SELF-REPORT CONCUSSION RELATED SYMPTOMS

by

SCOTT GENE PILAND

(Under the Direction of Michael S. Ferrara)

ABSTRACT

The purpose of this project was two-fold. The first purpose was to confirm the factorial and construct validity of responses to the Head Injury Scale (HIS) and a measure of symptom severity based upon the Post Concussion Symptom Scale (PCSS) and Graded Symptom Checklist (GSC). The second purpose was to examine the relationships between baseline responses to the self-report measures and variables that may serve to influence composite self-report scores. A priori models were tested using confirmatory factor analysis (CFA). Using an experimental design, scores were compared on each scale between concussed and non-concussed groups. Participants (N = 1065, male n = 805) were college athletes (age of 19.81 ± 1.53 years) from 7 NCAA institutions. Experimental analyses (N = 27, concussed n=17). Two day test-retest reliability was conducted with a sample of healthy college student (n=83). Participants completed baseline measures for two scales and health questionnaire. Experimental analysis was performed on Baseline and Days 1, 2, 3, and 10 post-injury.

Evidence for the reliability, factorial, and construct validity was provided for the 9-item HIS and the 9-item severity scale. Significant interaction on responses to the 9-item HIS and 9-item severity scale were found. Statistical differences between groups were observed on days 1 and 2 post-concussion. Previous concussion history and controllable conditions served to increase baseline responses to each measure. Daily fatigue, physical illness, and orthopedic injury can serve to increase self-report symptom scores. These variables need to be controlled prior to collecting non-concussed baseline measures on self-report symptoms.

INDEX WORDS: Self-Reported Symptoms, Confirmatory Factor Analysis, Asymptomatic, Factorial Validity, Construct Validity, Concussion.
SELF-REPORT CONCUSSION RELATED SYMPTOMS

by

SCOTT GENE PILAND

B.S. Valdosta State University, 1996

M.A. The University of North Carolina at Chapel Hill, 1998

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

2004
DEDICATION

To Aimee.

Philippians 4:13
ACKNOWLEDGEMENTS

This project was funded by a Doctoral Dissertation Grant awarded by the National Athletic Trainers Association-Research and Education Foundation. Funding was also provided by The Kindig Research Awards: The University of Georgia Department of Exercise Science.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ACKNOWLEDGEMENTS</td>
<td>v</td>
</tr>
<tr>
<td>1</td>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>Purpose and Rationale</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>Statement of the Problem</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>Hypothesis</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>REVIEW OF THE RELATED LITERATURE</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Sport-Related Concussion</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td>Neurophysiology of Concussion</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Sport-Related Concussion Defined</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>Post Concussion Syndrome</td>
<td>12</td>
</tr>
<tr>
<td></td>
<td>Sport-Related Concussion Self-Report Symptoms</td>
<td>14</td>
</tr>
<tr>
<td></td>
<td>Psychometric Evaluation of the HIS Self-Report Symptom Scale</td>
<td>18</td>
</tr>
<tr>
<td></td>
<td>Reference List</td>
<td>20</td>
</tr>
<tr>
<td>3</td>
<td>EVIDENCE FOR THE FACTORIAL AND CONSTRUCT VALIDITY OF</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>SELF-REPORT CONCUSSION RELATED SYMPTOM SCALES</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Abstract</td>
<td>30</td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Methods</td>
<td>38</td>
</tr>
</tbody>
</table>
4 INVESTIGATION OF BASELINE SELF-REPORT CONCUSSION

SYMPTOM SCORES .................................................................................67
Abstract .......................................................................................................68
Introduction .................................................................................................70
Methods .......................................................................................................74
Procedure .....................................................................................................76
Data Analysis ..............................................................................................76
Results .........................................................................................................76
Discussion ...................................................................................................78
Conclusion ...................................................................................................81
References ...................................................................................................81
Tables ..........................................................................................................85

5 SUMMARY AND CONCLUSIONS ..............................................................89
CHAPTER 1
INTRODUCTION

Sports-related concussion comprises 20% of the 1.54 million head injuries estimated to occur annually in the United States (Erlanger, Kutner, Barth, & Barnes, 1999). Among 15 to 24 year olds, sports-related concussions are second in frequency to head injuries incurred from motor vehicle accidents (Sosin, Sniezek, & Thurman, 1996). In an epidemiological study of concussion in collegiate and high school football participants, it was reported that 888 (5.1%) of 17,549 participants sustained at least one concussion during a single season of competitive play (Guskiewicz, Weaver, & Padua, 2000). Of the 888 athletes who suffered a concussion, 131 (14.7%) went on to experience a second head injury during the same season. Approximately 30% of the participants who experienced a concussion returned to participation on the same day of injury (Guskiewicz, et al., 2000). This statistic is a cause for concern because at least 17 deaths related to second impact syndrome occurred between 1992 and 1997 (Cantu, 1998).

Second impact syndrome results when consecutive concussive insults take place within a short period of time (days), before complete healing can occur from the initial incident. Most often the outcome of the second injury is more severe (Cantu, 1998). The percent increase in the risk of incurring a second impact syndrome is debated among allied healthcare professionals. Empirical evidence is necessary to accurately represent the relative risk. Until that information is readily available, clinicians should approach a concussed athlete's return to play with utmost caution, using scientifically supported measures on which to base return to play decisions.
Current "return to play" guidelines dictate that prior to returning to activity from a concussive incident, athletes be deemed asymptomatic in both a non-exerted and exerted state for at least 7 days (Kelly & Rosenberg, (AAN) 1997; Cantu, 1998; 2001). A common understanding of the term asymptomatic is that one is free of all symptoms prior to return to play. Concussion symptoms can be thought of as an umbrella term defined by a variety of categories: physical examination, posturography, neurocognitive function, self-reported symptoms, and neuroimaging. Three of these categories (posturography, neurocognitive function, and self-reported symptoms) garner the majority of the sport concussion research focus (Collins, et al., 1999; Erlanger, et al., 1999; Guskiewicz, 2001; Lovell & Collins, 1998; McCrea, et al., 1998).

Guskiewicz provided the analogy that each of these categories comprises a piece of a puzzle which, when put together, represents overall concussion. Though each of these categories can act to describe a concussive injury, no conclusive evidence can be found to support that any single one is an independent marker of sports-related concussion resolution (Guskiewicz, 2001). Hence, it has been recommended that clinicians use a multi-faceted approach, to include the three most common categories highlighted by the literature into their assessments of the injury and its follow up evaluations (Cantu, 1998; Lovell, et al., 1998; Guskiewicz, 2001; Bailes & Cantu, 2001; Aubrey, et al., 2002). With this recommendation, it is important to note the availability of each measure to allied healthcare clinicians. Balance diagnostic equipment and force plate instruments are expensive, time consuming, and limited to a few academic and professional organizations. Though these instruments result in reliable and valid responses, they are often unavailable to the average clinician (Guskiewicz, 2001;
Riemann & Guskiewicz, 2000). Use of neuropsychological tests to evaluate neurocognitive functioning requires that personnel be trained in test administration and interpretation of score meaning. The majority of instruments used in concussion testing batteries is not within the scope of practice for athletic trainers and require a trained neuropsychologist for interpretation (Barr, 2001). These approaches present clear barriers to the clinician (Grindel, Lovell, & Collins, 2001). Consequently, clinicians often base their assessment and return to play decisions upon self-reported symptomology (Lovell, et al., 2003).

Posturography and neuropsychological assessments have both shared the majority of attention by researchers with an ever-growing corpus of empirical evidence being accumulated for both approaches (Collins, et al., 1999; Echemendia, et al., 2001; Erlanger, et al., 1999; Lovell, et al., 1998; Macciocchi, et al., 1996). Conversely, the measurement of self-reported concussion related symptoms has lacked similar scientific investigation. Instruments that have been reported in the literature have not been developed adequately. The two most common self-report symptom instruments, the Post Concussion Symptom Scale (PCSS) and the Graded Symptom Checklist (GSC), have been purported as standards for clinical incorporation (Aubrey, et al., 2002; Guskiewicz, et al. 2003). However, the PCSS and the GSC have been presented with little to no information regarding psychometric development, measurement properties, or even basic instructions for its implementation and interpretation (Lovell, et al. 1998; Guskiewicz, et al. 2003; McCrea, et al. 2003). It is common knowledge within the profession of athletic training that clinicians often rely upon self-report measures of concussion related symptoms that have been developed based upon limited information and little scientific
investigation. (Lovell, et al., 1998; Matser, et al., 2000; Piland, et al., 2003). Recently, evidence has been provided to support a more thorough investigation of concussion related self-reported symptom methods (Piland, et al., 2003).

**Purpose and Rationale**

The purpose of this study is two-fold. Its first purpose is to confirm the factorial validity and construct validity of responses to the Head Injury Scale (HIS) and a measure of symptom severity based upon the PCSS and GSC. The general hypothesis regarding factorial validity is that baseline responses to self-report concussion symptom scales represent a three-factor higher order model. This *a priori* model, originally based upon the theoretical assumption reported in the literature by Maroon (2000), has been strongly supported by investigations of responses to the HIS instrument (Piland, et al., 2003). Construct validity evidence will be obtained via correlational analysis and an experimental condition existing of comparison of responses between injured and non-injured controls at baseline, Day 1, Day 2, Day 3, and Day 10 following the injury.

The second purpose of this study will be to examine the relationships between baseline responses to the self-report measures of concussion symptoms and variables that may serve to influence composite self-report score. Because concussion related symptoms are not singularly related to concussion and can be influenced by other non-concussive conditions we felt it important to evaluate the influences of two uncontrollable variables and three controllable variables. The two non-controllable variables were sex, and previous history of concussion and the three controllable variables were physical illness, orthopedic injury, and fatigue. Evaluation of these variables potential influence upon baseline responses to the HIS and severity scales could
offer invaluable information towards the collection of more representative baseline symptom scores.

   The use of self-reported concussion symptom scales among allied healthcare professionals in the assessment and evaluation of sport-related concussion is widespread and generally accepted by the medical community. However there is little evidence to demonstrate the reliability and validity of these measures.

**Statement of the Problem**

Two major problems in terms of symptom scales exist at the present time: 1) There is no consistent format or administration instructions provided in the literature for the use of a summative self-reported concussion related symptom scale. 2) There is very little to no scientific evidence of the psychometric development and soundness of such a scale. The presence of these two problems can and has lead to confusion among allied healthcare providers. Such confusion could ultimately result in making incorrect and misguided return to play decisions. We will attempt to rectify both of these problems by independently examining the two most contemporary forms of summative self-report symptoms scales: the HIS and severity scale based upon the PCSS and GSC. Recent investigations into the validity of the scales will be continued and expanded upon by evaluating the psychometric properties of the instruments. This will be the first time such information will be provided on a large scale.

**Hypotheses**

The following hypothesis will be independently tested using responses to the HIS and severity scale instruments.
• Test-retest reliability of the responses to the HIS measure of symptom
duration will be demonstrated by a strong correlation (R=.70) between scores
collected from a sample comprised of both football athletes and non-athlete
university students at two time periods.

• Test-retest reliability of the responses to the measure of symptom severity will
be demonstrated by a strong correlation (R=.70) between scores collected
from a sample comprised of both football athletes and non-athlete university
students at two time periods.

• The a priori 3-factor higher order model will be demonstrated, by commonly
accepted indices-of-fit, to provide a less than adequate fit to baseline
responses to the 16-item HIS measure of symptom duration.

• The a priori 3-factor higher order model will be demonstrated, by commonly
accepted indices-of-fit, to provide strong evidence of factorial validity of
baseline responses to the 9-item HIS measure of symptom duration.

• The a priori 3-factor higher order model will be demonstrated, by commonly
accepted indices-of-fit, to provide a less than adequate fit to baseline
responses to the 16-item measure of symptom severity.

• The a priori 3-factor higher order model will be demonstrated, by commonly
accepted indices-of-fit, to provide strong evidence of factorial validity of
baseline responses to the 9-item measure of symptom severity.

• There will be a statistically significant main effect between concussed and
non-concussed composite responses to the self-report symptom scales.
• There will be a statistically significant difference between composite baseline symptom scores between athletes reporting a previous history of concussion and those not reporting a previous history.

• There will be no statistically significant difference between composite baseline symptom scores between sexes.

• There will be a statistically significant difference between composite baseline symptom scores between athletes reporting feeling physically ill and those not reporting to be physically ill.

• There will be a statistically significant difference between composite baseline symptom scores between athletes reporting having an orthopedic injury and those not reporting an orthopedic injury.

• There will be a statistically significant difference between composite baseline symptom scores between athletes reporting feeling fatigued from the testing days activities and those not reporting to be fatigued from the testing days activities.
CHAPTER 2
REVIEW OF THE RELATED LITERATURE

Sport Related Concussion

Traditionally, concussion has been described as a brief transient loss of brain function due to electrical and chemical changes within the brain with no gross lasting structural damage (Alexander, 1995). Experimental animal models and some human research suggest that concussion may be attributed to diffuse focal injury due to shearing forces between white and gray brain matter (Shaw, 2001; Giza & Hovda, 2001; Cantu, 2001).

