2011-01-01

An Investigation of the Effect of Repeated Testing and Extra Incentive on Neurocognitive Performance in Non-Concussed Adults

Edina Reka Bene
University of Texas at El Paso, edinabene@yahoo.com

AN INVESTIGATION OF THE EFFECT OF REPEATED TESTING AND EXTRA INCENTIVE ON NEUROCOGNITIVE PERFORMANCE IN NON-CONCUSSED ADULTS

EDINA R. BENE
Interdisciplinary Health Sciences

APPROVED:

Anthony P. Salvatore, Ph.D., Chair

Bess Sirmon Fjordbak, Ph.D.

Joe Tomaka, Ph.D.

Michael P. Cannito, Ph.D.

Patricia D. Witherspoon, Ph.D.
Dean of the Graduate School

AN INVESTIGATION OF THE EFFECT OF REPEATED TESTING AND EXTRA INCENTIVE ON NEUROCOGNITIVE PERFORMANCE IN NON-CONCUSSED ADULTS

by

Edina R. Bene, MA

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

Interdisciplinary Health Sciences

THE UNIVERSITY OF TEXAS AT EL PASO

May 2011

Acknowledgements

I would like to express my deepest gratitude to my Dissertation Chair and advisor, Dr. Anthony P. Salvatore. His constant support, mentorship, ever enduring patience, and guidance have pushed me to the best. Thank you for giving me the opportunity to work with you, for being my mentor, and for standing beside me in these past 4 years.

I would also like to sincerely thank my Dissertation Committee members for their kind help that I could always count on over the years. Namely, to Dr. Bess Sirmon Fjordbak for her encouragement and assistance in my doctoral studies; to Dr. Joe Tomaka for his continuous support and witty demeanor while explaining the different statistical procedures, research designs, and all the tricks with SPSS; and to Dr. Michael P. Cannito for his willingness to participate in my Dissertation Committee, his patience and mentorship, and for having me as a doctoral student. I feel truly indebted to all of you for being in my corner and pushing me to succeed.

A special "Thank You" to all my colleagues and friends who helped and supported me tremendously over the years. I am truly blessed to have people like you in my life. I have learned a lot from all of you and I miss you all so much.

I am most grateful for all the students who were willing to give their time to participate in this research project; for all the professors and colleagues who helped me recruit and finish my data collection in time; and for the UTEP College of Health Sciences as well as the Graduate School to provide me with the essential financial support and funding of this research project.

And last but not least, my heartfelt appreciation belongs to my loving family for their unconditional love and faith in me that allowed me to pursue my dreams and kept me going during the most challenging times. Szeretlek titeket! Köszönöm, hogy vagytok nekem!

Thank you so much to all of you. I could never have come this far without every one of your contributions. It has been an amazing journey and wonderful adventure!

Abstract

Standardized neurocognitive testing is a commonly used assessment method to diagnose and manage sports-related concussion. This method involves testing athletes at preseason, and re-testing them if a concussion occurs. Athletes usually return to competition once they are symptom-free and their cognitive performance post concussion returns to their preseason cognitive performance. Although this approach in sports-related concussion management seems well-guided and clear-cut, there are a number of factors that can interfere with the accuracy of assessment. This study addressed three such factors: test-retest reliability, the effect of extra incentive on neurocognitive performance, and practice effects.

33 non-concussed UTEP collegiate students were recruited and tested serially on four occasions using the Immediate Post-Concussion Assessment and Cognitive Testing battery version 2.0 (ImPACT) with 7-8 days between testing sessions. In this single-group A-A-B-A pretest-posttest design, the dependent measures were the Verbal and Visual Memory, Reaction Time, Visual Processing Speed, and Cognitive Efficiency Index composite scores on the ImPACT battery. Participants received a one time-only extra monetary incentive on the third test session. Time and the presence or absence of extra monetary incentive were the independent variables. Pearson correlation was applied to measure test-retest reliability for each composite score across the repeated administrations of the ImPACT battery. Repeated measures Analysis of Variance (ANOVA) and Trend Analysis were used to measure whether there was a change in performance across test sessions and to examine whether there was a change in the individual composite scores within participants before and after the monetary incentive is offered. Several test-retest reliability coefficients did not meet criteria and varied across time. Results also showed practice effects in verbal memory, visual processing speed, reaction time, and cognitive efficiency index scores. A significant linear trend was also found in visual memory, reaction time, and visual processing speed indicating improvement of performance across test sessions. Extra monetary incentive did not show any significant change when performance was compared to the second and forth

test sessions. Although extra monetary incentive did not improve performance on the battery, changes of performance, practice effects, and variable test-retest reliability measures may result from other behavioral factors. Therefore, concussion assessment and management should not rely on a single neurocognitive measure or test. Multiple assessment protocol for sports-related concussion management is needed.

Table of Contents

Acknowledgements....iv

Abstract....v

Table of Contents.... vii

List of Tables....ix

List of Figures....x

CHAPTER ONE....1

Introduction....1

CHAPTER TWO....15

- Literature Review....15
- Pilot Studies....68
  - *Study 1*....72
  - *Study 2*....80
  - *Study 3*....83
- Statement of the Problem....92
- Statement of the Purpose....93
- Research Questions and Hyptheses....95

CHAPTER THREE....97

- Methods and Design....97
- Data Analysis....109

CHAPTER FOUR....114

- Results....114

CHAPTER FIVE....124

- Discussion....124
- Conclusions....129
- Limitations of the Study....131

References..133

Appendix A..146

Appendix B..161

Biosketch……………. 171

List of Tables

Table 1. : Comparison of concussion grading systems..154

Table 2. : Descriptive information (mean and SD) of age at baseline and at PC1 testing. 154

Table 3. : Time post between date of last concussion and testing sessions...............................155

Table 4. : Time post between PC1 and last PC assessment..155

Table 5. : Mean and Median values for each individual composite score at baseline and first post-concussion testing...155

Table 6. : Mean and Median values for each individual composite score at baseline and last post-concussion testing...156

Table 7. : Mean and Median values for each individual composite score across all testing sessions. 156

Table 8. : Descriptive statistics of the participants: Mean (SD), and range. 157

Table 9. : Descriptive statistics for ImPACT Total Symptom Scores by gender using ImPACT Normative Data..157

Table 10. : Means (standard deviations) for ImPACT composite scores and for the adjusted scores on test session 1.. 157

Table 11. : Average ImPACT composite scores with standard deviation (SD) at Test 1, 2, 3*, and 4. 158

Table 12. : Means and standard deviations of performance on the ImPACT battery by gender.159

Table 13. : Test-retest Pearson correlations for the ImPACT battery. 160

Table 14. : Standard error of measurement across test sessions...160

Table 15. : Comparison of test-retest reliability values on the ImPACT battery. 160

List of Figures

Figure 1.: Median performance on Verbal Memory across testing sessions in concussed collegiate athletes. 161

Figure 2.: Median performance on Visual Processing Speed in concussed collegiate athletes. .161

Figure 3.: Median performance on Reaction Time across testing sessions in concussed collegiate athletes. 162

Figure 4.: Self-reported Total Symptom Scores across testing sessions in concussed collegiate athletes. 162

Figure 5.: Median performance on Verbal and Visual Memory across testing sessions in concussed collegiate athletes....163

Figure 6.: Performance on Visual Processing Speed across testing sessions in concussed collegiate athletes 163

Figure 7-10.: Mean performance on each cognitive measure on the ImPACT battery across testing sessions. 164

Figure 11.: Mean performance on the ImPACT Verbal Memory composite score across the four test sessions in non-concussed college students....166

Figure 12.: Mean performance on the ImPACT Visual Memory composite score across the four test sessions in non-concussed college students....166

Figure 13.: Mean performance on the ImPACT Visual Processing Speed composite score across the four test sessions in non-concussed college students. 167

Figure 14.: Mean performance on the ImPACT Reaction Time composite score across the four test sessions in non-concussed college students....167

Figure 15.: Mean performance on the ImPACT Cognitive Efficiency Index score across the four test sessions in non-concussed college students....168

Figure 16.: Trend line in Verbal Memory performance across test sessions in non-concussed college students. 168

