, CT: Appleton & Lange, 1996.

95. Shumway-Cook A, Horak FB: Assessing the influence of sensory interaction of balance. Suggestion from the field. Phys Ther 1986; 66(10): 1548-50.

96. Peterson CL, Ferrara MS, Mrazik M, Piland S, Elliott R: Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. Clin J Sport Med 2003; 13(4): 230-7.

97. Iverson GL, Gaetz M, Lovell MR, Collins MW: Relation between subjective fogginess and neuropsychological testing following concussion. J Int Neuropsychol Soc 2004; 10(6): 904-6.

98. McCrory PR, Berkovic SF: Second impact syndrome. Neurology 1998; 50(3): 677-83.

99. Huettel SA, Song, A.W., McCarthy, G.: Functional Magnestic Resonance Imaging Sunderland, MA: Sinauer Associates, Inc., 2003.

100. Lovell MR, Pardini JE, Welling J, et al.: Functional brain abnormalities are related to clinical recovery and time to return-to-play in athletes. Neurosurgery 2007; 61(2): 352-9; discussion 359-60.

101. Ptito A, Chen JK, Johnston KM: Contributions of functional magnetic resonance imaging (fMRI) to sport concussion evaluation. NeuroRehabilitation 2007; 22(3): 217-27.

102. Davis GA, Iverson GL, Guskiewicz KM, Ptito A, Johnston KM: Contributions of neuroimaging, balance testing, electrophysiology and blood markers to the assessment of sport-related concussion. Br J Sports Med 2009; 43 Suppl 1: i36-45.

103. Ma M, Lindsell CJ, Rosenberry CM, Shaw GJ, Zemlan FP: Serum cleaved tau does not predict postconcussion syndrome after mild traumatic brain injury. Am J Emerg Med 2008; 26(7): 763-8.

104. Ingebrigtsen T, Romner B, Kongstad P, Langbakk B: Increased serum concentrations of protein S-100 after minor head injury: a biochemical serum marker with prognostic value? J Neurol Neurosurg Psychiatry 1995; 59(1): 103-4.

105. Townend W, Ingebrigtsen T: Head injury outcome prediction: a role for protein S-100B? Injury 2006; 37(12): 1098-108.

106. Molina P: Endocrine Physiology, 2nd ed. New York, New York: McGraw-Hill, 2006.

107. Baumgartner TA, Hensley, L.D.: Conducting and Reading Research in Health and Human Performance 4th ed. New York, NY: McGraw Hill, 2006.

108. Schatz P: Long-term test-retest reliability of baseline cognitive assessments using ImPACT. Am J Sports Med; 38(1): 47-53.

109. Maxwell SE, Delaney, H.D.: Designing Experiments and Analyzing Data, 2nd ed. Mahweh, New Jersey: Lawrence Elrbaum Associates 2003.

110. McNair DM, Heuchert, JW.P. : Profile of Mood States Technical Update North Tonawanda, NY: Multi-Health Systems Inc. , 2005.

111. Lovell M: ImPACT 2007 (6.0) Software User's Manual Pittsburgh, PA ImPACT Applications Inc. , 2007.

112. Lovell MR: ImPACT 2007 (6.0) Clinical Interpretation Manual. Pittsburgh, PA ImPACT, 2007.

113. Green P: Green's Word Memory Test for Windows: User's Manual and Program. . Edmonton, Canada: Green's Publishing Inc. , 2005.

APPENDIX A

Consent Form

I agree to participate in the research study entitled, Test- retest Reliability of a Computerized Neuropsychological Test of a Computerized Neuropsychological Test: Part 2, which is being conducted by Dr. Michael Ferrara (706.542.4801) and Jacob Resch (706.542.3273) of the Department of Kinesiology at the University of Georgia. I understand that this participation is entirely voluntary: I can refuse to participate or I can withdraw my consent at any time without penalty or loss of benefits to which I am otherwise entitled and have the results of the participation, to the extent that it can be identified as mine, returned to me, removed from the research records and destroyed.

The following points have been explained to me:

1. The research is being conducted to study the parallel form equivalence of a computerized neuropsychological testing platform (ImPACT) utilized for the evaluation and management of sport-related concussion. The primary purpose of the project is to determine if the five forms of the computerized test are indeed equivalent to one another adding validity to the examination.