In the evaluation of sport-related musculoskeletal injury, athletic trainers rely on a thorough understanding of the mechanism from which the injury was generated to assist in the assessment and treatment of the injured athlete. Understanding the cause of a concussive incident is no different. Differing mechanisms may result in a variety of injury presentations. A common mechanism for concussive injury is described as an accelerating - decelerating motion usually related to head to head contact, head to ground contact, or head to sturdy equipment (goal post) contact. The mechanism is associated with coupe and contra-coupe injuries that display shearing forces between the cortex and sub-cortical projections (Barth, 2001; Kelly, 2001). Though found in sport, this type of mechanism is most commonly affiliated with non-sports related concussion resulting from motor vehicle accidents (Mittenburg, 1997). Further, rotational forces, has been depicted as a more serious concussion mechanism. This injury mechanism can be caused by some indirect blow to the face or side of the head (temporal region), which quickly spins the cranium around the brain forcing the brain into a delayed motion. As the brain
is forced into rotation, it rotates with its most superficial structures moving initially followed by the deeper tissues. This mechanism of injury is directly responsible for shearing forces and possible damage between the white and gray matter (Barth, 2001). The first mechanism can be visualized when a player gets tackled or falls onto the playing surface slamming his or her head onto the ground, the second occurs when the player rotates or is spun violently around by the blow of an opponent or contact is made to the side or the face mask resulting in rotation of the head. At this point in concussion research, the relationship between levels of injury severity and causal mechanisms has not been investigated in humans. Animal models provide an understanding about the neurophysiology of concussion and a causal injury mechanism (Wojtys, et al., 1999).

**Neurophysiology of Concussion**

Upon injury, the brain experiences a sudden neuronal depolarization that leads to an indiscriminate triggering of voltage dependent channels which releases several excitatory amino acids, with the primary substance identified as glutamate. Glutamate initiates a neurometabolic cascade that leads to ionic shifts and chemical changes within the brain. 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% (Giza, 2001; Kelly, 2001; Shaw, 2001). The increase in extracellular $K^+$ triggers membrane sodium-potassium pumps to try and re-establish homeostasis. Increase pump action requires ATP, so glycolysis is activated to attempt to meet the energy demand. This occurs in an oxygen-starved environment since cerebral blood flow has been reduced. The increase in extracellular $K^+$ also 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 and has been affiliated with decreases in neuronal firing which leads to neurocognitive dysfunction (Giza, 2001; Kelly, 2001; Shaw, 2001).

In 2001, Giza reported that extracellular $K^+$ increases quickly in humans after the onset of injury but decreases within the following 30 minutes (Giza, 2001). Elevated levels of potassium have been related to decreased mitochondrial function and propagation of the energy deficit by triggering sodium-potassium pump response. $Ca^{2+}$ on the other hand remains at elevated levels from 2 - 4 days post injury (Giza, 2001). 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. Lactate increase can cause a change in the $pH$ of the surrounding area, which has been shown to damage cellular membranes increasing brain vulnerability. The decrease in cerebral brain flow has been documented in humans and has been shown to last up to 4 weeks (Giza, 2001; Kelly, 2001). This condition decreases brain metabolism, which has been related to neuronal dysfunction and decreases in brain function (i.e. neurocognitive function and possibly motor control). This condition is also thought to place the brain into a state of vulnerability where it cannot respond to injury or hyperactivity, therefore increasing the severity of the reaction to a second concussive injury. Lacking a more complete understanding of the biological causal links between injury mechanism and resulting symptom presentation it is difficult for scientist, and clinicians to convey a consensus of the definition of the injury. Until a consensus is reached, the methods for which
symptoms are measured needs to be psychometrically evaluated and evidenced to be representative of sport related concussion.

**Sport Related Concussion Definition**

In 2001, a group of sport-related concussion experts met in Vienna for the 1st International Symposium on Concussion in Sport (Aubrey, et al., 2002). The symposium was organized by the International Ice Hockey Federation (IIHF), the Federation Internationale de Football Association Medical Assessment and Research Centre (FIFA, F-MARC), and the International Olympic Committee Medical Commission (IOC). The group released a summary and agreement statement from this meeting. This statement included a revision of a 35-year-old "consensus" definition put forth by the Congress of Neurological Surgeons and accepted by the American Medical Association (AMA) and the International Neurotraumatology Association. The revised definition reads as follows:

"Concussion is defined as a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces. Several common features that incorporate clinical, pathological, and biomechanical injury constructs that may be used in defining the nature of concussive head injury include:

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

2) 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 largely 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 neuroimaging studies." (Aubrey, et al., 2002)

This revision of the classic concussion definition is important because it recognizes that a loss of consciousness is not a main predictor of injury severity, and that neuroimaging evaluations are not abnormal in individuals with sport related concussion.

The most commonly researched categories of concussion evaluation techniques are: neurocognitive performance, posturography, and self-reported symptoms. These categories have each been demonstrated to be useful in the monitoring of the resolution of concussion, but no single measurement method has been shown to be an independent measure of the injury (Guskiewicz, 2001). Current clinical recommendations suggest the use of a multi-faceted approach to concussion assessment (Cantu, 1998; Lovell, et al., 1998; Guskiewicz, 2001; Aubrey, et al., 2002).

**Post Concussion Syndrome**

Persons sustaining a concussion may develop a number of physical, cognitive, and behavioral symptoms following injury defined as post concussion symptoms. Self-reported symptoms have been recorded as early as the 1600's. The "Learned Dr. Read"
included descriptions within his published medical text on head injury (McCrory, 2001). Since that time, a corpus of authors has presented what have become commonly accepted symptoms of acute concussion. These common symptoms can include: headache, impaired memory, dizziness, fatigue, sensitivity to light, sensitivity to noise, loss of concentration, nausea, vomiting, feeling of being "in a fog", feeling of being "slowed down", sleep disturbances, vision problems, difficulty remembering, irritability, and balance difficulties (Barth, et al., 1983; Barth, et al., 1989; Binder, 1989; Alves, et al., 1993; Macciocchi, et al., 1996; Macciocchi, 1998; Iverson & Lange, 2003).

Persistent symptoms, which may linger following a concussion, fall under a defined syndrome called post concussion syndrome (PCS). This syndrome is defined by the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; DSM-IV; American Psychiatric Association, 1994) as the presence of concussion related symptoms for a minimum of 3 months following concussion. Patients who suffered a concussion from a motor vehicle accident and reported two or more persistent symptoms at 3 months post injury are likely to complain of PCS symptoms up to 12 months or even several years afterward (Alves, et al., 1993; Ferguson, et al., 1999). The cause of PCS and the conditions that maintain it are controversial (Iverson & Lange, 2003). DSM-IV criteria suggest that cerebral dysfunction and structural anomaly are primary causes of PCS. In contrast, the International Classification of Mental and Behavioral Disorders (ICD-10; World Health Organization, 1992) diagnostic criteria and other research suggest that psychological factors and contextual factors influence the prevalence and course of PCS (Binder, 1986; Gouvier, et al., 1992; Iverson & Lange, 2003). Rimel suggested that premorbid substance abuse and or monetary compensation play potential causative roles
in PCS (Rimel, et al., 1981) PCS is believed to be caused by either biological effects of the injury, psychological factors or both (Iverson & Lange, 2003).

In 1992 a study found no differences between MHI and normal groups on the frequency, intensity, and duration of PCS symptoms, although each of these covaried with fluctuating levels of daily stress (Gouvier, et al., 1992). The authors concluded that post concussion symptoms vary in accordance with an individual's level of psychosocial stress, and factors such as coping style or cognitive appraisal will affect the manifestation of PCS. Although the initial cause of PCS may be physiological, psychological factors appear to play a role in its maintenance (Mittenberg & Burton, 1994). An interesting finding of concussion in sports suggest that persistent symptoms occur no more frequently following sport-related MHI than in uninjured controls (Barth, et al., 1989). The rate for which sport related concussions lead to a diagnosis of PCS is unknown, but it has been shown to be substantially less than that of other populations who incur similar injuries. Ferguson and colleagues suggest that self-reported symptoms, like those observed with PCS, resolves rapidly in the majority of athletes because sport participants expect no persisting problems following their injury (Ferguson, et al., 1999).

**Sport Related Concussion Self-Report Symptoms**

The evaluation of self-reported sport-related concussion symptoms began with the landmark work of Barth (Barth, et al., 1983). The four-year prospective study utilized the sport environment as a laboratory to clinically investigate concussive injuries. The study collected non-injured preseason baseline measures of neurocognitive performance of 2,350 college football players. During the span of the study, 195 athletes incurred a concussion. Compared to an orthopedic injury group and uninjured control group, a
statistically significant higher number of the concussed athletes reported experiencing selected symptoms up to 10 days after their injury (Barth, et al., 1983). In 1993, the prevalence of post concussion symptomology in 587 patients with non sports-related uncomplicated, mild head injury was studied (Alves, et al., 1993). This group evaluated responses to a self-report measure comprised of 16 commonly accepted post concussion symptoms. This self-report scale, developed from the initial works of Barth, investigated the level of symptom severity experienced by respondents. A linear decrease in symptomology over a 1-year follow-up period was found. This instrument was also used to evaluate the prevalence of symptom severity in 138 athletes with a single incidence of concussion (Macciocchi, et al., 1996). The athletes completed a pre-season base-line report and upon injury completed the same self-report scale 24 hours after injury, 5 days after injury and 10 days after injury. It was found that concussed athletes tested 24 hours and 5 days post injury reported statistically significant increases in the severity of headaches, dizziness, and memory problems relative to controls compared to the injured subjects. Another study presented results of a statistically significant difference between injured and control group reports of symptom severity 2 hours following injury with no increased symptom severity reported following 48 hours (Echemendia, et al., 2001). In 2000, Guskiewicz and associates used a symptom severity scale to record commonly reported symptoms by injured players. The most common symptoms reported were headache followed by dizziness, and confusion. Each of these studies provides support to the value of assessing self-reported symptoms, but a lack of evidence to demonstrate the instruments psychometric measurement properties has not been provided.
An instrument similar to the original 16-item severity scale used by Barth and associates was presented as the Post-Concussion Symptoms Scale (PCSS) in a paper discussing the use of neuropsychological evaluation in collegiate football players in 1998 (Lovell, et al., 1998). The authors stated the importance of using a standardized measurement tool for the evaluation of self-reported symptoms but failed to provide any explanation of the psychometric development of the PCSS or any empirical evidence of the reliability and validity of the summative measure. The PCSS has been incorporated into the concussion assessment battery of the National Hockey League as well as the National Football League (Cantu, 1998). Recently, the summary and agreement statement of the First International Conference in Sport supported the PCSS as the standardized scale for use in the assessment of concussion (Aubrey, et al., 2002). In 2003 Collins and associates used concussed athlete responses to the PCSS to demonstrate that those athletes reporting a post concussive headache 7 days following the concussive incident also reported a larger number of other concussive symptoms on the same day compared to athletes who did not report a post concussive headache (Collins, Field, Lovell, Iverson, et al., 2003). Also, in a separate study, it was suggested that post concussive responses to the PCSS did not predict deficits in measures of neurocognitive performances (Collins, et al., 2003).

While the PCSS has been recommended for use by the 1st International Symposium on Concussion in Sport, there are variations in use (Lovell, et al., 1998; Collins, et al., 2003). The Graded Symptom Checklist (GSC) instrument has been presented in multiple publications (Guskiewicz, et al. 2003; McCrea, et al. 2003). Like the PCSS, the GSC requires respondents to rate individual symptoms on a 7-point Likert-
type scale with the response options of none (0) to severe (6), and the instructions of "how would you describe your symptoms". Other self-report concussion symptom scales differ from the PCSS by administering the scale with the anchors of never (0) to always (6) combined with the instructions to report how long the symptoms have been present within the previous 24-hour period. Both types of formats are found in the sport-concussion literature (Barth, 1983; Macciocchi, 1996; Aubrey, et al., 2002; Gouvier, 1992; Piland, et al., 2003; Peterson, et al., 2003; Lovell, et al., 2003).