Figure 17.: Trend line in Visual Memory performance across test sessions in non-concussed college students. 169

Figure 18.: Trend line in Visual Processing Speed performance across test sessions in non-concussed college students....169

Figure 19.: Trend line in Reaction Time performance across test sessions in non-concussed college students. 170

Figure 20.: Trend line in Cognitive Efficiency Index score across test sessions in non-concussed college students. 170

CHAPTER ONE

Introduction

Overview of the issues

Concussion assessment, diagnosis, and management all together impose a complex and challenging clinical process. First, not all concussions can be detected by neuroimaging techniques (Aubry, Cantu, Dvorak, Graf-Baumann, Johnston, Kelly, Lovell, McCrory, Meeuwisse, & Schamasch, 2002). Second, patients suffering a concussion present a wide variety of symptoms that can include somatic, psychological, and cognitive symptoms all together (Hall, Hall, & Chapman, 2005). Third, scope and the length of symptoms arising after a head injury may vary individually. Fourth, loss of consciousness may not necessarily accompany a concussion (McCrory, Johnston, Meeuwisse, Aubry, Cantu, Dvorak, Graf-Baumann, Kelly, Lovell, & Schamasch, 2005). Fifth, there is a tendency for athletes to hide their symptoms when they sustain a concussive head injury during practice and/or game (Barth, Alves, Ryan, Macciocchi, Rimel, Jane, & Nelson, 1989, Echemendia, Putukian, Mackin, Julian, & Shoss, 2001, Hunt, Ferrara, Miller, Macciocchi, 2007). Because athletes' tendency to underreport and hide symptoms may have a serious impact on their health, well-being, and academic-professional career, research has been focusing on developing more refined and sensitive tests and guidelines for concussion assessment and management. These assessment guidelines and management protocols include testing athletes at baseline (i.e., preseason, before a concussion occurs), assessing athletes serially after a concussion occurs, and how to decide when athletes can safely return to play.

Concussion management

Historically, Barth et al. (1989) were pioneers in establishing a return-to-play (RTP) decision protocol by using neurocognitive (NC) testing in a systematic manner (Echemendia &

Julian, 2001). First, they set up a preseason “baseline” NC performance of athletes which could be compared later on to post-concussion NC performance. Then these concussed athletes were tested serially until their cognitive performance returned to the level of their baseline performance. Third, post-concussion symptoms also had to resolve to baseline level. Even though testing cognitive-neurological functioning has been discussed in concussion research literature for a longer period of time, developing new protocols and techniques of sports-related concussion assessment and management is fairly recent. The basic notion behind examining NC functioning in athletes with a concussion was to make sure they can safely return to competition or practice and decrease the likelihood of developing a Post Concussion Syndrome or Second Impact Syndrome.

Despite recent developments in concussion assessment, diagnosis, and management, sports-related concussions in athletes are still receiving a lot of attention in the public media as well. For example, there have been reports about athletes hiding their head injuries and “play down” the effects of a concussion (CBS News, *Survey: NFL players hide, fear concussions*, 2009, The San Francisco Examiner, *Quotes from players interviewed recently by the AP about concussion*, 2009), high school athletes facing risks of premature return to play (USA Today, *High school athletes face serious concussion risks*, 2009), young athletes suffering brain damage due to improper concussion management (NBC Sports, *Concussion warnings too late for N. J. athlete*, 2009), or a young athlete dying after sustaining a concussion (The Huffington Post, *Dylan Steigers dead: Eastern Oregon football player dies from practice head injury*, 2010, *Brady Lee Frazier dead: 8th grader dies in baseball scrimmage*, 2010). Although the number of athletes suffering persistent brain damage or death due to sports-related concussion had decreased over the last 30 years, changes in concussion management policy have been continuous and still a significant concern in today’s sports.

The issue of "not reporting"

While the underlying rationale behind these efforts was to prevent premature return-to-play decisions and to eliminate the potential risk of the athlete to suffer long-lasting symptoms, a severe head trauma or death, this task faced difficulties. For one, suffering a concussion was not seen as a major problem historically. According to Echemendia, Putukian, Mackin, Julian, & Shoss (2001), many athletes often continue to play despite symptoms since there is a widespread expectation that mild head injuries are "part of the game" (p. 23). According to the quotes of the football players reported in The San Francisco Examiner (*Quotes from players interviewed recently by the Associated Press about concussion*, 2009), this tendency is evident today and likely to continue:

> I don't want to come out of games. I always feel that's some kind of weakness. Somebody hits me and takes me out of the game, I feel weak. So if something happens, I take a minute to try to re-gather myself and then go back. — New Orleans Saints defensive lineman Anthony Hargrove.
>
> You get a concussion, they've got to take you out of the game. So if you can hide it and conceal it as much as possible, you pay for it the next day, but you'll be able to … stay in the game. — Washington Redskins fullback Mike Sellers.
>
> I've known of players hiding concussions. ... Sometimes players aren't real sure. They hit their head, they get a little cuckoo for a little while. It happens all the time. — Kansas City Chiefs center Rudy Niswanger.

The reasons behind athletes "not complaining" range broadly. Athletes may feel that admitting a "small head injury" 1) would be the sign of weakness, 2) would make them feel that they let their team mates, coach, and trainer down, 3) would jeopardize their scholarship, 4) would lose their position in the team, 5) would miss the opportunity to demonstrate their skills in front of a professional agent, 6) would see themselves as a failure even by their family and team mates, but without question athletes 7) would definitely risk the opportunity to continue to play and/or removing the athletes from the upcoming competition (Barth et al., 1989, Echemendia et al., 2001 Hunt et al., 2007). Furthermore, even if athletes experience a concussion, they may underreport and minimize their symptoms in order to return to competition (Echemendia & Julian, 2001, Echemendia & Cantu, 2003, Broglio, Macchiocchi, & Ferrara, 2007b). Another crucial risk factor is that some athletes are unaware of the signs and symptoms that accompany a minor head injury (McCrea, Hammeke, Olsen, Leo, & Guskiewicz, 2004, LaBotz, Martin, Kimura, Hetzler, & Nichols, 2005, Valovich McLeod, Bay, Heil, McVeigh, 2008).

In summary, amateur athletes tend *not* to report previous injury for a number of reasons: 1) thinking that the injury was not serious enough for further medical attention; 2) do not want to be withdrawn from competition due to professional aspirations; and 3) lack of awareness of probable concussion (McCrea et al., 2004, LaBotz et al., 2005, Valovich McLeod et al., 2008). All these reasons behind “not reporting” may further explain why concussions are difficult to be detected.

The issue of detecting a concussion

One of the reasons why concussions are not recognized immediately is linked to loss of consciousness. Loss of consciousness (LOC) is not a necessary phenomenon that accompanies a concussion. Therefore, there is a high risk that concussions may go undetected. Today, the most conventional definition of concussion states that it “results in a graded set of clinical syndromes that may or may not involve loss of consciousness” (McCrory, Johnston, Meeuwisse, Aubry, Cantu, Dvorak, Graf-Baumann, Kelly, Lovell, & Schamasch, 2005, p. 196). Consequently, even if there is no LOC involved, a concussion can be just as severe. However, according to

Echemendia & Julian (2001), there is a lack of consistency in the definition of concussion, which causes problems further on. Since many people equate LOC with concussion, it is crucial to define the term "concussion" or "mild traumatic brain injury" clearly and in detail in order to collect accurate information for assessment and diagnosis (Echemendia & Julian, 2001).

Health risks

The attitude of athletes "not reporting" or "not recognizing" mild brain injuries involves serious health-related risks. First, without proper concussion management these athletes increase the likelihood of getting concussed again and/or develop a Post Concussion Syndrome (Guskiewicz, Weaver, Padua, & Garrett, 2000, Guskiewicz, McCrea, Marshall, Cantu, Randolph et al., 2003). Second, cumulative concussions could also lead to developing late-life cognitive impairments (Guskiewicz, Marshall, Bailes, McCrea, Cantu et al., 2005). Third, by increasing the likelihood of getting injured again while still being concussed, athletes may sustain a disabling and in some cases deadly Second Impact Syndrome (Cantu, 1996). Such attitude of athletes "not reporting" concussions may seriously impact their health academic and future professional career. Therefore, concussion management has to be sensitive enough to detect deficits due to concussion and preventive enough to detect whether athletes are truly aware of the severity and symptomology of sustaining concussion.