2. The procedures are as follows: I will receive a detailed explanation of the study, the benefits and risks for participation and then sign the informed consent form. At that time I will complete both a health questionnaire. At this time, if I do not meet the inclusion criteria I will be excluded from this study. Exclusion criteria for this study

includes English not being my primary language, if I have been diagnosed with a learning disability or attention deficit disorder, if I have been diagnosed with a concussive injury during the past six months or if I suffer a concussion throughout the duration of the study, if I engage in alcohol or drug use 24 hours prior to testing, or if I am unable to utilizing a computerize mouse with my right hand. If included, I will then complete the Profile of Mood States Brief Form (POMS-B) and will be randomly assigned to one of four groups. After completion of each form I will complete the Green's Verbal Memory Test and then complete initial demographical information required to complete the initial form of ImPACT for data organization and storage and then complete form 1 of ImPACT. I will then return to the Athletic Training Laboratory 45 days later to complete another form of ImPACT, the POMS-B inventory, and Green's Verbal Memory Test. The final testing session will occur five days after the second session and will include another form of ImPACT, the POMS-B inventory, and the Green's Verbal Memory Test.

3. Testing will consist of three parts: A health questionnaire, the POMS-B inventory, the Green's Verbal Memory Test, and computerized neuropsychological testing. I will complete the health questionnaire which will provide information regarding my history of any prior concussions. The POMS-B inventory will be completed to assess various mood states I may be experiences prior to computerized testing. The Green's Verbal Memory Test will assess my verbal memory. The neuropsychological portion will consist of a computerized assessment of my thinking, reasoning and memory abilities. Each of these parts will occur in the St. Mary's Athletic Training Laboratory

(Room 110B) in the Ramsey Center and will take approximately 30 minutes to complete.

4. I understand that I will be tested on the following schedule after consenting for my participating in this study. Initially I will report to the St. Mary's Athletic Training Laboratory to complete a health questionnaire, the POMS-B inventory, and the first form of ImPACT. After completion of ImPACT will be set a date and time (45 days after the initial testing) with the principal or co-investigator to complete another form of ImPACT, a POMS-B inventory, and the Green's Verbal Memory Test. After completion of the second session I will set a time five days from that time to take one more form of ImPACT, the POMS-B inventory, and the Green's Verbal Memory Test.

5. The benefit that I may expect from this research is to gain a better knowledge and understanding of concussion and the recovery from this type of injury. Others may benefit from your participation in this study by increasing the knowledge that exists regarding concussion. Upon request, at the conclusion of the study, I may receive information regarding my performance on ImPACT.

6. All individually-identifiable information and data collected will be kept under lock and key and password protected computers. Individual identifiers such as your name, birth date, address, etc… will be removed and replaced with a number which will be your sole identifier. Identifiers will be removed immediately after entry into the

study. At the completion of this study, any personal information linking my performance to a numerical code will be destroyed.

7. The investigators will answer any further questions about the research, now and during the course of the project and can be reached at 706.542.4801 for Dr. Michael Ferrara and 706.542.3273 for Jake Resch.

I understand that I am agreeing by my signature on this form to take part in this research project and understand that I will receive a signed copy of this consent form for my records.

Michael S. Ferrara_____ _____ _____

Telephone: 706-542-4801 Signature Date

Email: mferrara@uga.edu

_____ _____ _____

Name of Participant Signature Date

Please sign both copies, keep one and return one to the researcher.

Additional questions or problems regarding your rights as a research participant should be addressed to the Chairperson, Institutional Research Board, University of Georgia, 612 Boyd Graduate Studies Research Center, Athens, Georgia,

30602-7411Telephone: (706) 542-3199: E-mail Address IRB@uga.

APPENDIX B

Health Questionnaire

Please answer the following questions as accurately and as thoroughly as you can.

NAME: Last:_____Middle Initial:_____First:_____

Home Address: _____

Telephone number (Home) _____ (Work)_____

What is your birth date? Month: _____ Day: _____ Year:19_____

What is your current age? _____ Circle your sex: male female

What is your RACE/ETHNICITY? What is your current:

(Check one) Cumulative GPA: _____

- ❑ White (not of Hispanic origin) SAT Score: _____
- ❑ Black (not of Hispanic origin)
- ❑ Hispanic
- ❑ Asian or Pacific Islander
- ❑ American Indian or Alaskan Native
- ❑ Other

3. Have you ever had a concussion? (CIRCLE ONE) YES NO

4. How many times have you had a concussion?
   (CIRCLE ONE) 0 1 2 3 4 4+

5. What is the year of your most recent concussion? (Skip if not applicable) _____

6. Are you physically sick (cold, flu, allergies) today? (CIRCLE ONE) YES NO

7. Are you currently receiving treatments for any type of injury?
   (Example: ankle sprain, bruise, pulled muscle) (CIRCLE ONE) YES NO

8. Are you tired from any physical activities you have participated in today?
   (CIRCLE ONE) YES NO NOT APPLICABLE

9. Have you ever been diagnosed with a learning disability or Attention Deficit Disorder?
   (**CIRCLE ONE)** YES NO

10. Is English your primary language? **(CIRLCLE ONE)** YES NO

11. Have you had any caffeinated beverages during the past 24 hours, if so, how much (beverage and amount in ounces)? _____