Regardless of the format, strong recommendation for the collection and assessment of self-reported symptomology was made by Maroon (Maroon, et al., 2000). The authors describe the use of neuropsychological testing batteries to evaluate cognitive deficits and recommend compilation of player's reports of post concussion symptoms. They suggest that reported symptoms can be categorized by one of three groups: 1) somatic symptoms including, headaches, dizziness, balance problems, or nausea; 2) neuropsychological symptoms including, anxiety, depression, or irritability; and 3) cognitive symptoms including, impairment of attention, memory, processing speed. These groupings formulate a basis for establishing relationships between symptoms that may lead to the ability to understand the meaning of symptom reports. The author's recommendations helped to establish the theoretical underlying structure of baseline self-reported symptoms. It is theorized that baseline responses to a concussion symptom scale will form into a cohesive group of symptoms that can be described by a three-factor higher order latent variable model. This cohesive group of baseline symptoms will support the main effects found between the symptom reports of uninjured control and concussed athletes.
Evidence to support the reliability and validity of empirically driven scales must be present before strong inferences based upon the scores can be made (Muthen & Kaplan, 1985). The degrees to which strong inferences can be made depend upon the foundations of the reliability and validity of responses to an instrument (Messick, 1995). Currently clinicians are using symptom scales as empirical measures of either symptom severity or duration and they are basing return to play decisions upon these scores. Until appropriate empirical evidence to support the reliability and validity of such measures is presented, the inferences made from responses to self-report symptom instruments should be done so with caution.

Psychometric Evaluation of the HIS Self-Report Symptom Scale

In 2003, Piland and associates used cross sectional data from 270 non-injured university athletes (males = 223) to provide evidence to demonstrate the strengths and weaknesses of a 16-item self-report symptom scale developed to be representative of instruments described in the literature. The instrument was entitled the Head Injury Scale (HIS). The HIS is intended to be a measure of symptom duration, as reported on a 7-point Likert-type scale with the anchors of "never", "sometimes", and "always". To test for factorial and construct validity the data from the sample was used to evaluate the fit of a theoretically derived three-factor higher order model using Confirmatory Factor Analysis (CFA). The a priori structure of the underlying responses was based upon the previously reported theoretical assumption that concussion-related symptoms should be categorized into three areas (Maroon, et al., 2000). Based upon commonly accepted criteria results of the CFA indicated that the fit of the model to the data from the 16-item HIS was a good but not an excellent fit (Jöreskog, 1993; Jöreskog & Sörbom, 1996; Bentler & Bonett,
1980; Bentler, 1990). Hence, the 16-item HIS was modified by removing symptoms that offered poor support for the model based upon theoretical reasoning (i.e., did the symptom make clinical sense, did the item have a double meaning i.e. "sensitive to Light/Noise") and substantiated by large standardized residuals. The resulting 9-item HIS was then analyzed using the same method. It was found that the theoretically derived model was an excellent fit to the 9-item HIS (Jöreskog, 1993; Jöreskog, et al., 1996; Bentler, et al., 1980; Bentler, 1990). The findings supported the presence of a 3-factor a priori structure within the summative score responses of healthy male and female collegiate athletes. This suggested that the previously mentioned areas could best explain data obtained from a baseline self-reported symptom scale: somatic symptoms, cognitive symptoms, and neuropsychological symptoms (Maroon, et al., 2000).

Factorial validity and construct validity was further supported by the second part of the study which involved the analysis of data collected from concussed collegiate athletes (n=17) and paired healthy controls (n=16) using the 9-item HIS scale. Results supported the use of an empirically driven symptom scale by demonstrating statistically significant days by group interaction \([F(4, 28) = 6.38, p = .001, \eta^2 = .542, \varepsilon = .595]\). Statistical differences between groups were observed on Day 1 and Day 2, but not on Day 3 and Day following injury.

This initial investigation provided excellent evidence for the presence of the 3-factor higher order model, and hence, support for its factorial validity. But, due to the size of the sample, re-estimation of the re-specified model (9-item model) was not cross validated on another sample drawn from the same population for which the results are to be generalized. Thus, the same model needs to be specified in a much larger sample
Such a study could confirm the original and re-specified model and would provide exceptional evidence of factorial validity of the measure, which according to Messick (1995) is the highest form of validation evidence, being that all other types of validity rely on the presence of factorial validity for there calculation. If this evidence is realized then the inferences from the composite scores of the HIS instrument can be appropriately utilized.

Reference List


CHAPTER 3

EVIDENCE FOR THE FACTORIAL AND CONSTRUCT VALIDITY OF SELF-REPORT CONCUSSION RELATED SYMPTOMS SCALES

Piland, S.G. To be submitted to Neurosurgery.
Objective

This study was performed to replicate findings from a previous study that evaluated the factorial and construct validity of baseline responses to the 16-item Head Injury Scale (HIS). This study also evaluated the factorial and construct validity of a 16-item and 9-item severity scale among a sample of male and female college athletes.

Design and Setting

Using a cross-sectional design, we tested two competing models of the HIS and two competing models of the severity scale using confirmatory factor analysis (CFA). Construct validity was assessed using Pearson product-moment correlation analyses. Using an experimental design, we compared scores on the each scale between concussed and non-concussed groups using a 2 (groups) × 5 (time) mixed model analysis of variance.

Subjects

Participants (N = 1065) in the cross-sectional analyses were predominately male (n = 805) college athletes with a mean age of 19.81 ± 1.53 years from 7 NCAA institutions. Participants (N = 27) in the experimental analyses were concussed (n = 17) and non-concussed control (n = 10) college athletes with a mean age of 20.51 ± 1.92 years.
Measurements

Participants completed baseline measures for the two scales and a brief health history questionnaire. Concussed individuals and controls were evaluated on Baseline and Days 1, 2, 3, and 10 post-injury.

Results

Using CFA, the theoretically derived three-factor model did not provide a good fit to either the 16-item HIS or the 16-item severity scale, thus competing 9-item models were tested. The subsequent analysis indicated that the three-factor model provided a good fit to the 9-item HIS as well as the 9-item severity scale. The three-factors for each scale were best described by a single second-order factor, namely concussion symptoms. Scores from the 16-item and 9-item HIS and 16-item and 9-item severity scale were strongly correlated.

The $2 \times 5$ ANOVA on the 16-item HIS demonstrated a significant groups by days interaction. Statistical differences between groups were observed on days 1 and 2 post-concussion. The $2 \times 5$ ANOVA on the 9-item HIS demonstrated a significant groups by days interaction. Statistical differences between groups were observed on days 1 and 2 post-concussion.

The $2 \times 5$ ANOVA on the 16-item severity scale demonstrated significant groups by days interaction. Statistical differences between groups were observed only on day 1. The $2 \times 5$ ANOVA on the 9-item HIS demonstrated a significant groups by days interaction. Statistical differences between groups were observed on days 1 and 2 post-concussion.
Conclusions

This study confirmed previous findings by providing evidence for the factorial and construct validity of the HIS among collegiate athletes. This study also provided initial evidence for the factorial and construct validity of the 9-item severity scale. These scales might aid in the making of return to play decisions by physicians and athletic trainers.

Key Words:

Confirmatory Factor Analysis, Factorial Validity, Construct Validity, Self-Report Symptoms
Introduction

The assessment of sport-related concussion is becoming one of the most popular areas of sports medicine research. This is evidenced by the growing corpus of information that is being compiled and produced (15, 19, 22, 35, 36). The number of published scientific articles related to sport-related concussion has nearly tripled in the past 5 years. Advances towards the understanding of this injury and its proper assessment, improvements are being made. However, many questions related to the basic science of measurement remain unanswered in the evaluation paradigm.

As a result of increased research the recommendation that sports medicine clinicians utilize a multi-faceted approach in their assessment of concussion (3, 25, 27, 32, 35, 39, 43, 45). This suggestion reflects the complexity of the injury itself in that concussions symptoms (loss of consciousness, amnesia, headache, difficulties with balance, and neuropsychological deficits) could present in a variety of ways. Concussions are unlike most orthopedic injuries which can be isolated to a specific anatomical structure based upon mechanism of injury and evaluative tests. The location and level of concussive injury to the brain is highly variable due to the direction and magnitude of force (24, 48). Hence, the effects (symptoms) are not always predictable and no single measure is be able to capture the full ramifications of the concussion injury. Presently, evidence suggests that there is no single biological marker of the injury. Thus, it is essential that all available measures, including but not limited to, posturography, neurocognitive performance, and self-report symptoms, have adequate evidence of score reliability and validity.
Score validation, which is an on-going and continual process, has been predominately reported for posturography (27, 28, 46) and neurocognitive measures (1, 4, 5, 14, 16, 35, 37, 39) with only recent evidence being provided for a measure of self-report symptoms (45). Self-report symptoms can be evaluated either by administering a non-summative checklist or summative scale. Checklists are typically comprised of a list of concussion related symptoms and are used to provide basic information concerning the presence or absence of a specific symptom. Checklists should not be affiliated with any type of composite scoring structure. Summative scales do have an associated scoring structure from which inferences can be drawn (45). Known scales purport to measure either the length of time one has experienced the symptoms (duration) (43, 45) or how severe the symptoms have felt (severity) (4, 26, 35, 41). Inferences made from such composite scores are only as strong as the validity evidence available to support the scores meaning (42). Initially, experts suggested that clinicians incorporate a checklist type of self-report symptom evaluation for use in concussion assessment (12), but a composite score approach has gained in popularity (3, 26).

Measures of self-report symptoms associated with summative scores can potentially offer sports medicine clinicians greater information than that obtained from a simple checklist. This potential has lead to recent acceptance for their use, as well as driven the research to evaluate changes in summative self-report scores following a concussive injury (15, 16, 25, 26, 41). However, the psychometrics and measurement properties of the instruments must be thoroughly evaluated and tested prior to making inferences from obtained scores regarding return to play decisions (25, 26). An approach that strengthens the ability to make safer return to play decisions is the availability and
interpretability of pre-concussion baseline information. This is of tremendous importance to the development of a self-report symptom scale due to the fact that symptoms that are typically related to concussion (headache, nausea, difficulty concentrating, fatigue etc…) are also typically related to a host of other causes (illness, activity, non-concussive injury, environmental stressors, psychological disorders, etc…) (23, 30). Thus, psychometric and measurement properties must first be established in a non-concussed athletic population. This process allows researchers and clinicians to understand and interpret how athletes typically experience symptoms that are related to concussion when athletes are not concussed.

Scale development is a lengthy process, which should provide ample information as to the interpretability of obtained scores. This result is accomplished by first providing a theoretical assumption as to the underlying measurement structure of the observable items (symptoms) which comprise the scale. Performing an exploratory factor analysis (EFA) on non-concussed data to discover how responses to each of the variables vary or covary, is one approach while another requires researching the available literature to find how experts predict each of the variables should interact. Secondly, this theoretical assumption must be tested for evidence of factorial validity. Validity refers to the degree to which the evidence supports the inferences that are made from the scores. Though there are numerous ways to approach the issue of validation, validity is a unitary concept. The unified theory of validity outlines six types of evidence for score validity. One is factorial validity, which describes how well an underlying dimension can be used to summarize relationships between item responses. This distinguishable aspect of score
validity is of utmost importance because it provides a basis for the use of summative scores for all other tests of score validity (42).

In 2000 Maroon and associates suggested that individual concussion related symptoms are represented by one of three groupings: Cognitive, Neuropsychological, and Somatic (39). This recommendation provided the basis for the theoretical assumption of the underlying structure of responses to summative scales. Though it has remained untested in responses to the two most commonly reported summative self-report symptom scales, the Post Concussion Symptom Scale (PCSS) (3, 35, 39) and the Graded Symptom Checklist (GSC) (26, 41). Both scales incorporate a range of symptoms that requires responses to a 7-point Likert-type scale with response options ranging from [0] “no symptom”, [1] “mildly severe”, [3] “moderately severe” to [6] “severe.” Responses to each concussion related symptom are summed to produce an overall composite severity score. The numbers of symptoms included on either scale has varied with appearances in the literature. The PCSS has ranged from 16 to 22 items, while the GCS was most recently reported to consist of 17 concussion related symptoms (4, 14, 15, 16, 26, 37). Baseline composite scores to these similar instruments have also varied reaching as high as 10(±15.1) (16) to as low as .99 (±3.26) (26). Post concussion changes in score reports between injured athletes and non-injured controls have been reported to be as brief as 48 hours to as long as 7 days (16, 19, 20, 26, 27, 34, 45). Even with the great amount of variance in results, the First International Symposium for Concussion in Sport has endorsed the PCSS as the standard for clinical incorporation (3). To date, the PCSS and GSC have been presented with little to no information regarding their psychometric
development, measurement properties, or even basic instructions for implementation and interpretation.

The theoretical assumption of three underlying representations of individual symptoms has been formally tested against responses to a self-report measure of concussion symptom duration, using the Head Injury Scale (HIS) (45). The HIS is comprised of 16 symptoms (headache, nausea, vomiting, balance difficulty, sensitivity to light and noise, numbness and tingling, fatigue, drowsiness, trouble falling asleep, sleeping more than usual, sadness, nervousness, difficulty concentrating, feeling “in a fog”, feeling “slowed down”, difficulty remembering) known to be related to concussion. Responses are given for each symptom to a 7-point Likert-type scale with the response option representing the length of time during a 24-hour period that the athlete experienced the symptom. Using confirmatory factor analysis (CFA) the study reported that a three factor higher order model provided a good fit to the 16-item HIS. Then, based upon clinical reasoning and substantive evidence, 7 items were removed from the scale. The resulting 9-item HIS model (headache, nausea, balance difficulty, fatigue, drowsiness, trouble falling asleep, difficulty concentrating, feeling “in a fog”, feeling “slowed down”) was an excellent fit to the data, thus providing strong factorial evidence. However this study was limited by the sample size (N=279), and it did not provide direct evidence that scales measuring severity would possess the same underlying structure (45).