Approach for concussion management

Besides awareness and reliability of self-reports, another potential factor that may impact safe RTP decisions is linked to reliable, sensitive, and valid NC testing. The most current concussion assessment and management protocols apply NC assessment in the form of baseline/preseason testing and subsequent follow-up post-concussion assessments. Consequently, when joining a team a new athlete undergoes NC assessment before the season starts. After a concussion has occurred, the athlete is assessed at different time intervals until his or her symptoms and cognitive performance returns to the baseline level. Time between testing sessions may vary. For instance, experts and researchers at the University of Pittsburgh Medical Center assess concussed athletes within 36 hours, 4 days, and 7 days post-injury (Lovell, Collins, Iverson, Field, Maroon et al., 2003), while Echemendia and his colleagues at the Pennsylvania State University evaluate concussed athletes within 2 hours, 48 hours, 7 days and 30 days post-injury (Echemendia et al., 2001). Here at the UTEP Concussion Management Clinic, we follow the baseline assessment approach and evaluate athletes within 48-72 hours, 5 days, 7 days, and 10-14 days post injury. To date, we have screened more than 900 collegiate and high school athletes here in El Paso at baseline, and managed more than 45 athletes with a concussion using the baseline-serial assessment approach. To track the long-term effects of post-concussion symptoms, we also began scheduling our recently concussed UTEP athletes at 30 days, 3 months, 6 months and 12 months post-concussion. In addition, we always present our athletes some basic and essential information about concussion in order to raise awareness. Furthermore, we provide our concussed athletes with detailed feedback after completing NC assessment for preventive and management purposes.

The use of baseline and serial assessment is critical in the diagnosis and management of sports-related concussion. First, such clinical approach provides crucial information on symptoms and NC performance that helps form a decision whether the athlete is safe to return to play. Second, we are able to track the recovery curve of the concussed athlete. Third, by tracking recovery we are able to compare follow-up cognitive performance to the baseline cognitive functions. In addition, we are also able to see if there is any discrepancy between the athletes' self-reported symptoms and their cognitive performance since there is evidence suggesting that deficits in neurocognitive performance could persist even if the athlete no longer reports symptoms (Lovell et al., 2003, Broglio et al., 2007b).

The issue with reliable concussion assessment & management

Although baseline and follow-up concussion management protocols differ among clinical practices, the use of serial assessment also generates potential confounding issues in the form of practice effect (Barth et al., 1989, Collie, Maruff, Darby, & McStephen, 2003, Peterson, Ferrara, Mrazik, Piland, & Elliott, 2003, Iverson, Lovell, & Collins, 2003a). The confounding issue of learning could skew the results of NC performance, and consequently impact the clinical decision-making about the safe return-to-play of the athlete. Not only is such risk present in traditional paper-pencil NC tests (Barth et al., 1989, Collie et al, 2003, Peterson et al., 2003), but there is more research on computerized concussion assessment tools exhibiting low-to-moderate test-retest reliability and some practice effects (Broglio et al., 2007a, Iverson et al., 2003a). Still, computerized concussion assessment tools resolve more issues in concussion diagnosis and management than traditional paper-pencil tests do: 1) they provide more accurate measures; 2) they are able to control stimuli precisely, 3) have many alternate forms to reduce practice effect, 4) have better psychometric characteristics and more sensitive to subtle cognitive changes; 5) they offer more controlled and standardized administration, and 6) provide precise measurement for multiple cognitive domains (Hall et al., 2005). Therefore, computerized NC testing and clinical assessment forms the basis of today's concussion assessment and management.

Because NC testing has been used for a few decades as a way of detecting sports-related concussion, a lot of emphasis is put on developing more up-to-date techniques of concussion assessment. Traditional neuroimaging procedures (e.g., CT, MRI, EEG) may not detect the effects of a concussion because the acute clinical symptoms in most cases reflect a functional disturbance rather than a structural injury (Aubry et al., 2002). However, these techniques are very helpful in identifying more serious effects such as fracture and intra-cranial bleeding (Aubry et al., 2002). Since neuroimaging techniques may not necessarily detect a concussion, computerized NC tests became more refined for more accurate diagnostic purposes.

ImPACT

To address the limitations of traditional paper-pencil NC tests and neuroimaging techniques, Dr. Lovell and his colleagues at the Pittsburgh University developed a computerized NC test battery, called Immediate Postconcussion Assessment and Cognitive Testing (ImPACT, Lovell, 2006). ImPACT version 2.0 (ImPACT Applications Inc., Pittsburgh, 2000) is currently the most widely used test program specifically designed for the management of sports-related concussion. ImPACT has been a highly useful tool for our concussion management protocol as well along with other assessment tools we have been applying. It is suitable for concussion assessment across high school, collegiate, and professional levels of sport participation. It is also a computerized, Internet-based test program which makes the test easy to administer for both professionals and trainers or coaches. ImPACT 2.0 measures cognitive skills (verbal, visual, and working memory, learning, visual processing speed, reaction time) after concussion; it helps to evaluate the severity of the concussion and the post-injury condition of the athlete as well as to track recovery until the athlete can safely return to play (ImPACT, 2006).

There are numerous studies demonstrating that ImPACT is a sensitive measure of detecting cognitive decrements due to concussion (Iverson, Lovell, & Collins, 2003a, Iverson, Franzen, Lovell, Collins, 2003b, Iverson, Lovell, & Collins, 2005, Iverson, Gaetz, Lovell, & Collins, 2004, Lovell, Collins, Iverson, Field, Maroon et al., 2003, McClincy, Lovell, Pardini, Collins, Spore, 2006, Schatz, Pardini, Lovell, Collins, Podell, 2006). However, test-retest reliability coefficients and practice effects reports on the ImPACT battery are not consistent with one another. While some argue that ImPACT is a reliable assessment tool (Iverson et al., 2003a) and shows no improvement of performance due to learning (Lovell et al., 2003), others claim that it does not meet the necessary test-retest reliability criteria and that performance fluctuations across test sessions may be due to learning (Broglio, Ferrara, Macciocch, Baumgartner, Elliott, 2007a, Randolph, McCrea, & Barr, 2005). Therefore, additional research may be necessary to confirm statements about its reliability rates and whether or not practice effects are involved. These are two of the objectives of this dissertation.

The issue of effort during NC testing and its impact on concussion management

In addition to practice effect, there are studies indicating that there may be a discrepancy between the effort athletes demonstrate at baseline testing and the effort they show at post-concussion testing. Such discrepancy not only affects the NC performance at baseline, but poses risks to the athlete's health as well when the athlete sustains a concussion. For example, an athlete demonstrating lack of effort at baseline NC testing, but showing maximum effort at post-concussion NC testing may have an effect on NC test performance and how results are interpreted. As a result the athlete may seem that he or she is ready to return to play when in fact the athlete has not recovered yet from the actual concussion.

Bailey et al. (2006) described such change in effort as differential motivation in sports-related concussion. They defined "suboptimal motivation" as lack of adequate effort or demonstrating less effort that could occur typically at baseline testing. They defined "optimal motivation" when athletes demonstrated increased level of effort based on the proposition that athletes with a concussion get "highly motivated" to return to play after sustaining a concussion (Bailey, Echemendia, & Arnett, 2006, Bailey & Arnett, 2006). In this sense, the study suggests that athletes who do not demonstrate optimal effort at baseline testing and increase their effort greatly during post-injury assessment, their baseline performance will fall below their true cognitive abilities. Therefore, after a concussion occurs, these athletes may appear to return to their baseline by one week post-injury due to a change in their effort when in fact they may be still recovering. Such diagnostic issues may place athletes to further risks if they are permitted to return to play prematurely based on inaccurate baseline cognitive abilities and underreported concussion-related symptoms. More specifically, such a situation may have dangerous implications on RTP protocols, and therefore could result in premature RTP and in increased likelihood of getting injured again (Guskiewicz et al., 2000). Therefore, it is suggested that confounds of serial assessment and factors that motivate athletes such as love of the game, return to play/practice, endangering scholarship, role in the team, other professional aspirations could be related. Though investigating motivational factors is not the purpose of this dissertation study, the effort that is demonstrated during testing; however, could answer whether such “drive” (i.e., incentives) affects NC performance or not. This is the third aim of this dissertation study.

This second aim is based on research indicating that athletes tend not to complain about their concussion for various reasons, and even if they are diagnosed with a concussion they may not complain about their symptoms either because they are not aware of the severity of the injury or they hide them in the hope of getting back to play as soon as possible (Bailey et al., 2006, Bailey & Arnett, 2006, McCrea et al., 2004, LaBotz et al., 2005, Valovich McLeod et al., 2008). Although more studies address the confounding issue of change in effort in athletes on their NC performance and how it may affect the diagnosis and RTP decisions, only one investigation has been carried out so far in this issue. In this study non-concussed collegiate individuals are proposed to be tested to control for the effects of concussion on cognitive performance. Participants will be tested serially under two different conditions based on whether or not receiving extra monetary incentives.

Research on effort in other settings suggests that certain factors such as extra incentives, contingency management result in improving outcomes and/or modifying certain behaviors, especially in rehabilitation programs (Drebing, Ormer, & Krebs, 2005, Geller, Peterson, &

Talbott, 1982, Hagenzieker, 1991, Hermann, de Montes, Dominguez, Montes, & Hopkins, 1973, Petry, Pierce, Stitzer, Blaine, Roll et al., 2005, Phillips, 1968, Phillips, Phillips, Fixsen, & Wolf, 1971). This provides reason to presume that consequenting test taking behavior to increase participants' effort with extra monetary incentives (i.e., raising interest in performing well on the test) could result in better performance on neurocognitive tests; therefore, may affect the diagnostic process. Since athletes with a concussion often express their intention to get back to play as soon as they can, this drive may lead to changes in athletes' effort to complete the ImPACT battery. As a result, there may be some discrepancies in cognitive performance when comparing the pre- and post-concussion test results. In order to control for the effects of concussion on NC performance, non-concussed collegiate students are proposed to be tested serially in this dissertation study.

Our clinical observations

The assumption that change in effort could affect NC performance is also based on our clinical observation here, at the UTEP Concussion Management Clinic. We have been assessing and treating athletes since 2007. During these assessments we have observed that athletes do not necessarily demonstrate adequate effort when tested at baseline, especially when those testing sessions were organized in groups. This inadequate effort was often demonstrated by athletes often talking to one another during baseline assessment, using their cell phones, watching each others' screen, not always reading or following the instructions carefully, comparing each other's test stimulus and performance, and sacrificing speed for accuracy. For example, invalid baseline test results were often spotted among athletes in those test sessions when athletes were tested in large groups.

Furthermore, when sustaining a concussion some athletes tended to minimize and underreport their symptoms, especially when their injury occurred during the season with an upcoming sports event the following week. By the time these athletes recovered and were about to start the RTP protocol (i.e., the last post-concussion testing), we have noticed that some of them tended to exceed their baseline performance on the ImPACT battery on at least one of the four cognitive composite scores (i.e., scores on Verbal Memory, Visual Memory, Visual Processing Speed, and Reaction Time). These observations lead to the following pilot studies to determine if these observations could be substantiated statistically.

Pilot studies

Using our data three pilot studies were conducted to examine 1) which cognitive functions are impaired due to a concussion on the ImPACT battery; 2) whether our concussed athletes recover (i.e., return to baseline) and whether they exceed their baseline performance; and 3) whether there is a tendency to improve performance on those composite scores that were not impaired on the ImPACT battery. Twenty-two concussed collegiate athletes were administered to the Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT version 2.0, ImPACT Applications Inc., Pittsburgh, 2000) battery prior to and after sustaining a concussion to evaluate cognitive deficits and self-reported symptoms, and to evaluate whether athletes recover and exceed their baseline performance on the ImPACT battery when tested serially. Based on previous studies, performance was expected to decrease significantly on all the composite scores of ImPACT when comparing pre- to the first post-concussion test results. Self-reported symptom scores were expected to be significantly higher at the first post-concussion assessment than at baseline. Furthermore, we expected athletes to improve their composite scores on the ImPACT battery when comparing the first and the last post-concussion (i.e., when athletes no longer report symptoms and their performance returns to their baseline level) test results indicating that they recover from their injury. However, greater performance was expected on at least one cognitive composite score when comparing the last post-concussion test results to the baseline. This last comparison may also indicate that this group of athletes with a concussion might demonstrate a change in effort as reflected by their composite scores on the ImPACT battery.

Baseline NC performance was compared to the first and to the last post-concussion performance of the 22 athletes to investigate whether athletes with a concussion recover and exceed their baseline performance on ImPACT. In addition, the first and last post-concussion performance was compared for reliability purposes to examine whether pattern of cognitive recovery is consistent with the cognitive deficits marked on the first post-concussion evaluation. Verbal Memory, Visual Memory, Reaction Time, Visual Processing Speed, and Total Symptom Scores were measured as the dependent variables for this study. Non-parametric related-samples Friedman and Wilcoxon tests revealed that our sample of concussed athletes displayed deficits only in Reaction Time and reported significantly more concussion-related symptoms when comparing baseline to the first post-injury test results. However, the group of 22 athletes with a concussion performed significantly better on Visual Memory when compared their baseline to the last post-concussion performance. Our concussed athletes also displayed the tendency to improve on Verbal Memory and Visual Processing Speed when comparing their *median* cognitive performance across testing sessions (i.e., baseline, first, and last post-concussion assessment) on the ImPACT battery. This improvement; however, was not statistically significant. In addition, comparison between the performance of the first and last post-concussion indicated that athletes recover from their concussion. Still, this group's tendency to show higher composite scores in Verbal and Visual Memory, and Visual Processing Speed at the last post-concussion testing than at baseline has raised further questions. A full report of these findings will be covered in Chapter II.

These results suggest that our group of athletes may demonstrate change in effort as reflected by their NC performance when tested serially on the ImPACT battery despite their concussion. The data also proposes that effort (i.e., due to the expressed desire to get back to play as soon as possible) may affect cognitive performance and may impact the clinical evaluation and management of sports-related concussion. Future study is necessary to further examine the impact of change in effort on NC performance and its impact on concussion diagnosis and management. Because test-retest reliability measures of the ImPACT battery have been inconsistent with one another and did not take change in effort into consideration, it is crucial for clinical purposes to examine whether changes in effort result in changes in cognitive performance.

CHAPTER TWO

Literature Review

Definition of concussion

From the early 20th century until the end of the 1950's, concussion was described as an injury caused by a severe blow to the head that is accompanied by brief loss of consciousness. The first complex definition of concussion was introduced by Symonds (1962), who argued that even without loss of consciousness (LOC), the acute symptoms may be the same "with subsequent long-continued disturbance of consciousness often followed by residual symptoms". In addition, he concluded that concussion depends on diffuse injury to nerve cells and fibers occurring at the moment of the accident in which the injury may not even be reversible.

As research advanced in mild head injuries, so did the definition of concussion become more refined. A frequently applied definition was suggested by the American Academy of Neurology (AAN, 1997) by describing concussion as "a trauma induced alteration in mental status that may or may not involve loss of consciousness" (p. 582).

More recently, a consensus statement was provided on the definition of concussion at the 3rd International Conference on Concussion in Sports. In this statement concussion is defined "as a complex pathophysiological process affecting the brain, induced by traumatic biomechanical forces" (McCrory, Meeuwisse, Johnston, Dvorak, Aubry et al., 2009, p. 341). The nature of a concussive brain injury includes several common features including the following: 1) it may be caused by a blow or jolt to either the head or to the body that causes the brain to move rapidly inside the skull; 2) it typically results in the rapid onset of neurological impairments that are usually short-lived and resolve spontaneously; 3) it may result in neuropathological changes; however, the acute clinical symptoms largely reflect a functional disturbance rather than a structural injury; 4) it results in a set of somatic, psychological, and cognitive symptoms that may or may not involve loss of consciousness; 5) it typically results in grossly normal structural neuroimaging studies (McCrory et al., 2009).

Prevalence

According to the Center for Disease Control and Prevention (CDC), an estimated 1.6 to 3.8 million sports- and recreation-related concussions, or mild traumatic brain injuries (mTBI), occur each and every year in the United States (CDC, 2007). About 300,000 of the reported incidents are sports-related involving LOC (Thurman & Guerrero, 1999). Furthermore, 75% of all traumatic brain injuries are diagnosed with a concussion or with other forms of mTBI (CDC, 2007).

As reported further by the CDC (2007), children and adolescents between the ages 5-18 account for 2.6 million emergency department visits, of which 135,000 were diagnosed with concussions each year between 2001 and 2005. As for collegiate sport injuries, more than 9000 concussions were reported in the last 16 years; that is an average of 563 annually (Hootman, Dick, & Agel, 2007). Assuming the sample represents 15% of the total population, the number of concussions in collegiate sports would equate to 3753. Among all reported concussions, football is the leading cause (55% of all reported head injuries) for college males, while for college females soccer represents a relatively high rate of concussion. This pattern is consistent with other concussion surveys conducted in high schools.

According to Echemendia & Julian (2001), data collected from high schools, colleges, and professional sports teams indicate relatively high rates of mTBI in many sports. More specifically, in high school sports, concussions occur more often in competitive sports. As reported by Powell & Barber Foss (1999) and CDC (2007), the main cause of mild head injury for high school males is football, while for females the main cause is soccer. Particularly for children and adolescents (between the ages 5-18), bicycling, soccer, football, basketball, and other playground activities are the five primary causes that lead to mild head injuries (CDC, 2007). Based on these data, Echmendia and Julian (2001) estimated that mTBI for the US population project approximately 63,000 concussions occurring each year only in high school sports, 63% of which are football related head injuries. It must be noted, however, that these figures do not include data on high school ice-hockey, rugby, or lacrosse which also involve high risk of mTBI (Echmendia & Julian, 2001).

The epidemiology research of Guskiewicz, Weaver, Padua, & Garrett (2000) investigated concussion in football and other related issues such as incidence of injury and patterns of return-to-play decisions in high schools and colleges. Using the data gathered from 17,549 football players, the incidence was the greatest at the high school level (5.6%) and collegiate division III (5.5%) level. Furthermore, the results showed that the players who sustained one concussion in a season were three times more likely to sustain a second concussion in the same season when compared with uninjured players. In terms of concussion management, the authors reported that 30.8% of all players sustaining a concussion returned to participation on the same day of injury, which might leave the players in question highly vulnerable to sustain a second head injury. Because even concussions may have serious long-term effects, it is crucial to identify and manage head injuries properly in order to avoid the consequences of a second head injury. Suffering a second head injury before recovering from a previous concussion may leave the person severely disabled, or it could even cause death. To avoid such catastrophic consequences, emerging emphasis has been placed on proper concussion management.

Brief history of concussion diagnosis, assessment & management

Recognizing a concussion. Although concussion is well-documented across centuries, there was much confusion how to define it and, consequently, how to diagnose it even in the late 20th century. The issue is rooted in the detection of concussion since no separation was made between mild traumatic brain injury and severe TBI (McCrory & Berkovic, 2001). Documentation of concussion started very early; the first written records of concussion can be found among the works of Hippocrates (460-377 B.C.), who actually used the term “concussion” but most likely referring to brain injuries in general (Berner, 1935). Along with the documentation, the first assessments were based on salient and unambiguous observations. For example, Hippocrates in his observations made comments about concussions in which the patient becomes speechless and unconscious (McCrory & Berkovic, 2001). Second, not only the documentation but the etymology of the word “concussion” demonstrates how far its history goes back; it derives from the Latin verb *concutere* which means “to shake violently” (Cantu, 1996). Though concussion has been recognized since the ancient times, the term became widely used most probably in the Renaissance (between the 14^{th} – 16^{th} century; McCrory & Berkovic, 2001).

The first detailed description of concussion appeared in the 10^{th} century among the records of Rhazes (853-925 B.C.), a Persian physician. He distinguished concussion from severe head injury by explaining concussion as an abnormal physiological state but with no sign of gross traumatic lesions in the brain (Zillmer, Schneider, Tinker, & Kaminaris, 2006).

In the Roman period, Galen (129-216 A.D.), surgeon and anatomist, made the first reference to sports-related head injury that occurred during gladiator games. In the Roman times (and even in the Middle Ages) the Olympic Games were centered mainly around war games such as wrestling and gladiator fights (Zillmer et al., 2006).

In the Middle Ages, “concussion” became a widely used term. Physicians, such as the French Ambroise Paré and the Italian Jacopo Berengario de Carpi, were more focused on describing the symptoms of concussion and introduced the term “commotio cerebri” (Berner, 1935; Shaw, 2006). Commotio cerebri, another Latin term, literally meant “shaking of the brain” and it was widely used in Europe as another synonym term of concussion (Kruijk, Twijnstra, Meerhoff, & Leffers, 2001). Commotio cerebri later on stood for closed head injury with post-traumatic amnesia lasting less than an hour and/or loss of consciousness lasting less than 15 minutes. During the middle ages, surgeons observed that as a result of concussion, a short-lasting paralysis of cerebral function appeared with acute symptoms such as impaired memory, mental confusion. The second cutting-edge observation was that concussions may occur without any skull facture and may result in brain swelling and/or intracerebral hemorrhage (Shaw, 2006).

From the age of Enlightenment throughout the 19th century, concussion "research" made its first attempts to conduct systematic observations. It was characterized as the period of proposing hypotheses to explain the underlying mechanisms of concussion and find etiological relationship between the injury and its symptoms (McCrory & Berkovic, 2001). The first experimental animal studies took place in this period (Berner, 1935). In addition, the study of acute concussion expanded with the observations of post-concussion symptoms. Besides concussion research, athletics changed in the 19th century as well: Wimbledon was established, the Olympic Games were reinstituted, and organized football emerged in the U.S. (Zillmer et al., 2006). Later these events turned out to be great sources of more elaborate study of mild head injuries.

The 20th century brought the first major breakthrough in concussion research and sports, especially in the U.S. By that time football head injuries became so frequent that they drew more and more attention to the safety issues affecting the players. With athletes not wearing protective helmets and the sport lacking safety regulations, in a year 19 athletes got killed or suffering paralysis during playing football. Therefore, there was much chance that the sport would grow to further brutality. To save the sport and protect the players, President Roosevelt established the National Collegiate Athletic Association in 1906. As a result, the rules of the game were changed and more attention was paid to protective outfit. Unfortunately, even after establishing the

Association and the new rules, football deaths still continued to be a constant phenomenon until 1973 when helmets actually became mandatory for football games reducing the number of fatal and concussive head injuries (Zillmer et al., 2006). This crucial step in protecting athletes also may have been due to the increasing assessment approaches in concussion and sports-related head injuries.

However, not only was the assessment at that time was at its early stage, but the definition of concussion was still somewhat unrefined. The first definitions of concussion in the 20th century were so broad that they lead to confusion in terms of defining different types of head injuries. Though the symptomology was somewhat more firmly established, distinction between head injuries was not explored that elaborately. Papers on concussion distinguish between serious and slight concussion based on the length of loss of consciousness (Berner, 1935). For example, at the beginning of the 20th century concussion was regarded as an injury that is caused by:

> "a severe blow on the head and is marked by unconsciousness, shallow respiration, slow pale, cold skin and pallor, often followed by partial stupor, vomiting and severe headache. The respiration may occasionally be so shallow that the patient gives the appearance of being dead. In most cases, however, consciousness is regained within a few minutes and these are therefore called "slight concussion." When consciousness is not regained for several days the case is described a "serious concussion" (Berner, 1935, p. 273).

Until the end of the 1950's, concussion was described as an injury that is accompanied by brief loss of consciousness. A more complex definition of concussion was introduced by Symonds (1962). He was also the first scientist to propose that loss of consciousness does not necessarily accompany all concussions. He argued that even without loss of consciousness the acute symptoms may be the same "with subsequent long-continued disturbance of consciousness often followed by residual symptoms”. In addition, he concluded that concussion depends on diffuse injury to nerve cells and fibers occurring at the moment of the accident in which the injury may not even be reversible.

Concussion & early neurocognitive (NC) assessment. In the 1980’s clinical and epidemiological studies revealed more specific neurocognitive deficits such as confusion, attention and concentration problems, memory impairments, slow information processing, decreased verbal expression, slow reaction time, and decreased problem solving (MacFlynn, Montgomery, Fenton, Rutherford, 1984, Conboy, Barth, & Boll, 1986, Armstrong, 1987, Stuss, Stethem, Hugenholtz, Picton, Pivik, Richard, 1989). In addition, a more elaborate analysis of symptoms appeared in concussion research that refined the understanding of concussion. Symptoms were distinguished among somatic such as headache, visual problems, dizziness; neurological and psychological dysfunctions such as depression, irritability, and anxiety.

The second major breakthrough of the 1980’s was the use of neurocognitive tests that emerged in this period for further scientific exploration in concussion and for careful assessment and identification of the injury (Conboy et al., 1986). In addition, more studies were focusing on the epidemiology and concussive incidents in young athletes at schools (Goodwin Gerberich, Priest, Boen, Straub, & Maxwell, 1983).

Although previous papers showed confusion about how to manage concussions and under what conditions can athletes return to play (Murphy & Simmons, 1959), by the 1980’s there was much understanding about the risks of premature return to competition. Studies on concussive incidents were gaining more and more attention because of the rate of fatal injuries and the long-lasting effects of the symptoms. Therefore, much emphasis was put on the actual identification, careful assessment and management of mild head injuries (MacFlynn et al., 1984, Stuss et al., 1989).

Identification of concussion, however, appeared to be problematic since most concussions were recognized to appear without loss of consciousness. Because the signs and symptoms were often left unrecognized as a concussion and because the risk of progressive debilitation was potential, the identification of concussion became extremely crucial. For example, Goodwin et al. (1983) shared the views of Maroon, Steele, & Berlin (1980) and proposed that even those with suspected concussion must not return to play "until they can move with usual dexterity and speed and are perfectly oriented as to the time, place, and their own identity and are able to identify the activities in which they were engaged just prior to the injury incident" (Goodwin et al,, 1983, p. 1370). Consequently, these studies included the first return-to-play decisions to be suggested. As a result of early NC testing, return-to-play decisions became much more developed with the advancement of concussion research and with more understanding of the relationship between injury and its consequences (e.g., length of symptoms). For further management of concussions and to avoid premature return-to-play decisions, Goodwin et al. (1983) proposed to continue observing the cases that became persistent and making sure that the symptoms changed over time. Still, the research that was done mainly focused on measuring post-concussion syndrome and the effects of head injuries in general. Even though distinctions between head injuries became more elaborate in this period, studies were not yet consistent in how they define concussion. As an example, Stuss et al. (1989) measured reaction time with varying time intervals (within 3 months post-injury) in people with concussion which they defined concussion as a head injury with loss of consciousness but without demonstrable focal neurological deficit. On the other hand, the study of MacFlynn et al. (1984) lacked the definition of mild head injuries and simply included patients whose post-traumatic amnesia was less than 24 hours. Whether participants suffered loss of consciousness was unreported.

The third major breakthrough came with the study of Barth et al. (1989). These authors were among the first to introduce an assessment and management approach that differed from previous clinical approaches. First, they set up a preseason "baseline" NC performance of athletes which could be compared later on to post-concussion NC performance. Then these concussed athletes were tested serially until their cognitive performance returned to the level of their baseline performance. Third, post-concussion symptoms also had to resolve to baseline level. The approach of *baseline assessment of athletes* was based on previous suggestions that athletes should not be allowed to RTP until cognitive functioning resolves to normal and athletes become symptom-free. In addition, the "baseline approach" also offered a solution for the issue that head injuries vary across individuals. Even though testing cognitive-neurological functioning has been discussed in concussion research literature for a longer period of time, Barth et al. introduced a more refined RTP guideline by assessing athletes in a systematical manner: pre- and post-concussion. The basic notion behind examining cognitive functioning in athletes with a concussion was to make sure they can safely return to competition or practice and decrease the likelihood of developing a Post Concussion Syndrome or a disabling and in some cases deadly Second Impact Syndrome.

Moving towards a new era in concussion assessment & management.

During the 1990's concussions in both high school and collegiate contact sports became the major focus in mild head injury research. More and more research was focusing on 1) the neurometabolic changes that occur during a concussive head injury (Wojtys, Hovda, Landry, Boland, Lovell et al., 1999); 2) classifying and grading concussions (Colorado Medical Society, 1991, AAN, 1997); 3) exploring the recovery pattern following a concussion, 4) developing more refined RTP guidelines, 5) post concussion syndrome and second impact syndrome (Cantu &

Voy, 1995, Cantu, 1996, Erlanger, Kutner, Barth, & Barnes, 1999) and 6) developing traditional NC tests to computer programs for assessing sports-related concussion.

Neurometabolic research. Research on the neurometabolic cascade advanced considerably. These studies observed that immediately after a biomechanical insult to the brain, a blunt, sudden, and unsystematic release of neurotransmitters and unchecked ionic fluxes occur. This indiscriminant and abrupt process leads to further ionic shifts due to neuronal depolarization with the discharge of potassium and influx of calcium. These ionic shifts end in changing cellular physiology. Parallel to this process, axonal damage occurs as well with fibers becoming swollen and stretched (Wojtys et al., 1999; Giza & Hovda, 2001). The findings were extremely crucial since they showed some evidence that cognitive signs and symptoms of concussion may be strongly related to the neurometabolic cascade as a result of an insult to the head.

Grading systems. The classification of concussion improved so much that by 2000 there were numerous grading systems categorizing concussion severity (e.g., Cantu, 1986; Colorado Medical Society, 1991; American Academy of Neurology, 1997). These grading systems included loss of consciousness, amnesia, confusion, and the length of symptoms for characterizing a concussion as grade I (mild), grade II (moderate), and grade III (severe). Although some grading systems differed from one another, for the first time loss of consciousness as an issue in identifying a concussion seemed to reach consensus among researchers by recognizing that even without loss of consciousness an athlete can still suffer a concussion. The second major improvement in developing grading systems was that they helped forming a diagnosis (Cantu, 2001). Thirdly, they also lead to further concussion management and treatment by proposing new return-to-play (RTP) guidelines. Such advanced RTP guideline was proposed by the American Academy of Neurology (1997). Today the three major grading systems used for concussion diagnosis were developed by Robert Cantu, American Academy of

Neurology (AAN), and the Colorado Medical Society (Collins, Lovell, & McKeag, 1999, see Table 1 in Appendix A).

Nature of recovery. Although the grading systems assisted in the diagnosis and distinction among the severity of head injuries, much research started to focus on the recovery pattern of athletes with a concussion. The purpose behind such research practices lay in the importance of developing a safe, patient-oriented concussion management protocol. Therefore, recent studies have been investigating how long it takes to recover from a sports-related concussion. According to research, there is evidence that symptoms and cognitive deficits usually resolve within 7-14 days post injury in high school and collegiate athletes with a concussion (Lovell, Collins, Grant, Iverson, Field, Maroon et al., 2003, McClincy, Lovell, Pardini, Collins, & Spore, 2006). Furthermore, self-reported symptoms of headache, dizziness, and nausea tend to resolve by Day 4. As an example, Lovell et al. (2003) found that high school athletes with amnesia and disorientation that last longer than five minutes did not fully recover within seven days post-injury and showed decline on NC performance when compared to their baseline scores. In contrast, high school athletes with mental status changes that last less than five minutes returned to their baseline level of NC performance within four days post-injury.

Duration of on-field mental status changes (amnesia and confusion) showed a relationship with the length of recovery. High school athletes who suffered amnesia, confusion, and disorientation lasting longer than 5 minutes did not fully recover in a week. They demonstrated longer-lasting post concussion symptoms and a decline in memory performance. On the other hand, athletes with on-field mental status changes lasting less than 5 minutes returned to their baseline level of performance within 7 days post-concussion (Lovell et al., 2003).

The results of McClincy et al. (2006) suggest that the RTP protocol of AAN (1997) may not be suitable to apply. AAN guideline permits grade I concussion (i.e., mild, no LOC) RTP the same day on injury if symptoms resolve in 15 minutes and allows athletes with grade II concussions (i.e., moderate, no LOC) RTP within 1 week. Their study, however, suggests that athletes with a concussion (but with no LOC) recover in 2 weeks. This is a slower recovery pattern with cognitive deficits that usually tend to resolve by day 14. Another finding of the study was that cognitive deficits may be present even though the athlete does not report any symptoms. In addition, recovery times may differ significantly from one athlete to another; therefore individualized NC assessment is recommended when making RTP decisions (Echemendia & Cantu, 2003; Collins et al., 2004). Furthermore, NC testing has to be sensitive enough to demonstrate cognitive deficits, even if athletes do not report symptoms.

Age-related differences in recovery. Field, Collins, Lovell, & Maroon (2003) suggest that age-related differences between recovery patterns exist. The authors showed that high school and college athletes differ significantly in their recovery from a concussion. High school athletes with a concussion demonstrated that they recover slower than college athletes with a concussion. As an example, high school athletes with concussion had significant memory impairments at least 7 days post-injury, whereas college football players showed significant memory deficits for only 24 hours after injury. Findings in self-reporting symptoms were consistent with McClincy et al (2006). These findings suggest that recovery among athletes sustaining a concussion is highly diverse. Therefore, individually-based NC assessment is a crucial perspective in concussion evaluation and management. These people may vary in the time course of physical and cognitive recovery depending on the number and severity of their symptoms, health history, and age. Such factors have to be considered when evaluating a concussion.

Field et al. (2003) also found that self-reported symptoms resolved by day 3 after injury, but athletes still demonstrated significant cognitive deficits at least 7 days after a concussion. This may suggest several assumptions. One may be that athletes with a concussion may desire to return to play; hence, they would underreport their symptoms without even knowing what the consequences may be. Secondly, symptoms may resolve prior to cognitive deficits, indicating that the brain is still at a vulnerable phase and needs time to heal even if athletes no longer experience symptoms. Such discrepancies between symptom-reporting and cognitive recovery call for a carefully structured RTP guideline.

Return-to-play (RTP) guidelines. With the improvement of concussion assessment and severity grading system, the nature of neurocognitive recovery became more observable. These advances in concussion management contributed to more specific return-to-play (RTP) guidelines to protect athletes' health and well-being. AAN was among the first to develop RTP guidelines for all 3 grades of concussion (American Academy of Neurology, 1997). Therefore the RTP decision was much more guided by the severity of concussion and the sideline examination: AAN suggested removing the player from competition right after the injury and examine the symptoms and mental status immediately after the injury. The RTP decision was based on whether the athlete shows signs or symptoms of a concussion, suffered LOC, the results of the CT or MRI scan is, and his mental status is clear. These guidelines became crucial in order to prevent severe or even catastrophic incidents following a head injury. Consequently, AAN put much emphasis on allowing athletes return to play only if 1) the athlete is asymptomatic, 2) has a normal CT or MRI scan if symptoms may last longer than a week and/or the athlete suffered a severe concussion, and 3) the neurologic examinations come back as normal.

Making accurate RTP decisions is crucial for a number of reasons. First, according to research, athletes sustaining a concussion are extremely vulnerable for sustaining a second a concussion (Cantu & Voy, 1995, Guskiewicz et al., 2000). If a head-injured athlete returns to play before his or her symptoms have resolved completely, he or she may be at an increased risk to suffer a disabling or even fatal injury if the previous concussion has not healed yet. Research and clinical practice names such incidences as Second Impact Syndrome (SIS, Cantu & Voy, 1995, Echemendia & Julian, 2001). Therefore, management of first concussion is crucial in order to avoid the devastating consequences of SIS. Second, recurring concussions may lead to chronic, long-lasting cognitive, physical, and psychological deficits (Iverson et al., 2004). If symptoms get worse over time and last longer than three months, literature identifies such mental status as Post-Concussion Syndrome (PCS) (Echemendia & Julian, 2001, Hall et al., 2005). Therefore, there has been much emphasis put on NC assessment based on the fact that most clinicians use it to avoid premature RTP decisions and make sure the athlete has healed from his/her concussion completely.

Tracking recovery through serial NC assessment is a basic and necessary method to make safe RTP decisions. However, the variability in recovery among people sustaining a concussion is highly diverse. Therefore, individually-based NC assessment is a crucial perspective in concussion evaluation and management. These people may vary in the time course of physical and cognitive recovery depending on the number and severity of their symptoms, health history, and age. A study examining how age may play a role in recovery from sports-related concussion compared NC performance between high school and collegiate athletes sustaining a concussion (Field et al., 2003). The results showed that high school athletes with a concussion demonstrated a more prolonged memory dysfunction than did collegiate athletes with a concussion. Consequently, high school athletes experiencing a mild head injury may demonstrate slower and more diverse recovery from the injury when compared to concussed collegiate athletes. This is, again, another finding that suggests individualized assessment contributes to better concussion management since research could not demonstrate a protocol that could fit all concussions.

For example, in the case of children and adolescents, normal developmental process and how it may interact with and affect the outcome of a concussion should be taken into consideration when making RTP decisions (Patel et al., 2005). More specifically, reaching the baseline level of performance after a concussion may not necessarily indicate the resolution of symptoms and the readiness to return to competition in children and adolescent athletes. Since normal maturation and cognitive development are ongoing in these years, the authors suggest more frequent baseline testing to avoid confusion in the RTP decision making. Also, future research is necessary to develop RTP guidelines for children and adolescent athletes and to better clarify the potential differences between adults and children with regards to recovery from mild head injury (McCrory et al., 2005, Patel et al. 2005).

As for symptoms, most concussion specific NC assessment tools include a symptom inventory in which subjects can mark and rate the presence of a symptom. Tracking symptoms of concussion and how they resolve is, again, a key factor in RTP decisions. According to the First International Conference on Concussion in Sport (Aubry et al., 2001), whenever an athlete shows any signs or symptoms of a concussion, s/he “should never return to play” (Aubry et al., 2002, p. 8). Only when the symptoms resolve, is the athlete allowed to return to practice or competition. However, even then, the athlete must follow a strict RTP protocol.

Based on the Consensus Statement on RTP that was approved at the International Conference on Concussion in Sport (McCrory et al., 2004), today’s protocol follows a stepwise strategy for RTP and consists of 6 steps. Step 1: During the time of recovery the athlete is recommended to do no activities other than having rest, while the athlete may be seen within 36 hours, 4 and 7 days after sustaining a concussion for NC assessment to follow track of recovery (Lovell et al., 2003). Step 2: Only if the athlete becomes asymptomatic is he allowed to do some light aerobic exercise (i.e., walking, but no training). If the athlete experiences any symptoms, s/he should rest until symptoms resolve and start the protocol again. Step 3: If the athlete does not experience any symptoms after light aerobic exercise, the athlete can start to do some sport-specific exercise. Step 4 is when the athlete is allowed to do some non-contact training drills. Step 5: If still symptom free, the athlete can proceed to full contact training after being medically cleared. Step 6 is when the athlete can return to competition. In addition, if the athlete experiences any concussion symptoms, he athlete has to start the protocol again from step 1.

The RTP protocol of the University of Pittsburgh Medical Center Sports Concussion Program is basically the same as the RTP protocol approved by the Consensus Statement of International Conference Concussion in Sport (McCrory et al., 2008). However, Lovell et al. (2004) strongly recommend that especially high school and younger athletes should not be returned to play once they sustain a head injury during game no matter how mild it is. They argue that keeping them out of the game would allow for closer evaluation of evolving signs and symptoms and will help to prevent more severe injury. The authors also recommend formal NC testing (e.g., ImPACT) within 24 hours post-injury to assess acute neurocognitive status followed by regular re-evaluation of the athlete.

Neurocognitive (NC) assessment

Although traditional NC testing was conducted in concussion for research purposes at first, the growing scientific findings gradually led to the clinical use of NC testing to manage the injury. As a result, sideline and baseline assessment evolved from NC testing and became significant parts of actual concussion management besides post-injury assessment by the mid 1990's (Lovell & Pardini, 2006).

Before computerized assessment, the identification and evaluation of concussion was mainly organized around traditional paper-pencil tests. These tests were usually used for three reasons: first, to evaluate the acute effects, secondly to evaluate the long-term effects of concussion; and thirdly to evaluate the effects of multiple concussions.

With the introduction of computers there was a better chance that NC testing could change and become more effective and precise in measuring performance. Such test development started in the 1970's, when test developers started working on developing traditional paper-pencil tests into computer-based batteries (Schlegel & Gilliland, 2007). In this period, concussion assessment both at baseline and post-injury levels were done by traditional paper-pencil tests. These tests included short batteries that measured memory and attention (e.g., Hopkins Verbal Learning Test, Brief Visuospatial Memory Test, WAIS-III Digit Span Test, WAIS-III Letter-Number Sequencing Test) processing speed and executive functioning (e.g., WAIS-III Digit Symbol subtest, Symbol Digit Modalities Test, Trail Making Test, Controlled Oral Word Association Test, Stroop Color Word Test), and symptoms (Post-Concussion Symptom Checklist).

For example, Barth et al. (1989) used the Trail Making Test A & B (Reitan, 1955), the Paced Auditory Serial Addition Task (PASAT, Gronwall, 1977), and the Symbol Digit Test (Smith, 1973) to measure cognitive performance of athletes at baseline and within 24 hours, 5, and 10 days post-injury. The athletes' performance demonstrated that these tests are sensitive measures of testing cognitive functioning when athletes sustain a concussion. A significant finding was that athletes by day 10 returned to their baseline performance indicating that mild concussions may resolve in two weeks.

Following the baseline approach of Barth and colleagues, Hinton-Bayre, Geffen, & McFarland (1997) conducted a study with professional rugby players on both baseline and post-injury levels. They investigated the effect of sensitivity in tests of processing speed of information and the effect of practice. The authors used short tests measuring speed of information such as the Symbol Digit Modalities Test, the Digit Symbol Substitution Test, and the Speed of Comprehension Test. The results showed that these tests were sensitive to cognitive deficit in the speed of information processing; however, the results on 2 tests (i.e., the Speed of Comprehension and Digit Symbol Substitution tests) actually improved with practice, whereas the Symbol Digit Modalities test proved to provide reliable test results.

As a result, another assessment issue emerged with the use of neurocognitive tests: patients' performance after a number of follow-up assessments exceeded the performance of the non-concussed control group. As a result, the question whether tests show reliable and stable results over time started a chain reaction of ideas as researchers were thinking which neurocognitive tests they could use over time for concussion management and for monitoring true recovery (MacFlynn et al., 1984).

Although NC testing became useful for concussion assessment and for characterizing the injury, there were concerns about their reliability measures (i.e., whether a test provides stable and reliable results over time) about some tests having poor test-retest reliability and significant practice effects as well due to the lack of alternate forms (Cernich, Reeves, Sun, & Bleiberg, 2007). For the majority of these paper pencil tests, test-retest reliabilities ranged between 0.39 and 0.78 (Broglio et al., 2007). In addition, concussion assessment (especially baseline/preseason assessment) started to receive more and more interest, even though the assessment itself was lengthy and time-consuming. Since usually one test measured only one cognitive function such as memory, processing speed, reaction time, attention, more tests had to be administered to athletes to measure more cognitive functions at both baseline and post-injury levels. These batteries, however, proved to be expensive as well. Because conventional paper-pencil tests may take several hours to administer, require a neuropsychologist to administer and evaluate the results, and are not available for high school and collegiate levels of sport participation, there has been growing need for cheaper, less time-consuming, more available assessment batteries (Collie et al., 2001; Patel et al., 2005). In addition, these traditional tests were not specifically designed for concussion assessment and management. All these issues led to developing computerized assessment batteries that are specifically designed for concussion assessment and management.

Computerized NC assessment. As a response to administering traditional tests, more traditional tests were upgraded and developed into computerized test forms. Specifically in concussion assessment, computerized baseline assessment of athletes started around the early 1990's. The earliest test battery that was designed particularly for sports-related concussion assessment and management is called the Immediate Post Concussion Assessment and Cognitive Testing battery (ImPACT, 1989). ImPACT was developed by Drs Mark Lovell, Joseph Maroon, and Micky Collins at the University of Pittsburgh Medical Center (Lovell, 2006). The battery that was based on adopting certain traditional paper-pencil tests got expanded and modified later on. Recently other concussion assessment programs have been developed. These test batteries are the Automated Neuropsychological Assessment Metrics (ANAM, Reeves, Winter, Bleiberg, & Kane, 2007), CogSport (CogState Ltd), Concussion Sentinel (CogState Ltd.), HeadMinder Concussion Resolution Index (CRI, Headminder Inc., 1999), and the above mentioned Immediate Post Concussion Assessment and Cognitive Testing (ImPACT Applications Inc., Pittsburgh) battery. According to Patel et al. (2005) and Cernich et al., (2007), computer-based assessment programs have many advantages as opposed to traditional testing; they are cheap, easy to administer, short, collect and store data automatically, and more available for amateur sports teams such as high schools and colleges. Furthermore, they provide more accurate measures: they are able to control stimuli precisely, have many alternate forms to reduce practice effect, have better psychometric characteristics, are more sensitive to subtle cognitive changes; they offer more controlled and standardized administration and provide precise measurement for multiple cognitive domains. Therefore, computerized NC testing and clinical assessment forms the basis of today's concussion assessment and management.

The importance of NC assessment in athletes was first emphasized by the First International Conference on Concussion in Sport (Aubry et al., 2002). A consensus was reached that “neurocognitive testing is one of the cornerstones of concussion evaluation and contributes significantly to both understanding of the injury and management of the individual” (Aubry et al., 2002, p. 8). However, most traditional paper-pencil tests are not specifically designed for testing the effects of mild head injuries, and they still involve practice effects. Computerized tests using different stimuli randomly may overcome the issue of practice effect, and have the advantage to measure reaction time and processing speed more accurately (with ImPACT reaction time is measured to 1/100th of second) than conventional tests (ImPACT , 2006).

Here, at the UTEP Concussion Management Clinic, we have been using the ImPACT version 2.0 battery (ImPACT Applications Inc., Pittsburgh) since 2007 for baseline and post-concussion assessment. This dissertation project is designed to apply ImPACT as the assessment instrument for measuring NC performance; therefore, the next sections introduce and describe all the essential properties of ImPACT that is necessary to understand 1) how the battery was developed and designed; 2) what it measures; 3) its structure and parts of the test; 4) what its psychometric properties are (how reliable/valid/sensitive it is); and 5) its clinical value.

ImPACT

According to Lovell (2006), ImPACT was developed in the 1990’s by the Pittsburgh Steelers concussion program 1) to bridge the gap between traditional testing and reliable/sensitive assessment, 2) to supplement the limitations of neuroimaging technique, and 3) to contribute to concussion management specifically. At first, Version 1.0 of ImPACT was available for testing in 2002, while the latest version of ImPACT (i.e., version 2.0) was released in 2004. The difference between the first and the second version is the addition of the ‘Design Memory’ test, which measures visual memory. Adding this subtest clearly separated memory into 'verbal' and 'visual' memory; therefore it calculates separate composite score (Lovell, 2006).

Version 2.0 of ImPACT battery is one of the computerized concussion assessment tools that has been used most often currently. According to Lovell (2006), the majority of National Football League teams and NASCAR, Formula 1, and other motor sports apply ImPACT. Furthermore, 125 Colleges and 300 high schools utilize ImPACT across the United States. As noted, the staff of the UTEP Concussion Management Clinic has been administering ImPACT 2.0 to athletes for both baseline and post-injury testing since 2007.

Considering its many advantages, ImPACT has become very popular since its development for many reasons: First, it is specifically designed for concussion management by measuring cognitive skills (i.e., verbal and visual memory, reaction time, processing speed of information, and total symptom score) after head injury. Second, it is easy to access and administer. Third, it does not necessarily require a neuropsychologist to administer the test. Fourth, it allows both individual and group assessment across high school, collegiate, and professional levels. Fifth, it automatically stores data and calculates results immediately after the test. