THE PSYCHOMETRIC PROPERTIES OF CONCUSSION ASSESSMENT TOOLS IN HIGH

SCHOOL ATHLETICS

by

TAMERAH NICOLE HUNT

(Under the Direction of Michael S. Ferrara)

ABSTRACT

Recent leaders in concussion assessment have suggested the use of baseline test scores to

measure premorbid function and allow athletes to serve as their own control. There are,

however, several factors that may affect the psychometric properties of baseline scores. The

factors investigated in this study were effort and reliability of the revised Balance Error Scoring

System protocol.

This research is two fold. First we examined the level of effort in an athletic population

and determined if sub-optimal effort effects neuropsychological test scores. One hundred and

ninety-nine high school athletes were administered a brief neuropsychological test battery

including the Dot Counting Test (DCT) and the Rey 15-item test with recognition trial (R15-R).

We found that Sub-optimal effort existed in an athletic population. Moreover, significant

differences existed between effort groups on several of the neuropsychological tests.   The

presence of sub-optimal effort utilizing a gross measure of effort suggests a conservative

estimate of sub-optimal effort. The addition of objective measures of effort will identify sub-

optimal effort and thus providing validity evidence of interpretations regarding baseline

neuropsychological test scores.

Second we calculated the reliability of the revised BESS protocol. One hundred and forty-four high school athletes performed the revised protocol which consisted of three trials of four conditions (firm and foam surfaces and single leg and tandem stance). Statistically significant differences existed between trial one and trial two ($F_{(1.65, 286)} = 4.890$, p=.013). Further, an intraclass reliability coefficient was obtained for three trials of four conditions ($R = 0.88$), two trials of four conditions ($R = 0.84$) and one trial of four conditions ($R = 0.73$). The revised protocol increased the reliability of the BESS scores however the presence of a practice effect suggests allowing additional familiarization trials prior to administration.

Validating interpretations of test scores can not be accomplished without a reliable measure. Further, a thorough knowledge of the effect of confounding variables will clarify interpretations from test scores. Psychometrically sound instruments support the ability to make and interpret clinical decisions regarding injury and return to participation. Continual evaluation of the psychometric properties of commonly used concussion assessment tools will enable validation of interpretations from test scores.

INDEX WORDS: Concussion, High School, Athletes, Psychometric Properties, Reliability

THE PSYCHOMETRIC PROPERTIES OF CONCUSSION ASSESSMENT TOOLS IN HIGH

SCHOOL ATHLETICS

by

TAMERAH NICOLE HUNT

BACHELOR OF SCIENCE, UNIVERSITY OF DELAWARE, 2000

MASTER OF SCIENCE, JAMES MADISON UNIVERSITY, 2002

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

2006

THE PSYCHOMETRIC PROPERTIES OF CONCUSSION ASSESSMENT TOOLS IN HIGH

SCHOOL ATHLETICS

by

TAMERAH NICOLE HUNT

Major Professor:	Michael S. Ferrara
Committee:	Ted Baumgartner
	L. Stephen Miller
	Stephen Macciocchi

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2006

DEDICATION

This entire process is dedicated to my family and friends. You pushed me to reach for the stars and enjoy life.

To Dad: Yes, they do teach grammar and the best editor in the world. Retire, no more papers to review.

To Mom: Thank you for your love and support during periods of both frustration and despair. Your unconditional love for me turned into true friendship for us.

To Paula: My big sis, thank you for reminding me to be spontaneous, free spirited and be true to myself. Thank you for your friendship and love.

To Trudy: A shoulder to cry on, the epitome of true friendship, your support and our weekly outings to Kellys maintained my sanity.

I love you all and could not have accomplished this without any of you.

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

x

CHAPTER 1

INTRODUCTION

Introduction

There are no universal agreements on the terminology and definition of concussion. Mild head injury (MHI), mild traumatic brain injury (MTBI) and concussion have been names used interchangeable for the same condition. The two most popular cited definitions are based on functional status and the nature of medical signs and symptoms present at the time of injury (Aubry et al., 2002; Cantu, 1997). The term concussion previously was defined by the committee of Head Injury Nomenclature of the Congress of Neurological Surgeons as "a clinical syndrome characterized by immediate and transient post-traumatic impairment of neural function, such as alteration of consciousness and disturbance of vision or equilibrium due to brain stem involvement" (Esselman & Uomoto, 1995). However, this definition was found to be vague and ambiguous, which caused confusion for physicians and allied medical practioners associated with concussion assessment. More recently, concussion has been defined as "a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces" (Aubry et al., 2002). Concussion may be caused either by a direct blow to the head, face, neck or elsewhere on the body with an "impulsive" or rotational force transmitted to the head, which may result in the rapid onset of short-lived impairment of neurological function that resolves spontaneously.

The Centers for Disease Control and Prevention recently proclaimed that concussions in athletes have reached an epidemic proportion in the United States. Each year in the United

States, an estimated 20 to 30 million youth between 5 and 17 years of age participate in community-sponsored athletic programs. The number of injuries reported for this age group has increased proportionally to the participation rate, with some studies reporting a higher incidence rate than collegiate athletes. There are approximately 1.5 million high school and middle school football players, while there are only 75,000 college players. With the greater number of exposures in middle school and high school, the risk of concussion is greater than that of their collegiate counterparts.

An estimated 63,000 concussions occur annually in high school sports (Guskiewicz, Weaver & Padua, 2000). The high school rate was more than twice that of the Division I collegiate rate. McCrea (2004) recently found a higher prevalence rate (15.3 %) in high school athletics than previously reported in the literature, which was three times that originally suggested in high school athletics. Additionally, he found that 80% of youth concussions are mild concussions and go unreported to medical personnel. The potential for catastrophic injury following concussion requires an understanding of the physiology and sequelae following a concussive injury. Furthermore, the education of the common signs and symptoms of concussion to athletes, parents and coaches is essential.

Repeated mild head injury over a short period of time can be catastrophic or fatal (Centers for Disease Control and Prevention, 1997). From 1945-1999, there were 712 fatalities resulting from football, with head injuries accounting for 69%. These fatalities have been attributed to Second Impact Syndrome, epidural and subdural hematoma, and fractures. High school football appears to produce a greater number of fatalities compared to its collegiate and professional counterparts. Although the incidence of life threatening brain injury has decreased

2

in sports, there is evidence that suggests that concussion may be more common and more serious than previously thought.

Second Impact Syndrome (SIS) is a serious complication following MHI with mortality rates near 50% and morbidity rates of 100% (Cantu, 1997). SIS results when a second blow to the head occurs to an athlete returning to a game before the symptoms associated with the first injury have been cleared. It begins with persisting post-concussion symptoms that may produce visual, motor, or sensory changes (Cantu, 1992).

Following the initial MHI, brain cells survive in a vulnerable state (Wojtys et al., 1999). After the second impact, a disruption of the brain's autoregulatory system leads to increased intracranial vasodialation and pressure (Cantu, 1998; Cantu, 1992; Kelly, Lissel, Rowe, Vincenten, & Voaklander, 2001). The intracranial pressure decreases the perfusion to the brain and can lead to ischemia (Kelly et al., 2001) and herniation of the uncus of the temporal lobes through the foramen magnum (Cantu, 1998). Within seconds to minutes, the athlete collapses with rapidly dilating pupils, loss of eye movement, and respiratory failure (Cantu, 1992).

Almost all reported cases of SIS have occurred in athletes under the age of 18 (Cantu, 1998; Cantu, 1992; McCrory & Berkovic, 1998). It is theorized that the apparent higher risk of cerebral swelling in children following MHI is due to different autoregulatory responses to trauma in these younger individuals (Bruce et al., 1981; Schnitker, 1949; Snoek, Minderhound, & Wilmink, 1984). Therefore, the prevention of SIS in younger population through the adequate assessment of MHI is critical.

Following a concussive injury, individuals typically display transient deficits in cognitive functioning that can often be detected through neuropsychological assessment. The cognitive areas of immediate and working memory, attention, information processing speed and

problem solving abilities are affected following concussion (Barth et al., 1983; Grindel, Lovell, & Collins, 2001; Guskiewicz, Ross, & Marshall, 2001; Hinton-Bayre, Geffen, & McFarland, 1999; Lovell, 2001). The assessments chosen, therefore, must adequately evaluate these constructs.

Investigations into younger populations widely use neuropsychological testing in children sustaining non-sports-related head injuries (Anderson et al., 1997; Donders & Strom, 2000; Silver, 2000). Further, research has been funded to examine sports-related concussion and fMRI in adolescent athletes (Lovell, 2001 grant number 5R01HD042386-04): pediatric mild head injury (Kay, 2004 grant number 1R49CE000284); and, an investigation of the outcomes associated with pediatric traumatic brain injury (Dise-Lewis, 2004 grant number 1R49CE000394-01). However, most of these studies have included children with moderate to severe head injury. The studies lacked uninjured controls, involved equipment not accessible to allied health professionals working with adolescents, and were limited in the number of investigations specifically for sports-related MHI.

Researchers have suggested that the younger the brain at initial injury, the more resilient it is, and a faster recovery would occur (Anderson et al., 1997; Donders & Strom, 2000). High school athletes, on average, take longer to recover and report more symptoms following concussion than adults (Collins et al., 2003; Daniel, Olesniewicz, & Reeves 1999; Field, Collins, Lovell, & Maroon, 2003; Lovell & Collins, 2000; Lovell et al., 2003). Lovell (2000) reported that neuropsychological test scores were statistically significant between the injured and control groups on measures of memory and post concussion symptoms in high school athletes for up to seven days post injury. Moser found that adolescents with two or more concussions were statistically different from those with zero to one previous concussion on the Repeatable Battery

for the Assessment of Neuropsychological Status (Moser & Shatz, 2002). Field et al. (2003) also reported 161 high school athletes that sustained concussions had increased post concussion symptoms and delayed recovery using the Controlled Oral Word Association Test and Brief Visuospatial Memory Test compared to the concussed college athletes and their matched controls.

While these findings provide some basic descriptive information about youth concussion, many of these studies lack sufficient controls and have numerous threats to internal and external validity. The reliability and validity of the concussion assessment tools provide a greater understanding about deficits following post-injury and comparisons to baseline measures. Accurate interpretations from test scores in adolescents provide information necessary to make safe return to participation decisions.

Due to the potential complications and long-term consequences of returning to competition too early, several assessments of cognition have been used to help clinicians make safe return to participation decisions. Recent published guidelines for the assessment of concussion suggest baseline testing as a means of comparing premorbid and post-injury function (Aubry et al., 2002; Guskiewicz et al., 2004; McCrory et al., 2005). The guiding capacity to interpret deficits following injury requires that all clinical interpretation of data be valid.

Test scores are used to draw inferences about examinee behavior in a situation. Validity is the process used to collect evidence to support the inferences drawn from the test score. Test scores reflect two components; the "true score" and "error score" described in classical test theory (Baumgartner & Jackson, 2003).

The addition of error score in test scores compromises the validity of clinical interpretations. To obtain high validity evidence, a test developer would try to eliminate as many

extraneous variables associated with error as possible. An example of an extraneous variable would include an individual's effort. Sub-optimal effort would result in the inability to estimate an individual's "true score" ability.

The importance of specific examination of effort in neuropsychological examination is crucial, as poor test-taking effort may show up as an absence of practice effects for subjects who have demonstrated some learning ability. Additionally, unaccountable ups and downs of scores or wide variations in intratest response patterns are difficult to interpret clinically due to poor effort at baseline testing (Lezak, Howieson, & Loring, 2004).

Much of the data regarding effort and malingering revolves around adult populations involving workman's compensation or litigation (Lees-Haley, 1997; Youngjohn, Burrows, & Erdal, 1995a). When approached with the same problem in adolescent, it is attributed to effort. The inability to determine whether optimal effort was performed is more difficult when clinicians are not suspecting the adolescent to exert sub-optimal effort (Faust, Hart, Guilmette, & Arkes, 1988). Therefore, the addition of procedures to assess effort during test-taking might help to correctly identify adolescents with deficits following injuries.

An individual's effort may impact neuropsychological scores on baseline performances, decreasing the ability to make adequate clinical interpretations regarding the post-injury assessment following a concussive injury. Currently, there is no research in sports-related concussion assessment that has examined whether an individual gave maximal effort during baseline testing to assist in the interpretation of the scores.

Postural stability has been recognized as an important assessment following concussion (Guskiewicz & Perrin, 1996; Ingersoll & Armstrong, 1992; Riemann & Guskiewicz, 2000) and can be assessed using several different methods. Two techniques are commonly used: the

Sensory Organization Test (SOT) within the Smart Master Balance System (Clackamas, OR), and the Balance Error Scoring System (BESS). The SOT consists of 3 trials of 6 conditions (18 total trials), with each trial lasting 20 seconds. The test systematically alters visual and somatosensory referencing in an attempt to individually evaluate the three components of the balance mechanism (visual, vestibular, and somatosensory). The SOT is the gold standard for postural stability in concussion however; the SOT is not portable and is very expensive.

The BESS was developed as an objective assessment tool for clinicians with minimal cost and training for the evaluation of postural stability following concussion. During the test, the athlete stands on two surfaces (firm and foam), with their eyes closed performing 3 stances (double leg stance, single leg stance, tandem or heal to toe stance). In previous studies, the BESS has been correlated to the SOT composite scores and demonstrated similar recovery patterns (Guskiewicz, Riemann, Perrin, & Nashner 1997; Riemann & Guskiewicz, 2000; Susco, Valovich, Gansneder & Shultz, 2004; Valovich, Perrin, & Gansneder, 2003; Valovich et al., 2004). Further, the BESS has the advantage of being administered in a short period of time with little equipment and cost (Guskiewicz & Perrin, 1996; Guskiewicz et al., 1997).

Considerable research has been conducted to examine the psychometric properties of the BESS. In studies following concussion, subjects exhibit acute postural stability alterations up to 5 days post-injury (Guskiewicz et al., 2001; Riemann & Guskiewicz, 2000) with recovery usually occurring within 4 to 7 days post-injury to pre-injury baseline values. Although the BESS appears to be sensitive to subtle deficits following concussion, it has several drawbacks. Recently, one researcher found that administration of multiple trials of the BESS results in practice effects, with the number of errors decreasing with each consecutive trial (Valovich et al.,

2003). Further, the effects of fatigue increase the number of errors acutely, but the athlete recovered after 20 minutes of rest following an exercise session (Susco et al., 2004).

As the BESS is used in a multifaceted approach to concussion assessment, these findings suggest slight alterations of the BESS may increase the reliability of scores. Within our pilot data, we found that the reliability of the BESS is $r = 0.60$. Most psychometricians would say the current reliability would not be adequate for a measurement with implications for return to participation decisions, especially when these findings suggest that simple changes can increase the reliability of the measure.

Since athletic trainers are required to make important clinical decisions, the reliability and validity of the assessment tool are essential. Removing stances within the BESS that provide limited variance would benefit the reliability of the measure and thus improve clinical interpretations. Based on the findings of this study, it appears that removal of the double leg stance would increase the reliability of the measure and decrease the amount of time to administer the test. Moreover, adding additional trials of the single leg and tandem stances would further increase the reliability, while keeping the administration time brief.

Although the reliability coefficients increase to $r \geq 0.91$ (excellent reliability), it is not practical for Certified Athletic Trainers to administer seven trials of 4 conditions (28 tests). The addition of that many trials introduces practice effects, fatigue during test administration, as well as increased administration time. Therefore, the protocol was revised to consist of at least three trials of the four conditions to obtain the optimal psychometric properties of the revised protocol.

Purpose

The purpose of this research was two-fold. First, this research evaluated the presence of sub-optimal effort in a high school population. Griffin (1997) found 20 to 40% of patients are

capable of malingering in clinical and sub-optimal effort groups. Although sub-optimal effort

was found in compensation claims, the presence of sub-optimal effort in an athletic population

has not been advocated during baseline testing. If sub-optimal effort exists in this population, it

is hypothesized that the percentage should be low. Previous research has been conducted in

clinical and sub-optimal effort populations, but there has never been a study of this magnitude in

sports-related concussion research. The general hypothesis is that sub-optimal effort exists in an

athletic population.

If sub-optimal effort exists in an athletic population, does it indeed affect

neuropsychological test scores? Determining whether sub-optimal effort effects test scores

provides information to validate the clinical interpretations of the neuropsychological test scores

post injury. Effort tests provide an objective measure to determine if the test results are accurate.

Beneficial information can be gained by including an objective effort test into all concussion

assessments to validate clinical interpretations of neuropsychological tests. Utilizing invalid

baseline data would prevent comparative interpretations between premorbid and post-injury

status.

The second purpose of this research was to examine the psychometric properties of the

revised Balance Error Scoring System protocol. The test-retest reliability and intratest reliability

of BESS scores, using the traditional protocol, have been investigated in collegiate and high

school populations. A pilot study utilizing a generalizability study offered evidence that the

double leg stance provided little to no variance within the system. Moreover, the percentage of

variance within each facet of the traditional protocol of the BESS defined the important

components of the BESS. Performing a stance that provides little information is futile and

therefore, the stance was removed. Previous researchers (Hunt & Ferrara, in review) theorized

that changing the model would result in an intraclass reliability coefficient of 0.88; but, would that indeed occur?

The use of concussion assessment tools has been suggested by numerous organizations (Aubry et al., 2002; Guskiewicz et al., 2004) during baseline concussion assessment. The psychometric properties of several of the instruments have not been thoroughly investigated. Poor psychometric properties of tests increase the potential for invalid interpretations within the assessment battery. Increasing the validity of interpretations of the test results will aid important decisions such as return to participation and recovery status, especially in high school athletics.

Statement of the Problem

Concussion in an adolescent population can be a potentially catastrophic and a costly injury. Higher numbers of adolescent athletes participate in sports compared to their adult counterparts, which indicate that adolescence might be a greater risk for concussion. Researchers investigating adolescent athletes reveal a longer recovery time and an increase in the number of symptoms reported compared to their adult counterparts. The long-term consequences of mild head injuries in children are, to date unknown, indicating the need for adequate assessments and interpretations of clinical tests to assess when a child can resume normal activities, especially athletics.

Recent guidelines suggest the need for baseline concussion assessment to measure premorbid function. Baseline scores are invalid when the results and interpretations are not reliable and valid during assessment. Although many cognitive and motor tests are currently used to assess outcomes in adolescence with a variety of conditions, the applicability of the clinical sports-related head injury assessments patterns have not been extensively investigated. The problem occurs when interpretations from baseline testing are not validated; because

10

evaluations of obtained test scores cannot be interpreted accurately. Identifying psychometrically sound tests for use with adolescent athletes during assessment periods may reduce the risk of returning an athlete too soon, thus decreasing the risk for serious injuries.

The suggestion of baseline testing to establish premorbid function has focused on neuropsychological, balance and symptomology testing within concussion assessment. Neuropsychological assessment has been suggested as a part of a multifaceted approach in concussion assessment. Neuropsychological testing provides information regarding cognitive abilities affected by concussion. In forensic neuropsychology, clinical interpretations from test scores are validated by the use of objective effort tests. This technique, however, has not been incorporated into athletic training and concussion assessment. While sub-optimal effort affects neuropsychological test scores in clinical population, the presence and effect of sub-optimal effort has not been determined in a high school athletic population.

The psychometric properties of the Balance Error Scoring System (BESS) have recently been investigated in a high school population during serial administrations (Valovich, 2004), after fatigue (Susco, 2004), and following exertion (Wilkins, 2004). These studies presented test-retest reliability of BESS scores as 0.71-0.77 and intratest reliability as 0.60, which is poor to moderate reliability. To obtain premorbid function, the measure must have great reliability and thus the validity of the clinical interpretations can be justified. To improve the reliability of the measure, the double leg stance on both the firm and foam surface was removed, and additional trials of the single leg and tandem stances were added. The reliability of the revised model, however, has not been investigated.

Specific Aims and Hypotheses:

1. Assess the presence of sub-optimal effort in high school athletes using a brief concussion assessment battery.

> Hypothesis 1: Sub-optimal effort will be present in at least 10% of the high school athletes.

2. Determine the effect of sub-optimal effort on neuropsychological test scores in high school athletes.

> Hypothesis 2: There will be differences in total symptoms reported and neuropsychological test scores between optimal effort and sub-optimal effort during baseline testing.

3. Estimate the reliability of the revised model of the Balance Error Scoring System.

> Hypothesis 3: Internal consistency reliability of the revised BESS protocol will be demonstrated by an intraclass reliability coefficient of at least 0.85.

The following chapters will outline the literature related to concussion, and two original research manuscripts. A manuscript will be submitted to Archives of Clinical Neuropsychology. The manuscript will describe the effect of effort on neuropsychological testing in high school athletes. This chapter will be followed by a manuscript to be submitted to the Journal of Athletic Training. This manuscript focuses on the reliability of a revised Balance Error Scoring System protocol. General conclusions and summary and references complete the dissertation.

CHAPTER 2

REVIEW OF LITERATURE

Epidemiology

The frequency of all sports-related concussion has been estimated to be up to 300,000 occurrences annually in the United States (Thurman, Branch, & Sniezek, 1998). The Centers for Disease Control and Prevention recently proclaimed that concussions in athletes have reached an epidemic proportion in the United States. Although the incidence of life-threatening brain injury has decreased in most sports, there is new evidence suggesting that concussion may be more common and more serious than previously thought (Guskiewicz, Weaver, & Padua, 2000; McCrea, Hammeke, Olsen, Leo, & Guskiewicz, 2004). The numbers attributed to concussion has recently been questioned as studies reveal that athletes may underreport the incidences of concussion (McCrea et al., 2004). This may be directly influenced by the vague definition of concussion.

Definition of Injury

Research describes a group of clinical syndromes as mild traumatic brain injury (MTBI), traumatic brain injury and concussion. MTBI was conceptually defined by a CDC workgroup as " an injury to the head as result of blunt trauma or acceleration or deceleration forces that results in one or more of the following conditions:

1. Any period of observed or self-reported transient contusion, disorientation, impaired consciousness, dysfunction of memory around the time of injury, or loss of consciousness lasting less than 30 minutes.

2.	Observed signs of neurological or neuropsychological dysfunction, headache, dizziness, irritability, fatigue or poor concentration. (CDC, 1997)

MTBIs in athletic populations appear to result in transient symptoms with short duration. In athletic research, the term concussion is the most common. Concussion comes from the Latin term *concutere*, which means to strike together. There are no universal agreements on the definition of concussion. The two most popular cited definitions are based on functional status and the nature of medical signs and symptoms present at the time of injury. The term concussion was previously defined as "a clinical syndrome characterized by immediate and transient post-traumatic impairment of neural function, such as alteration of consciousness and disturbance of vision or equilibrium due to brain stem involvement (Bailes & Cantu, 2001). This definition was, however, found to be vague and ambiguous, creating confusion for those associated with concussion assessment.

More recently, concussion has been defined as "a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces" (Aubry et al., 2002). This consensus from the 1st international conference for concussion suggested five conditions for concussion.

1.	Concussion may be caused either by a direct blow to the head, face, neck or elsewhere on the body with an "impulsive" or rotational force transmitted to the head.

2.	Concussion typically results in the rapid onset of short-lived impairment of neurological function that resolves spontaneously.

3.	Concussion may result in neuropathological changes, but the acute clinical symptoms rarely reflect a functional disturbance rather than structural injury.

4.   Concussion results in a graded set of clinical syndromes that may or may not involve loss of consciousness, resolution of the clinical and cognitive symptoms typically follows a sequential course.

5.   Concussion is typically associated with grossly normal structural neuroimaging studies.

Classification of Injury

In 2001, 25 published grading scales had been developed (Johnston, McCrory, & Mohtadi, 2001). The grading scales, however, are not based on empirical data but subjective opinions of physicians working with concussed athletes. The arbitrary recommendation of an athlete being asymptomatic for seven days under rest and exertion has not been empirically justified. Common treatment and return-to-play guidelines, therefore, have been based purely upon clinical experience.

Of the 20 published grading scales and RTP guidelines, three are most commonly used; Cantu, Colorado and American Academy of Neurologist (AAN). When surveyed, certified athletic trainers, on average, report the use of the AAN more than any other grading scale (Ferrara, McCrea, Peterson, & Guskiewicz, 2001).These grading scales are typically based on several on-field markers such as loss of consciousness, post-traumatic amnesia, and/or concussion related symptoms. The most common grading scales are presented below.

American Academy of Neurology

Grade 1: Transient confusion, no LOC, symptoms resolve in < 15 minutes

Grade 2: Transient confusion, no LOC, symptoms last > 15 minutes

Grade 3: Any LOC, either brief or prolonged

Colorado

>*Grade 1: No LOC, confusion or amnesia in < 15 minutes*
>
>*Grade 2: No LOC, confusion, amnesia*
>
>*Grade 3: Any LOC*

Cantu- Revised

>*Grade 1: No LOC, PTA or post concussive signs and symptoms lasting less than*
>
>>*30 minutes*
>
>*Grade 2: LOC less than 1 minute, PTA or Post concussive signs and symptoms*
>
>>*lasting > 30 minutes, but less than 24 hours*
>
>*Grade 3: LOC > 1 minute, PTA > 24 hours*

Recently, Cantu revised his grading scale to reflect the opinion that an athlete be asymptomatic. The scale now includes the disappearance of symptoms and classification of concussion after symptoms resolution. All guidelines follow the same basic premise with loss of consciousness (LOC) being graded the most severe and requiring the longest time for recovery. The major change in the scale was the inclusion of posttraumatic amnesia (retrograde and anterograde) as an indicator of concussion severity (Cantu, 1998).

Typically the management protocol for a grade 1 is that athletes may return to play the same day if normal sideline assessments are present at rest and with exertion. For a grade 2 concussion, they recommend deferment from contact for one full week without symptoms at rest and with exertion. With a grade 3 concussion, deferred contact for one week without symptoms; and, prolonged loss of consciousness is deferred for two weeks without symptoms.

Mechanism of Concussion

There are several noted mechanisms for concussion. The most common mechanism in sports-related concussion is the coup/contrecoup model (Bailes & Cantu, 2001; Cantu, 1992; Cantu, 1997). In the coup/contrecoup model, the cranium moves in a direction until it is stopped by the physical limitations of the neck tissues. When this happens, the brain continues to move until it is acted upon by an outside force. While the cerebral spinal fluid will offer some resistance to motion, it is often not enough. As such, the brain will make contact with the interior of the cranium. The initial site of contact is called a coup injury. This can result in mechanical damage to the brain at the point of contact.

If the brain were to then rebound into the opposite side of the cranium, a contrecoup injury would result. Again, structural deformation can result to the point of impact at both locations. This type of injury commonly results from a direct impact to the head, either to the frontal or occipital bones.

Another mechanism that could lead to more severe injuries is caused by a rotational component. A diffuse axonal injury will result with uncontrolled head rotation. In this model, the lateral rotation of the head results in a stretching of the cerebral neurons. Ommaya (2002) proposed the centripetal theory of injury which stated that the most superficial neurons would sustain the greatest level of insult, and the degree of injury would decrease toward the rotational center of the brain. The result of a diffuse axonal injury is a stretching of the neurons, but contact of the brain with the internal cranium does not occur. The stretching, if severe enough, can tear some tissues and vessels, but is more likely to result in a depolarization of the neurons.

Metabolic Cascade

After a blow to the head, the neurons of the brain are stretched to the point where mechanical deformation takes place and the voltage gated channels within the membrane are opened. The brain experiences a sudden neuronal depolarization which leads to an indiscriminate triggering of voltage dependent channels which release several excitatory amino acids, the main one being glutamate. Glutamate begins a neurometabolic cascade which leads to ionic shifts and chemical changes within the brain (Giza & Hovda, 2001). Glutamate binds with NDMA receptors, which in turn open channels for which potassium (K^+), and calcium (Ca^{2+}) efflux and influx (respectively). As a response to injury, the brain blood flow decreases by up to 50% (Hovda & Villablanca, 1998; Kelly et al., 2001). The increase in extracellular K^+ triggers membrane sodium-potassium pumps to try and re-establish homeostasis. Increased pump action requires ATP, so glycolysis is sent into "overdrive" condition to attempt to meet the energy demand. This takes place in an oxygen-starved environment because cerebral blood flow has been reduced.

The increase in extracellular K^+ has a negative effect on the mitochondrial ability to perform normal oxidative capacities. This, in turn, further elevates the energy crisis within the brain. Once extracellular K^+ has begun to decrease, the brain begins to experience further metabolic slowing. This slowing has been termed spreading depression. The condition has been affiliated with decreases in neuronal firing which leads to neurocognitive dysfunction.

Giza et al. reported in 2001 that extracellular potassium increases quickly after the onset of injury, but decreases within the following 30 minutes. Elevated levels of potassium have been related to decreased mitochondrial function and propagation of the energy deficit by triggering sodium-potassium pump response. Calcium, on the other hand, remains at elevated levels from

two to four days post injury. Elevated levels of calcium have been attributed to neuronal

disconnection and cell death.

A by-product of the energy crises is the over production and build-up of lactate. Too

much lactate can cause a change in the ph of the surrounding area, which has been shown to

damage cellular membranes increasing brain vulnerability. Over time this process will correct

itself.

It appears that adults return to normal following a sport-related concussion in about three

days (Guskiewicz et al., 2001). The decrease in cerebral brain flow, however, has been

documented in humans and it has been shown to last up to ten days. The condition decreases

brain metabolism, which has been related to neuronal dysfunction and decreases in brain

function (i.e. neurocognitive function and possibly motor control). The condition is also thought

to place the brain into a state of vulnerability. It is believed that this condition places the brain

into a situation where it cannot respond to injury or hyperactivity, therefore increasing the risk

for a second concussive injury. It has been reported that concussed athletes are up to six times

more likely to get a second concussion than non-concussed athletes (Cantu, 2001).

Concussion Assessment Tools

Physical Evaluation

Physical evaluation is typically the initial source of information. Physical examination

typically occurs on the sideline or within the first few minutes of injury. During a physical

evaluation, a complete history should be obtained. This provides information regarding the

likelihood of previous concussion, symptoms occurring not related to current concussive injury,

or post-concussive symptoms.

When surveyed, more than 95% of ATCs used the clinical examination as the primary assessment tool for concussion. The increase percentage of certified athletic trainers utilizing physical examinations (33% reported by Ferrara et al., 2001) shows a trend for concussion assessment. Clinical examinations may be in combination with additional concussion assessment tools (Notebaert & Guskiewicz, 2005).Typical assessments in a physical evaluation include evaluation of nervous, motor, and sensory systems, and signs and symptoms of injury. From this phase, the athlete is diagnosed with concussion and additional testing may be warranted.

Self-Report Symptoms

Another commonly used tool to assess an individual is self-report symptoms. Several self-report symptom scales and checklists exist. Common scales are the Post Concussion Symptom Scale (PCSS), the Graded Symptom Checklist (GSC) and the Head Injury Scale (HIS). When surveyed, 85% of certified athletic trainers used symptom checklists as part of the concussion assessment battery. (Notebaert & Guskiewicz, 2005)

Several symptoms have been linked to the occurrence of concussion. These symptoms include headache, nausea, vomiting, dizziness, fatigue, difficulty sleeping, and drowsiness. The symptoms can be divided into somatic, cognitive and sensory domains (Piland, Motl, Ferrara, & Peterson, 2003). Attention must be paid to concussion related symptoms that naturally occur, such as headache. Several investigators found that during baseline examination, athletes self-report the presence of symptoms (Field et al., 2003; Piland et al., 2003).

Neuroimaging

Upon arrival in the Emergency Room following an injury, a Computerized Topography scan (CT) of the brain is typically ordered. The CT scan is the choice of test to evaluate and exclude intracranial bleeding (Bazarian, Wong, & Harris, 2006; Mendez, Hurley, Lassonde,

Zhang, & Taber, 2005). Jagoda (2002) found three to ten percent of CT scans following concussion had a traumatic abnormality. Only one percent of those cases, however, required neurosurgical intervention.

Magnetic Resonance Imaging constructs detailed anatomic images that are sensitive to traumatic lesions. Magnetic Resonance Imaging (MRI) has been used acutely after injury, but concussion patients could not be differentiated with MRI use alone. Most abnormalities identified using MRI consisted of contusions and hematomas. Sport-related concussion, however, results in functional deficits rather than structural deficits. At a recent meeting of the leaders in concussion assessment, it was reported that imaging techniques such as x-ray, Magnetic Resonance Imaging and CT would be unable to detect concussion unless there were gross structural changes in the brain (Aubry et al., 2001). CT and traditional MRI, therefore, may not be as useful as initially thought. Newer technologies may offer different techniques to evaluate brain function using imaging.

Functional MRI (fMRI) may provide a more thorough assessment. fMRI provides information regarding neural function during task performance using a noninvasive technique (Bazarian et al., 2006; Mendez et al., 2005). This technique can be incorporated with the new dual task paradigms being researched in concussion assessment. (Broglio, Tomporowski, & Ferrara, 2005) A new approach, Magnetic Source Imaging (MSI) includes the use of MRI to obtain anatomic information while investigating the electrophysiology data from magnetoencephalography (MEG) (McAllister, Sparling, Flashman, & Saykin, 2001; Mendez et al., 2005). MSI offers tracking of real-time brain activity, without distorting, by conduction through the brain, skull and scalp.

With the metabolic changes following concussion, utilization of positron emission tomography (PET) and single-photon emission computed tomography (SPECT) may provide more accurate information regarding the nature of the injury (Bazarian et al., 2006; McAllister et al., 2001; Mendez et al., 2005) . PET and SPECT can evaluate metabolism of the specified region and blood flow (respectively) associated with activation of that region. These measures attempt to quantify correlations between metabolic flow and injury severity, post-concussion symptoms, and recovery. (McAllister et al., 2001) As technology continues to improve, understanding the effects and consequences of concussion should also improve.

Neuropsychological Tests

Neuropsychological testing batteries have been researched and used by trained Neuropsychologists for many years. Recently, they have gained favor in the profession of athletic training as tools for assessing cognitive function prior to and following concussion. Neuropsychological testing is used to provide a sensitive index of higher brain functioning by measuring functions such as memory, attention, speed and flexibility of cognitive processing. These functions have been determined to become sensitive during impairments associated with concussion.

Recent guidelines suggest a multifaceted approach to concussion assessment to capture the variability of deficits following injury. These guidelines are accompanied by the suggestion of baseline testing all athletes prior to the competitive season to compare premorbid function to that of post-injury. Neuropsychological tests for sports-related concussion can be administered prior to injury. Pre-testing provides the medical staff evaluating the athlete post-injury a baseline, so that subtle changes in cognitive and motor functions can be detected. Pre-testing would allow

a direct comparison of what is "normal" should an athlete be injured. In the event that baseline scores were not available, clinicians should use published norms.

The use of NP test batteries for the assessment and follow-up care of athletes originated from work in the early 1980's when sports became a laboratory setting for the study of acute mild head injuries (Barth et al., 1989). These functions include memory, cognitive processing speed, working memory, and/or executive function (Lovell, Iverson, Collins, McKeag, & Maroon, 1999; Macciocchi, Barth, Alves, Rimel, & Jane, 1996; Peterson, Ferrara, Mrazik, Piland & Elliott, 2003; Pellman, Lovell, Viano, Casson, & Tucker, 2004b). Commonly used assessments include Trail Making Test (Guskiewicz et al., 1997; Guskiewicz et al., 2001; Macciocchi et al., 1996) Digit Symbol Substitution Test (Macciocchi et al., 1996; Maddocks & Saling, 1996), Controlled Oral Word Association Test (COWAT) (Guskiewicz et al., 2001), Hopkins Verbal Learning Test (HVLT), Stroop Word Color Test (Guskiewicz et al., 2001) and recently, computerized versions such as ImPact (Collins et al., 2003; Collie, Makdissi, Maruff, Bennerll, & McCrory, 2003).

Although these tests are considered to be the "gold standard" in concussion assessment, they have never been validated for use on concussed athletes (Grindel et al., 2000). Further, there has been no consensus among researchers as to which neuropsychological tests within the battery are the most sensitive in detecting change following concussion (Randolph, McCrea, & Barr, 2005).

Researchers have investigated recovery patterns following concussion from hours (Echemendia, Putukian, Mackin, Julian & Shoss, 2001; Macciocchi et al., 1996) up to seven days (Guskiewicz et al., 2001; McCrea et al., 2003) in collegiate and professional populations. Following a concussive injury, individuals typically display transient deficits in cognitive

functioning that can often be detected through neuropsychological assessment. Further, concussed athletes exhibit an increase in symptoms reported acutely, but follow a non-linear recovery. (Guskiewicz et al., 2003; Macciocchi et al., 1996; McCrea et al., 2003)

The long-term consequences of concussion are still under investigation. Guskiewicz (2003) found that athletes reporting three of more previous concussions are three times more likely to have additional concussions with an increase in recovery time and symptoms reported. Repeated head trauma has been associated with dementia pugilistica or punch drunk syndrome (Roberts, Allsop & Bruton, 1990). One study of collegiate football players who have sustained two or more prior concussions reported significant long-term deficits in the domains of executive functioning and speed of information processing, as well an increase in self-report symptoms. Further, retired professional football players with a history of concussion were found to have increased rates of cognitive dysfunction. Within the same study, players with a history of three or more concussions had a fivefold prevalence of being diagnosed with mild cognitive impairment, and a threefold report of memory deficits (Guskiewicz et al., 2005). Although longitudinal research is being conducted regarding the long-term consequences of concussion, much is still not understood.

Posturography

The construct of postural stability is comprised of the postural control system. The postural control system uses three sensory systems to operate a feedback control circuit between the brain and the musculoskeletal system. Under normal circumstances, postural stability is maintained by integrating the somatosensory, visual, and vestibular systems. If one system is not operating correctly, the other systems should compensate (Ingersoll & Armstrong, 1992; Riemann & Guskiewicz, 1997).

Typically, what is seen following a concussion is that the athlete will have difficultly integrating information from the three components of the balance mechanism. While the somatosensory aspect appears to remain normal, integration between the visual and vestibular components does not function properly (Guskiewicz et al., 2001).

Postural stability has been recognized as an important piece of the concussion assessment battery (Ingersoll & Armstrong, 1992; Guskiewicz et al., 1997; Guskiewicz & Perrin, 1998; Guskiewicz et al., 2001; Riemann & Guskiewicz, 1997; Shumway-Cook & Horak, 1986; Riemann & Guskiewicz, 2000; Thompson, Sebastianelli, & Slobounov, 2005). Thompson (2005) found postural deficits following injury after symptoms resolution. Guskiewicz (2001) found deficits in postural ability when cognitive deficits dissipated.

The balance assessment may be more sensitive to cognitive changes following injury than some pen and paper tests. Deficits in posturography, without deficits in other domains, suggest the balance mechanism operates on different pathways in the brain. Further, this phenomenon establishes the need for a multifaceted approach to assessment specific to various areas of brain functioning. Posturography provides the clinician with valuable information in making a return-to-play decision.

Youth Concussion

The developing brain of children and adolescents is thought to respond differently than adults following concussive injury. Typically, young children are at increased risk for closed head injury due to car accidents, falling down stairs, abuse, and more frequently sports-related injuries (Aldrich, Eisenberg & Saydjari, 1992; Anderson, 2001). It is widely recognized that the effects of head trauma in children may be considerably different than in adults. However, with

the recent inclusion of pee-wee leagues and youth travel squads, the risk for concussion has increased in an immature population.

There have been limited investigations on the short-and long-term consequences of concussions. Recently, researchers have focused on young adults (ages 18-30), with some investigations of retired athletes. From these studies, it has been determined that most acute concussive episodes in sports have neurocognitive and motor deficits lasting less than five days. Children with concussions showed personality changes, headaches, irritability, school learning difficulties, memory and attention deficits. These conditions were also characterized by the older and adolescent groups (Simon, Letourneau, Vitorino, & McCall, 2001; Roman et al., 1998; Levin et al., 1997). However, it is also difficult to assess developmental deficits in children after concussion, as signs of overt neurologic dysfunction and loss of developmental potential may only be lacking. The loss of developmental potential may only be demonstrable at a later time or under specific circumstances.

Epidemiology

There have been an estimated 63,000 cases of concussion in American high school sports (Guskiewicz et al., 2000). There are approximately 1.5 million high school and middle school football players, compared to only 75,000 college players. With the greater number of participant exposure in middle school and high school, the risk of concussion is greater than that of their collegiate counterparts.

Few sports-related concussion epidemiological investigations have been reported in children and adolescents. One investigation reported that head injuries, as a percentage of all sports-related injuries were 2.8% for all children <10 years, 3.7% 10-14, and 4.2% 15-19 years old. (Kelly et al., 2001)   McCrae (2004) recently found a higher prevalence rate (15.3 %) in high

school athletics than previously reported in the literature. This finding was three times that of the original rates suggested in high school athletes, indicating that athletes may under-report concussive incidences.   Guskiewicz (2000) found an incidence rate of 5.6% in the high schools. Guskiewicz found that players who sustained one concussion in a season were three times more likely to sustain a second concussion.

Metabolic Cascade

The metabolic cascade has been presented for the adult model; however, there are apparent significant differences for the immature and developing brain. The difference between the adult and adolescent brain is attributed to brain plasticity and metabolic differences. Plasticity refers to the ability of the CNS to change in response to environmental influences. Dennis (2000) defines plasticity as a functional and structural plasticity. Structural plasticity refers to anatomical, chemical and electrical properties of the brain that recover following injury. Structural plasticity is assessed by imaging to determine location of lesion and the relationship between the lesion location and behavioral changes following injury. Functional plasticity refers to the cognitive and behavioral competencies that recover following brain injury.

Functional plasticity is assessed by examining deficits in cognitive processes. Recovery has also been explained on the basis of "behavioral substitution", the use of similar neurophysiological processes to achieve the original end by different means. One theory is that the period of neuronal plasticity is one where organizational changes can occur as a compensatory response following mild head injury, and therefore long-term changes are not noted (Chugani, Phelps, & Mazziotta, 1987). Woods and Yueber (1973) examined the related plasticity of both hemispheres. They found that when the injury was located in the left

hemisphere, children showed subsequent deficits in tests of both verbal and nonverbal skills relative to neurologically normal children.

Spatial learning was impaired in the adult rats following concussion, while the immature rats displayed no impairment (Prins & Hovda, 1998). Weight drop models utilizing young rats result in injury only in severities that result in high mortality rates. (Prins, Lee, Cheng, Becker, & Hovda, 1996) In young brains, cells are assumed to be less committed to specific functions so that they can take over the functions of damaged parts of the brain more easily. These theories, however, are older and have not been demonstrated in recent literature.

Metabolic differences, however, may be the main reason for differences between adult and adolescent brain injury physiology. Rates of glucose metabolism in humans have demonstrated significant differences in the local cerebral metabolic rate of glucose consumption between children younger than 15 years and adults (Chugani et al., 1987). Experimental models of concussion found the youngest rats became hypotensive after mild injuries. During severe fluid percussion rates, the mortality rate approached 100% (Prins et al., 1996). Decreased metabolic rates correlates with a decrease in plasticity and diminished recovery following injury.

Another hypothesis is that the developing brain has a lower metabolic rate which, in turn, makes it more resilient to hypoxic events. (Robinson, 1981) Although there is no consensus on the effects of mild head injury on the developing brain, researchers believe that recovery depends on the severity of insult, age at injury, and developmental period when the concussion occurred (Cantu, 1997).

Finally, metabolic and pathophysiological differences have been found between mature and immature brains. This may be contributory to the differences noted between adults and children following brain injury (Capruso & Levin, 1996). Studies using immature animals and

pediatric patients suggest unique differences in post-traumatic cerebral metabolism that may make the developing brain more venerable to acidosis. Young rats have a shortened period of cerebral metabolic depression in response to injury compared to adult rats. Both exhibit substantial post-traumatic hyperglycolysis (Robertson, Bucci, & Fiskum, 2004). Although there is no consensus on the effects of MHI on the developing brain, researchers believe that the severity of the injury and the age at the time of injury contribute to changes in both the acute and long-term neurobehavioral responses following MHI (Cantu, 98).

Second Impact Syndrome

Repeated mild head injury over a short period of time can be catastrophic or fatal (Centers for Disease Control and Prevention, 1997). From 1945-1999, there were 712 fatalities resulting from football, with head injuries accounting for 69%. These fatalities have been attributed to Second Impact Syndrome, epidural and subdural hematoma, and fractures. High school football appeared to produce a greater number of fatalities compared to their collegiate and professional counterparts. Although the incidence of life threatening brain injury has decreased in sports, there is evidence that suggests that concussion may be more common and more serious than preciously thought.

Second Impact Syndrome (SIS) is a serious complication following MHI, with mortality rates near 50% and morbidity rates of 100% (Bowen, 2003; Cantu, 1997). SIS is a rapid cerebral edema and herniation after a second head injury. SIS occurs when a second blow occurs when the symptoms associated with the first blow have not cleared. It starts with persisting post-concussion symptoms when the athlete collapses with symptoms that include visual, motor, or sensory changes (Cantu, 1992).

29

It appears that SIS originates from disordered autoregulation of the brain's blood supply. Following the initial injury, brain cells survive in a vulnerable state (Wojtys et al., 1999), and after the second impact, a disruption of the brain's autoregulatory system leads to increased intracranial vascular engorgement (Bowen, 2003; Cantu, 1998; Cantu, 1992; Kelly et al., 2000). The intracranial pressure decreases the perfusion to the brain and can lead to ischemia (Kelly et al., 2000) and herniation of the uncus through the foramen magnum (Bowen, 2003; Cantu, 1998). Within seconds to minutes, the athlete collapses with rapidly dilating pupils, loss of eye movement, and respiratory failure (Bowen, 2003; Cantu, 1992). Brain stem failure occurs within an estimated two to five minutes.

Almost all reported cases of SIS have occurred in athletes under the age of 18 (Cantu, 1998; Cantu, 1992; McCrory & Berkovic, 1998). Cerebral swelling in children following MHI increases the risk for SIS due to different autoregulatory responses to trauma in these younger individuals (Bruce et al., 1981; Schnitker, 1949; Snoek et al., 1984). The prevention of SIS through concussion evaluation of younger athletes is critical to prevent catastrophic consequences (Meuller 2001).

Youth Concussion Assessment Tools

Concussion assessments in youth athletes typically begin with limited availability of medical care. Similar to their adult counterparts, youth research and evaluation tools have filtered into the pediatric population. Similarly, there are several methods for concussion assessment in youth. In the adolescent population, a conservative approach tends to be the more common thread. However, the most common assessment techniques and findings will be highlighted in the following sections.

Typical concussion assessment begins with an Emergency Room visit. Following brain injuries in youth athletes, most parents, and those responsible for the athlete, take them to the Emergency Room for care. Upon arrival in the Emergency Room following a concussion adolescent athletes are given a CT scan. Neuroimaging, such as MRI and CT scan, can identify gross deformities and lesions that may result from the injury; however, it does not increase knowledge during the evaluation of concussions (Aubry et al., 2002; Homer & Kleinman, 1999). The disadvantages of these techniques have been discussed previously. This is an important step in adolescents to rule out more serious and life threatening conditions such as hematomas, gross deformity etc, prior to treating the athlete.