The purpose of this project was to evaluate the factorial and construct validity of the HIS and a measure of self-report symptom severity based upon the GSC and PCSS using a strictly confirmatory approach CFA on baseline data from preseason concussion
testing of non-injured male and female collegiate athletes. Further evaluation of the construct validity of the 16 and 9-item models of the two scales were based upon comparisons of scores between groups of concussed and non-concussed collegiate athletes from before to after a concussive episode.

**Methods**

**Participants**

All participants signed an informed consent document and the study was conducted in accordance with rules and regulations designed for protection of human subjects. Participants (N = 1065) in the cross-section analysis were sports participants attending one of 7 NCAA institutions (Table 1). Participants ranged from 18 to 27 years of age (Mean age = 19.81 ± 1.53 years). The sample was predominantly male (n= 805).

Participants (N = 27) in the experimental analyses consisted of concussed (n = 17) and non-concussed (n=10) NCAA collegiate athletes. Participants’ ages ranged from 18 to 27 years of age, (Mean age = 20.51 ± 1.9 years). We used the American Academy of Neurology guidelines in which the injured group was comprised of three Grade I, 12 Grade II, and two Grade III concussions (3). Two concussions were from women’s basketball with the remaining incurred in football.

**Instrumentation**

*Symptom Duration*

The Head Injury Scale (HIS) was comprised of 16 symptoms commonly affiliated with concussion (45). Instructions to the respondent read: “Here is a list of symptoms that people often feel when they have a concussion. After reading each symptom please circle the number that best describes how long you have experienced the symptom during the
previous 24-hour period (today).” Items were rated on a 7-point Likert-type scale with the response options: [0] I have never experienced this symptom, [1] I have experienced this symptom very briefly today, [3] I have experienced this symptom sometimes during today (about half the day long), [6] I have always experienced this symptom today (all day long). Item responses were then summed to provide an overall composite score.

Symptom Severity

The scale we used to measure severity was designed to represent the two severity related scales common to the literature (PCSS and GSC) (26, 35, 41). The severity scale incorporated the same 16 symptoms from the HIS. Instructions to the respondent read: “The previous scale asked you to report how long you have experienced each of the listed symptoms during the previous 24-hour period (today). Now we would like for you to please circle the number that best describes the way you have been feeling during the previous 24-hour period (today).” As with the HIS, items were rated on a 7-point Likert-type scale, but response option were written to reflect item severity: [0] I have not experienced this symptom, [1] I have felt mild problems with this symptom today, [3] I have felt moderate problems with this symptom today, [6] I have felt severe problems with this symptom today. Item responses were then summed to provide an overall composite score.

Procedure

The baseline measures were collected by certified athletic trainers (ATC) trained in the administration of this self-report battery along with a brief health questionnaire. Administrators were provided with pre-assigned packets of the self-report battery. To address the possibility of ordering effects associated with responses to the lists of 16
items, each packet contained three alternate forms for each scale. There was a high
correlation (r >.95) found between each of the three forms (N=57). The front of each
form included the 9-item grouping of symptoms reported to have the strongest factorial
validity evidence (45). The back of the form was comprised of the remaining 7
symptoms. Because initial validity evidence has been provided for the HIS and not the
severity scale, we ordered the forms so that respondents would always complete the HIS
prior to completing the severity scale.

Participants in the experimental part of this study were members of the original
sample collected during baseline testing. The participants either sustained a concussion or
were non-concussed individuals who were paired with concussed athletes based upon
administration site, gender, sport, and/or age. Upon the assessment of concussion by the
site administrator (ATC), athletes were administered the self-report battery for 10
consecutive days, beginning 24 hours after the concussion episode. Control participants
were assessed at the same intervals. Instructions for the scales were similar to those given
at baseline for both symptom scales.

Data Analysis

Confirmatory Factor Analysis

Each scale and item grouping was independently tested using a strictly
confirmatory approach confirmatory factor analysis (CFA) with maximum likelihood
estimation in AMOS 5.0 (Small Waters Corp., Chicago, IL) (2). Maximum likelihood
(ML) estimation assumes that the data represent a multivariate normal distribution; this
was not the case in the present sample based on the ordinal nature of the items and the
inflated normalized multivariate estimates of skewness and kurtosis for the 16 and 9-item
groupings of the HIS and 16 and 9-item groupings of the severity scale (Mardia’s multivariate coefficient= 574.84 [CR= 390.82] 16-item HIS, 161.17 [CR= 186.90] 9-item HIS, 764.69 [CR=519.90] 16-item severity, 204.16 [CR= 236.79] 9-item severity) (11, 38). Violation of multivariate normality inflates the computed chi-square value and has a tendency to moderately to severely underestimate standard errors (43, 47). Thus, analysis was conducted employing the Bootstrapping technique offered by the AMOS software. Bootstrapping is a way of estimating standard error and significance based not on assumptions of normality, but on empirical re-sampling with replacement of the data (11, 20, 49). The AMOS software also calculates confidence intervals for which to test how the estimates are biased by non-normality. Statistically significant confidence intervals suggest that the estimate is significantly different than zero which justifies appropriateness of the drawn paths of the model (11, 18, 21).

**Model Specification**

Model specification is important for subsequent replication of our results with new data sets. Hence, the initial measurement model underlying the HIS and severity scale contained three latent variables of cognitive, neuropsychological, and somatic symptoms as seen in Figure 1. The measurement model was specified for input into AMOS 5.0 using standard procedures for establishing parameters in matrices containing factor loadings, factor variances and covariances, and item uniqueness. The matrix containing factor loadings was specified to reflect simple structure such that items on the scale being analyzed loaded on only one of the latent variables. The factor loading for the first indicator on each latent variable was constrained to be 1.0 to establish the metric of each latent variable (31).
We also tested the fit of a higher-order model to the scales that is similar to Figure 2. The higher-order model was specified such that a single second-order factor (i.e., concussion symptoms) described the covariances among the three first-order factors (i.e., cognitive, neuropsychological, and somatic symptoms). The higher-order model was specified for input into AMOS 5.0 using standard procedures for establishing parameters in matrices containing factor loadings, factor variances and covariances, item uniqueness, and path coefficients. The matrix containing factor loadings was specified to reflect simple structure. The factor loading for the first indicator on each latent variable was constrained to be 1.0. The matrices of factor variances and covariances and item uniqueness were specified to be diagonal. The matrix of path coefficients was specified to link a single second-order factor to the three first-order factors. The path coefficient for the first latent variable was constrained to be 1.0 to establish the metric of the second-order factor (31).

**Model Fit**

We employed the chi-square ($\chi^2$) statistic, the Non-Normed Fit Index (NNFI), Comparative Fit Index (CFI), and Root Mean Square Error of Approximation (RMSEA) to evaluate the fit of the models (6, 7, 8, 9, 10, 29, 50). The chi-square statistic assessed absolute fit of the model to the data, but it is sensitive to sample size and assumes the correct model (9, 31). The NNFI is a type-II incremental fit index and tests the proportionate improvement in fit by comparing the target model to a baseline model with no correlations among observed variables (6). The CFI is a type-III noncentrality based fit index and tests the relative improvement in fit by comparing the target model to a baseline model with no correlations among observed variables (7). The NNFI and CFI
values should approximate 0.90 and 0.95 to be indicative of reasonably acceptable (6, 9) and good model-data fit (29). The RMSEA represents closeness of fit or “error per degree of freedom” (10). The RMSEA value should approximate .08, .05, and 0 to indicate reasonably close, moderately close, and exact fits, respectively (29, 50). ML and Bootstrapped parameter estimates and standard errors were calculated by the AMOS software. This information was used to determine the effects of non-normality and the correctness of the drawn paths.

**Experimental Condition**

The second part of this study utilized a 2 (Groups: injured and non-injured control) x 5 (Time: testing days) mixed model ANOVA based on the multi-variate F-statistic (Pillai-Bartlett) to independently evaluate differential changes in HIS and severity scale scores between the injured and control groups across days. Effect sizes associated with the F-statistics were expressed as eta-squared ($\eta^2$). The Greenhouse-Geisser epsilon ($\varepsilon$) was reported when the sphericity assumption was violated (i.e., if Mauchly’s test of sphericity was statistically significant at $p < .05$). Post-hoc analysis of differences between injured and control athletes on individual testing days were performed using independent samples t-tests with one-tailed p-value. The $\alpha$ value was adjusted for family-wise comparison using Bonferroni method. Effect sizes between groups were expressed as Cohen’s d (i.e., injured mean minus non-injured mean divided by the pooled SD). Data analysis was performed using Statistical Package for the Social Sciences, *SPSS 11* (Chicago, IL. 2003).
Results

Descriptive Statistics

The means, standard deviations, skewness, and kurtosis values for the individual items on the HIS and severity scales at baseline are reported in Table 2. The means and standard error of the means for the 16 and 9-item HIS (Figure 3 and 4) and the 16 and 9-item severity (Figure 5 and 6) scales between concussed and non-concussed groups across days are graphically represented in Figures 3 - 6.

Reliability

Investigation of the test-retest reliability of the two scales was conducted over a period of two-days with a sample (N= 83) of athletes and non-athletes. The two-day test-retest reliability of the 16 and 9-item HIS was R=.91, and .85. The two-day test-retest reliability of the 16 and 9-item severity scale was R = .94, and .86. Internal consistency as measured by Cronbach’s coefficient alpha ranged from .94 to .97.

CFA on the 16-item HIS

The 3-factor measurement model did not represent an acceptable fit to the 16-item HIS ($\chi^2 = 942.02, \text{df} = 101, \text{RMSEA} = .08 \ [90\% \text{ CI} = .08 - .09], \text{NNFI} = .83, \text{CFI} = .86$). The chi-square statistic was significant ($p < .05$), but this statistic assumes the correct model and is affected by sample size. The RMSEA value met the reasonable threshold value of 0.08 but the 90% CI around the RMSEA point estimate did not improve below the acceptable threshold. The CFI and NNFI values were well below the minimally accepted threshold value of .90 (29). The standardized covariances between latent constructs were .82 (Somatic and Neuropsychological), .85 (Somatic and Cognitive), and .95 (Neuropsychological and Cognitive). The mean of the standardized factor loadings
was .60, with a median of .59; the standardized factor loadings ranged between .35 and .72. Therefore, there was little support for the 3-factor model for the 16-item HIS.

**CFA on the 9-Item HIS**

The 3-factor measurement model represented a good fit to the 9-item HIS ($\chi^2 = 216.28$, df = 24, RMSEA = .08 [90% CI = .07 - .09], NNFI = .92, CFI = .94). The chi-square statistic was significant ($p < .05$), but this statistic assumes the correct model and is affected by sample size. The RMSEA value was 0.08, and the lower bound of the 90% confidence interval was below the threshold. The NNFI and CFI provided moderate support for the 9-item, three-factor model. The NNFI and the CFI values exceeded the 0.90 standard (29) and the CFI approached .95. The standardized covariances between latent constructs were .78 (Somatic and Neuropsychological), .87 (Somatic and Cognitive), and .95 (Neuropsychological and Cognitive). The mean of the factor loadings for the 9-item HIS was .67, with a median of .67; the factor loadings ranged between .48 and .79. Consequently, there was moderate support for the three-factor model to the 9-item HIS.

We tested the fit of a single second-order factor to describe the covariances among the three first-order factors. The higher-order model represented a good fit to the 9-item HIS ($\chi^2 = 216.28$, df = 24, RMSEA = .08 [90% CI = .07 - .09], NNFI = .92, CFI = .94), and fit identically compared with the correlated, three-factor measurement model. This was expected as the higher-order model contained only three first-order latent variables (9). Therefore, there was moderate support for a single, second-order factor underlying the three, first-order factors underlying the 9-item HIS.
Internal Consistency

The internal consistency of the 16-item HIS and 9-item HIS was estimated using Cronbach’s coefficient alpha (17). The estimates of internal consistency for the 16-item HIS and the 9-item HIS were 0.88 and 0.85, respectively. The reason for the slight difference in estimates of internal consistency between the 16-item and 9-item HIS involves the number of test items; the number of items on a self-report scale positively biases coefficient alpha.

CFA on the 16-Item measure of severity

The 3-factor measurement model failed to represent an acceptable fit to the 16-item severity scale \( (\chi^2 = 1110.18, \text{df} = 101, \text{RMSEA} = .09 [90\% \text{ CI} = .09 - .10], \text{NNFI} = .79, \text{CFI} = .82) \). The chi-square statistic was significant \( (p < .05) \). The RMSEA value, including the lower bound of the 90% confidence interval, exceeded the acceptable threshold value of 0.08. The CFI and NNFI values were well below the minimally accepted threshold value of .90 (6). The standardized covariances between latent constructs were .81 (Somatic and Neuropsychological), .88 (Somatic and Cognitive), and .92 (Neuropsychological and Cognitive). The mean of the standardized factor loadings was .59, with a median of .58; the standardized factor loadings ranged between .37 and .71. Hence, adequate support for the 3-factor model for the 16-item severity scale was not provided.

CFA on the 9-Item measure of severity

The 3-factor measurement model represented a good fit to the 9-item severity scale \( (\chi^2 = 239.82, \text{df} = 24, \text{RMSEA} = .09 [90\% \text{ CI} = .08 - .10], \text{NNFI} = .90, \text{CFI} = .93) \). The chi-square statistic was significant \( (p < .05) \). The RMSEA value exceeded the
reasonable threshold of .08, but the lower bound of the confidence interval approached an acceptable level. The NNFI and CFI provided support for the 9-item, three-factor model. The NNFI and the CFI values met and exceeded the acceptable threshold of .90. The standardized covariances between latent constructs were .85 (Somatic and Neuropsychological), .92 (Somatic and Cognitive), and .93 (Neuropsychological and Cognitive). The mean of the factor loadings for the 9-item severity scale was .65, with a median of .64; the factor loadings ranged between .50 and .77. Consequently, there was good support for the three-factor model to the 9-item severity scale.

We tested the fit of a single second-order factor to describe the covariances among the three first-order factors. The higher-order model also represented an acceptable fit to the 9-item severity scale ($\chi^2 = 239.82$, df = 24, RMSEA = .09 [90% CI = .08 - .10], NNFI = .90, CFI = .93) and fit identically compared with the correlated, three-factor measurement model. This was expected as the higher-order model contained only three first-order latent variables (9). Therefore, there was support for a single, second-order factor underlying the three, first-order factors underlying the 9-item measure of severity.

**Internal Consistency**

The internal consistency of the 16-item and 9-item severity scale was estimated using Cronbach’s coefficient alpha (17). The estimates of internal consistency for the 16-item and the 9-item severity scale were 0.87 and 0.84, respectively. As stated before, the reason for the slight difference in estimates of internal consistency between the scales involves the number of test items.
**Correlation Analysis**

Composite scores to the 16-item and 9-item HIS were strongly correlated with Pearson r value of .96. Responses to the 16-item and the 9-item severity scale were also strongly correlated (r=.96). HIS 16-item composite score shared a strong relationship with the composite score of the 16-item severity scale (r=.89), and a Pearson r value of .87 was demonstrated between responses to the 9-item duration and severity scales. A paired samples t-test demonstrated that responses to the HIS were statistically higher than responses to the severity scale (t (1064) =9.07, p<.05, d = .2).

**Experimental Condition Analysis**

*Duration Scale (HIS)*

The 2 x 5 mixed model ANOVA on the 16-item HIS demonstrated a significant group by time interaction [F (4, 20) = 13.86, p = .026, η²=.395, , = .525]. There was not a statistically significant difference between concussed and non-concussed groups on baseline HIS scores (t (25) = .55, p = .582, d = .22). There were large, statistically significant differences between groups on 16-item HIS scores on Day 1 (t (25) = 3.08, p = .008, d = 1.28) and Day 2 (t (25) = 2.59, p = .016, d = .99). There were non-significant differences between groups on Day 3 (t (25) = 1.65, p = .113, d = .63) and Day 10 (t (24) = -.749, p = .461, d = -.28) (see Figure 3).

The 2 x 5 mixed model ANOVA on the 9-item HIS demonstrated similar results with a significant group by time interaction [F (4, 20) = 4.93, p = .006, η²=.484, , = .599]. There was not a statistically significant difference between concussed and non-concussed groups on baseline HIS scores (t (25) = .753, p = .459, d = .32). There were large, statistically significant differences between groups on 9-item HIS scores on Day 1
(t (25) = 3.90, p = .001, d = 1.54) and Day 2 (t(25) = 2.72, p = .012, d = 1.15), but not on Day 3 (t (25) = 1.44, p = .161, d = .61) or Day 10 (t(24) = -.522, p = .607, d = -.20) (see Figure 4).

Severity scale

The 2 × 5 mixed model ANOVA on the 16-item severity scale demonstrated a significant group by time interaction [F (4, 20) = 3.85, p = .017, η² = .423,  = .550]. There was not a statistically significant difference between concussed and non-concussed groups on baseline severity scale scores (t (25) = .131, p = .897, d = .05). There were large, statistically significant differences between concussed and non-concussed groups on 16-item severity scores for only Day 1 (t(25) = 3.12, p = .005, d = 1.22). There were non-significant differences between groups Day 2 (t (25) = 1.98, p = .058, d = .82), Day 3 (t (25) = 1.25, p = .221, d = .50), and Day 10 (t(24) = -.920, p = .367, d = -.34) (see Figure 5).

The 2 x 5 mixed model ANOVA on the 9-item severity scale demonstrated similar results with a significant group by time interaction [F (4, 20) = 3.93, p = .015, η² = .429,  = .559]. There was not a statistically significant difference between concussed and non-concussed groups on baseline HIS scores (t (25) = .099, p = .922, d = .04). There were large, statistically significant differences between groups on 9-item severity scores on Day 1 (t(25) = 3.37, p = .002, d = 1.36) and Day 2 (t(25) = 2.11, p = .045, d = .88), but not on Day 3 (t (25) = 1.35, p = .189, d = .55) or Day 10 (t(24) = -.825, p = .418, d = -.32) (see Figure 6).
Discussion

Increased research in the area of sport related concussion is providing sports medicine clinicians with more tools to incorporate into their concussion assessment protocols. Being that a multi-faceted approach is needed for concussion assessment it is crucial that all available tools undergo rigorous reliability and validity testing. Once evidence is made available, the discussion of the underlying reasons for the implementation of each type of measurement approach is also important. Self-report symptoms are taken as a part of the general assessment of all athletic injuries and have historically played an important role in the assessment of concussion. Current return to play guidelines require that an athlete be asymptomatic at both rest and exertion prior to being cleared for return back into competition (12, 13, 33). Most concussion grading scales utilize the presence of self-report symptoms as markers for a more severe level of concussion (12, 13, 33). Obtaining self-report symptom reports from uninjured and injured athletes, whether by checklist or summative scale, is the most accessible and cost effective concussion assessment tool available to sports medicine clinicians. This need for accessible, clinically relevant assessment tools has led to the acceptance and increased use of summative scales by sports medicine clinicians without adequate evidence as to the reliability and validity of such approaches.

Factorial Validity

Our study is the second of its kind to provide evidence of factorial validity for responses to the 9-item HIS and the first study to provide evidence for the factorial validity of responses to a 9-item scale measuring symptom severity. In order to understand the importance and relevance of our findings, one must first examine the
clinical reasoning behind conducting such an investigation. As mentioned previously, self-report symptoms related to concussion could be related to other conditions. This is evidenced by the presence of symptom reports at non-injured baselines (30). Baselines are one of the most important parts of the concussion assessment because they provide sports medicine clinicians with a benchmark for making return to play decisions. Thus, we hypothesized that the concussion related symptoms reported at non-injured baseline would fit the theoretical assumption that 3 underlying factors (cognitive, neuropsychological, and somatic) describe concussion related symptoms. If a group of concussion related symptoms reported by non-injured athletes fit the theoretical assumption, then that group should remain cohesive following concussive injury. Hence, any effects realized between groups of non-injured and concussed athletes are supported. In other words, the group of concussion related symptoms that are reported at some level by non-injured athletes should be the same cohesive symptoms that are reported to increase in symptom duration and symptom severity following injury due to the fact they share a predictable pattern of relationships.

Our initial study demonstrated that 9-items from an original 16-item self-report scale measuring duration responded in this manner (45). The 9-items tapped into one of three underlying constructs, which in turn tapped into a single higher order construct representing concussion symptoms. An experimental condition demonstrated that effects between concussed and control group responses to the full 16-item scale were not lost when the 7 extraneous items were removed from the analysis (45). This served to indicate that the cohesiveness of responses to the 9-items remain following a concussion. These results demonstrated both factorial and construct validity of the 9-item HIS, but prior to
making suggestions from this evidence, it was necessary to re-test our theories in a larger sample.

The initial findings were supported and extended by our current results. We tested the fit of a 3 factor higher order model to data from 1065 baseline responses to the 16 and 9-item HIS self-report measure of symptom duration, and a 16 and 9-item self-report scale measuring symptom severity. Findings from the present study strongly supported previous research in that the three correlated factor model represented a better fit to the 9-item scale responses than the 16-item scale responses. Our study undertook a strictly confirmatory approach in that the posited model either fit or it did not fit, hence either providing convergent or divergent evidence of the 9-item factorial validity. Results of this study provided strong convergent evidence for the factorial validity of the 9-items.

This evidence is important because it confirms the main point of the initial investigation. The concussion related 9-items of the HIS and severity scales (headache, nausea, balance difficulty, fatigue, drowsiness, trouble falling asleep, difficulty concentrating, feeling “in a fog”, feeling “slowed down”) form a cohesive grouping at non-injured baseline. This does not purport that the remaining 7 items are not related to concussion, just that they do not relate to the other symptoms at baseline in a way that provides a meaningful underlying structure. Lack of structural validity greatly decreases the level of inferences clinicians may draw from obtained scores, therefore potentially undermining interpretation. Evidence to support the continued cohesiveness of the 9 symptoms following concussion is demonstrated by the findings from the experimental condition.
Construct Validity

There was a strong relationship between scores from the 9-item and 16-item models of both scales, and the strong relationship between the versions of the scales provided convergent evidence for the construct validity of the 9-item HIS and 9-item severity scores. These relationships indicate that the 9-item HIS and the 9-item severity scale taps the construct of concussion symptoms comparably to the 16-item HIS and 16-item severity scale.

The 9-item and the 16-item models of both versions of scales were comparable in monitoring self-report concussion symptom resolution. With both models of the HIS there was a non-significant difference between groups at baseline, but scores were significantly increased on days one and two in the concussed group compared to the non-concussed group. No significant differences in HIS scores between groups on day three were observed, but the moderate effect size on this day suggests the presence of substantial group difference. Group differences on Day 10 were not supported by either statistical group differences or elevated effect sizes. We found similar results with the 9 and 16-item models of the severity scale. Though the 16-item severity scale did not have a significant p value on day 2, the moderate effect size suggests substantial group differences. This reflects the decreased statistical power of our small number of concussed subjects. It is also of value to report that a statistical difference was demonstrated for Day 2 when the 9-item severity scale was utilized.

These differences in symptom report between concussed and non-injured athletes are similar to those reported previously (1, 15, 19, 34, 40, 45). Typically a pattern of change ranging between 2 and 7 days following a concussion has been demonstrated
using either duration or severity as the scale descriptor (19, 25, 34, 36, 45). Our findings are relevant because they demonstrate the cohesiveness of the 9 symptoms that have factorial validity during non-injured baselines. The between group effects demonstrated by responses to the 16-item scales were not lost or decreased when the 7 symptoms not found to be structurally valid at baseline were removed.

**Conclusion**

Our study has confirmed evidence for the reliability, factorial validity, and construct validity of baseline responses to the 9-item HIS, and it has provided initial evidence for the reliability, factorial validity, and construct validity of baseline responses to the 9-item severity scale. Baseline reports are a crucial aspect to the proper assessment of the concussive injury. Return to play decisions are based upon an athlete presenting as asymptomatic for 7 days, both at rest and with exertion, following a concussive injury (13, 33). Summative scales that lack measurement principles and structural validity evidence do not assist in the goal of providing the best possible environment for athletic safety. Thus, gaining a thorough understanding of concussion-related symptoms, which can be reported by non-concussed individuals, can lead to improved decision making by sports medicine clinicians.

A validated measure of baseline concussion-related symptoms is crucial due to concern for scores obtained via self-report methodology. The evidence presented here allows for the evaluation of scores obtained from the 9-item HIS and the 9-item severity scale. Factors such as sex, previous concussion, daily fatigue, and orthopedic injury may alter the level of symptom reports at non-concussed baseline. These intrinsic and extrinsic factors are common to the athletic population and require investigation. Future
research should not only evaluate factors that may influence baseline symptom reports, but also continue to investigate the purpose of what symptom scales intend to measure. Symptom duration has clinical relevance and may provide very different information than symptom severity. Both 9-item duration and severity measures tap into the higher order construct of concussion symptoms, but they may represent different outcomes. Our findings allow for these research questions to be further evaluated.

References


Tables

Table 1. Descriptives for testing sites.

Southeast Administration Sites

<table>
<thead>
<tr>
<th>School</th>
<th>n</th>
<th>sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>409</td>
<td>Males n=217</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females n=192</td>
</tr>
<tr>
<td>Site 2</td>
<td>52</td>
<td>Males n=50</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females n=2</td>
</tr>
<tr>
<td>Site 3</td>
<td>123</td>
<td>Males n=108</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females n=115</td>
</tr>
<tr>
<td>Site 4</td>
<td>99</td>
<td>Males n=99</td>
</tr>
<tr>
<td>Site 5</td>
<td>55</td>
<td>Males n=55</td>
</tr>
</tbody>
</table>

Midwest Administration Sites

<table>
<thead>
<tr>
<th>School</th>
<th>n</th>
<th>sex</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 6</td>
<td>244</td>
<td>Males n=193</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Females n=51</td>
</tr>
<tr>
<td>Site 7</td>
<td>83</td>
<td>Males n=83</td>
</tr>
</tbody>
</table>
Table 2. Descriptive statistics for responses to the 16-items on the HIS and 16-items on the severity scale baseline sample (N = 1065)

<table>
<thead>
<tr>
<th>Duration Item (HIS)</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headache</td>
<td>0.76</td>
<td>1.19</td>
<td>1.70</td>
<td>2.47</td>
</tr>
<tr>
<td>Nausea</td>
<td>0.26</td>
<td>0.73</td>
<td>3.42</td>
<td>12.62</td>
</tr>
<tr>
<td>Vomiting</td>
<td>0.08</td>
<td>0.40</td>
<td>6.78</td>
<td>55.67</td>
</tr>
<tr>
<td>Difficulty balancing</td>
<td>0.17</td>
<td>0.57</td>
<td>4.50</td>
<td>24.11</td>
</tr>
<tr>
<td>Fatigue</td>
<td>0.95</td>
<td>1.30</td>
<td>1.29</td>
<td>0.80</td>
</tr>
<tr>
<td>Trouble falling asleep</td>
<td>0.58</td>
<td>1.15</td>
<td>2.27</td>
<td>4.94</td>
</tr>
<tr>
<td>Sleeping more than usual</td>
<td>0.37</td>
<td>.091</td>
<td>2.78</td>
<td>7.89</td>
</tr>
<tr>
<td>Drowsy</td>
<td>0.77</td>
<td>1.17</td>
<td>1.57</td>
<td>1.78</td>
</tr>
<tr>
<td>Sensitivity to light/noise</td>
<td>0.18</td>
<td>0.63</td>
<td>4.54</td>
<td>23.02</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.34</td>
<td>0.82</td>
<td>2.99</td>
<td>10.25</td>
</tr>
<tr>
<td>Nervousness</td>
<td>0.45</td>
<td>0.99</td>
<td>2.65</td>
<td>7.48</td>
</tr>
<tr>
<td>Numbness/Tingling</td>
<td>0.15</td>
<td>0.61</td>
<td>4.89</td>
<td>26.48</td>
</tr>
<tr>
<td>Feeling “slowed down”</td>
<td>0.51</td>
<td>0.97</td>
<td>2.18</td>
<td>4.49</td>
</tr>
<tr>
<td>Feeling like “in a fog”</td>
<td>0.23</td>
<td>0.67</td>
<td>3.59</td>
<td>14.78</td>
</tr>
<tr>
<td>Difficulty concentrating</td>
<td>0.50</td>
<td>0.99</td>
<td>2.43</td>
<td>6.35</td>
</tr>
<tr>
<td>Difficulty remembering</td>
<td>0.30</td>
<td>0.80</td>
<td>3.47</td>
<td>14.51</td>
</tr>
<tr>
<td>Composite Score 16-item HIS</td>
<td>6.59</td>
<td>8.70</td>
<td>2.20</td>
<td>4.32</td>
</tr>
<tr>
<td>Composite Score 9-item HIS</td>
<td>4.72</td>
<td>6.07</td>
<td>1.88</td>
<td>4.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Severity Item</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headache</td>
<td>0.69</td>
<td>1.18</td>
<td>1.92</td>
<td>3.38</td>
</tr>
<tr>
<td>Nausea</td>
<td>0.19</td>
<td>0.60</td>
<td>4.05</td>
<td>19.40</td>
</tr>
<tr>
<td>Vomiting</td>
<td>0.06</td>
<td>0.37</td>
<td>7.93</td>
<td>73.68</td>
</tr>
<tr>
<td>Difficulty balancing</td>
<td>0.13</td>
<td>0.50</td>
<td>5.68</td>
<td>43.05</td>
</tr>
<tr>
<td>Fatigue</td>
<td>0.81</td>
<td>1.19</td>
<td>1.53</td>
<td>1.69</td>
</tr>
<tr>
<td>Trouble falling asleep</td>
<td>0.51</td>
<td>1.07</td>
<td>2.43</td>
<td>5.78</td>
</tr>
<tr>
<td>Sleeping more than usual</td>
<td>0.27</td>
<td>0.78</td>
<td>3.56</td>
<td>14.09</td>
</tr>
<tr>
<td>Drowsy</td>
<td>0.59</td>
<td>1.06</td>
<td>1.90</td>
<td>3.03</td>
</tr>
<tr>
<td>Sensitivity to light/noise</td>
<td>0.14</td>
<td>0.52</td>
<td>5.46</td>
<td>37.68</td>
</tr>
<tr>
<td>Sadness</td>
<td>0.27</td>
<td>0.77</td>
<td>3.70</td>
<td>15.86</td>
</tr>
<tr>
<td>Nervousness</td>
<td>0.39</td>
<td>0.94</td>
<td>3.03</td>
<td>10.02</td>
</tr>
<tr>
<td>Numbness/Tingling</td>
<td>0.12</td>
<td>0.52</td>
<td>5.54</td>
<td>33.48</td>
</tr>
<tr>
<td>Feeling “slowed down”</td>
<td>0.41</td>
<td>0.90</td>
<td>2.60</td>
<td>6.95</td>
</tr>
<tr>
<td>Feeling like “in a fog”</td>
<td>0.19</td>
<td>0.63</td>
<td>4.25</td>
<td>20.49</td>
</tr>
<tr>
<td>Difficulty concentrating</td>
<td>0.37</td>
<td>0.82</td>
<td>2.77</td>
<td>8.71</td>
</tr>
<tr>
<td>Difficulty remembering</td>
<td>0.23</td>
<td>0.69</td>
<td>4.11</td>
<td>20.02</td>
</tr>
<tr>
<td>Composite Score 16-item severity scale</td>
<td>5.35</td>
<td>7.78</td>
<td>2.70</td>
<td>10.87</td>
</tr>
<tr>
<td>Composite Score 9-item severity scale</td>
<td>3.88</td>
<td>5.48</td>
<td>2.38</td>
<td>7.85</td>
</tr>
</tbody>
</table>

Note. M = Mean; SD = Standard deviation; Skew = Skewness; Kurt = Kurtosis
Figures

Figure 1. 3-correlated factors theoretical measurement model.
Figure 2. 3-factor higher-order theoretical measurement model.

- Headache
- Nausea
- Vomiting
- Balance problems
- Sensitivity to L/N
- Numbness/tingling
- Sleeping more
- Drowsiness
- Fatigue
- Sadness
- Nervousness
- Tired falling asleep
- Feel "slowed down"
- Feel "in a fog"
- Diff concentrating
- Diff remembering
Figure 3. Mean responses for injured and control groups using 16-item HIS across testing days.

*p ≤ .05
Figure 4. Mean responses for injured and control groups using 9-item HIS across testing days.

* $p \leq 0.05$
Figure 5. Mean responses for injured and control groups using 16-item Severity scale across testing days.

* p ≤ 0.05
Figure 6. Mean responses for injured and control groups using 9-item Severity scale across testing days.

* $p \leq 0.05$
CHAPTER 4

INVESTIGATION OF BASELINE SELF-REPORT
CONCUSSION SYMPTOM SCORES

2Piland, S.G. To be submitted to the Journal of Athletic Training.
ABSTRACT

Objective

This study investigated the influence of previous history of concussion, sex, fatigue, physical illness, and orthopedic injury upon the baseline responses to two reliable and valid self-report summative concussion related symptom scales.

Design

A cross-sectional study.

Subjects

Participants (N = 1065) in the cross-sectional analyses were predominately male (n = 805) college athletes with a mean age of 19.81 ± 1.53 years from 7 NCAA institutions.

Measurements

All participants completed baseline measures for 9-item measure of symptom duration, a 9-item measure of symptom severity, and a 27-item brief health questionnaire.

Results

A relatively high number of respondents reported concussion related symptoms at baseline, including: fatigue (45%), drowsiness (39%) headache (38%). Sex effect was not significant for responses to either scale. Respondents reporting a previous concussion had statistically higher scores on both measures of symptom duration and symptom severity. The presence of fatigue, physical illness, and orthopedic injury were found to increase composite scores for both measures. CFA provided strong evidence for the factorial validity of responses of people reporting no fatigue, physical illness, and/or orthopedic injury to both scales.
Conclusions

This study demonstrates that variables such as daily fatigue, physical illness, and orthopedic injury can serve to increase self-report symptom scores from two reliable and valid scales. These variables need to be controlled by clinicians prior to collecting non-concussed baseline measures on self-report symptoms.

Key Words:

Base Rate, Self-Reported Symptoms, Factorial Validity
Introduction

Investigations into the role of self-report concussion related symptoms following a concussive injury began with the landmark work of Barth (1). The four-year prospective study utilized the sport environment as a laboratory to clinically investigate concussive injuries. Since this study, authors have presented what have become commonly accepted self-report symptoms of acute concussion. These usually include: headache, impaired memory, dizziness, fatigue, sensitivity to light, sensitivity to noise, loss of concentration, nausea, vomiting, feeling of being "in a fog", feeling of being "slowed down", sleep disturbances, vision problems, difficulty remembering, irritability, and balance difficulties (1, 2, 3, 4, 5, 6, 7). The majority of self-report concussion related symptom research has been concerned with the rates at which these symptoms occur following the concussive injury. This focus is a result of two driving factors: 1) there is no single biological marker of concussion and 2) concussion related symptoms do not always resolve quickly following a concussive injury.

The severity and resolution of a concussion is not adequately represented by just a single sign or symptom; it requires a multi-faceted approach to evaluate this injury. Signs and symptoms related to changes in cognition, memory, reaction time, attention, information processing, balance, and self-report symptoms are all included as potential markers of the concussive injury and resolution from this injury (7, 8, 9, 10, 11, 12, 13, 14, 15, 16). Accurate and representative measurements of each of these markers can potentially provide the sports medicine clinicians with information on which to make safe return to play decisions. In regard to self-report symptoms, instruments incorporating summative scales are becoming the measures of interest (17). This would involve the use
of a summative 7-point Likert-type scale to score the severity (13, 17, 18) or duration (19) of self-report symptoms.

Typically, following sport-related concussion self-report symptoms have been demonstrated to resolve in 10 days or less (7, 15, 18, 19, 20), but this trend is not always consistent, especially in non-sport related concussion (6). Persistent symptoms, which may linger following a concussion, are associated with post concussion syndrome (PCS). This syndrome is defined by the Diagnostic and Statistical Manual of Mental Disorders (21) as the presence of concussion related symptoms for a minimum of 3 months following concussion. The cause of PCS and the conditions that maintain it are controversial (6). DSM-IV criteria suggest that cerebral dysfunction and structural anomaly are primary causes of PCS. In contrast, the International Classification of Mental and Behavioral Disorders (22) diagnostic criteria and other research suggest that psychological factors and contextual factors influence the prevalence and course of PCS (5, 6). PCS is believed to be caused by either biological effects of the injury, psychological factors or both and should not be considered synonymous with mild traumatic brain injury or concussion (6).

Due to its controversial existence and difficulty in diagnosis, investigations into PCS have focused upon the base-rates of concussion related self-report symptoms experienced in the healthy population. Several studies have strongly demonstrated high endorsement rates for these symptoms in normal populations. Iverson and Lange (6) recently reported that of the 104 healthy young male and female community volunteers, 89.3% reported fatigue, 76.7% reported poor concentration, 74.7% reported poor sleep, 55.3% reported headache, and 41.7% reported nausea. These reports were similar to those
found in a sample of 85 males and females without any identifiable history of head injury, neurological disease, or psychiatric illness (23). Fifty-eight percent of the sample reported fatigue, 58% reported poor concentration, 50.6% reported sleep disturbances, 40% reported headache, and 13% reported nausea. Neither study found differences in self-report symptom reports between the sexes (6, 23). These findings related to the prevalence of concussion related symptoms among the normal population are consistent with others found in the literature (24,25). Implications towards PCS are that many of the normal healthy respondents who have never experienced a concussion meet the criteria for PCS diagnosis. These findings are not solely related to PCS research, but have substantial implications for sport-related concussion.

Recommendations towards appropriate concussion assessment protocols agree that non-concussed baseline measures of athletes are an integral part of the assessment paradigm (17, 18, 39). In measures of self-report symptoms, the presence of concussion related symptoms prior to injury lacks thorough investigation. Similar to reports from the PCS literature, a relatively high proportion of healthy athletes report concussion related symptoms at baseline. In a study of the baseline symptom responses of 279 Division I males and females, headache was the most commonly reported symptom (58.8%) followed by fatigue (51.3%), trouble falling asleep (28.4%), difficulty concentrating (26.9%), nausea (25.4), and difficulty balancing (23.3%). However, there has been a paucity of information related to pre-existing physical conditions that could influence baseline scores. It is common for baselines to be obtained at the most convenient time. Athletes who may have had to participate in workouts, or may be experiencing an acute orthopedic injury, or may be physically ill might have a tendency to report higher level of
base rate symptoms than those without these experiences. The presence or absence of each of these variables can be established prior to making the baseline assessment, but to date, no study has evaluated the potential influence of these variables upon self-report concussion related symptoms.

Until recently, studies of this nature could not be appropriately applied because of the lack of information regarding self-report scale standardization, scale reliability and scale validity. Each of these areas is of utmost importance due to the fact that inferences drawn from summative scores are only as strong as the validity evidence provided for the scores (19). Strong evidence for the reliability, factorial validity and construct validity of responses to a summative self-report concussion related symptom measure of duration has been demonstrated (19). Using confirmatory factor analysis and a follow-up experimental condition, the study results demonstrated that responses to a 9-item scale provided a strong fit to a 3-factor higher order model representing concussion symptoms. In an experimental condition, the 9-items demonstrated cohesiveness in their ability to capture changes in self-report symptoms between concussed athletes and non-concussed athlete controls (19). Findings from this initial investigation were confirmed by a study involving a much larger sample of 1065 collegiate athletes. The follow-up project also provided initial evidence to the reliability, factorial validity and construct validity of 9-item summative scale measuring symptom severity (26). With the demonstration of strong evidence of validity for the 9-item measure of symptom duration and the 9-item measure of symptom severity it was appropriate and necessary to investigate influences upon baseline responses to self-report measures of concussion related symptoms.
Given the lack of information of base rate self-report symptoms, the purpose of this study was to evaluate the variables that may influence baseline responses to two reliable and valid symptom scales. This information will assist in determining the proper methodology for administration of self-report symptom scales and ultimately assist in making better clinical decisions.

**Methods**

**Participants**

All participants signed an informed consent document, and the study was conducted in accordance with rules and regulations designed for protection of human subjects. Participants (N = 1065) in the cross-section analysis were sports participants in NCAA institutions (Table 1). Subjects ranged from 18 to 27 years of age (Mean age = 19.81 ± 1.53 years) and were predominantly male (n= 805).

**Instrumentation**

*Symptom Duration*

The 9-item Head Injury Scale (HIS) is a summative 7-point Likert type scale comprised of concussion related symptoms that has been demonstrated to produce reliable and valid responses (19, 26). Two day test-retest reliability for the 9-item HIS was R=85 and internal consistency as measured by Cronbach’s alpha was .94 (27). Instructions to the respondent read: “Here is a list of symptoms that people often feel when they have a concussion. After reading each symptom please circle the number that best describes how long you have experienced the symptom during the previous 24-hour period (today)”. Items were rated on a 7-point Likert-type scale with the response options: [0] I have never experienced this symptom, [1] I have experienced this symptom
very briefly today, [3] I have experienced this symptom sometimes during today (about half the day long), [6] I have always experienced this symptom today (all day long). Item responses were then summed to provide an overall composite score.

**Symptom Severity**

The 9-item severity scale was designed to represent the two severity related scales common to the literature (PCSS and GSC) (11, 13, 17, 18, 28). The summative scale comprised of concussion related symptoms has been demonstrated to produce reliable and valid responses (19, 26). Two day test-retest reliability for the 9-item severity scale was R=.86 and internal consistency as measured by Cronbach’s alpha was .97.

Instructions to the respondent read: “The previous scale asked you to report how long you have experienced each of the listed symptoms during the previous 24-hour period (today). Now we would like for you to please circle the number that best describes the way you have been feeling during the previous 24-hour period (today).” As with the HIS, items were rated on a 7-point Likert-type scale, but response options were written to reflect item severity: [0] I have not experienced this symptom, [1] I have felt mild problems with this symptom today, [3] I have felt moderate problems with this symptom today, [6] I have felt severe problems with this symptom today. Item responses were then summed to provide an overall composite score.

**Brief History Questionnaire**

This two-page 27-item questionnaire is designed to collect descriptive information for each participant. It includes information related to previous history of sport-related concussion, acute physical illness, acute orthopedic injury, presence of fatigue from daily activities prior to baseline testing, type of sport played, years of exposure to the played sport, and position.
Procedure

The baseline measures were collected by certified athletic trainers (ATC) trained in the administration of this self-report battery. The HIS was administered in conjunction with the severity scale, and a brief health questionnaire. Each form included the 9-item grouping of symptoms reported to have factorial validity (19).

Data Analysis

A Pearson r value and paired samples t-test were calculated to evaluate the relationship between the composite scores of both scales. Multiple Kruskal-Wallis One-way Analysis of Variance (ANOVA) by ranks were conducted to evaluate the influences of pre-existing variables upon the responses to each scale (29). This nonparametric technique was selected due to the violations of multivariate normality demonstrated by skewness and kurtosis values presented in Table 2.

Results

Descriptive Statistics

The means, standard deviations, skewness and kurtosis values for responses to the 9-item HIS and severity scale are reported in Table 2. Responses to both scales were found to violate the assumption of data normality.

9-item HIS and 9-item severity scale

A Pearson r value of .87 was calculated between the responses to the 9-item HIS and 9-item severity scale. Paired sample t-test suggested that composite responses to the HIS measure of symptom duration were statistically higher than the composite responses to the measure of severity (Table 2).
Baseline Responses

Statistically significant differences were demonstrated between the baseline responses of athletes reporting a previous history of concussion and those reporting no history on both scales ($\chi^2 = 20.71$, df=1, $p \leq .001$ for the 9-item HIS, and $\chi^2 = 31.41$, df=1, $p \leq .001$ for the 9-item severity scale). No significant differences were realized by sex ($\chi^2 = .45$, df=1, $p > .001$ for the HIS, and $\chi^2 = 2.49$, df=1, $p > .001$ for the severity scale).

However, those athletes reporting the presence of fatigue ($\chi^2 = 161.13$, df=1, $p \leq .001$ for the HIS, and $\chi^2 = 166.36$, df=1, $p \leq .001$ for the severity scale), acute orthopedic injury ($\chi^2 = 22.01$, df=1, $p \leq .001$ for the HIS, and $\chi^2 = 23.51$, df=1, $p \leq .001$ for the severity scale), or physically ill ($\chi^2 = 15.91$, df=1, $p \leq .001$ for the HIS, and $\chi^2 = 19.27$, df=1, $p \leq .001$ for the severity scale) reported higher composite scores on both the 9-item HIS and 9-item severity scale than those without these conditions.

Based upon these findings we removed all variables deemed to be clinically controllable (fatigue, orthopedic injury, physically ill) from the analysis and re-valuated the effects of the two non-controllable variables (sex and previous history of concussion). We found that removal of respondents possessing the controllable variables did not change the statistical relationships between sex ($\chi^2 = .45$, df=1, $p > .001$ for the HIS, and $\chi^2 = 2.49$, df=1, $p > .001$ for the severity scale) and previous history of concussion ($\chi^2 = 20.71$, df=1, $p \leq .001$ for the 9-item HIS, and $\chi^2 = 31.41$, df=1, $p \leq .001$ for the 9-item severity scale). We also conducted a post hoc examination of the factorial validity of both scales after respondents reporting influencing variables were removed from the data set. Our previous study demonstrated strong evidence for the factorial validity of both 9-item
symptom scales (26). The data set included all of the aforementioned influencing variables. We postulated that though these variables acted to elevate overall composite scores, their removal would not have detrimental effects to the 3-factor higher order model shown to fit the previous data. We tested our assumptions using a previously reported protocol (19, 26).

The 3-factor higher order measurement model represented a good fit to the 9-item HIS ($\chi^2 = 118.37$, df = 24, RMSEA = .07 [90% CI = .06 - .09], NNFI = .93, CFI = .95) and the 3-factor higher order measurement model represented a good fit to the 9-item severity scale ($\chi^2 = 140.85$, df = 24, RMSEA = .08 [90% CI = .07 - .09], NNFI = .91, CFI = .94). For both scales, the chi-square statistic was significant (p <.05), but this statistic assumes the correct model and was affected by sample size (30, 31, 32). The RMSEA value was 0.07 and 0.08, which exceeds and meets the threshold respectively (33, 34). Also, the lower bound of the 90% confidence interval was below the threshold (.08) and approached or met .06. The NNFI and CFI provided good support for the 9-item, three-factor model for both scales. The NNFI and the CFI values exceeded the 0.90 standard (30, 33, 35), and the CFI met the .95 threshold for the 9-item HIS representing a strong model/data fit. These findings confirm the factorial validity of both self-report scales in the absence of influential variables.

Discussion

Baseline scores to measures of concussion serve as assistive criteria for sports medicine clinicians making return to play decisions. This is of importance because return to play guidelines recommend that athletes be deemed asymptomatic prior to returning to
full activity (36, 37, 38). Thus, the benchmarks for determining asymptomatic require
definition and understanding prior to score interpretation.

We demonstrated that non-concussed athletes report concussion related self-report
symptoms at baseline. In Table 3 we presented the frequencies of which our baseline
respondents reported concussion related self-report symptoms. Our findings were
consistent with those presented in previous research (19, 6, 24). A high percentage of
athletes, along with healthy, normal people report some level of concussion related
symptoms at non concussed baseline (Table 4). Post concussion, the most common
symptoms (fatigue, headache, difficulty concentrating, drowsiness, and trouble falling
asleep) are also among the most common symptoms expressed at non-injured baseline
(10). Though much emphasis has been placed upon symptom reports following a
concussive injury, these findings suggest the need for investigations into the base rates
for concussion related symptoms. It is also relevant to examine how symptom reports are
influenced by pre-existing variables that are common to sport.

We demonstrated that there was no difference in the composite scores of the 9-
item HIS and the 9-item severity scale between males and females. This was consistent
with previous base rate research (6, 24). We also reported that athletes reporting “fatigue
from today’s activities”, being “physically ill”, and experiencing an” orthopedic injury”,
all variables that are common to sport, reported increased composite scores of symptom
duration and symptom severity when compared to those athletes not reporting these
variables. Inclusion of these conditions at baseline may not appropriately represent how
an athlete typically presents with concussion related symptoms, and would not providing
the clinician with the most appropriate baseline to evaluate when comparing to post-injury scores.

A novel aspect to this study was the evaluation of those athletes reporting a previous history of concussion compared to those without. Previous history of concussion has been shown to be related to decreased neurocognitive performance, a pre-disposition to incur another concussive episode, and a resulting slower recovery from that episode (18, 20). We found that those reporting a previous history of concussion reported higher composite scores for symptom duration and symptom severity than those without a previous history. The implications being that all high respondents should be questioned about their history of concussion. Obtaining a thorough medical history can provide additional information about the athlete’s typical signs and symptoms for concussion which can support the decision to conduct a more thorough medical evaluation.

Based upon our findings that fatigue, physical illness, and orthopedic injury serve to increase overall composite scores measured by the 9-item HIS and 9-item severity scale, we conducted a CFA of the data with the respondents reporting any of these variables removed from the data set. Our previous studies have not controlled for such influential variables, thus it was important to demonstrate the structural validity of responses to our two measures of self-report symptoms. With the influential controllable variables removed we found strong evidence for continued support of the factorial validity of the 9-item HIS and 9-item severity scale. This evidence suggests that the 9-items of the HIS and severity scale (headache, nausea, balance difficulty, fatigue, drowsiness, trouble falling asleep, difficulty concentrating, feeling “in a fog,” feeling
“slowed down”) remain a cohesive grouping of symptoms in the presence of variables that can artificially increase baseline self-report symptom scores.

**Recommendations and Conclusion**

It is our recommendation that clinicians employ the 9-item HIS and 9-item severity scale as reliable and valid measure of baseline self-report symptoms. Along with this we recommend that clinicians control for variables that are common to sport that may serve to increase normal baseline responses. Clinicians should provide athletes with a mechanism to report their previous history of concussion as well as the presence of fatigue, physical illness, and/or orthopedic injury prior to being allowed to complete a pre-season baseline self-report symptom test. Baselines from athletes reporting any of the three controllable variables should be obtained at a later date. Composite scores from athletes reporting a previous history of concussion should be given a thorough clinical evaluation to provide the sports medicine clinician with appropriate information to base return to play decisions.

**References**


Table 1. Descriptives measures for testing sites. (N=1065)

<table>
<thead>
<tr>
<th>School</th>
<th>n</th>
<th>9-item</th>
<th>9-item</th>
<th>sex</th>
<th>variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>Site 1</td>
<td>409</td>
<td>3.7(5.3)</td>
<td>3.1(4.9)</td>
<td>Males n=217</td>
<td>PC n=98</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Females n=192</td>
<td>PI n=58</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=49</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=44</td>
</tr>
<tr>
<td></td>
<td>244</td>
<td>6.8(5.9)</td>
<td>5.8(5.2)</td>
<td>Males n=193</td>
<td>PC n=64</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Females n=51</td>
<td>PI n=26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=52</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=109</td>
</tr>
<tr>
<td>Site 3</td>
<td>83</td>
<td>5.8(5.9)</td>
<td>4.9(5.6)</td>
<td>Males n=83</td>
<td>PC n=19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PI n=11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=40</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=20</td>
</tr>
<tr>
<td>Site 4</td>
<td>52</td>
<td>9.4(10.33)</td>
<td>8.3(9.8)</td>
<td>Males n=50</td>
<td>PC n=19</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PI n=9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=8</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=18</td>
</tr>
<tr>
<td>Site 5</td>
<td>55</td>
<td>3.4(2.6)</td>
<td>1.9(2.0)</td>
<td>Males n=55</td>
<td>PC n=27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PI n=2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=2</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=3</td>
</tr>
<tr>
<td>Site 6</td>
<td>123</td>
<td>4.5(6.5)</td>
<td>3.1(5.1)</td>
<td>Males n=108</td>
<td>PC n=27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Females n=115</td>
<td>PI n=11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=11</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=5</td>
</tr>
<tr>
<td>Site 7</td>
<td>99</td>
<td>1.2(3.2)</td>
<td>1.2(3.5)</td>
<td>Males n=99</td>
<td>PC n=22</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>PI n=7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>OI n=4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>F n=0</td>
</tr>
</tbody>
</table>

Note: PC= “Have you ever had a previous concussion”, PI= “Have you been physically ill this week?”, OI= “Have you incurred an orthopedic injury this week?”, F= “Are you feeling fatigued from today’s activities?”
Table 2. Descriptive statistics for responses to the 9-items on the HIS and 9-items on the severity scale baseline sample (N = 1065)

<table>
<thead>
<tr>
<th>Duration Item (HIS)</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headache</td>
<td>0.76</td>
<td>1.19</td>
<td>1.70</td>
<td>2.47</td>
</tr>
<tr>
<td>Nausea</td>
<td>0.26</td>
<td>0.73</td>
<td>3.42</td>
<td>12.62</td>
</tr>
<tr>
<td>Difficulty balancing</td>
<td>0.17</td>
<td>0.57</td>
<td>4.50</td>
<td>24.11</td>
</tr>
<tr>
<td>Fatigue</td>
<td>0.95</td>
<td>1.30</td>
<td>1.29</td>
<td>0.80</td>
</tr>
<tr>
<td>Trouble falling asleep</td>
<td>0.58</td>
<td>1.15</td>
<td>2.27</td>
<td>4.94</td>
</tr>
<tr>
<td>Drowsy</td>
<td>0.77</td>
<td>1.17</td>
<td>1.57</td>
<td>1.78</td>
</tr>
<tr>
<td>Feeling “slowed down”</td>
<td>0.51</td>
<td>0.97</td>
<td>2.18</td>
<td>4.49</td>
</tr>
<tr>
<td>Feeling like “in a fog”</td>
<td>0.23</td>
<td>0.67</td>
<td>3.59</td>
<td>14.78</td>
</tr>
<tr>
<td>Difficulty concentrating</td>
<td>0.50</td>
<td>0.99</td>
<td>2.43</td>
<td>6.35</td>
</tr>
<tr>
<td>Composite Score</td>
<td>4.72</td>
<td>6.07</td>
<td>1.88</td>
<td>4.32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Severity Item</th>
<th>M</th>
<th>SD</th>
<th>Skew</th>
<th>Kurt</th>
</tr>
</thead>
<tbody>
<tr>
<td>Headache</td>
<td>0.69</td>
<td>1.18</td>
<td>1.92</td>
<td>3.38</td>
</tr>
<tr>
<td>Nausea</td>
<td>0.19</td>
<td>0.60</td>
<td>4.05</td>
<td>19.40</td>
</tr>
<tr>
<td>Difficulty balancing</td>
<td>0.13</td>
<td>0.50</td>
<td>5.68</td>
<td>43.05</td>
</tr>
<tr>
<td>Fatigue</td>
<td>0.81</td>
<td>1.19</td>
<td>1.53</td>
<td>1.69</td>
</tr>
<tr>
<td>Trouble falling asleep</td>
<td>0.51</td>
<td>1.07</td>
<td>2.43</td>
<td>5.78</td>
</tr>
<tr>
<td>Drowsy</td>
<td>0.59</td>
<td>1.06</td>
<td>1.90</td>
<td>3.03</td>
</tr>
<tr>
<td>Feeling “slowed down”</td>
<td>0.41</td>
<td>0.90</td>
<td>2.60</td>
<td>6.95</td>
</tr>
<tr>
<td>Feeling like “in a fog”</td>
<td>0.19</td>
<td>0.63</td>
<td>4.25</td>
<td>20.49</td>
</tr>
<tr>
<td>Difficulty concentrating</td>
<td>0.37</td>
<td>0.82</td>
<td>2.77</td>
<td>8.71</td>
</tr>
<tr>
<td>Composite Score</td>
<td>3.88</td>
<td>5.48</td>
<td>2.38</td>
<td>7.85</td>
</tr>
</tbody>
</table>

Note. M = Mean; SD = Standard deviation; Skew = Skewness; Kurt = Kurtosis
Table 3. Percent Frequency of Individual Symptom Reports. (N=1065)

<table>
<thead>
<tr>
<th>Symptom</th>
<th>9-item HIS</th>
<th>9-item severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>fatigue</td>
<td>45%</td>
<td>42%</td>
</tr>
<tr>
<td>drowsiness</td>
<td>39%</td>
<td>35%</td>
</tr>
<tr>
<td>headache</td>
<td>38%</td>
<td>31%</td>
</tr>
<tr>
<td>trouble falling asleep</td>
<td>28%</td>
<td>25%</td>
</tr>
<tr>
<td>feeling “slowed down”</td>
<td>28%</td>
<td>23%</td>
</tr>
<tr>
<td>difficulty concentrating</td>
<td>28%</td>
<td>22%</td>
</tr>
<tr>
<td>nausea</td>
<td>16%</td>
<td>12%</td>
</tr>
<tr>
<td>feeling like “in a fog”</td>
<td>14%</td>
<td>12%</td>
</tr>
<tr>
<td>difficulty balancing</td>
<td>11%</td>
<td>9%</td>
</tr>
</tbody>
</table>
### Table 4. Percent Frequency of Individual Symptom Reports from previous studies.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>fatigue</td>
<td>52%</td>
<td>89%</td>
<td>59%</td>
</tr>
<tr>
<td>headache</td>
<td>59%</td>
<td>55%</td>
<td>40%</td>
</tr>
<tr>
<td>trouble falling asleep</td>
<td>39%</td>
<td>75%</td>
<td>51%</td>
</tr>
<tr>
<td>difficulty concentrating</td>
<td>37%</td>
<td>77%</td>
<td>59%</td>
</tr>
<tr>
<td>nausea</td>
<td>25%</td>
<td>42%</td>
<td>13%</td>
</tr>
<tr>
<td>feeling like “in a fog”</td>
<td>15%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>difficulty balancing</td>
<td>24%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>drowsiness</td>
<td>39%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>feeling “slowed down”</td>
<td>29%</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CHAPTER 5

SUMMARY AND CONCLUSIONS

Sports medicine clinicians are charged with the task of prevention, recognition, and treatment of athletic related injuries. To fulfill this task appropriately sports medicine clinicians must stay abreast of the latest recommendations and changes. The area of sport-related concussion presents many difficulties in its prevention, assessment, and treatment. Originally assessment and treatment recommendations and protocols were developed from the experiences of experts within the field, but recently rigorous scientific investigations have begun to shed light upon the most appropriate paradigms to enlist for safer athlete care. Concussion assessment and treatment is made complex by the fact that no one individual biological marker of concussion is available in which to measure occurrence, severity, and resolution of the injury. Rather, sports medicine clinicians must rely upon a multi-faceted approach that incorporates multiple instruments to measure signs and symptoms. Even with a multi-faceted approach, concussion assessment can be very difficult. The most sensitive measures of concussion have limitations to their availability and uses.

Common measures which assess posturography can be expensive and out of the reach of the typical sports medicine clinician. Measures of neurocognitive performance are limited in their use to those that are professionally trained. Though gains have been made towards making these approaches more accessible, sports medicine clinicians typically rely upon a good physical medical exam and self-report symptoms in making assessment and return to play decisions.
Though commonly used, measures of self-report concussion related symptoms, especially those that involve a summative score, have not been subjected to the rigorous evaluations of psychometric and measurement properties that other instruments, which involve summative scores (i.e., psychological tests), are expected to be subjected to. The scores from summative scales that capture the duration or severity of which an athlete experiences concussion related symptoms are used to track the resolution of the injury. This means that inferences as to the well-being of an athlete are made based upon these scores. Inferences made from scores are only as strong as the evidence of validity provided for them. Based upon this, our project was conducted to support the available evidence for the factorial and construct validity of a self-report measure for duration, and provide initial factorial validity and construct validity evidence for a self-report measure of symptom severity.

Using a cross-sectional design, we tested two competing models of the HIS and two competing models of the severity scale using confirmatory factor analysis (CFA). Construct validity was assessed using Pearson product-moment correlation analyses. Using an experimental design, we compared scores on the each scale between concussed and non-concussed groups using a 2 (groups) × 5 (time) mixed model analysis of variance. Pre-season baseline data was collected on 1065 predominately male (n = 805) college athletes with a mean age of 19.81 ± 1.53 years from 7 NCAA institutions. Participants (N = 27) in the experimental analyses were concussed (n = 17) and non-concussed control (n = 10) college athletes with a mean age of 20.51 ± 1.92 years.
Participants completed baseline measures for the two scales and a brief health history questionnaire. Concussed individuals and controls were evaluated on Baseline and Days 1, 2, 3, and 10 post-injury.

Using CFA, the theoretically derived three-factor model did not provide a good fit to either the 16-item HIS or the 16-item severity scale, thus competing 9-item models were tested. The subsequent analysis indicated that the three-factor model provided a good fit to the 9-item HIS as well as the 9-item severity scale. The three-factors for each scale were best described by a single second-order factor, namely concussion symptoms. Scores from the 16-item and 9-item HIS and 16-item and 9-item severity scale were strongly correlated. The $2 \times 5$ ANOVA on the 16-item HIS demonstrated a significant groups by days interaction. Statistical differences between groups were observed on days 1 and 2 post-concussion. The $2 \times 5$ ANOVA on the 9-item HIS demonstrated a significant groups by days interaction. Statistical differences between groups were observed on days 1 and 2 post-concussion. The $2 \times 5$ ANOVA on the 16-item severity scale demonstrated significant groups by days interaction. Statistical differences between groups were observed only on day 1. The $2 \times 5$ ANOVA on the 9-item HIS demonstrated a significant groups by days interaction. Statistical differences between groups were observed on days 1 and 2 post-concussion.

This study confirmed previous findings by providing evidence for the factorial and construct validity of the HIS among collegiate athletes. This study also provided initial evidence for the factorial and construct validity of the 9-item severity scale. With such evidence we felt it appropriate to evaluate the influence of pre-existing conditions
that are common to athletics, to baseline responses to the validated 9-item versions of the two scales.

We found that a relatively high number of respondents reported concussion related symptoms at baseline, including: fatigue (45%), drowsiness (39%) headache (38%). Sex effect was not significant for responses to either scale. Respondents reporting a previous concussion had statistically higher scores on both measures of symptom duration and symptom severity. The presence of fatigue, physical illness, and orthopedic injury were found to increase composite scores for both measures.

It is concluded that responses to the 9-item HIS measure of symptom duration and responses to the 9-item measure of symptom severity are both reliable and valid. Clinicians now also know that pre-existing, controllable conditions common to athletes, can serve to increase baseline scores of self-report symptoms. But if controlled, scores that are representative of how athletes typically experience these types of symptoms can be obtained. Having appropriate baseline measures will allow sports medicine clinicians to make better decisions related to the pre-season assessment as well as the acute assessment of sport related concussion. Reliable and valid measures of self-report symptoms that represent both the duration and severity of concussion related symptoms are important because many questions concerning the presentation of concussion related symptom both before and after a concussive episode remain to be answered. These instruments can now be used more effectively in the investigation of sport-related concussion.