Once the athlete is diagnosed with a concussion, additional problems occur with follow-up care and home care instructions. A survey of Emergency Room physicians demonstrated improper follow-up instructions following concussion for greater than 69.7% of youth athletes hospitalized for concussion. (Genuardi & King, 1995) The multifaceted approach allows for a more comprehensive evaluation that may be accomplished by allied health care practitioners to ensure the best medical care for youth athletes.

Self-Report Symptoms

The most important factor on the field, for young athletes, in assessing concussions is self-report symptoms (dizziness, confusion, headache, amnesia, etc.) These markers provide an idea that the athlete has incurred a concussive incident. Careful consideration of self-report symptoms must be accompanied by other measures to ensure the safety of the athlete. Self-Report Symptoms (SRS) are important because most guidelines still state that no athlete returns to participation while still symptomatic. SRS can be administered in either a checklist or scale. It is important to identify baseline measures for comparison following injury. Also, one must

emphasize symptoms that are commonly experienced during premorbid examination to eliminate premorbid symptoms that would confound data post injury.

Recently, Collie (2006) found that athletes who report symptoms still demonstrate cognitive deficits. Yeates (1999) found that children who suffer from MHI were more likely to display Post Concussion Syndrome (PCS) than controls; however, these increases were seen only in certain somatic and cognitive symptoms rather than in the entire range. Children who displayed these increases in PCS had poorer performances on neuropsychological testing compared to baseline, and had only partially resolved themselves three months post-injury. Field (2003) conducted a study with 161 high school football players and 22 male soccer players. Evaluations of 24 hours, three, five, and seven days post-injury, found that high school athletes who sustained a concussion had increased post-concussion symptoms compared to the injured college athletes and their matched controls.

Some grading scales use PTA as a marker of concussion severity and return-to-participation guidelines. Amnesia, however, is difficult to document and unreliable with this population because of limitations in their abilities to give verbal reports and to describe subjective symptoms (Hynd & Willis, 1987; Levin et al., 1988). Therefore, although symptomology is important, many self-reports from the adolescent may be unreliable and a more comprehensive approach should be taken.

Neuropsychological Assessment

Neuropsychological testing in children has only recently been fully developed. Typically, these neuropsychological test batteries are comprised of downward extensions of tests used in adult populations (Spreen & Strauss, 2000). Some of these tests, however, may not be appropriate because the domains being tested might require different demands on the child's

cognitive system (Spreen & Strauss, 2000). Additionally, different factors affect neuropsychological test scores inclusive of previous concussion history; educational background, premorbid level of function, cultural background, age, test anxiety, distractions, sleep deprivation, medication, psychiatric disorders, learning disabilities, and previous neuropsychological testing.

The applicability of these tests in a pediatric population has not been studied. It is important to determine whether these assessments are clinically appropriate for use in a younger population. Additionally, it is imperative these assessment tools be studied in children and adolescent athletes to determine the effectiveness and adequacy in identifying neurological deficits in this population following MHI. Furthermore, factors relating to the development and maturation of this young age group, which could confound the results obtained by these assessments, must be understood.

Neurocognitive recovery curves following mild traumatic brain injury have been the focus of recent research for pediatric and high school athletes. They are likely to deviate from that of adult curves because of dissimilar intracranial responses to severe brain injuries, not mild traumatic brain injury as seen in the literature. These changes include prolonged cerebral edema, delayed dysautoregulation, and improved cortical plasticity. (Grindel et al., 2001)

Researchers investigating recovery patterns in young athletes found, on average, that they take longer to recover, and report an increase in symptoms compared to their matched controls and collegiate athletes (Guskiewicz et al., 2000; Lovell et al., 2003; McCrea, 2001). Lovell et al. (2003) reported that neuropsychological test scores were statistically significant between the injured and control groups on measures of memory and post concussion symptoms in high school athletes for up to seven days post injury. McCrae (2001) assessed 141 high school

football athletes utilizing the Standardized Assessment of Concussion. Concussed athletes demonstrated decreases on post-concussion screenings when compared to baseline data.

Field (2003) conducted a study with 161 high school football players and 22 male soccer players. Evaluations of 24 hours, three, five, and seven days post-injury using the Controlled Oral Word Association Test (COWAT) and Brief Verbal Memory Test (BVMT) found that high school athletes that sustained a concussion had increased post-concussion symptoms and recovery. Using COWAT and BVMT, concussed athletes took longer (up to seven days post-injury) than that of the injured college athletes and their matched controls. Further, Moser (2002) evaluated the previous history of concussion in youth. He found significant differences for athletes with recent concussion on measures of Repeatable Battery for the Assessment of Neuropsychological Status (RBANS). Additionally, those with two or more concussions were statistically significant from those with 0-1 previous concussions.   These findings contrast with other studies (Barth et al., 1989) that demonstrate collegiate athlete's recovery within five to ten days. These measures are, however, not normed for youth. The sample sizes were small and there were only a few people with one concussion.

Long-term Consequences of Concussion

It is well established that the likelihood of permanent neurological deficits increase in children who suffer moderate to severe traumatic brain injury, but the long-term consequence of sports-related concussion is still under investigation. Long-term consequences following concussive injury have been investigated at four months (Gulbrandsen, 1984) to years (Klonoff, Clark, & Klonoff, 1993; Mueller, 2001). A follow-up study of adults who suffered a mild head injury during childhood found neurocognitive deficits of 23.7% when tested five years following the injury (Klonoff, 1977). A subsequent follow-up 23 years post-injury found that 31% of the

sample reported subjective sequelae (Klonoff et al., 1993). These complaints were typically

related to the severity of the head injury. These long-term studies might indicate that MHI in

childhood can have lasting consequences.

Additional researchers found deficits incurred by young athletes may last up to one year

post-injury. Mueller (2001), in his study, pointed out that 23.7% of youth with head injuries

continued to demonstrate measurable neuropsychological deficits up to one year post-injury.

Anderson (2001) tested children that had head injuries between the ages of three and seven.

They were tested six months and 30 months post injury. Anderson found that the mild head

injured group revealed significant deficits months after the event on memory and language skills.

While the long-term consequences are unknown, researchers have investigated the effect

of cumulative concussions in high school athletes (Collins et al., 2002; Lovell et al., 1999; Moser

& Shatz, 2002). Lovell (1999) evaluated the effect of previous history of concussion and

learning disabilities. He found that Impact scores were significantly different between the

injured and control groups on measures of memory and post-concussion symptoms at 36 hours,

four days and seven days post injury respectively. Collins (2003) examined 163 athletes with

sports-related concussion. He found that high school athletes with three or more previous

concussion are more likely to have a more severe on-field presentation of concussion markers

(LOC, PTA). Further, athletes with three or more concussions were 9.3 times more likely than

athletes with no previous history to demonstrate three or four abnormal on-field markers.

This research corroborates previous studies that found high school athletes, with a self-

reported history of three prior concussions, nine times more likely to exhibit three or four on-

field abnormal markers of injury when they experienced subsequent injuries than those with no

prior concussion. The long-term consequences of concussion in the youth population have not

been thoroughly investigated, and continued longitudinal studies are necessary to thoroughly understand the nature of concussion in a developing brain.

Adolescent athletes provide an additional concern as injuries typically occur during times of rapid development (Aldrich et al., 1992; Prins & Hovda, 1998). There is a growing body of evidence that pre-injury conditions, based on family environment and the level of child's development prior to injury, can predispose the child to deficits. (Donders & Strom, 2000) Children who suffer TBI and live in a maladjusted family environment will incur negative effects on recovery. Injury related variables were the most important influences on neurobehavioral recovery during the first year after TBI. Lovell (2004) found that athletes with no on-field markers and symptoms resolving within 15 minutes of injury were still statistically different, but recovered to control levels by day six.

Factors aside from environmental concerns have been evaluated. Collins found that athletes with previous histories of concussions and/or learning disabilities were significantly different from those with no previous histories and/or learning disabilities. Further, Collins (2003) divided 109 high school athletes based upon the presence or absence of headache. He found significant differences between the headache and the no-headache group using Impact, a computerized neuropsychological test battery, on measures of memory and reaction time.

Postural Stability Assessments in Children

The development of postural stability and the postural control system result from complex interactions between the nervous and musculoskeletal systems (ShumwayCook & Horat, 1986). Despite continual growth and maturation during the pubertal years, the maturation of postural stability typically reaches adult-like values between the ages of seven to ten (Assainte

& Amblard, 1995). Hence, the maturation of postural stability takes place prior to assessments in high school.

Tests of postural stability are commonly used to assess balance in children, with a variety of underlying pathologies such as cerebral palsy and obesity. Most of these measures of postural stability in children were made using force plates. Other common measures of postural stability in children include tilt board tests (r= 0.45-0.54; Atwater, Crowe, Deitz, & Richardson, 1991), the Sensory Organization Test (NeuroCom), and more recently the BESS.

The Sensory Organization Test within the Neurocom Smart Balance Master consists of three trials of each of the six conditions (18 total trials), with each trial 20 lasting seconds. This system works from a force plate that has the ability to measure angles and forces being generated at the ankle, knee and hip. The test systematically alters visual and somatosensory referencing in an attempt to individually evaluate the three components of the balance mechanism (visual, vestibular, and somatosensory). A typical composite balance score is 79 (Guskiewicz et al., 2001). The SOT is the gold standard for postural stability in concussion, however, the SOT is not portable and is very expensive.

Guskiewicz (1997) demonstrated balance deficits in athletes sustaining a mild head injury in a sport using this test. Guskiewicz was unable, however, to demonstrate cognitive deficits using the brief battery of pen and paper tests described above. He indicated that the balance assessment may be more sensitive to changes in cognitive functioning following MHI. While the SOT appears to be a valid measure of postural control, it is not practical for every setting.

Prior to the use of computerized force platform systems, the Rhomberg test was commonly used to evaluate postural stability following an injury. The premise behind the Rhomberg test was to subjectively evaluate somatosensory impairment by having the patients

37

close their eyes to eliminate vision as a sensory source (Black, Wall, Rockette, & Kitch, 1982; Cohen, Blatchy, & Gombash, 1993; Riemann & Guskiewicz, 1999). Since the inception of the test,, there have been various interpretations and modifications to increase the sensitivity and specificity for balance disorders. Several judgment criteria were established, such as time to first touch down, or compensatory event (Cohen et al., 1993; Jansen, Larson, & Oleson, 1982). Most modifications have occurred in foot position (stance) to alter the participant's base of support. However, each modification of the original Rhomberg test was criticized for the lack of sensitivity in situations where postural sway increases without the complete loss of postural stability and clinical populations, such as concussive injuries. Continued modification of the Rhomberg test resulted in the traditional BESS protocol.

The BESS was developed as an objective assessment tool to be used by clinicians with minimal cost and training for the evaluation of postural stability following concussion. During the test, the athlete stands on two surfaces (firm and foam); with eyes closed performing three stances (double-leg stance, single-leg stance, tandem, or heal-to-toe stance). Riemann (1999) found intertester reliabilities for each condition within the BESS compared to long force plate sway measures ranging from 0.78-0.96, dependent upon the stance. Moreover, the BESS has been correlated to the SOT composite scores and demonstrated similar recovery patterns. (Guskiewicz et al., 1997; Riemann & Guskiewicz et al., 2000; Susco et al., 2004 Valovich et al., 2003; Valovich et al., 2004)

Considerable research has been conducted to examine the psychometric properties of the BESS. In studies following concussion, subjects exhibit acute postural stability alterations up to five days post-injury (Guskiewicz et al., 2001; McCrea et al., 2003; Riemann & Guskiewicz, 2000) with recovery usually occurring within four to seven days post-injury to pre-injury

baseline values. Furthermore, Guskiewicz evaluation found deficits for up to seven days following the injury and return to normal values.

Although the BESS appears to be sensitive to subtle deficits following concussion, it has several drawbacks. Recently, Valovich (2004) found athlete's scores on the BESS were less reliable during subsequent trials. Moreover, administration of multiple trials of the BESS results in practice effects, with the number of errors decreasing with each consecutive trial. (Valovich et al., 2003) Further, the effects of fatigue increase the number of errors acutely, but the athlete recovered after 20 minutes of rest following an exercise session (Susco et al., 2004).

Effort

It is important for athletic trainers to discriminate between improvements due to learning, practice effects, poor baseline measures and improvements due to recovery from MHI. Both neuropsychological and postural stability assessments are used repeatedly to assess initial status following MHI, and as a means to track an athlete's progress following MHI. Typically the athletic trainer will administer these tests at the time of injury, for several days following the injury, and up to three months following the injury (Guskiewicz & Perrin, 1996; Guskiewicz et al., 1997; Lovell et al, 1999; Macciocchi et al., 1996; McCrea, 2001). Since athletes are administered these assessments multiple times, there is a chance that a learning effect will take place.

Understanding the extent of this learning effect and quantifying the learning curves are essential if athletic trainers are to accurately interpret the results of a given assessment and make appropriate decisions regarding returning the athlete to play (Guskiewicz et al., 2001; McCaffrey et al., 2000). To reduce this effect equivalent, multiple forms have been established for repeat administrations of the test. One approach involves identifying the magnitude of practice effects

on the various assessments which allows this source of error to be removed from the scores in which one is interested. One must try, however, to replicate the clinical situation (Franzen, 1999) and the population that it tests (Lovell et al., 1999) so that the learning effects are studied in a manner using the same time intervals, number of test sessions, and subject pool, as one would utilize clinically.

Another approach involves establishing the level of effort within the population during baseline testing. In concussion assessment, it is traditionally assumed that maximal effort is given during all test sessions. This may not, however, be the case with baseline testing. If an athlete does not put forth their best effort, then baseline scores are not comparable and, therefore, a direct interpretation of post-injury data cannot be made. There is no research in concussion testing that has examined whether maximal effort is given during baseline testing, or whether those not giving maximal effort will affect test score comparisons post-injury.

Effort is defined as "a usually earnest attempt" or "the use of physical or mental energy to do something". Objective effort tests have been used to determine the advertant and inadvertent effort to fake bad or malinger. Additionally, effort tests adequately characterize patients that demonstrated sub-optimal effort or purposely performed poorly (malingering) (Hall & Pritchard, 1996; Rogers, Ernest, & Liff, 1993; Tombaugh, 1996). The importance of specific examination of effort in neuropsychological assessment is elicited by the high base rate for malingering, deception, or suboptimal effort among individuals claiming cognitive impairment ranging from 20% (Griffin, Glassmire, Henderson, & McCann, 1996) to 40% (Greiffenstein, Baker, & Gola, 1994).

Standardized effort assessments provide an objective method to identify invalid test results. Subjects that exhibit suspect effort using an objective effort test have been proven to

have invalid neuropsychological test scores (Green, Rohling, Lees-Haley, & Allen, 2001a; Green & Iverson, 2001b; Moss, Jones, Fokias, & Quinn, 2003; Ruff & Parker, 1993). Several studies have examined whether mild head injury patients score differently from compensation claimants. Green, et al. (2001a) found a greater proportion of exaggerators in the mild head injured group compared to those in the compensation/ litigation groups. (Green et al., 2001a; Green & Iverson, 2001b, Rohling, Green, Allen, & Iverson, 2002) This research alludes to the need for effort testing within the mild head injured population. Much of the data regarding effort and malingering revolves around adult populations involving workman's compensation or litigation (Lees-Haley, 1997; Rohling, Myers, & Millis, 2003; Youngjohn, 1995b).When approached with the same problem in adolescence however, it is attributed to effort.

Faust (1988) found that when adolescents were asked to perform poorly, clinicians could not differentiate between those who were asked to perform poorly and those with head injuries. The disability classifications given to each adolescent, due to inaccurate scores, were incorrect for those that were instructed to perform poorly. The adolescents instructed to perform poorly were classified into the same group as those with clinical disabilities (Faust, 1988). The inability to determine whether maximum effort was performed is more difficult when clinicians are not suspecting the adolescent to perform poorly. Therefore, the addition of procedures to assess the effort of test-taking might help to correctly identify suspect effort and the inability to make valid interpretations regarding deficits following injury.

There are several measures available to identify those not putting forth optimal effort during testing. These measures are designed to give the illusion that the test becomes harder when in reality the test is simple. Those subjects trying to "fake bad" or demonstrating poor

effort would overestimate poor performance, thus giving scores far beyond the normative rate. These measures include the Rey 15 item test and the Dot Counting test.

The Rey 15 item test with recognition trial was developed to examine the subjects' cooperation. The participant who consciously or unconsciously wishes to appear impaired will fail at a task that all but the most severely brain injured or mentally retarded participants perform easily. A person who cannot recall and recognize for a total score of 20 would be considered to have sub-optimal effort (Boone, Salazarm, Lu, Warner-Chacon, & Razani, 2002a). Specificity of a cut score of less than 20 ranged from 61.1% to 100% while sensitivity of the measure ranged from 61.2% to 77.6% while administration takes less than five minutes.

The DCT is a brief task that assesses lack of effort either intentional or unintentional by examining an over learned skill and by measuring time to completion as items become more difficult. A cut score of 14 for optimal controls establishes sensitivity 92% of and specificity of 99% (Boone et al., 2002b). Further, a brief administration time less than 10 minutes is adequate for inclusion into a short neuropsychological test battery.

Sub-optimal effort may impact neuropsychological scores on baseline performances, decreasing the ability to make adequate clinical interpretations regarding the post-injury assessment following a concussive injury. Currently, there is no research in sports-related concussion assessment that has examined whether an individual exerted optimal effort during baseline testing to assist in the interpretation of the scores. The addition of a brief gross measure of effort may be used as a screen to obtain conservative estimates of sub-optimal effort. If sub-optimal effort is found, the participant should be retested with a more comprehensive effort test within the neuropsychological test battery.

Measurement Techniques

The statistical tools used to quantify a "good" measure have been addressed in several

newsletters (Atkinson, 1995; Nevill, 1996). The increased awareness created an influx of studies

concerning measurement issues within the literature. The most commonly described topics

involve the reliability and validity of a specific measurement tool.

Reliability

Reliability can be defined as the consistency of measurements, or of an individual's

performance on a test, or the absence of measurement error (Baumgartner & Jackson, 2003;

Morrow, 1995; Safrit & Wood, 1989). Reliability can also be explained in terms of observed

scores, true scores, and error scores. Reliability can be calculated as the ratio of the true score

variance to the observed score variance.

Understanding these issues, reliability is said to be achieved when measurement error is reduced

and differences among the test subjects can be identified. Measurement error will always be a

concern and it is seen as a lack of agreement between scorers, a lack of consistent performance

by the test subjects, or even as a failure of the measures to be consistent.

Additional factors affecting reliability include: scoring accuracy, number of test trials, the

test environment, presentation/format of the test instructions, and the degree of difficulty of the

test itself. An acceptable amount of reliability is, therefore, established when the test subjects are

heterogenous in their abilities and the test can discriminate ability levels within the group. There

are different methods with which reliability can be estimated.

There are several types of reliability, although only a few will be described. Reliability

coefficients are typically calculated in terms of the sources of measurement error. The first type

of reliability is stability reliability. Stability reliability is defined as the day-to-day variability in

measurements. This is the most common type of reliability analysis as the interest is in what changes occur if testing is done on multiple days.

The second type is objectivity. Objectivity is the degree upon which multiple scorers agree based on the scores given within a measure. This is also known as rater reliability. Rater reliability is contingent upon a clearly defined scoring system, and the degree to which the scorers can assign a given score accurately.

The third type of reliability is internal consistency. This type of reliability measures the consistency of the parts of the measure. The extent to which the parts of the test measure the same idea would indicate high or low internal consistency.

Obtaining valid interpretations is dependent upon having test scores that are highly reliable. However, a researcher can have a highly reliable measure without the ability to validly interpret test scores. The process of validating interpretations from test scores must include both reliability and validity studies.

Validity

Validity evidence is for the inferences made about a test score. Prior to the 1950's researchers only investigated criterion and content related validity. However, for most social sciences there was no clear cut, concise criterion procedure for obtaining validity evidence for constructs. A construct is something constructed by mental synthesis. Therefore, in 1955 Cronbach and Meehl published a paper regarding the poor state of affairs with validity evidence. They originally described four types of validation.

The first type of validity they described was criterion oriented validity which includes predictive and concurrent validity. This type of validation compares test scores with some

measure which is the criterion (Cronbach & Meehl, 1955; Huck, 2000). If the criterion measure is obtained some time after the test is given, this would be considered predictive validity.

The second type of validity is concurrent validity. If the test score and the criterion score are determined at essentially the same time it is considered concurrent validity (Cronbach & Meehl, 1955; Huck, 2000). For this method of validation the criterion behavior is of utmost concern. However, this type of validity is difficult to assess if there is no criterion available.

The third type of validity they described was content validity. This type of validity is observed when a test is a sample of the universe in which the researcher is interested (Cronbach & Meehl, 1955; Huck, 2000). This type of validation is evaluated by deductive reasoning. In this validity, the tester is concerned with the type of behavior involved with the test performance. Normally, this type of validity is determined by having experts evaluate the measure. However, this is subjective and difficult to observe if there is no acceptable universe available.

The first three types of validity were readily accepted by academia; however, these types of validity failed to provide adequate information when there was no criterion or acceptable universe. This process was extremely difficult in the social sciences, especially psychology. Therefore, they introduced construct validity.

A construct is "some type of attribute of people, assumed to be reflected in test performance". Therefore, construct validity is involved when the test is to be interpreted as a measure of some attribute or quality which is not "operationally defined". Construct validity is described as a process required when "no criterion or universal content is accepted as entirely adequate to define the quality to be measured". The quality underlying the test is the most important aspect. To provide construct validity evidence, a researcher must assemble evidence about what the construct meant. In this process, the criterion is defined as the testing proceeds.

Evaluating construct validity can be accomplished in several ways. The first was the examination of group differences. This technique requires the researcher to calculate sample statistics to evaluate statistical differences between groups on one or more specified constructs. The second was to examine the correlation matrices. This matrix was a correlation between tests thought to measure the same construct. This also involved factor analysis (which is a mathematical technique used to reduce a difficult system of correlations into fewer dimensions) which was developed by Spearman (1904). The theoretical premise was if two tests were thought to measure the same construct, they should be correlated or, using factor analysis, should factor into one group. This proved a daunting task since there were no precedents for this measure. This type of validation was based upon a significant amount of theory which had to translate to something measurable.

For a researcher, Cronbach and Meehl ideas were revolutionary. However, the ideas were difficult to grasp and incorporate into clinical use. Therefore, Campbell and Fiske (1959) introduced a method to provide information regarding construct validity by examining both convergent and discriminate validity. Based upon construct and convergent, validity tests that measure the same construct should be correlated, and those that are not correlated do not measure the same construct. The correlations between the measures give information regarding the trait, method, and thus, the construct. This method is very useful if performed correctly as it provides information about reliability as well as validity.

To obtain data from which you can make valid inferences, you must first have a reliable (consistent) measure. If the reliability coefficients are large enough, you continue to examine the validity information. The validity information is the correlation between the same traits across multiple methods. Convergent validity measures the correlation of a test to other measures.

If convergent correlations are high, examination of discriminate validity is warranted. Discriminate validity occurs when the test measures something different from those used in convergent studies. Furthermore, discriminate validity is demonstrated by exhibiting low correlations with measures of unrelated constructs or tests.

A psychometrically sound test has high reliability (consistency). Reliability must be established before collecting validity evidence. Currently, construct validity is a very commonly used validity evidence in the sciences. Gathering evidence of reliability and validity is an ongoing process, although no single set of observations provides critical evidence. Many observations over time can gradually clarify what the construct means, how the test measures the construct, and what group of tests measure the construct best. Adequate evidence to support the reliability and validity of test scores must be obtained prior to interpretations.

CHAPTER 3

MATERIALS AND METHODS

Subjects

Subjects came from six northeast Georgia high schools. The sample consisted of athletes participating in both varsity and junior varsity football at those schools. Control subjects were taken from the same subject pool and matched by age, ethnicity, Hopkins Verbal Learning Test scores and position. All subjects ranged between the ages of 13 to 19 years. All subjects and their guardians were notified of the testing sessions and signed informed consent and assent forms.

Recruitment of subjects

The subjects were recruited from schools that participate in the Georgia High School Association and have a Certified Athletic Trainer or team physician present during practice and competitions. Each school was notified of the project. Authorization from administration to conduct this study in each school was obtained prior to any contact with coaches, athletic directors and athletes. After authorization from the school administration, consent forms were sent to the parents of each athlete. Minor assent forms were obtained. Once the parental consent form was received, the athlete was enrolled in the study. Each athlete and guardian was informed that participation in this study was strictly voluntary and did not affect the player's academic or team status. The participant also had the right to withdraw from the study at any time and have their information destroyed or returned.

Definition of Concussion

A concussion was defined as an evaluation of the athlete by a Certified Athletic Trainer or team physician that finds any alteration in mental status (marked by confusion), PTA, LOC, report of symptoms associated with concussion, or one point deficits compared to baseline SAC scores.

Inclusions of the Study

Subjects between 13 and 19 years old.

Male Football players

All ethnicities

Exclusions of the study

Athletes who have had a concussion less than one month prior to baseline testing.

Athletes with English as a second language.

Outcome Measures

Neuropsychological test batteries have been extensively used in previous research to assess various neurocognitive functions including information processing speed, immediate memory, concentration, attention, cognitive set shifting and verbal fluency.[22-24] This battery covers domains most likely to be affected by brain injury, has established norms for this age group, has clear instructions, and is reliabile.[21, 25-27] The motor test battery includes the Balance Error Scoring System (BESS) and Digital Finger Tapping Test. The Rey 15 item test and the Dot counting tests were used to assess effort. A total of ten dependent measures were obtained.

Standardized Assessment of Concussion (SAC) (McCrea, 2001)

The SAC is a neurocognitive examination that was specifically intended for the assessment of concussed athletes on the sideline of play. The test components include orientation, memory, concentration and delayed recall. Alternate forms (A through D) of the test are used to decrease practice and learning effects.

Concussed athletes scored significantly different than non-concussed athletes immediately following injury to two days post-injury (McCrea, 2001). Scores 48 hours post-injury returned to baseline values for the injured group. A decline in SAC score at the time of injury is 95% sensitive and 76% specific in accurately classifying injured and uninjured subjects. Reliability analysis demonstrated a test-retest reliability of 0.53.

Balance Error Scoring System (BESS) (Riemann & Guskiewicz, 1997)

The BESS consists of two conditions (single leg and tandem) and two surfaces (foam and firm). The firm surface will be the high school gymnasium floor. The foam surface consisted of a 61 by 61 by 10 cm thick block of 0.88 kilogram density polyester open cell foam (load deflection of 17.5 to 19.3). Each stance is held for 20 seconds with the subjects' hands on hips and eyes closed. Total number of errors is calculated for each condition. These errors consisted of opening eyes, stepping, stumbling, balance checks, falling from the test position; moving their hips into an angle >30 degrees of flexion or abduction; lifting their toes or heels from the test surface; or remaining out of the test position for greater than five seconds. The BESS will be video recorded to increase accuracy and efficiency of scoring.

Concussed athletes have shown deficits in postural stability using the BESS for up to five days post-injury in a collegiate population with recovery usually occurring within four to seven days to pre-injury values. Similar recovery curves have been demonstrated in youth athletes.

BESS scores were compared to long force plate measures and demonstrated a reliability coefficient of r =0.30-0.78 (Riemann & Guskiewicz, 1999) and test-retest coefficient r = 0.67.

Trail Making Test (Reitan & Wolfson, 1958)

The Trail Making Tests, Parts A and B, from the Halstead Reitan Neuropsychological Test Battery (Tucson, AZ; Reitan and Wolfson) is a test of speed of attention, sequencing, mental flexibility, and visual search and motor function (Spreen, 2000). The test consists of parts, A and B. Part A consists of encircled numbers from 1 to 25 selectively spread across a sheet of paper to insure no line drawing overlap. The object of the test is for the subject to connect the numbers in order, beginning with 1 and ending with 25 in as little time as possible. Part B is more complex than part A because it requires the subject to connect numbers and letters in an alternate pattern (1-A-2-B-3-C etc.), in as little time as possible.

Trail Making Test performances by patients with mild traumatic brain injury are slower than those of control subjects (Leininger, Gramling, & Farrell, 1990). Concussed athletes were found to have an increased number of errors as well as statistically significant differences between injured and control participants (Guskiewicz et al., 2000; Guskiewicz et al., 2001; Macciocchi et al., 1996; Echemendia et al., 2001). Errors for the mild head injured group appear as impulsivity and perseverative errors. The test is appropriate for ages 11 and older with moderate practice effects and test-retest reliability coefficients for Part A, r= 0.81 and for Part B, r = 0.67.

Symbol Digit Modalities Test (SDMT) (Smith, 1982)

SDMT assesses complex scanning, visual tracking, sustained attention, and motoric speed. This test involves substituting numbers for random symbols using a reference key. The subject is given 90 seconds to pair specific numbers with a given symbol.

Concussed athletes have lower scores for SDMT. (Echemendia et al., 2001; Collins et al., 1999). Test-retest reliability coefficient was 0.74 in young athletes tested one to two weeks apart (Hinton- Bayre et al., 1997). There does appear to be moderate practice effects if administered over a relatively short period of time; however, this test is appropriate for ages 8 and older.

Digit Span (Wechsler, 1981)

Digit Span is a part of the Wechsler battery that measures immediate memory, attention, and concentration. This test is comprised of both a forward and a backward version. Both tests consist of seven pairs of random number sequences that require the athlete to repeat the series verbatim. The examiner begins with the smallest number sequence and increases number sequences with the correct pattern. Digit Forward has been demonstrated to measure efficiency of attention, while Digit Backwards measures working memory necessitating visual and spatial abilities.

Patients with mild traumatic head injuries are apt to repeat the correct digits, but mix up the order. Mild head injured patients have spans that drop below recall of four or five spans, but return to normal levels within a year (Lezak et al., 2004). Digit Forward Reliability coefficient range from 0.84-0.91, while Digit Backward Reliability coefficients range from r = 0.69-0.75, but appears to be sensitive to practice effects.

Hopkins Verbal Learning Test (HVLT) (Brandt, 1991)

HVLT assesses verbal learning and recall of words. This test requires an individual to learn a list of 12 words from three semantic categories. There are three learning trials in which the subject is read the complete list of words and required to repeat as many words as they can following each trial. After a brief delay of approximately 15 minutes, subjects are again required to produce as many words as can be remembered.

HVLT has been useful in mild head injured participants. Concussed athletes demonstrate fewer words recalled over trials. Further, delayed recall and recognition trials are significantly different than matched controls (Guskiewicz et al., 2001; Echemendia et al., 2001). Of those with scores of 25 or greater on the summed learning trials, 92% did not have a postconcussive syndrome one month after injury. Reliability coefficients range from approximately R= 0.50 (after 9 months in healthy adult over age of 60) to R= 0.77 (serial testing of athletes).

Digital finger tapping Test (Reitan, 1993)

This test measures motor speed and manual dexterity. The digital finger tapping tests consists of five 10-second finger tapping trials. The average taps during five trials (more if necessary) within a five-tap range is used for data collection. Test-retest reliability coefficients range from 0.58 to 0.94 for both normal and neurologically impaired.

Brain-disordered participants tend to have a slowing effect on their finger tapping rate. Traumatic brain injury impedes the rate of tapping (Haaland, Temkin, Randahl, & Dikman, 1994; Ruff, Wylie, & Tennant, 1993). However, it appears that slowed processing underlies slow finger tapping, and not changes in motoric ability.

Rey 15 item test with Recognition Trial (Boone et al., 2000a)

This test, also known as the Rey Memory Test, is used to test the validity of a memory report. The examiner has 15 items marked on a piece of paper in five rows with three characters per line. Patients are exposed to the characters for 10 seconds, then the characters are withdrawn for 15 seconds and the patient is asked to recall the characters.

Participants with sub-optimal effort have been found to deny recall or to make more omission errors, while those with traumatic brain injury can report at least nine items with excellent specificity (97-100%), although sensitivity was modest (47%) (Bernard & Fowler,

1990; Greiffenstein et al., 1994; Boone et al., 1995; Frederick, 2002). Sensitivity ranges from 5-86% and specificity ranges from 56-100%.

The Dot Counting Test (DCT) (Boone et al., 2002b)

The DCT is a brief task that assesses lack of effort either intentional or unintentional by examining an over-learned skill, and by measuring time to completion as items become more difficult. The test consists of six serially numbered 3x5 cards with randomly arranged dots. The cards are shown to the patient one at a time. The patient is told to count and identify the number of dots as quickly as possible.   The patient's time will increase gradually with the increased number of dots. One deviation away from the mean would indicate a person is not offering their best effort. Test-retest reliability coefficients range from 0.75-0.96.

Self-Report Symptoms (Piland et al., 2003)

This scale includes measurements of presence or absence of symptoms as well as duration and severity of symptoms. This Head Injury Scale contains 16 items derived from the most commonly described in the literature. This self-report measure reports on a likert scale for both duration and severity. Three scores are obtained: duration, severity and total score.

Self-report symptoms are taken as a part of the general assessment of all athletic injuries and have played an important role in the assessment of concussion. A non-significant difference between the concussed group and the matched controls exists during post-injury examinations. Further the concussed patients reported more symptoms at higher durations and severities than matched controls. Self-report symptoms are important because most guidelines still say that no athlete returns to participation while still symptomatic.

Pilot Study

Preliminary data was collected by administering a brief neuropsychological test battery which included SDMT, Digit Span, SAC, Digital Finger Tapping Test, Grooved Pegboard, and the BESS to 78 high school football players. Neuropsychological tests for children tend to be downward extensions of the adult batteries (Spreen, 1984). Within each neuropsychological domain, (attention, memory, etc.), specific tests have been examined for their psychometric properties. However, normative values for tests are dependent upon the population from which they are based (Keppel, 1991). Examination of the preliminary data revealed that subjects tested performed equitable for the age dependent normative values. All measures were within one standard deviation of the normative values set for each specified test within the population. Therefore, this sample was comparable and representative of all youth athletes.

Moreover, within the pilot study, a generalizability study was conducted utilizing the facets of the BESS. An advantage of this study is that it employs generalizability theory, which allows variance in the BESS to be partitioned and quantified. A completely balanced five facet crossed model was utilized to determine the percent variance attributed to each facet in the G-study. This study confirmed that the variance associated with the double leg stance was extremely low. Although this project found the intraclass reliability of the BESS to be 0.60, most psychometricians would say the current reliability would not be adequate for a measurement with implications for return to participation decisions. When the double leg stance for both the firm and foam surfaces was removed the reliability of the BESS increased to R= 0.71. Since the single leg and tandem stances accounted for most of the variance within the stance facet, it was suggested that the BESS protocol consist of two trials using firm and foam conditions for the single leg and tandem stances.

Data Collection Procedures

All data was collected during summer and fall (May-August) of the 2005 football season. All evaluations were conducted by individuals with training in the administration of the test instruments by a neuropsychologist (Dr. Stephen Macciocchi). Testing included the aforementioned neuropsychological tests, Self-Report Symptoms, the Rey 15-item test with recognition trial, the Dot Counting Test, and a health questionnaire. All testing took place at the athlete's high school at times that did not interfere with class schedules.

Baseline scores on the R15-R and the DCT were examined. A cut score for suspect effort for the R15-R was set as a total score less than 20 (Boone et al., 2002a), while a cut score for the DCT was set at 14 (Boone et. al., 2000b). All data was examined to eliminate outliers and the remaining subjects were divided into groups based on normal and suspect effort scores. All tests were scored by the primary investigator.

The Balance Error Scoring System was also administered. The traditional model which consists of three stances (double leg, single leg and tandem stance) on two surfaces (firm and foam) was revised. The new model eliminated the double leg stance on both the firm and foam surfaces. The new model consisted of two stances (single leg and tandem stance) on two surfaces (firm and foam surfaces).

Prior to the test, each participant was allowed to familiarize themselves with the BESS by practicing each stance on the different surfaces for ten seconds. For all conditions, the participants were then instructed to stand quietly, as still as possible, with their hands on their hips and eyes closed. Further, they were instructed that if they moved during the test, to return to the test position as quickly as possible. The participant was positioned 10 to 15 feet away from the tester and video camera. Each participant was videotaped and scored at a later date, which

allowed for repeated and precise scoring. All tapes were scored twice by the primary

investigator. The intratester reliability coefficient was calculated and revealed good intratester

reliability (r = 0.98).

Statistical Analysis

All data analysis was conducted using SPSS 13.0 (SPSS Inc., Chicago, IL.). The

significant levels were set *a priori* at p$\leq$ 0.05. Tests of homogeneity of variance and normality of

score distribution were conducted. Further descriptive statistics were examined, looking for

outliers.

Subjects were divided into groups based upon effort scores. Individuals with suspect

effort (cut score of 20 or less on the Rey 15 item test with Recognition Trial and 14 or less on the

Dot Counting Test). A One-Way ANOVA was run to determine differences between groups for

each dependent measure.

Hypothesis Testing

Hypothesis 1: Sub-optimal effort will be present in at least 10% of the high school athletes.

Hypothesis 1 was tested using a z-test for one sample difference.

Hypothesis 2: There will be differences in total symptoms reported and neuropsychological test

scores between optimal effort and sub-optimal effort during baseline testing.

Hypothesis 2 was tested using a one-way ANOVA design.

Hypothesis 3: Internal consistency reliability of the revised BESS protocol will be demonstrated

by an intraclass reliability coefficient of at least 0.85.

Hypothesis 3 was tested by calculating the intraclass reliability coefficient.

Power Analysis

   The number of subjects necessary was estimated using a significance level set at $\acute{\alpha} = 0.05$, power = 0.80 and a moderate effect size of 0.75 for the repeated measures (three trials) ANOVA design sample size ranged from 11-21 subjects for the selected tests. The number of subjects necessary for the statistical differences between groups, One-way ANOVA, was calculated using a significance level set at $\acute{\alpha} = 0.05$, power = 0.80 and a moderate effect size of 0.75 using procedures from Keppell (1991). A sample size of 26 was estimated. Previous research has found a 15% concussion rate in high school football athletes; therefore, 199 subjects were tested at baseline in hopes of obtaining an sub-optimal effort sample of 21-30 athletes.

CHAPTER 4

THE EFFECT OF EFFORT ON BASELINE NEUROPSYCHOLOGICAL TEST SCORES IN

HIGH SCHOOL FOOTBALL ATHLETES [1]

Abstract

Objective: Sports-related concussion assessment has not examined an individual's effort level during baseline testing. This research examined the level of effort in an athletic population and determined if sub-optimal effort effects neuropsychological test scores. **Design:** 199 high school athletes, mean age (15.54+ 1.17 years) were administered a brief neuropsychological test battery including the Dot Counting Test (DCT) and the Rey 15-item test with recognition trial.

Measurement: One-way analyses of variance were calculated for each outcome measure to examine differences between effort groups. **Results:** 177 athletes exerted optimal effort while 22 exerted sub-optimal effort. Significant differences existed between effort groups on several of the neuropsychological tests. **Conclusions:** Sub-optimal effort does exist in the athletic population during baseline testing with statistically significant differences between groups on neuropsychological test scores. Adding an objective effort test is necessary to improve interpretations of scores within concussion assessment.

Keywords: Effort, concussion, adolescent, neuropsychological tests

Introduction

Neuropsychological testing has been researched and used by trained Neuropsychologists for many years for a variety of medical conditions. The use of these neuropsychological test batteries for the assessment and follow-up evaluation of athletes originated from work in the early 1980's when sports became a laboratory setting for the study of acute mild head injuries (Barth et al., 1989). Neuropsychological testing has gained additional favor as an assessment tool during the evaluation and treatment of concussed athletes with the publication of several mild traumatic brain injury position statements. (Aubry et al., 2002; Guskiewicz et al., 2004; McCrory et al., 2005) Their primary suggestion emphasizes using neuropsychological testing during baseline assessment to attain a premorbid estimate of domains that are likely to be affected by concussion.

Neuropsychological testing is used to provide a sensitive index of higher brain functioning by measuring functions such as memory, attention, and speed of cognitive processing. In sports-related concussion assessment, the most common approach has been to administer a brief battery, less than 40 minutes, which measures functions usually affected by brain injury. These functions include memory, cognitive processing speed, working memory, and/or executive function (Lovell, Iverson, Collin, McKeag, & Maroon, 1999; Macciocchi, Barth, Alves, Rimel, & Jane, 1996; Pellman, Viano, Casson, Arfken, & Powell, 2004). Commonly used assessments include Trail Making Test (Guskiewicz, Riemann, Perrin, & Nashner, 1997; Guskiewicz, Ross, & Marshall, 2001; Macciocchi, Barth, Alves, Rimel, & Jane, 1996), Digit Symbol Substitution Test (Macciocchi et al., 1996; Maddocks & Saling, 1996), Controlled Oral Word Association Test (COWAT) (Guskiewicz et al., 2001), Hopkins Verbal Learning Test (HVLT), Stroop Word Color Test (Guskiewicz et al., 2001), and recently

computerized versions such as ImPact (Collie, Makdissi, Maruff, Bennerll, & McCrory, 2003; Collins et al., 2003). However, there has been no consensus among researchers as to which neuropsychological tests within the battery are the most sensitive in detecting change following concussion (Barr, 2001; Randolph, McCrea, & Barr, 2005).

Research involving collegiate and professional athletes suggests subtle deficits on neuropsychological tests that return to baseline values five to seven days post-injury (Guskiewicz et al., 2001; McCrea et al., 2003). Deficits following injury have been illustrated with increased time to completion for Trails A and B (Guskiewicz & Perrin, 1998; Guskiewicz et al., 2001; Macciocchi et al., 1996; McCrea et al., 2003). Further, decreased word recall for HVLT (Echemendia, Putkian, Mackin, Julian, & Shoss, 2001; Guskiewicz et al., 2001) and lower scores for SDMT (Collins et al., 1999; Echemendia et al., 2001) were obtained for concussed participants. High school athletes on average take longer to recover following concussion (greater than seven days) (Field, Collins, Lovell, & Maroon, 2003; Lovell et al., 2003; Lovell, Collins, Iverson, & Johnston, 2004) than their collegiate counterparts. Studies even suggest cognitive deficits in adolescents for one year following injury (Mueller, 2001) compared to baseline values.

Baseline neuropsychological testing comparisons to pre-injury function are solely based upon the accumulation of valid baseline measures. If baseline scores are questionable, comparisons to premorbid function are not acceptable. Several aspects of test administration can alter effectiveness of neuropsychological test scores. These may include, but are not limited to,: distractions during testing, misunderstood directions, leading the participant to respond, effort etc. However, health care practitioners are unaware as to whether effort exerted during baseline concussion assessment plays a factor in poor test scores.

Initial assessments of effort were subjective clinical impression during neuropsychological evaluations. However, these subjective claims were prone to error due to absence of corrections for administration error or suboptimal effort. (Faust, Hart, Guilmette, & Arkes 1998; Green, Rohling, Lees-Haley, & Allen, 2001a; Green & Iverson, 2001b; Iverson & Binder, 2000; Sweet, 1999) The popularity of forensic neuropsychological practice began in the 1980s in both the civil and criminal realm. As a result, neuropsychologists started using objective effort tests as compensation claims and cases of malingering increased.

Effort is defined as "a usually earnest attempt" or "the use of physical or mental energy to do something". Standardized effort assessments provide an objective method to identify invalid test results. Patients with sub-optimal effort, in general, perform below the usual range for persons who complain of symptoms or disorders on an organic basis (Lezak, Howieson, & Loring, 2004). Additionally, sub-optimal test-taking effort may show up as an absence of practice effects for participants who have demonstrated some learning ability in follow-up testing. Furthermore, unaccountable variations or extreme high and low intratest response patterns are difficult to interpret clinically when effort is poor at baseline testing (Lezak et al., 2004; Green et al., 2001a).

Several brief measures are available to identify sub-optimal effort. These measures are designed to give the illusion that the test progresses in difficulty when the difficulty actually never changes. Participants trying to "fake bad" or demonstrating sub-optimal effort would overestimate poor performance producing scores far beyond the normative values. Popular brief effort measures include the Rey 15 Item Test (Rey) (Boone, Salazarm, Lu, Warner-Chacon, & Razani, 2002a) and the Dot Counting Test (DCT) (Boone et al., 2002b; Frederick, 2002).

The Rey 15-item memory test is a widely used brief test to evaluate gross effort. (Boone et al., 2002a; Frederick, 2002; Philpot, 1993) Participants with sub-optimal effort have been found to deny recall or make more omission errors, while those with traumatic brain injury can report at least nine items with excellent specificity (97-100%), although sensitivity was modest (47%) (Bernard & Fowler; Boone et al., 1995, 1990; Greiffenstein, Baker, & Gola, 1994). With high test specificity, performance below the cut score is indicative of sub-optimal effort. The traditional scoring and interpretation techniques reported excellent specificity but poor sensitivity of the measure (Arnett, Hammeke, & Schwartz, 1995; Griffin, Glassmire, Henderson, & McCann, 1997; Vickery, Berry, Inman, Harris, & Orey, 2001).

In response to previous research reporting poor sensitivity, a recognition trial was added to the Rey 15 Item Memory Test (R15-R) to improve the sensitivity of the measure (Boone et al., 2002a). In addition to the recognition trial, the use of the combination score (i.e., free recall + [recognition - false positives] <20) substantially increased sensitivity (71%) while maintaining high specificity (> or=92%). The lower ranges of specificity and sensitivity came from mental retardation, learning disabled and demented clinical populations.

The DCT has also been used to detect sub-optimal effort. Overall, the DCT is a brief task that assesses lack of effort, either intentional or unintentional, by examining an over-learned skill and measuring time to completion as items become more difficult. Four scores can be obtained, a mean ungrouped dots time, a mean grouped dots time, total numbers of errors and a combination score. The time to count group dots should be faster than ungrouped dots. However, Frederick (2002) found that a faster time to count the ungrouped dots had only 5% sensitivity among simulated malingerers.

Boone et al. (2002b) recently reported increased sensitivity and specificity using the combination score (total time to count ungrouped dots plus time to complete grouped dots plus number of errors committed). Using the combination score, the specificity of the DCT increased (72.1%-86.0%) in all clinical groups except moderate dementia (Boone et al., 2002b). Utilizing the combination score and recognition trial, sensitivity increased to greater than 85% while high specificity was maintained.

There is no previous research in sports-related concussion assessment that has examined whether an individual provided optimal effort during baseline testing. The addition of an objective effort assessment could assist in the interpretation of neuropsychological test scores post-injury to pre-injury baseline scores. The guiding capacity to determine deficits following injury requires that all interpretation of data, both pre-morbid and post-injury be valid. If an athlete does not put forth his/her best effort, baseline scores may be irrelevant. Further, the validity of interpretations following deficits may be questionable when validity of baseline data is subjective.

The purpose of this research was to examine if sub-optimal effort is present during baseline testing in high school athletes. Additional purposes include whether effort effects neuropsychological test scores in high school athletes during baseline concussion testing. We hypothesize that sub-optimal effort will be present in more than ten percent of high school athletes. Further, those participants with sub-optimal effort will perform worse than those participants with optimal effort on baseline neuropsychological test scores.

Procedures

Participants: A total of 206 high school football players (15.54 + 1.17 years) from five Northeast Georgia high schools were tested prior to the 2005 competitive football season. Any

participant that had sustained a musculoskeletal injury or head injury three months prior to testing were excluded from the study. All participants and guardians read and signed the informed consent and assent forms prior to inclusion into the study according to the University of Georgia Institutional Review Board.

Data Collection Procedures

All participants underwent a pre-season baseline evaluation on a brief neuropsychological battery and self-report symptom assessment measures during pre-season. A health questionnaire was also administered at baseline to obtain demographic information, concussion history, pre-existing neurological conditions and other medical conditions. Assessments were conducted by individuals, trained by a Neuropsychologist on test administration, at the participants' high school in an isolated room. All tests were scored by the primary investigator.

Main Outcome Measures

Brief Neuropsychological Battery

A brief neuropsychological battery incorporating tests that have been extensively studied in sport-related concussion and neuropsychological literature was administered to each participant (Collins et al., 1999; Field et al. 2003; Green et al., 2001a; Guskiewicz et al., 2001; McCrea et al., 2003; Pellman et al., 2004 Rohling, 2003 et al) (See Table 1). Participants were administered the tests in the same order to decrease test interference. The testing order consisted of: the Self-Report Symptom Checklist, Standardized Assessment of Concussion, R15-R test, Hopkins Verbal Learning Test, Finger Tapping Test, Trails Making Test A, Trails Making Test B, Symbol Digit Modalities Test, Digit Span Test, Dot Counting Test, and Hopkins Verbal Learning Test delayed recall and recognition.

Effort Tests

Rey 15-Item Test with Recognition Trial (Boone et al., 2002b).

This test consists of two pages. The first page contains 15 items arranged in five rows. The recognition task consists of a page containing 15 items from the original stimulus page interspersed with 15 spoils similar to the target items. Participants are exposed to the stimulus for 10 seconds and the stimulus is withdrawn.   Then 10-15 seconds elapse and the participant is asked to recall the characters. The word 15 is emphasized in the instructions. When the participants finish, the reproductions are removed and the recognition form placed in front of them with the instructions "On this page are 15 things I showed you as well as 15 items that were not on the page. I want you to circle the things you remember from the page I showed you". Scores used in this analysis were the combination score derived by recall correct + (recognition correct-false positives). The participant who consciously or unconsciously wishes to appear impaired will fail at a task that all but the most severely brain injured or mentally retarded participants perform easily. A person who cannot recall and recognize for a total score of 20 would be considered to have sub-optimal effort (Boone, 2002). Specificity of a cut score of 20 or less ranged from 61.1% to 100% while sensitivity of the measure ranged from 61.2% to 77.6% while administration takes less than five minutes.

The Dot Counting Test (Boone et al., 2002a)

The DCT is a brief task that assesses lack of effort either intentional (malingering) or unintentional (unconscious) by examining an over-learned skill and by measuring time to completion as items become more difficult. The test consists of six serially numbered 3x5 cards with randomly arranged dots. The cards are shown to the participant one at a time. The participant is told to count the number of dots as quickly as possible and tell the administrator the

answer. The participant's time for each card is recorded on the answer sheet. The average time for the ungrouped dots, average time for the grouped dots, and the total number of errors are recorded on the answer sheet. The participant's time will increase gradually with the increased number of dots. A cut score of 14 or less for optimal controls establishes sensitivity of 92% and specificity of 99%. Further, a brief administration time less than 10 minutes is adequate for inclusion into a short neuropsychological test battery.

Statistical Analysis

Data was initially examined for outliers. Neuropsychological test data that were greater than two standard deviations from the sample mean on a multiple test within the battery were considered outliers and participants with outlier scores were removed. Seven participants were eliminated from the analysis and the remaining 199 participants were used in the statistical analyses. Two measurements (R15-R and DCT) were used to assess effort. Each test was administered individually and analyzed separately.

Baseline scores on the R15-R and the DCT were examined. A cut score for sub-optimal effort for the R15-R was set as a combination score less than 20 (Boone et al., 2002)**,** while a cut score for the DCT was set at 14 or less (Boone et al., 2000). The remaining 199 participants were divided into groups based on optimal and suboptimal effort scores.

All statistical techniques were run using SPSS 13.0 (SPSS Inc., Chicago, IL). To examine differences between groups, a One-way ANOVA was calculated for each of the outcome measures.

Results

Collins et al.(1999) suggested differences at baseline for those with a previous history of concussion and learning disabilities. Therefore, additional analyses were completed on these

68

subgroups. Thirty-six percent of the learning-disabled group demonstrated sub-optimal effort, while 17.6% of the participants with previous history of concussion demonstrated sub-optimal effort (Figure 1). A One-way ANOVA was calculated. There were no statistically significant differences between those with sub-optimal effort in the learning disabled, previous history of concussion, and the optimal group. With no difference between groups, the learning disabled and previous history of concussion group were included into the full analysis between those with sub-optimal and optimal effort (See Table 4.2).

Eleven percent of the total sample had sub-optimal effort on the DCT while three percent of the total sample demonstrated sub-optimal effort on the Rey 15-Item Test with recognition trial. For each test, the athletes were divided into two groups (optimal effort and sub-optimal effort) based upon suggested cut scores. Optimal effort (n= 177) and sub-optimal effort (n=22) was found for the DCT while optimal effort (n= 195) and sub-optimal effort (n= 4) occurred for the R15-R.

Less than three percent of the participants had sub-optimal effort using the R15-R. Due to the small sample size for the R15-R sub-optimal effort group, traditional assumptions in statistical analysis were violated. Therefore, only evaluation of test means were completed (See Table 4.3). Group means displayed a general trend for the sub-optimal effort group to perform worse than the optimal effort group.

Significant differences existed between groups formed using the DCT scores on several neuropsychological tests which include Trails B ($F_{(1, 196)}$, =14.31, $p<.01$); SDMT, ($F_{(1, 196)}$, = 10.112 $p<.01$); DS_{tot} ($F_{(1, 196)}$, = 12.48, $p<.01$); DS for ($F_{(1, 196)}$, = 5.101, $p= 0.02$); DS back ($F_{(1, 196)}$, =14.45, $p<.01$); HVLT ($F_{(1, 196)}$, = 6.96, $p= 0.02$); Self-Report Symptoms ($F_{(1, 196)}$, = 4.143,

p= 0.04); SAC ($F_{(1, 196)}$, =17.918, p<.01); Dominant Hand Tapping Test ($F_{(1, 196)}$, = 5.319, p=0.02).

Discussion

Eleven percent of the total sample (n=22) demonstrated sub-optimal effort on the DCT during baseline testing. The assumption that high school athletes exert optimal effort during all assessments is incorrect. The presence of athletes with sub-optimal effort justifies that an objective effort test be included in concussion assessment to identify those with invalid test scores. This study utilized brief measures that only identify gross effort. Although eleven percent may appear menial, this research suggests that one out of ten athletes has sub-optimal effort during baseline testing.

Additional examination of whether sub-optimal effort decreases neuropsychological test scores was examined. Statistically significant differences existed between those with sub-optimal and optimal effort on neuropsychological test scores. Participants that exhibit sub-optimal effort using an objective effort test have shown invalid neuropsychological baseline test scores. (Constantinou, Bauer, Ashendorf, Fisher, & McCaffrey, 2005; Green et al., 2001) If sub-optimal effort is present, baseline scores may be inaccurate and resultant data should be ignored. Our participants with sub-optimal effort appeared to perform below the optimal effort group on tests measuring information processing, memory, attention, concentration, learning, and motoric speed.

Previous history of concussion and/or self report of learning disabilities samples were not statistically significantly different from the normal sample in regards to effort assessment. There was a trend for those with both learning disabilities and previous history of concussion to perform more poorly on the neuropsychological tests compared to the other groups.

Although the percentages of those with sub-optimal effort were twice the rate of those with no LD and PHC (36.4% and 17% respectively), the small number of participants within each group may account for the lack of statistical significance. However, the higher percentage of those with sub-optimal effort in the learning disabled and previous history of concussion group requires the need for careful attention of these groups while administering baseline neuropsychological tests. Careful observation and evaluation of athletes with these conditions are recommended within concussion evaluation to obtain valid test scores, especially at baseline.

Concussion assessment tests focus on enhancing the discrete deficits involving cognitive domains, which included processing speed and memory. Typical concussion assessment involves serial testing, forcing the potential for at least three to seven testing sessions in a short period of time. Previous research has been completed examining practice effects during testing (Barr, 2001). However, subjective opinions during serial testing of athletes suggest athletes become bored and do not put forth high effort during multiple test sessions. Testing protocols may create situations for poor or lackadaisical performance due to boredom or poor effort. The addition of an objective measure for effort within a concussion assessment test battery will validate neuropsychological test scores.

The assumption that an athlete's desire and motivation to perform well at all times, including a concussion assessment, has not been proven. Athletes do not respond in a socially desirable manner (Piland, Motl, Ferrara, & Peterson, 2003); however, effort has never been assessed in the literature. The inability to determine whether optimal effort was exerted is especially difficult when clinicians are not expecting participants to perform poorly (Faust et al., 1988). The expectation of optimal effort could impair the clinician's judgment regarding the interpretation of neuropsychological test scores. The addition of procedures to assess effort

might help to correctly identify sub-optimal effort and increase the validity of interpretations regarding baseline neuropsychological test scores.

Subjective opinions of researchers describe multiple reasons for poor effort during testing. Athletes with prior history of concussion assessment, or who are knowledgeable about the purpose of baseline testing, may perform poorly to mask any cognitive changes attributed to a concussive injury. Their awareness provides the opportunity to exert suboptimal effort on baseline tests, so post-concussion deficits do not appear as dramatic. Further, obtaining reliable and valid baseline data is potentially confounded by mental or physical fatigue, illness, and poor test environment which might result in poor baseline scores. Another required meeting or test when the athlete may be tired would prove motive for sub-optimal effort and lack of motivation to perform well on the test. Since athletic trainers typically assume that athletes are exerting optimal effort, having an objective measure to determine effort is a necessary gauge to understand potentially contradictory information. Effort assessments have always been suggested in neuropsychology; however, this practice has not been developed and included in concussion assessment for athletic populations.

This study contained several limitations that need to be examined in the future. The number of participants who had sub-optimal effort was not equally distributed between the two effort tests. First, the Rey 15-Item Test with recognition trial has not been shown to be very sensitive to suboptimal effort. The measures in this research were used due to vast use in previous effort assessment research the administration techniques (traditional neuropsychological testing), and they could easily be administered within the protocol battery in a short period of time (brief effort test). The determination of sub-optimal effort utilizing these brief gross measures of effort would indicate conservative estimates of suboptimal effort. Further,

additional participants with sub-optimal effort may not have been captured using these tools. More sensitive tests that are short and highly specific would allow for valid interpretations of test scores.

Second, administration times within the battery may have affected outcome. Potentially, a participant's effort level may change within the battery administration time. Therefore, by the end of the battery (when the DCT was administered) effort levels may have decreased. Counter-balancing the test order would eliminate testing effects.

Third, the suggestion of using a cut score for the DCT of 17 (Boone et al., 2002b) was excluded because this cut score was based on clinical populations. Our baseline testing involves relatively heterogenous sample of high school athletes without serious clinical impairment diagnoses. Therefore, a cut score of 14 which was recommended by the test developers for non-clinical samples was incorporated into the test protocol. Determining an appropriate age specific cut score for effort is suggested.

This research assessed two hundred athletes. Obtaining a higher number of athletes with sub-optimal effort might ensure that the findings are reproducible. Further, examining sub-optimal and optimal effort with the inclusion of injured athletes to determine sensitivity and specificity of the measures within an athletic population would aid in decisions regarding which effort test to use.

With the recommendation of baseline testing for concussion assessment, obtaining valid baseline scores has never been so important (Aubry et al., 2002; Guskiewicz et al., 2004). If the athletes' baseline scores are not interpreted correctly due to sub-optimal effort, the accuracy of pre-morbid function and recovery from injury would be questionable. Further, direct comparison to an athletes' baseline might indicate an improvement following post-concussion injuries when

no change is present. Thus, the opportunity to make proper decisions regarding post concussion assessment would be difficult.

The use of neuropsychological tests has been advocated as a baseline measure. Research involving neuropsychological testing without an objective measure of effort might produce invalid interpretations regarding comparisons to baseline scores. Obtaining an objective measure of effort would allow the removal of acceptance of invalid baseline tests to increase the accuracy of scores within concussion assessment.

It is always assumed that athletes provide maximal effort during any testing session. If eleven percent of the sample has sub-optimal effort, eleven percent of the sample has invalid neuropsychological test scores. Any comparisons made on invalid data would be interpreted incorrectly. Researchers should include an objective effort test to validate interpretations of findings during concussion assessment.

Although additional effort measures are available, these brief gross measures of effort should be used as a screening test. If an athlete exerts sub-optimal effort using these brief measures of effort, utilization of a more comprehensive effort battery is suggested. Further, if sub-optimal effort is present during baseline neuropsychological testing, we recommend administering the test battery again when an objective effort test reveals optimal effort to validate baseline scores.

Further studies need to replicate the presence of sub-optimal effort in different athletic populations. Increasing the sample sizes to include clinical samples, regional and/or national samples will strengthen the findings of this study. Moreover, determining an effort test that is sensitive, specific, and time relevant to an athletic population would aid in the interpretation of the test scores.

References

Arnett, P.A., Hammeke, T.A., Schwartz, L. (1995). Quantitative and qualitative performance on Rey's 15-item test in neurological patients and dissimulators. *The Clinical Neuropsychologist*, 9, 17-26.

Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., Lovell, M, McCrory, P., Meeuwise, W., Schamasch, P. (2001). Summary and agreement statement of the First International Conference on Concussion in Sport, Vienna 2001. Recommendations for the improvement of safety and health of athletes who may suffer concussive injuries. *British Journal of Sports Medicine*, 36(1), 6-10.

Barr, W. (2001). Methodological Issues in Neuropsychological Testing. *Journal of Athletic Training,* 36(3), 297-302.

Barth, J., Alves, W., Ryan, T., Macciocchi, S., Rimel, R., Jane, J., Nelson, W. (1989). *Mild Head Injury in Sports: Neuropsychological Sequelae and Recovery of Function.* In: Levin, H., Eisenber, H., Benton, A. eds. Mild Head Injury. New York, New York: Oxford University Press, 257-275.

Beetar, J.T, Williams, M. (1995). Malingering Response Styles on the Memory Assessment Scales and Symptom Validity Tests. *Archives of Clinical Neuropsychology,* 10 (1), 57-72.

Bernard, L.C., Fowler, W. (1990). Assessing the validity of memory complaints: Performance of brain-damaged and normal individuals on Rey's task to detect malingering. *Journal of Clinical Psychology*, 46, 432-436.

Boone, K.B, Savodnik, I., Ghaffarian, S., Lee, A., Freeman, D. Berman, N. (1995). Rey 15-item

Memorization and Dot Counting Scores in a "stress" claim worker's compensation

population: Relationship to personality (MCMI) scores. *Journal of Clinical psychology,*

51, 457-463.

Boone, K.B, Salazarm, X., Lu, P., Warner-Chacon, K., Razani, J. (2002). The Rey 15-Item

Recognition Trial: A Technique to Enhance Sensitivity of the Rey 15-Item Memorization

Test. *Journal of Clinical and Experimental Neuropsychology*, 24(5), 561-573.

Boone, K.B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., Shamieh, E., Warner-Chacon, K.

(2002). Sensitivity and specificity of the Rey Dot Counting Test in patients with sub-

optimal effort and carious clinical samples. *Archives of Clinical Neuropsychology*, 17,

625-642.

Boone, K.B., Lu., P., Sherman, D., Palmer, B., Back, C., Shamieh, E., Warner-Chacon, K.,

Berman, N. (2000). Validation of a new technique to detect malingering of cognitive

symptoms: The b Test. *Archives of Clinical Neuropsychology,* 15, 227-241.

Collins, M.W., Grindel, S.H., Lovell, M.R., Dede, D.E., Moser, D.J., Phalin, B.R., Nogle, S.,

Wasik, M., Cordry, D., Daugherty, K.M., Sears, S.F., Nicolette, G., Indelicato, P.

McKeag, D.B. (1999). Relationship between concussion and neuropsychological

performance in college football players. *JAMA,* 282(10), 964-70.

Collins, M.W., Iverson, G.L., Lovell, M.R., McKeag, D.B., Norwig, J., Maroon, J. (2003). On-

field predictors of neuropsychological and symptom deficit following sports-related

concussion. *Clinical Journal of Sports Medicine,* 13(4), 222-9.

Constantinou, M., Bauer, L., Ashendorf, L., Fisher, J.M., McCaffrey, R.J. (2005). Is poor

performance on recognition memory effort measures indicative of generalized poor

performance on neuropsychological tests? *Archives of Clinical Neuropsychology*, 20,

191-198.

Echemendia, R.J., Putukian, M., Mackin, R.S., Julian, L., Shoss, N. (2001). Neuropsychological

test performance prior to and following sports-related mild traumatic brain injury.

Clinical Journal of Sports Medicine, 11(1), 23-31.

Faust, D., Hart, K., Guilmette, T., Arkes, H. (1988). Neuropsychologists' capacity to detect

adolescent malingerers. *Professional Psychology: Research and Practice,* 19(5), 508-

515.

Faust D., Hart, K., Guilmette, T.J., Arkes, H.R. (1998). Neuropsychologists' Capacity to Detect

Adolescent Malingerers. *Professional Psychology: Research and Practice,* 19(5), 508-

515.

Field, M., Collins, M.W., Lovell, M.R., Maroon, J. (2003). Does age play a role in recovery from

sports-related concussion? A comparison of high school and collegiate athletes. *Journal

of Pediatrics,* 142(5), 546-53.

Flaro, L., Green, P. (2000). Symptom Validity test results with children: CARB and WMT.

Archives of Clinical Neuropsychology, 15, 840.

Frederick, R.I. (2002).A Review of Rey's Strategies for Detecting Malingered

Neuropsychological Impairment. *Journal of Forensic Neuropsychology*, 2(3), 1-22.

Gervaus, R.O., Rohling, M.L, Green, P., Ford, W. (2004). A comparison of WMT, CARB, and

TOMM failure rates in non-head injured disability claimants. *Archives of Clinical

Neuropsychology,* 19, 475-487.

Goldberg, T.O., Miller, H.R. (1986). Performance of psychiatric inpatients and intellectually

deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology,* 42, 792-795.

Green, P., Rohling, M.L, Lees-Haley, P.R., Allen, L.M. (2001). Effort has a greater effect on test

scores than severe brain injury in compensation claimants. *Brain Injury*, 15(12), 1045-1060.

Green, P., Iverson, G.L. (2001). Effects of injury severity and cognitive exaggeration on

olfactory deficits in head injury compensation claims. *NeuroRehabilitation*, 16, 237-243.

Griffin, G.A., Glassmire, D.M., Henderson, E.A., McCann, C. (1997). Rey II: Redesigning the

Rey Screening Test of malingering. *Journal of Clinical Psychology,* 53, 757-766

Greiffenstein, M., Baker, W., & Gola, T. (1994). Validation of malingered amnesia measures

with a large clinical sample. *Psychological Assessment,* 6, 218–224.

Guskiewicz, K.M., Ross, S.E., Marshall, S.W. (2001). Postural Stability and Neuropsychological

Deficits After Concussion in Collegiate Athletes. *Journal of Athletic Training*, 36(3), 263-27 3.

Guskiewicz, K., Bruce, S., Cantu, R., Ferrara, M., Kelly, J., McCrea, M., Putukian M., Valovich

McLeod, T. (2004). National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion. *Journal of Athletic Training,* 39, 280-297.

Guskiewicz, K.M., Riemann, B.L., Perrin, D.H, Nashner, L.M. (1997). Alternative approaches to

the assessment of mild head injury in athletes. *Medicine and Science in Sport and Exercise*. 7(supplement): S213-221.

Iverson, G.L., Binder, L.M. (2000). Detecting exaggeration and malingering in

neuropsychological assessment. *Journal of Head Trauma Rehabilitation*, 15, 829-858.

Lezak, M. D., Howieson, D.B., Loring, D.W. (2004). *Neuropsychological Assessment.* Oxford, New York: Oxford Press.

Lovell, M.R., Iverson, G.L., Collins, M.W., McKeag, D., Maroon, J.C. (1999). Does loss of consciousness predict neuropsychological decrements after concussion? *Clinical Journal of Sports Medicine,* 9(4), 193-8.

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

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

Lynch, W. (2004). Determination of Effort Level, Exaggeration, and Malingering in Neurocognitive Assessment. *Journal of Head Trauma Rehabilitation*, 19(3), 277-283.

Macciocchi, S.N., Barth, J.T., Alves, W., Rimel, R.W., Jane, J.A. (1996). Neuropsychological functioning and recovery after mild head injury in collegiate athletes. *Neurosurgery.* 39(3), 510-4.

Maddocks, D., Saling M. (1996). Neuropsychological deficits following concussion. *Brain Injury*, 10(2), 99-103.

Moss, A., Jones, C., Fokias, D., Quinn, D. (2003). The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury,* 17 (5), 377-387.

McCrea, M., Guskiewicz, K.M., Marshall, S.W., Barr, W., Randolph, C., Cantu, R.C., Onate, J.A., Yang, J., Kelly, J.P. (2003) Acute effects and recovery time following concussion in collegiate football players: the NCAA Concussion Study. *JAMA*, 290(19), 2556-63.

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

Mueller, F. (2001). Catastrophic head injuries in high school and collegiate sports. *Journal of Athletic Training,* 36(3), 307-311.

Pellman, E.J., Lovell, M.R., Viano, D.C., Casson, I.R., Tucker, A.M. (2004). Concussion in professional football: neuropsychological testing--part 6. *Neurosurgery,* 55(6), 1290-303

Philpot, L.M. (1993). The effects of severity of cognitive impairment and age on two malingering tests: An investigation of the Rey Memory Test and Dot Counting Test in Alzheimers patients and optimal middle-aged and older adults. *Dissertation Abstracts International,* 53, 5001.

Piland, S.G., Motl, R.W., Ferrara, M.S., Peterson, C.L. (2003). Evidence for the factorial and construct validity of a self-report concussion symptom scale. *Journal of Athletic Training,* 38(2), 104-112

Randolph, C., McCrea, M., Barr, W. (2005). Is neuropsychological testing useful in the management of sport-related concussion? *Journal of Athletic Training*, 40(3), 139-152.

Rohling, M.L, Meyers, J.E., Millis, S.R. (2003). Neuropsychological Impairment Following Traumatic Brain Injury: A Dose-Response Analysis. *The Clinical Neuropsychologist*, 17(3), 289-302.

Rohling, M.L., Green, P., Allen, L.M., Iverson, G.I. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology,* 17, 205-222.

Ross, S.R., Krukowski, R.A., Putnam, S.H., Adams, K.M. (2001). The Memory Assessment Scales in the Detection of Incomplete Effort in Mild Head Injury. *The Clinical Neuropsychologist,* 17(4), 581-591.

Ruff, R.M., Wylie, T., Tennant, W. (1993). Malingering and malingering-like aspects of mild closed head injury. *Journal of Head Trauma Rehabilitation,* 8(3), 60-73.

Schretlen, D., Brandt, J., Krafft, L., Van Gorp, W. (1991). Some caveats in using the Rey 15-item memory test to detect malingered amnesia. *Psychological Assessment: A Journal of Consulting and Clinical Psychology,* 3, 667-672.

Sherman, D.S., Boone, K.B., Razan, J. (2002). Re-examination of a Rey Auditory Verbal Learning Test/Rey Complex Figure Discriminate Function to Detect Suspect Effort. *The Clinical Neuropsychologist,* 16(3), 242-250.

Sweet, J.J. (1999). *Malingering differential diagnosis.* In: J.J. Sweet (editor) Forensic Neuropsychology Fundamentals and Practice (Lisse: Swets and Zeitlinge).

Vickery, C.D., Berry, D.T, Inman, T.H., Harris, M.J, Orey, S.A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology,* 16, 45-73.

Youngjohn, J., Burrows, L., Erdal, K. (1995). Brain damage or compensation neurosis? The controversial post concussion syndrome. *The Clinical Neuropsychologist*, 9, 112-123.

Table 4.1: Neuropsychological Test Domains, Descriptions, Scoring and Administration Time

Measure	Functional Domain	Description	Score Range	Time needed to administer
Self-Report Symptom Checklist	Concussion related symptoms	Self-rated duration and severity of 22 symptoms	Total score duration and severity. Likert scale 0 (no severity and duration) to 6 (severe); higher score indicates more severe symptoms	2-3 minutes
Standardized Assessment of Concussion	Cognitive Function (orientation, immediate and delayed memory, and concentration	Brief neurocognitive assessment and neurological screening	Total score range, 0-30, lower score indicates more severe cognitive impairment	5 minutes
Neuropsychological Test Battery	Cognitive functioning (attention,	Hopkins Verbal Learning Test	Total score ranged for each individual measure; lower	20 minutes

	concentration, processing speed, motoric speed, anterograde memory, verbal memory.	(memory), Finger Tapping (motoric speed) Test, Trails A (cognitive processing), Trails B (cognitive processing), Symbol Digit Modalities Test (cognitive processing), Digit Span (attention)	score indicates more sever impairment except for trails (time to completion)	
Dot Counting Test Rey-15 item	Effort	Rey 15-Item Test with Recognition Trial, Dot Counting Test	Rey total score to 27. Dot Counting Test range	5 minutes

Table 4.2: Comparison of Neuropsychological Test Means for Optimal Effort Group, Learning Disabled (LD), with Sub-optimal Effort and Previous History of Concussion (PHC) with Sub-optimal Effort.

Tests	Sub-optimal effort group (n=22)	No PHC/ LD (n=9)	PHC (n= 5)	LD Sub-optimal effort (n=4)	PHC and LD (n=4)
Trails A	39.79+11.44	33.71 +6.59	46.72 +12.67	42.88 + 14.03	48.71+20.41
Trails B	84.21+ 15.21	86.22 + 17.75	83.12+14.85	81.81 + 9.10	83.94 + 1.17
SDMT	45.86+ 11.23	46.66 + 8.23	46.33 + 11.59	45.37 + 14.58	44.67 + 13.05
DS total	12.22 + 3.20	12.66 + 3.64	11.33 + 3.39	12.75 + 3.32	11.67 + 5.03
DS forward	7.54 + 1.96	7.88 + 2.02	6.88 + 2.14	7.75 + 2.25	7.00 + 3.00
DS backwards	4.68 + 1.46	4.77 + 1.78	4.44 + 1.51	5.00 + 1.19	4.67 + 2.08
Non-dominant Tap Test	49.99 + 6.92	52.91 + 6.94	49.02 + 5.04	46.65 + 10.15	47.8 + 4.66
Dom. Tap Test	53.58 + 6.21	55.64 + 6.10	52.44 + 5.49	50.30 + 9.51	48.93 + 2.61
HVLT tot	21.63 + 4.89	23.88 + 4.25	20.11 + 4.83	19.00 + 5.07	19.00 + 3.60
HVLT delayed	7.68 + 2.31	8.88 + 1.96	6.67 + 2.06	6.87 + 2.58	6.67 + 2.08
HVLT recognition	10.77 + 3.02	12.11 + 4.04	9.78 + 1.56	9.25 + 2.25	8.67 + 1.53
SAC	24.54 + 2.38	25.44 + 1.33	23.88 + 3.14	24.37 + 2.55	24.00 + 3.46
Symptoms	3.5 + 3.21	2.44 + 3.16	3.67 + 2.96	5.87 + 2.64	6.00 + 1.00

Table 4.3: Descriptive Data for Optimal and Sub-optimal Effort Group Using R15-R.

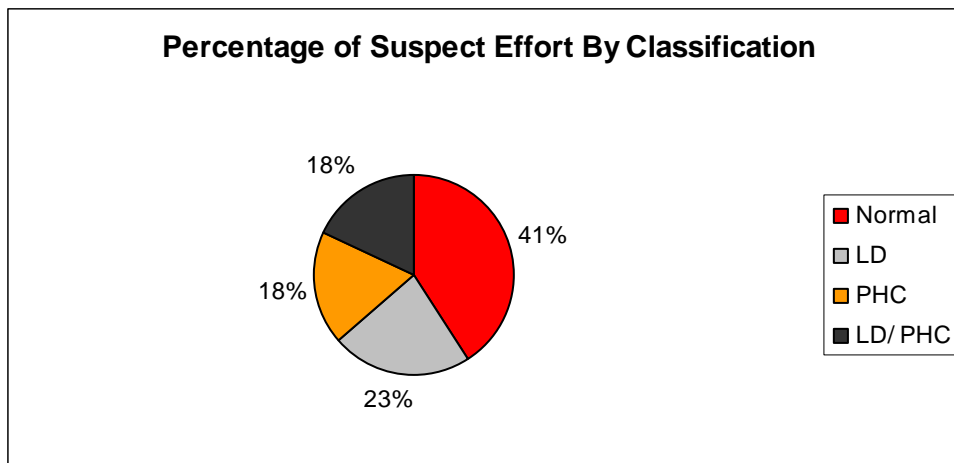| Test	Optimal Effort Mean + SD	Sub-optimal Effort Mean + SD
Self Report Symptoms	2.21 + 2.89	3.00 + 3.56
SAC Total	26.24 + 2.09	25.25 + 0.50
HVLT total	23.76 + 4.31	20.25 + 6.18
HVLT delayed	8.31 + 2.09	7.00 + 2.94
HVLT recognition	10.91 + 1.51	10.25 + 2.22
Dominant Tapping Test	56.89 + 8.22	58.50 + 3.89
Non-dominant Tapping Test	51.52 + 7.10	55.65 + 8.59
Trails A	34.60 + 11.03	40.09 + 9.65
Trails B	69.58 + 19.52	90.79 + 34.67
SDMT	52.26 + 9.63	42.50 + 13.08
DS total	14.63 + 3.54	13.75 + 2.21
DS forward	8.31 + 1.82	8.50 + 1.73
DS Backwards	6.31 + 2.19	5.25 + 0.96

Figure 4.1- Percentage of Sub-optimal Effort by Classification

Learning Disabled (LD), Previous History of Concussion (PHC)

CHAPTER 5

THE RELIABILITY OF THE BALANCE ERROR SCORING SYSTEM: A REVISED

PROTOCOL [1]

Abstract

Objective: The Balance Error Scoring System (BESS) is considered a part of the standardized assessment of concussion. The BESS has shown to be sensitive to the effects of concussion. However, the BESS has not undergone extensive psychometric testing to examine the optimal protocol. The purpose of this study was to determine the reliability of the revised BESS protocol. **Design and Setting:** The BESS consists of three trials on two surfaces (firm and foam) and two stances (single leg stance, tandem stance). Total number of errors was calculated for each condition. **Participants** 144 high school athletes with a mean age of 15.57 ± 1.15 years from six high schools. **Measurement:** An intraclass reliability coefficient and a repeated measures analysis of variance were calculated using SPSS 13.0. **Results:** The mean number of errors for each trial was 6.69 ± 0.36. Statistically significant differences existed between trial one and trial two ($F_{(1.65, 286)} = 4.890$, p=.013). Further, we found an intraclass reliability coefficient of 0.88 for three trials of four conditions, 0.84 for two trials of four conditions, and 0.73 for one trial of four conditions. **Conclusions:** The revised protocol increased the reliability of the BESS; however, the presence of a practice effect suggests allowing additional familiarization trials prior to administration. Valid interpretations of test scores cannot be accomplished without a reliable measure. Psychometrically sound instruments support the ability to make and interpret clinical decisions regarding injury and return to participation. Further analysis of the protocol should be performed on concussed participants to determine the sensitivity and specificity of the protocol.

Key Words: Psychometric Properties, BESS, Youth, Reliability

Introduction

Validity evidence is collected concerning the goodness of inferences made from test scores. Validation is an ongoing process which involves continuous revisions of a test to ensure that the test is psychometrically sound. The process of validating interpretations from test scores must include both reliability and validity studies. This process is, however, often time consuming and tedious.

Most measurement tools in athletic training attempt to identify the location of individuals on a quantitative continuum with respect to a particular construct. This is considered subject-centered measurement. Crocker and Algina described a ten-step process that should be followed in test construction for subject–centered measurement.[1] Each step defines, creates, alters or revises items within a test protocol. During test construction, each step should be performed to determine the test items to optimize test psychometrics and interpretations of scores. During the revision process, the researcher can omit or combine steps that were previously defined. Additionally, the revision process can identify cases where revision might be unnecessary.

The first step is to **identify the primary purpose for which the test score is used.** A variety of tools have been developed to measure postural stability. Devices such as force platforms and motion analysis equipment (i.e. The Neurocom Smart Balance Master) are used with objective postural stability tests. These devices, however, are expensive and their costs exceed the budgetary constraints of the clinical athletic trainer.

Prior to the use of computerized force platform systems, the Rhomberg test was commonly used to evaluate postural stability following injury. The purpose of the Rhomberg test was to subjectively evaluate somatosensory impairment by having the patient close his or her eyes to eliminate vision as a sensory source.[2-4] Since the inception of the test, there have been

various interpretations and modifications of the test to increase the sensitivity and specificity for balance disorders. Several judgment criteria were established, such as time to first touch down, or compensatory event.[3,5] Most modifications have occurred in foot position (stance) to alter the participant's base of support. However, each modification of the original Rhomberg test was criticized for the lack of sensitivity in situations where postural sway increases without the complete loss of postural stability within clinical populations, such as concussed athletes. Continued modification of the Rhomberg test resulted in the development of the BESS.

The BESS was developed as an objective assessment tool to be used by clinicians for the evaluation of postural stability following concussion. The BESS is a modified static protocol designed to evaluate postural stability in concussed athletes, with better sensitivity and specificity than the original Rhomberg test.[5-7]

The second step is to **identify behaviors to represent the construct.** The construct of postural stability is comprised of the postural control system. The postural control system uses three sensory systems to operate a feedback control circuit between the brain and the musculoskeletal system. Under normal circumstances, postural stability is maintained by integrating the somatosensory, visual, and vestibular systems. If one system is not operating correctly, the other systems should compensate.[5, 8-9]

Following a concussive injury, the participant has difficultly integrating information from the three components of the balance mechanism. Concussive trauma has been shown to lead to changes in the normal combination of sensory cues. Therefore, the person must recognize the pattern of remaining sensory cues that establish postural control.[10] While the somatosensory aspect appears to remain normal, integration between the visual and vestibular components does not function properly.[12]

The third step is to **prepare a set of test specifications, delineating the proportion of items that should focus on each aspect of the construct.** Riemann described the traditional protocol for the BESS by placing the subject in three stances on two surfaces. The BESS progresses in difficulty by placing the participant on a firm surface using a double leg stance, single leg stance and tandem leg stance. The more narrow the base of support, the more difficult it is to maintain postural control [4,8,11]. The difficulty is further increased by having the participant repeat the stances on a foam surface which creates an unstable surface. Each stance is performed for twenty seconds. The athletes are instructed to close their eyes and keep their hands on their iliac crests. Closing the eyes eliminates vision as a sensory source and creates a significant increase in sway in patients with cerebral or vestibular dysfunction.[2-4, 8] The number of errors committed in each condition (the stance on the surface) is representative of the ability to integrate the balance components eliminating the visual component.

The fourth step is to **construct an initial pool of items** and the fifth step is to **have items reviewed and revised.** A total of six conditions were determined to be the best protocol. The six conditions included one trial of the double leg stance on a firm surface, double leg stance on a foam surface, single leg stance on a firm surface, single leg stance on a foam surface, tandem stance on a firm surface and tandem leg stance on a foam surface.

Once the items (stances and surfaces) were determined, the protocol was reviewed by other researchers. Utilizing the six conditions, low scores on the Sensory Organization test paralleled high error scores found with the BESS.[4] Thus, the BESS was sensitive to acute postural stability abnormalities after concussion. Moreover, deficits on postural stability utilizing the BESS were found up to seven days post-injury.[12,13]

However, the number of errors committed, the reliability of the BESS, and thus the validity of the interpretations from the scores are dependent upon several previously investigated conditions. Administration of multiple trials of the BESS results in practice effects, with the number of errors decreasing with each consecutive trial.[14-16] Further, the effects of fatigue increase the number of errors acutely, but the athlete recovered after 20 minutes of rest following an exercise session.[17-18]

Determining the statistical properties of item scores and, when appropriated, eliminating items that do not meet pre-established criteria is the sixth step. Using the traditional protocol, Riemann found intratest correlation reliabilities for each condition within the BESS compared to long force plate sway measures ranging from 0.78-0.96 dependent upon the stance.[4] Interestingly, the double leg stance on the firm surface was the only condition which did not significantly correlate to long force plate measures. Objectivity of the investigator's ability to score repeated observations for the same BESS test ranged from 0.87 to 0.98.[16] Additionally, test-retest reliability (0.673) of the BESS has been reported,[14] while intraclass reliability scores for the six conditions have been reported as 0.60.[19] The reported reliability scores are from poor to good dependent upon what type of reliability (test-retest, intraclass, objectivity etc.) is being estimated. However, poor reliabilities scores decrease the ability to obtain valid interpretations regarding the postural stability of the participant.

The seventh and eighth steps are to **examine preliminary item and revise if necessary.** While the ninth step is to **field-test the conditions on a large sample representative of the population for whom the test is intended.** These three steps are typically performed simultaneously. All items are tested, typically, in a large population and revised if psychometric properties are not adequate. Although significant research has been conducted identifying

various reliabilities for the measure, most researchers have not found much difference among subjects for the double leg stance.[4, 13, 15] Riemann did not find significant differences between injured and control participants using the double leg stance on either surface, or sufficient variance associated with the stance to complete statistical analyses.[4] Additionally, the number of errors performed during the double leg stance does not increase like the other stances after exertion, during fatigue.[17-18]

Variance associated with the double leg stance on both the firm and foam surfaces was low in collegiate [4, 17] and high school populations.[16, 19] When variance for an item is low, it does not add any additional information within the test and when removed should increase the reliability of the measure.

Validation of a test protocol is an ongoing process. The purpose of this research was to determine if modifying the BESS protocol would increase the reliability of the measure. We hypothesize that removal of the double leg stance will increase the reliability of the measure to greater than 0.85 in a high school population.

Procedures

Participants

One hundred and forty-four high school football athletes (mean age 15.57 ± 1.15 years) were tested prior to the fall competitive season. Any participant that had sustained a lower extremity musculoskeletal injury and/or head injury within the three months prior to testing was excluded from the study. All participants and guardians read and signed the informed consent and assent forms prior to inclusion into the study as required by the University of Georgia Institutional Review Board.

Balance Error Scoring System

The protocol utilized for BESS administration was changed from the traditional protocol [4, 12, 16] based on our previous findings.[19] The double leg stance on both the firm and foam surfaces was removed. The revised protocol was comprised of four conditions, single leg and tandem leg stances on firm and foam surfaces. Completion of four conditions was considered a trial. Each participant completed three trials of the four conditions. The single leg stance was performed on the nondominant leg, as determined by which limb the participant would not preferentially use to kick a ball. The dominant leg was positioned in 20 degrees of hip flexion and 40 degrees of knee flexion, leaving the foot approximately 6 to 8 inches off the ground. The participant was instructed not to lean the dominant leg on the non-dominant leg. During the tandem position, the dominant leg was placed in front of the nondominant leg. The participants were instructed to touch the great toe of the nondominant leg to the heel of the dominant leg. They were also instructed to distribute the weight equally over both feet.

The firm surface was the hardwood basketball court of the high school gymnasium. The foam surface consisted of a 61 by 61 by 10 cm thick block of 0.88 kilogram density polyester open cell foam (load deflection of 17.5 to 19.3). A stopwatch was used to time each of the 20-second conditions.

Errors were determined following the same protocol described in previous studies. [4, 12, 16] An error was scored if the participant did any of the following: lifted the hands off the iliac crest; opened their eyes, stepped, stumbled, fell; moved the hip into more than 30 degrees of hip flexion; moved the hip into more than 30 degrees of flexion or abduction; lifted the forefoot or heel; remained out of the test position for more than five seconds. Each condition was scored

separately and total errors per condition were used in the analyses. Errors were summed to create a BESS total score.

Prior to the test, each participant was allowed to familiarize themselves with the BESS by practicing each stance on the different surfaces for ten seconds. For all conditions, the participants were then instructed to stand as still as possible with their hands on their hips and eyes closed. Further, they were instructed that if they moved during the test, to return to the test position as quickly as possible. The participant was positioned 10 to 15 feet away from the tester and video camera. Each participant was videotaped and scored at a later date, which allowed for repeated and precise scoring. All tapes were scored twice by the primary investigator. The intratester reliability coefficient was calculated using a one-way ANOVA model for a mean criterion score.

Statistical Analysis

Intraclass reliability coefficients within a one-way analysis of variance protocol were calculated for the revised protocol. Additionally, a repeated measures analysis of variance was calculated to examine differences between trials. The intraclass reliability coefficients and repeated measures ANOVA were obtained using SPSS 11.3.1 (SPSS Inc., Chicago, IL).

Results

We found the mean total number of errors was 20.08 ± 11.76 while the mean number of errors for each trial was 6.69 ± 0.36. The number of errors for each surface, stance and trial can be found in Table 1. The participants made the greatest amount of error in the single leg stance foam surface conditions. The tandem leg stance had slightly less errors than the single leg stance on either surface.

To evaluate whether there were differences between trials within the protocol, a repeated measures ANOVA was calculated. There were statistically significant differences between the trials. To determine whether the statistical significance was equal across all trials, simple contrasts were run. Statistically significant differences existed between trial one and trial two ($F_{(1.65, 286)} = 4.890$, p=.013) however; there were no significant differences between trials two and three (Figure 1).

Finally, an intraclass reliability coefficient was calculated for the revised protocol. The intraclass reliability measures internal consistency, the extent to which the parts of the test measure the same construct.[20, 21] A reliability coefficient for the revised protocol was calculated (0.88) using three trials of four conditions. Due to the presence of a practice effect for trial one, trial one was eliminated from the analysis. This resulted in a reliability coefficient for trials two and three of 0.84. Consistency of scores was obtained between trial two and trial three. This would suggest that the use of one trial might be appropriate for estimation of postural stability. However, the intraclass reliability coefficient for one trial was 0.73 (Table 2) which negates the change in protocol as only moderate reliability was obtained.

Discussion

We found that making minor modifications to the BESS protocol improved the reliability of the measure and reduced the practice effects. The average number of errors per trial was 6.69 $\pm$ 0.36. The revision of the protocol makes comparison to previous reports inappropriate.

Differences between trials suggest that a practice effect exists within the revised protocol. Previous researchers revealed that a practice effect existed with decreased errors committed with each subsequent session.[16] We found that a practice effect was present between the first and second trial; however, scores stabilized for subsequent trials.

The current protocol administered a familiarization trial prior to test administration. Given that a practice effect still existed for the first trial, we recommend that the first trial should not be scored or used in the analysis. It should be administered as a practice trial only. Additionally, administering and scoring three trials after the practice trial would, theoretically, increase the intraclass reliability coefficient. We obtained reliability coefficients of 0.84 and 0.88 using two and three trials respectively. Further, analyzing trial one with the practice effect might increase the reliability but would not increase validity evidence. Therefore, it is recommended to administer at least three, preferably four trials without scoring the first trial to obtain good intraclass reliability.

Intraclass reliability coefficients for the traditional BESS protocol (3 stances, 2 surfaces) were previously found to be 0.60.[19] The lack of variance in the double leg stance did not support inclusion of this stance in subsequent protocols. Theoretically, removal of the double leg stance in the traditional protocol estimated an increase in the intraclass reliability coefficient to 0.71.[19] Reliability coefficients are a function of the number of items within the protocol. Usually as the number of good items increases, the reliability of the measures increases. This assumes the additional items are good. Further, the removal of bad items will increase the reliability of the measure. The increase of the reliability coefficient when the double leg stance was removed occurred because the double leg firm and double leg foam were "bad items". The additional increase in intraclass reliability was, therefore, a function of further increasing the number of items within the protocol. Therefore, increasing the number of "good items" increases the reliability of the measure.

Relevance to the BESS

Revision of a test is a tedious process that should occur frequently. Once the purpose, construct, and items have been defined during the original test construction, the examination of the psychometric properties of the test can be examined on a routine basis. The revision process aids in the validation process of test score interpretations when changes in population, test administration, protocols and equipment are modified. However, validating interpretations relevant to test scores cannot be accomplished without a reliable measure.

Reliability estimates typically measure internal consistency of scores, stability of scores, and/or equivalence of alternate forms. Reliability can be impacted by many things, including the number of items within a measure, number of testers, sample size, etc. When the aforementioned factors change, the reliability of the test scores are subject to change.

Reliability coefficients quantify reliability by summarizing the consistency (or inconsistency) among the measurement.[1] Obtaining poor intraclass reliability coefficients would suggest that items within the test are flawed. Intratester reliability coefficients of 0.60 that can be interpreted as at least 60% of the total score variance is due to true score variance. Since reliability coefficients are calculated by estimating the ratio of true-to-observed score variance within a measurement tool.[22] If the flawed items are removed, the reliability of the measure should increase. Based upon subjective opinions of the BESS researchers have questioned the need and adequacy for the double leg stance in the BESS protocol. This research confirmed that the revised protocol increased the reliability of the measure and thus should improve the interpretations of the measure.

Analyzing and revising measures should be completed periodically to ensure that the measure is still reliable and interpretation of scores is valid in different populations, settings,

administration protocols, etc. Reliability studies in athletic training typically take the form of stability (test-retest) reliability. However, very little work is being done to establish the reliability of the items (conditions) within the test.

Clinical Implications

As standardized concussion assessment tools are suggested within the field of athletic training, having psychometrically sound instruments is a necessity.[23-25] The recommendation for a multifaceted approach including physical examination, self-report symptoms, neuropsychological and posturography assessment has created a significant amount of interest in the BESS.

Differences between the trials established the presence of a practice effect with higher number of errors for trial one than trials two and three. It appears that the practice effect only lasts through the first trial and the scores stabilize between trials two and three. The suggestion that clinicians eliminate the double leg stance is reaffirmed with this research. However, the recommendation to perform three trials of four conditions might be a challenge. It is suggested to administer one trial as a practice trial without scoring. Second, it is suggested administering at least two additional trials and preferably three trials for a total of three and four trials to obtain good reliability. It may appear that four trials of four conditions might be too time consuming, or introduce fatigue; however, scoring three trials after the practice trial will ensure excellent reliability.

Performing three trials within the BESS protocol takes approximately five minutes to administer within an athletic population. The revised protocol provides greater reliability of the BESS scores, which aids in the interpretation of postural abilities and, thus, deficits in balance

following injury. Psychometrically sound instruments support the ability to make and interpret clinical decisions regarding injury and return to participation.

Limitations

The research was performed using only non-concussion healthy athletes. Examining the revised protocol with concussed athletes will allow for evaluation of the specificity and sensitivity of the revised BESS protocol. Additionally, the methodology of this project included a familiarization trial prior to testing. Although the familiarization was similar to standard administration (10 seconds instead of 20 seconds), it should not be considered a practice trial. The presence of a practice effect with the inclusion of this familiarization would suggest that additional practice trials should be administered prior to baseline test sessions. Moreover, the number of trials used was based on the number of trials used in similar protocol measures with the Neurocom Sensory Organization Test. Investigating fatigue and stability of scores after administering four or more trials for both baseline and injured data may be helpful to determine the best protocol for use in concussion assessment. The final step of test construction (validation studies for the test) of the revised protocol was not completed. The study evaluated only the reliability of the measure. A researcher can have a highly reliable measure without the ability to validate interpretations from the test. Additional studies should be conducted to examine the construct and content validity of the revised protocol. Finally, the inclusion of collegiate and adult populations should be performed to ensure the generalizability to different populations.

References

1. Crocker, L., Algina, J. *Introduction to Classical and Modern Test Theory.* New York, NY: Holt, Rinehart and Winston Inc. 1986.

2. Black, F.O., Wall, C., Rockette, H.E., Kitch, R. Normal subject postural sway during the Rhomberg test. *Am. J. Otolaryngol.* 1982; 3:309-318.

3. Cohen, J., Blatchly, C., Gombash, L. A study of the clinical test of sensory interaction and balance. *Phys. Ther.* 1993; 73: 346-354.

4. Riemann, B.L., Guskiewicz, K.M., Shields, E.W. Relationship between clinical and forceplate measures of postural stability. *J Sport Rehabil.* 1999; 8: 71-82.

5. Jansen, E.R., Larsen, R., Oleson, M. Quantitative Rhomberg's test. *Acta Neurol. Scand.* 1982; 66: 93-99.

6. Riemann, B.L., Guskiewicz, K.M. Assessment of mild head injury using measures of balance and cognition: A Case Study. *J Sport Rehabil.* 1997; 6:283-289.

7. Ekdahl, C., Jarnlo, G., Anderson, S. Standing balance in healthy subjects. *Scand. J. Rehabil. Med.* 1989; 21: 187-195

8. Ingersoll C.D., Armstrong, C.W. The effects of closed-head injury on postural sway. *Med Sci Sport Exerc.* 1992; *24(7): 739-743.*

9. Shumway-Cook, A., Horak, F. Assessing the influence of sensory interaction on balance. *Phys. Ther.* 1986; 66:1548-1550.

10. Rubin, A.M., Wooley, S.M., Dailey, V.M., Goebel, J.A. Postural stability following mild head or whiplash injuries. *Am. J. Otol. 1995; 16(2):216-221.*

11. El-Kashlan, H.K., Shepard, N.T., Asher, A.M., Smith-Wheelock, M., Telian, S.A. Evaluation of Clinical Measures of Equilibrium. *The Laryngoscope.* 1998; 108: 311-319.

12. Guskiewicz, K.M., Ross, S.E., Marshall, S.W. Postural Stability and Neuropsychological Deficits after concussion in collegiate athletes. *J Athl Train.* 2001; 36 (3): 263-273.

13. Riemann, B.L., Guskiewicz, K.M. Effects of Mild Head Injury on Postural Stability as Measured Through Clinical Balance Testing. *J Athl Train.* 2000; 35(1): 19-25.

14. Valovich, TC. The use of the Standardized Assessment of Concussion and Balance Error Scoring System and Learning Effects in Youth Sports Participants [dissertation]. Charlottesville, VA: University of Virginia; 2002.

15. Valovich, T.C., Perrin, D.H., Gansneder, B.M. Repeat Administration Elicits a Practice Effect With the Balance Error Scoring System but Not With the Standardized Assessment of Concussion in High School Athletes. *J Athl Train.* 2003; 38(1): 51-56.

16. Valovich-McLeod, T., Perrin, D.H., Guskiewicz, K.M., Shultz, S.J., Diamond, R., Gansneder, B.M. Serial Administration of Clinical concussion assessments and learning effects in healthy young adults. *Clin J Sport Med.* 2004; 14(5): 287-295.

17. Susco, T.M., Valovich McLeod, T.C., Gansneder, B.M., Shultz, S.J. Balance Recovers Within 20 Minutes After Exertion as Measured by the Balance Error Scoring System. *J Athl Train.* 2004; 39 (3): 241-246.

18. Wilkins, J.C., Valovich McLeod, T.C., Perrin, D.H., Gansneder, B.M. Performance on the Balance Errors Scoring System Decreases after Fatigue. *J. Athl. Train.* 2004; 39(2): 156-161.

19. Hunt, T.N., Ferrara, M.S. The Generalizability of the Balance Error Scoring System. *J. Athl. Train.* In Review

20. Baumgartner T.A., Jackson, A.S., Mahar, M.T., Rowe, D.A. *Measurement for Evaluation*. Boston, Massachusetts: McGraw-Hill; 2003: 96-119.

21. Morrow, J.R., Jr. Generalizability theory. In M.J. Safrit and T.M. Wood (Eds.), *Measurement concepts in physical education and exercise science*. Champaign, IL: Human Kinetic; 1995.

22. Huck, S.W. *Reading statistics and research.* New York, NY: Addison Wesley Longman, Inc. 2000.

23. Aubry M., Cantu, R., Dvorak, J., Graff-Baumann, T., Johnston, K., Kelly, J., Lovell, M., McCrory, P., Meeuwisse, W., Schamasch, P. Summary and Agreement Statement of the 1st International symposium on Concussion in sport, Vienna 2001. *Clin J Sport Med*. 2002;12:6-11.

24. Guskiewicz, K., Bruce, S., Cantu, R., Ferrara, M., Kelly, J., McCrea, M., Putukian M., Valovich McLeod, T. National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion. *J Athl Train.* 2004;39: 280-297.

25. McCrory, P., Johnston, K., Meeuwisse, W., Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Kelly, J., Lovell, M., Schamasch, P. Summary and agreement statement of the Second International Conference on Concussion in Sport, Prague 2004. *Clin J Sport Med.* 2005; 15(2): 48-55.

Table 5.1: Trial, Surface and Stance Means and Standard Deviations.

Trial	Stance	Surface	Mean $\pm$ S.D
1	Single Leg	Floor	1.75 $\pm$ 1.74
1	Tandem Leg	Floor	0.44 $\pm$ 0.88
1	Single Leg	Foam	3.73 $\pm$ 2.39
1	Tandem Leg	Foam	1.27 $\pm$ 1.35
2	Single Leg	Floor	1.69 $\pm$ 1.68
2	Tandem Leg	Floor	0.42 $\pm$ 0.69
2	Single Leg	Foam	3.24 $\pm$ 2.00
2	Tandem Leg	Foam	1.06 $\pm$ 1.21
3	Single Leg	Floor	1.77 $\pm$ 1.60
3	Tandem Leg	Floor	0.52 $\pm$ 0.82
3	Single Leg	Foam	3.17 $\pm$ 1.86
3	Tandem Leg	Foam	1.02 $\pm$ 1.24

Table 5.2: Intraclass Reliability Coefficients by the Number of Conditions Analyzed.

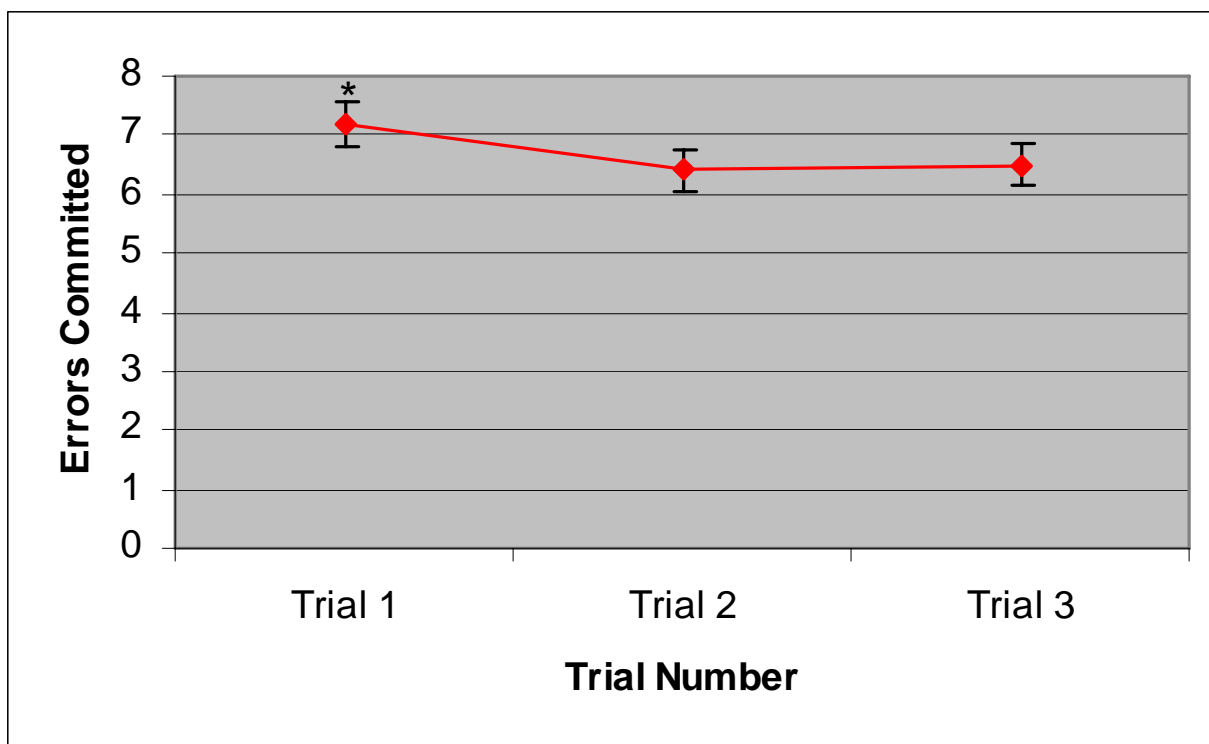| Trials	Number of conditions	Intraclass reliability coefficient
1, 2 & 3	12	0.88
2 & 3	8	0.84
2 only	4	0.73

Figure 5.1: Mean Number of Errors Per Trial for the Revised BESS Protocol.

*Statistically Significant Differences Existed Between Trial 1 and Trial 2.

CHAPTER 6

SUMMARY AND CONCLUSIONS

The Center for Disease Control and Prevention (CDC) recently proclaimed concussion as an epidemic (CDC, 2003). There have been an estimated 62,000 cases of concussion in American high school sports (Guskiewicz et al., 2001; Hinton-Bayre et al., 2000). This is a conservative estimation, as it is suggested that fifty-three percent of high school athletes do not report sports-related concussion (McCrea et al., 2004). The individual variability of symptoms, severity and recovery curves for each injury is specific to that individual making no two injuries identical.

The seriousness of concussion makes assessment and return to participation decisions even more critical. Allied health professionals have taken the task of assessment and treatment of these concussive injuries. Assessment of concussion has recently evolved into a multifaceted approach which includes physical examination, imaging, self-report symptoms, neuropsychological test, and postural stability. The multifaceted approach captures the variability of deficits following injury. The recent addition of neuropsychological and balance assessment to the concussion assessment battery provides the clinician with additional information specific to those areas of brain functioning.

Evaluation and return to participation decisions following a concussive injury leave athletic trainers the daunting task of making quick return-to-play decisions. With the potential for catastrophic consequences, assessments need to be accurate. The only way to ensure accurate decisions is to utilize reliable and valid measures.

Only recently have athletic trainers invested into the psychometric properties of tests. Score validation, which is an ongoing and continual process, has been reported for concussion assessment tools. Several investigations have evaluated the Head Injury Scale (Piland et al., 2002), the Standardized Assessment of Concussion (Valovich, 2002; Valovich et al., 2003), and the Balance Error Scoring System (Hunt & Ferrara, in review; Susco et al., 2004;Valovich et al., 2003; Wilkins et al., 2004) to improve the psychometric properties of each measure.

Several aspects of test administration were evaluated in this research. The first aspect of test administration that was evaluated was the effect of effort on neuropsychological test scores in high school athletes. Effort in athletic assessment has always involved subjective opinions of test performance. However, if sub-optimal effort is present in baseline testing, the score interpretations are not valid (Faust et al., 2001; Green et al., 2001). Eleven percent of the sample exerted sub-optimal effort in the high school athletic population. This obtained percentage is extremely conservative as a brief gross measure of effort was utilized. The presence of athletes who exert sub-optimal alludes to the need to evaluate baseline scores prior to post-injury comparisons. The presence of sub-optimal effort in a high school population justifies the need to include an objective measure of effort to validate test scores to ensure adequate comparisons to post-injury scores.

The second aspect revised the protocol of a commonly used clinical test, the Balance Error Scoring System. Scale development and revision is a lengthy process, which should provide ample information as to the interpretability of obtained scores (Cronbach, 1984). Poor reliability attests to inconsistency across persons, trials, and the measure itself. Calculated reliability coefficients for the traditional model of the Balance Error Scoring System revealed poor to good reliability (Guskiewicz et al., 2000, Guskiewicz et al., 2001; Riemann &

Guskiewicz, 1999; Susco et al., 2004;Valovich et al., 2003). However, the double leg stance

does not provide additional information to make adequate interpretations from the obtained

scores (Hunt).

The BESS protocol was revised by removing the double leg stance and adding additional

trials of the single leg stance and the tandem stance. The intraclass reliability of the revised

protocol was found to be 0.85. This increase in reliability justifies incorporating the revised

protocol of the BESS into common concussion assessment batteries. Making minor

modifications to the BESS protocol improves the psychometric properties of the test, which, in

theory, should improve the interpretations developed from the obtained scores.

The importance of obtaining baseline data to determine premorbid function has been

stressed in numerous positions statements. (Aubry et al., 2002; Guskiewicz et al., 2004; McCrory

et al., 2005) However, these comparisons can only be made if test scores are reliable and the

interpretations from the scores are valid. Elimination of the possible confounding aspects of

baseline testing is important to obtain reliable and valid interpretations of test scores. This

research should be continued in an attempt to understand potential threats to reliability and

validity within concussion assessment tools.

REFERENCES

Aldrich, E.F., Eisenberg, H.M., & Saydjari, C. (1992). Diffuse brain swelling in severely head injured children: A report from the NIH Traumatic Coma Data Bank. *Journal of Neurosurgery*, 76, 450-454.

Anderson, V.A., Morse, S.A, Klug, G, Catroppa, C., Haritou, F., Rosenfeld, J., & Pentland, L. (1997). Predicting recovery from head injury in young children: A prospective analysis. *Journal of the International Neuropsychological Society,* 3, 568-580

Anderson, S.J. (2001). On the importance of collecting local neuropsychological normative data. *South African Journal of Psychology,* 31(3), 29-34.

Arnett, P.A., Hammeke, T.A., & Schwartz, L. (1995). Quantitative and qualitative performance on Rey's 15-item test in neurological patients and dissimulators. *The Clinical Neuropsychologist,* 9: 17-26.

Assaiante, C., & Amblard,B. (1995). An ontogenetic model for the sensorimotor organization of balance controls in humans. *Human Movement Science*, 14, 13-43.

Atkinson, G. (1995). British Association of Sports Sciences Newsletter.

Atwater, S.W., Crowe, T.K., Deitz, J.C., & Richardson, P.K. (1991). Interrater and test-retest reliability of two pediatric balance tests. *Physical Therapy*, 70, 79-87.

Aubry M, Cantu R, Dvorak J, Graf-Baumann T, Johnston K, Kelly J, Lovell M, McCrory P, Meeuwisse W, & Schamasch P. (2002). Summary and agreement statement of the First International Conference on Concussion in Sport, Vienna 2001. Recommendations for the improvement of safety and health of athletes who may suffer concussive injuries. *British Journal of Sports Medicine,* 36(1), 6-10.

Bailes, J.E., & Cantu, R.C. (2001). Head Injury in Athletes. *Neurosurgery,* 48(1), 26-42.

Barr, W. (2001). Methodological Issues in Neuropsychological Testing. *Journal of Athletic Training*, 36(3), 297-302.

Barth, J., Alves, W., Ryan, T., Macciocchi, S., Rimel, R., Jane, J., & Nelson, W. (1989). Mild Head Injury in Sports: Neuropsychological Sequelae and Recovery of Function. In: Levin, H., Eisenber, H., Benton, A. eds. *Mild Head Injury.* New York, New York: Oxford University Press 257-275.

Baumgartner T.A., & Jackson, A.S. (2003). *Measurement for Evaluation.* Boston: McGraw-Hill

Bazarian, J., Wong, T., & Harris, M. (1999). Epidemiology and predictors of post-concussive syndrome after minor head injury in an emergency population. *Brain Injury,* 13, 173-189.

Bazarian, J.J., Blyth, B., & Cimpello, L. (2006). Bench to Bedside: Evidence for Brain Injury after Concussion- Looking beyond the Computed Tomography Scan. *Academic Emergency Medicine*, 13, 199-214

Beetar, J., & Williams, J. (1995). Malingering response styles on the Memory Assessment Scales and Symptom Validity Tests. *Archives of Clinical Neuropsychology,* 10, 57-72.

Bernard, L.C., & Fowler, W. (1990). Assessing the validity of memory complaints: Performance of brain-damaged and normal individuals on Rey's task to detect malingering. *Journal of Clinical Psychology,* 46, 432-436.

Black, F.O., Wall, C., Rockette, H.E., & Kitch, R. (1982).Normal subject postural sway during the Rhomberg test. *American Journal of Otolarynlogy,* 3,309-318.

Boone, K.B, Savodnik, I., Ghaffarian, S., Lee, A., Freeman, D. & Berman, N. (1995). Rey 15-item Memorization and Dot Counting Scores in a "stress" claim worker's compensation population: Relationship to personality (MCMI) scores. *Journal of Clinical psychology,* 51, 457-463.

Boone, K.B, Lu., P., Sherman, D., Palmer, B., Back, C., Shamieh, E., Warner-Chacon, K., & Berman, N. (2000). Validation of a new technique to detect malingering of cognitive symptoms: The b Test. *Archives of Clinical Neuropsychology,*15, 227-241.

Boone, K.B, Salazarm, X., Lu, P., Warner-Chacon, K., & Razani, J. (2002a).The Rey 15-Item Recognition Trial: A Technique to Enhance Sensitivity of the Rey 15-Item Memorization Test. *Journal of Clinical and Experimental Neuropsychology,*24(5), 561-573.

Boone, K.B., Lu, P., Back, C., King, C., Lee, A., Philpott, L., Shamieh, E., & Warner-Chacon, K. (2002b). Sensitivity and specificity of the Rey Dot Counting Test in patients with sub-optimal effort and carious clinical samples. *Archives of Clinical Neuropsychology,* 17, 625-642.

Bowen, A.P. (2003). Second Impact Syndrome: a rare, catastrophic, preventable complication in young athletes. *Journal of Emergency Nursing,* 29(3), 287-289.

Brandt, J. (1991).The Hopkins Verbal Learning Test: Development of a new verbal memory test with six equivalent forms. *The Clinical Neuropsychologist*, 5, 125-142.

Broglio, S.P., Tomporowski, P.D., & Ferrara, M.S. (2005). Balance performance with a cognitive task: a dual task testing paradigm. *Medicine and Science in Sports and Exercise,* 37(4), 689-695.

Bruce, D. A., Alve, A., Bilaniuk, L., Dolinskas, C., Obrist, W., & Uzzell, B. (1982). Diffuse cerebral swelling following head injuries in children: the syndrome of 'malignant brain edema. *Journal of Neurosurgery,* 54, 170-178.

Campbell, D.T., Fiske, D.W. (1959). Convergent and discriminate validation by the multitrait-multimethod matrix. *Psychological Bulletin,* 56(2), 81-105.

Cantu, R.C. (1992) Cerebral Concussion in Sport. *Journal of Sports Medicine*, 17, 64-74.

Cantu, R.C. (1997). Athletic Head Injuries. *Clinics in Sports Medicine,* 17(1), 37-43.

Cantu, R.C. (1998). Return to play guidelines after a head injury. *Clinical Journal of Sports Medicine*, 17, 45-60.

Cantu, R.C. (2001). Posttraumatic Retrograde and Anterograde Amnesia: Pathophysiology and implications in grading and safe return to play. *Journal of Athletic Training,* 36(3), 244-248.

Capruso, D., Levin, H. (1996). Neurobehavioral Outcome of Head Trauma. In: Evans R, ed. *Neurology and Trauma.* Philadelphia: W.B Saunders Co.

Centers for Disease Control and Prevention. (1997). Sports-related recurrent brain injuries-United States. *JAMA,* 277(15), 1190-1191.

Centers for Disease Control and Prevention. (2003). National Center for Injury Prevention and Control. Report to Congress. *Mild traumatic brain injury in the United States: steps to prevent serious public health problem.* Atlanta, GA: Centers for Disease Control and Prevention.

Chugani, H.T., Phelps, M.E., & Mazziotta, J.C. (1987). Positron Emission Tomography study of human brain functional development. *Annals of Neurology,* 1, 243-249.

Cohen, J., Blatchly, C., & Gombash, L. (1993). A study of the clinical test of sensory interaction and balance. *Physical Therapy,* 73, 346-354.

Collie, A., Makdissi, M., Maruff, P., Bennerll, K., & McCrory, P. (2006). Cognition in the days following concussion: comparison of symptomatic versus asymptomatic athletes. *Journal of Neurological and Neurosurgery Psychiatry*, 77, 241-245.

Collins, M.W., Grindel, S.H., Lovell, M.R., Dede, D.E., Moser, D.J., Phalin, B.R., Nogle, S., Wasik, M., Cordry, D., Daugherty, K.M., Sears, S.F., Nicolette, G., Indelicato, P., & McKeag, D.B. (1999). Relationship between concussion and neuropsychological performance in college football players. *JAMA*, 282(10), 964-70.

Collins, M.W., Lovell, M.R., Iverson, G.L., Cantu, R.C., Maroon, J.C., & Field, M. (2001). Cumulative Effects of Concussion in High School Athletes. *Neurosurgery,* 51, 1175-1181.

Collins, M.W., Iverson, G.L., Lovell, M.R., McKeag, D.B., Norwig, J., & Maroon, J. (2003). On-field predictors of neuropsychological and symptom deficit following sports-related concussion. *Clinical Journal of Sports Medicine*, 13(4), 222-9.

Constantinou, M, Bauer, L., Ashendorf, L., Fisher, J.M., & McCaffrey, R.J. (2005). Is poor performance on recognition memory effort measures indicative of generalized poor performance on neuropsychological tests? *Archives of Clinical Neuropsychology*, 20, 191-198.

Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory.* New York: Holt, Rinehart and Winston Inc.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin,* 52(4): 281-302.

Daniel, J., Olesniewicz, M., & Reeves, D. (1999). Repeated measures of cognitive processing efficiency in adolescent athletes: Implications for monitoring recovery from concussion. *Neuropsychiatry, Neuropsychology and Behavior Neurology,* 12(3), 167-169.

Dennis, M. (2000). Developmental plasticity in children: the role of biological risk, development, time, and reserve. *Journal of Communication Disorders,* 33(4), 321-331.

Donders, J., & Strom, D. (2000). Neurobehavioral recovery after pediatric head trauma: injury, pre-injury, and post-injury issues. *Journal of Head Trauma and Rehabilitation,* 15(2), 792-803.

Echemendia, R.J., Putukian, M., Mackin, R.S., Julian, L., & Shoss, N.. (2001). Neuropsychological test performance prior to and following sports-related mild traumatic brain injury. *Clinical Journal of Sports Medicine*, 11(1), 23-31.

Ekdahl, C., Jarnlo, G., & Anderson, S. (1989). Standing balance in healthy subjects. *Scandinavian Journal of Rehabilitative Medicine,* 21, 187-195

El-Kashlan, H.K., Shepard, N.T., Asher, A.M., Smith-Wheelock, M., &Telian, S.A. (1998). Evaluation of Clinical Measures of Equilibrium. *The Laryngoscope,* 108, 311- 319.

Esselman, P.C., & Uomoto, J.M.. (1995). Classification of the spectrum of mild traumatic brain injury. *Brain Injury,* 9, 417-424.

Faust, D., Hart, K., Guilmette, T., & Arkes, H. (1988). Neuropsychologists' capacity to detect adolescent malingerers. *Professional Psychology: Research and Practice*, 19(5), 508-515.

Ferrara M.S., McCrea, M., Peterson, C.L., & Guskiewicz, K.M. (2001). A survey of practice patterns in concussion assessment and management. *Journal of Athletic Training*, 36, 145-149.

Field, M., Collins, M.W., Lovell, M.R., & Maroon, J. (2003). Does age play a role in recovery from sports-related concussion? A comparison of high school and collegiate athletes. *Journal of Pediatrics*, 142(5):546-53.

Frazen, M.D. (1999). Reliability and validity of serial assessments. In Bailes, J.E., Lovell, M.R., Maroon, J.C (Eds). *Sport Related Concussion*. St Louis: Quality Medical Publishing Inc.

Frederick, R.I. (2002). A Review of Rey's Strategies for Detecting Malingered Neuropsychological Impairment. *Journal of Forensic Neuropsychology,* 2 (3), 1-22.

Genuardi, F.J, & King, W.D. (1995). Inappropriate discharge instructions for youth athletes hospitalized for concussion. *Pediatrics,* 95(2), 216-218.

Gervaus, R.O., Rohling, M.L, Green, P., & Ford, W. (2004).A comparison of WMT, CARB, and TOMM failure rates in non-head injured disability claimants. *Archives of Clinical Neuropsychology,* 19, 475-487.

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

Goldberg, T.O., & Miller, H.R. (1986). Performance of psychiatric inpatients and intellectually deficient individuals on a task that assesses the validity of memory complaints. *Journal of Clinical Psychology*, 42, 792-795.

Goldstein, G., & Watson, J. (1989). Test-retest reliability of the Halstead-Reitan Battery and the WAIS in a neuropsychiatric population. *The Clinical Neuropsychologist*, 3, 265-272.

Green, P., Rohling, M.L, Lees-Haley, P.R., & Allen, L.M. (2001a). Effort has a greater effect on test scores than severe brain injury in compensation claimants. *Brain Injury,* 15(12), 1045-1060.

Green, P., & Iverson, G.L. (2001b). Effects of injury severity and cognitive exaggeration on olfactory deficits in head injury compensation claims. *NeuroRehabilitation,* 16, 237-243.

Greiffenstein, M., Baker, W., & Gola, T. (1994). Validation of malingered amnesia measures with a large clinical sample. *Psychological Assessment,* 6, 218–224.

Griffin, G.A., Glassmire, D.M., Henderson, E.A., & McCann, C. (1997). Rey II: Redesigning the Rey Screening Test of malingering. *Journal of Clinical Psychology,* 53, 757-766

Grindel S.H, Lovell, M.R., & Collins, M.W. (2001). The assessment of sport-related concussion: the evidence behind neuropsychological testing and management. *Clinical Journal of Sports Medicine,* 11, 134-143.

Gulbrandsen, G.B. (1984). Neuropsychological sequelae of light head injuries in older children 6 months after trauma. *Journal of Clinical Neuropsychology,* 6, 257-268.

Guskiewicz, K.M., & Perrin, D.H. (1996). Research and Clinical Applications of Assessing Balance. *Journal of Sport Rehabilitation*, 5, 45-63.

Guskiewicz, K.M., Riemann, B.L., Perrin, D.H., & Nashner, L.M. (1997). Alternative approaches to the assessment of mild head injury in athletes. *Medicine and Science in Sports and Exercise*, 29(suppl 7), S213-S221.

Guskiewicz, K.M., & Perrin, D.H. (1998). Effect of mild head injury on cognition and postural stability [abstract]. *Journal of Athletic Training*, 33(suppl), S-8.

Guskiewicz K. Weaver, N., & Padua, D. (2000). Epidemiology of concussion in collegiate and high school football players. *American Journal of Sports Medicine*, 28(5), 643-650.

Guskiewicz, K.M., Ross, S.E., & Marshall, S.W. (2001). Postural Stability and Neuropsychological Deficits after concussion in collegiate athletes. *Journal of Athletic Training,* 36 (3), 263-273.

Guskiewicz, K.M., McCrea, M., Marshall, S.W., Cantu, R.C., Randolph, C., Barr, W. Onate, J., & Kelly, J.P. (2003). Cumulative Effects associated with recurrent concussion in collegiate football players. The NCAA Concussion Study. *JAMA*, 290 (19), 2549-2556

Guskiewicz, K., Bruce, S., Cantu, R., Ferrara, M., Kelly, J., McCrea, M., Putukian M., & Valovich McLeod, T. (2004). National Athletic Trainers' Association Position Statement: Management of Sport-Related Concussion. *Journal of Athletic Training,* 39, 280-297.

Guskiewicz, K.M., Marshall, S.W., Bailes, J., McCrea, M., Cantu, R.C., Randolph, C., & Jordan, B.D. (2005). Association between recurrent concussion and late-life cognitive impairment in retired professional football players. *Neurosurgery,* 57(4), 719-726.

Haaland, K.Y., Temkin, N., Randahl, G., & Dikmen, S. Recovery of simple motor skills after head injury. *Journal of Clinical and Experimental Neuropsychology,* 16, 448-456.

Hall, H., & Pritchard, D. (1996). *Detecting malingering and deception: Forensic distortion analysis.* Florida: St. Lucie Press.

Hinton-Bayre, A., Geffen, G., & McFarland, K.(1997). Mild head injury and speed of information processing: A prospective study of professional rugby league players. *Journal of Clinical and Experimental Neuropsychology*, 19, 275-289.

Hinton-Bayre, A. D., Geffen, G. M., Geffen, L. B., McFarland, K. A., & Friis, P. (1999). Concussion in contact sports: reliable change indices of impairment and recovery. *Journal of Clinical and Experimental Neuropsychology,* 21(1), 70-86.

Homer, C.J., & Kleinman, L. (1999). Technical report; minor head injury in children. *Pediatrics*, 104(6), 78.

Hovda, D.A., & Villablanca, J.R. (1998). Cerebral metabolism following neonatal or adult hemineodecortication in cats: effect on oxidative capacity using cytochrome oxidase histochemistry. *Brain Research: Developmental Brain Research,* 110(1), 39-50.

Huck, S.W. *Reading statistics and research.* (2000). New York: Addison Wesley Longman, Inc.

Hunt, T.N., & Ferrara, M.S. (In Review). The Generalizability of the Balance Error Scoring System. *Journal of Athletic Training.*

Hynd, G.W., & Willis, W.G. (1987). *Pediatric Neuropsychology.* Orlando: Grune & Stratton.

Ingersoll C.D., & Armstrong, C.W. (1992). The effects of closed-head injury on postural sway. *Medicine and Science in Sports and Exercise,* 24(7), 739-743.

Iverson, G.L., & Binder, L.M. (2000). Detecting exaggeration and malingering in neuropsychological assessment. *Journal of Head Trauma Rehabilitation,* 15, 829-858.

Jagoda, A.A., Cantrill, S.V., & Wears, R.L. (2002). Clinical policy: neuroimaging and decision-making in adult mild traumatic brain injury in the acute setting. *Annals of Emergency Medicine,* 40, 231-249.

Jansen, E.R., Larsen, R., & Oleson, M. (1982). Quantitative Rhomberg's test. *Acta Neurologica. Scandinavia,* 66, 93-99.

Johnston, K.M., McCrory, P. & Mohtadi, N. (2001). Evidence-based review of sport-related concussion: clinical science. *Clinical Journal of Sport Medicine,* 11, 150-159.

Kelly, K. D., Lissel, H. L., Rowe, B. H., Vincenten, J. A., & Voaklander, D. C. (2001). Sport and recreation-related head injuries treated in the emergency department. *Clinical Journal of Sports Medicine,* 11, 77-81.

Keppel G. (1991). *Design and Analysis: A researchers Handbook.* New Jersey: Prentice Hall.

Klonoff, H., Clark, C., Klonoff, P.S. (1993). Long-term outcome of head injuries: a 23 year follow-up study of children with head injuries. *Journal of Neurology, Neurosurgery, and Psychiatry,* 40, 1211-1219.

Klonoff, H., & Lowe, M.D. (1997).Head injuries in children: a prospective five year follow-up. *Journal of Neurology, Neurosurgery and Psychiatry,* 40, 1211-1219.

Lees-Haley, P. (1997). Attorneys influence expert evidence in forensic psychological and neuropsychological cases. *Assessment,* 4, 321-324.

Leininger, B.E., Gramling, S.E., & Farrell, A.D. (1990). Neuropsychological deficits in symptomatic minor head injury patients after mild concussion. *Journal of Neurology, Neurosurgery and Psychiatry,* 53, 293-296.

Levin, H.S., Song, J., Scheibel, R.S., Fletcher, J.M., Harward, H., Lilly, M., & Goldstein, F. (1997). Concept formation and problem-solving following closed head injury in children. *Journal of the International Neuropsychological Society,* 3, 598-607.

Lezak M, D., Howieson, D.B., & Loring, D.W. (2004). *Neuropsychological Assessment.* Oxford: Oxford Press.

Lovell, M.R., Iverson, G.L., Collins, M.W., McKeag, D., & Maroon, J.C. (1999). Does loss of consciousness predict neuropsychological decrements after concussion? *Clinical Journal of Sports Medicine*, 9(4); 193-198.

Lovell, M., & Collins, M. (2000). Neuropsychological assessment of the college football player. *Journal of Head Trauma and Rehabilitation*, 13(2), 9-26.

Lovell, M.R., Collins, M.W., Iverson, G.L., Field, M., Maroon, J.C., Cantu, R., Podell, K., Powell, J.W., Belza, M., & Fu, F.H. (2001). Recovery from mild concussion in high school athletes. *Journal of Neurosurgery*, 98(2), 296-301.

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

Lynch, W. (2004). Determination of Effort Level, Exaggeration, and Malingering in Neurocognitive Assessment. *Journal of Head Trauma Rehabilitation,* 19(3), 277-283.

Macciocchi, S.N., Barth, J.T., Alves, W., Rimel, R.W., & Jane, J.A. (1996). Neuropsychological functioning and recovery after mild head injury in collegiate athletes. *Neurosurgery,* 39(3), 510-4.

Maddocks, D., & Saling, M. (1996). Neuropsychological deficits following concussion. *Brain Injury*, 10(2), 99-103.

McAllister, T.W., Sparling, M.B., Flashman, L.A., & Saykin, A.J. (2001). Neuroimaging Findings in Mild Traumatic Brain Injury. *Journal of Clinical and Experimental Neuropsychology,* 23 (6), 775-791.

McCaffrey, R., Duff, K., & Westervelt, H. (2000). Practitioner's guide to evaluating change with neuropsychological assessment instruments. *Journal of Neurosurgery*, 98(2), 296-301.

McCrea, M., Guskiewicz, K.M., Marshall, S.W., Barr, W., Randolph, C., Cantu, R.C., Onate, J.A., Yang, J., & Kelly, J.P. (2001a). Acute effects and recovery time following concussion in collegiate football players: the NCAA Concussion Study, *JAMA,* 290(19), 2556-63.

McCrea M. (2001b). Standardized mental status assessment of sports concussion. *Clinical Journal of Sport Medicine*, 11(3), 176-81.

McCrea, M. Hammeke T, Olsen G, Leo P, & Guskiewicz K. (2004). Unreported concussion in high school football players: implications for prevention. *Clinical Journal of Sport Medicine*, 14(1), 13-7.

McCrory, P. R., & Berkovic, S. F. (1998). Second impact syndrome. *Neurology,* 50, 677-683.

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

Mendez, C.V., Hurley, R.A, Lassonde, M., Zhang, L., & Taber, K.H. (2005). Mild Traumatic Brain Injury: Neuroimaging of Sports-Related Concussion. *Journal of Neuropsychiatry Clinical Neuroscience*, 17 (3), 297-303.

Morrow, J.R., Jr. (1995). Generalizability theory. In M.J. Safrit & T.M. Wood (Eds.), *Measurement concepts in physical education and exercise science*. Champaign, IL: Human Kinetic

Moser, R.S., & Schatz, P. (2002). Enduring effects of concussion in youth athletes. *Archives of Clinical Neuropsychology*, 17(1), 91-100.

Moss, A., Jones, C., Fokias, D., & Quinn, D. (2003). The mediating effects of effort upon the relationship between head injury severity and cognitive functioning. *Brain Injury,* 17 (5), 377-387.

Mueller, F. (2001). Catastrophic head injuries in high school and collegiate sports. *Journal of Athletic Training,* 36(3), 307-311.

Nevill, A.M. (1964). Validity and measurement agreement in sports performance. *Journal of Sports Science,* 14, 199.

Notebaert, A.J., & Guskiewicz, K.M. (2005). Current Trends in Athletic Training Practice for Concussion Assessment and Management. *Journal of Athletic Training,* 40(4), 320-335.

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

Pellman, E.J., Viano, D.C., Casson, I.R., Arfken, C., & Powell, J. (2004a). Concussion in Professional Football: Injuries Involving 7 or More Days Out-Part 5, *Neurosurgery*, 55(5), 1100-1119.

Pellman, E.J., Lovell, M.R., Viano, D.C., Casson, I.R., & Tucker, A.M. (2004b). Concussion in professional football: neuropsychological testing--part 6. *Neurosurgery,* 55(6),1290-303

Peterson, C.L., Ferrara, M.S., Mrazik, M., Piland, & S., Elliott, R. (2003). Evaluation of neuropsychological domain scores and postural stability following cerebral concussion in sports. *Clinical Journal of Sports Medicine,* 13(4), 230-237

Philpot, L.M. (1993). The effects of severity of cognitive impairment and age on two malingering tests: An investigation of the Rey Memory Test and Dot Counting Test in Alzheimers patients and optimal middle-aged and older adults. *Dissertation Abstracts Internationa,* 53, 5001.

Piland, S.G., Motl, R.W., Ferrara, M.S., & Peterson, C.L. (2003). Evidence for the factorial and construct validity of a self-report concussion symptom scale. *Journal of Athletic Training*, 38(2), 104-112.

Prins, M.L, Lee, S.M., Cheng, C.L., Becker, D.P, & Hovda, D.A. (1996). Fluid percussion brain injury in the developing and adult rat: a comparative study of mortality, morphology, intracranial pressure and mean arterial blood pressure. *Brain Research: Developmental Brain Research*, 95, 272-282.

Prins, M.L., & Hovda, D.A.(1998). Traumatic brain injury in the developing rat: effects of maturation on Morris water maze acquisition. *Journal of Neurotrauma,* 15, 799-811.

Randolph, C., McCrea, M., & Barr, W. (2005). Is neuropsychological testing useful in the management of sport-related concussion? *Journal of Athletic Training,* 40(3), 139-152.

Reitan, R.M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual and Motor* Skills, 8, 271-276.

Reitan, R.M., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation.* (2nd ed.) Tuscon, AR: Neuropsychology Press.

Riemann, B.L., & Guskiewicz, K.M. (1997). Assessment of mild head injury using measures of balance and cognition: A Case Study. *Journal of Sport Rehabilitation,* 6, 283-289.

Riemann, B.L., & Guskiewicz, K.M., Shields, E.W. (1999). Relationship between clinical and forceplate measures of postural stability. *Journal Sport Rehabilitation*, 8, 71-82.

Riemann, B.L., & Guskiewicz, K.M. (2000). Effects of Mild Head Injury on Postural Stability as Measured Through Clinical Balance Testing. *Journal of Athletic Training,* 35(1), 19-25.

Roberts, G.W., Allsop, D., & Bruton, C. (1990). The occult aftermath of boxing. *Journal of Neurology Neurosurgery and Psychiatry,* 53,373-383.

Robertson, C.L., Bucci, C.J., & Fiskum, G. (2004). Mitochondrial response to calcium in the developing brain. *Developmental Brain Research,* 151,141-148.

Robinson, R.O. (1981). Brain death in children. *Archives of Disease in Childhood,* 56(9), 657-658.

Rogers, R., Ernest H., & Liff, C. (1993). Feigning neuropsychological impairment: A critical review of methodological and clinical considerations. *Clinical Psychology Review,* 13, 255–274.

Rohling, M.L., Green, P., Allen, L.M., & Iverson, G.I. (2002). Depressive symptoms and neurocognitive test scores in patients passing symptom validity tests. *Archives of Clinical Neuropsychology,* 17, 205-222.

Rohling, M.L, Meyers, J.E., & Millis, S.R. (2003). Neuropsychological Impairment Following Traumatic Brain Injury: A Dose-Response Analysis. *The Clinical Neuropsychologist,* 17(3), 289-302.

Roman, M.J., Delis, D.C., Willerman, L., Magluc, M., Demadura, R.L., de la Pena, J.L., Loftis, C., Walsh, J., & Kracun, M. (1995). Impact of pediatric traumatic brain injury on components of verbal memory. *Journal of Clinical and Experimental Neuropsychology,* 20(2), 245-258.

Rubin, A.M., Wooley, S.M., Dailey, V.M., & Goebel, J.A. (1995). Postural stability following mild head or whiplash injuries. *American Journal of Otolology*, 16(2):216-221.

Ruff, R.M., Wylie, T., & Tennant, W. (1993).Malingering and malingering-like aspects of mild closed head injury. *Journal Head Trauma Rehabilitation*, 8(3), 60-73.

Safrit, M.J., & Wood, T.M., (Eds.). (1989). *Measurement concepts in physical education and exercise science.* Champaign, Il: Human Kinetics.

Schnitker, M.T.(1949). A syndrome of cerebral concussion in children. *Journal of Pediatrics,* 35, 557-560.

Schretlen, D., Brandt, J., Krafft, L., & Van Gorp, W. (1991). Some caveats in using the Rey 15-item memory test to detect malingered amnesia. *Psychological Assessment: A Journal of Consulting and Clinical Psychology,* 3, 667-672.

Shumway-Cook, A., & Horak, F. (1986). Assessing the influence of sensory interaction on balance. *Physical Therapy,* 66, 1548-1550.

Silver, C.H. (2000). Ecological validity of neuropsychological assessment in childhood traumatic brain injury. *Journal of Head Trauma and Rehabilitation,*15(4), 973-988.

Simon, B., Letourneau, P., Vitorino, E., & McCall, J. (2001). Pediatric minor head trauma: indications for computed tomographic scanning revisited. *Journal of Trauma: Injury, Infection and Critical Care,* 51(2), 231-238.

Smith, A. (1982). *Symbol Digit Modalities Test (SDMT).* Manual (revised). Los Angeles: Western Psychological Services.

Snoek, J. W., Minderhound, J. M., & Wilmink, J. T. (1984) Delayed deterioration following mild head injury in children. *Brain,* 107, 15-36.

Spearman, C. (1904). "General Intelligence" objectively determined and measured. *American Journal of Psychology*, 15, 201-293.

Spreen, O., & Strauss, E. (2000). *A compendium of neuropsychological tests.* New York: Oxford University Press, Inc.

Susco, T.M., Valovich McLeod, T.C., Gansneder, B.M., & Shultz, S.J. (2004). Balance Recovers Within 20 Minutes After Exertion as Measured by the Balance Error Scoring System. *Journal of Athletic Training*, 39 (3), 241-246.

Sweet, J.J. (1999). Malingering differential diagnosis. In: J.J. Sweet (editor) *Forensic Neuropsychology Fundamentals and Practice* (Lisse: Swets and Zeitlinge)

Thompson, J., Sebastianelli, W., & Slobounov, S. (2005). EEG and postural correlates of mild traumatic brain injury in athletes. *Neuroscience Letters,* 377, 158-163.

Thurman, D., Branche, C., & Sniezek, J.E. (1998). The epidemiology of sports-related traumatic brain injuries in the United States: Recent developments. *Journal of Head Trauma Rehabilitation,* 13, 1-8.

Tombaugh, T. (1996). *Test of Memory Malingering*. Toronto: Multi-Health Systems.

Valovich, T.C. (2002). *The use of the Standardized Assessment of Concussion and Balance Error Scoring System and Learning Effects in Youth Sports Participants [dissertation].* Charlottesville, VA: University of Virginia

Valovich, T.C., Perrin, D.H., & Gansneder, B.M. (2003). Repeat Administration Elicits a Practice Effect With the Balance Error Scoring System but Not With the Standardized Assessment of Concussion in High School Athletes. *Journal of Athletic Training*, 38(1), 51-56.

Valovich-McLeod, T., Perrin, D.H., Guskiewicz, K.M., Shultz, S.J., Diamond, R., & Gansneder, B.M. (2004). Serial Administration of Clinical concussion assessments and learning effects in healthy young adults. *Clinical Journal Sports Medicine,* 14(5), 287-295.

Vickery, C.D., Berry, D.T, Inman, T.H., Harris, M.J, & Orey, S.A. (2001). Detection of inadequate effort on neuropsychological testing: A meta-analytic review of selected procedures. *Archives of Clinical Neuropsychology,*16, 45-73.

Wechsler, D. (1981). *WAIS-R manual.* New York: The Psychological Corporation.

Wilkins, J.C., Valovich McLeod, T.C., Perrin, D.H., & Gansneder, B.M. (2004). Performance on the Balance Errors Scoring System Decreases after Fatigue. *Journal of Athletic Training,* 39(2), 156-161.

Wojtys, E. Hovda, D., Landry, G., Boland, A., Lovell, M., McCrea, M., & Minkoff, J. (1999). Concussion in Sports. *American Journal of Sports Medicine,* 27(5), 676-687.

Yeates, K., Taylor, H., Barry, C., Drotar, D., Wade, S., & Stancin, T. (2001). Neurobehavioral symptoms in childhood closed head injuries: changes in prevalence and correlated during the first year postinjury. *Journal of Pediatric Psychology,* 26(2), 79-91.

Yeates, K.O., Luria, J., Bartkowski, H., Rusin, J., Martin, L., & Bigler, E.D. (1999). Post-concussive symptoms in children with mild closed head injuries. *Journal of Head Trauma Rehabilitation,* 14(4), 337-50.

Youngjohn, J., Burrows, L., & Erdal, K. (1995a). Brain damage or compensation neurosis? The controversial post concussion syndrome. *The Clinical Neuropsychologist,* 9, 112-123.

Youngjohn, J. (1995b). Confirmed attorney coaching prior to neuropsychological evaluation. *Assessment,* 2, 279-183.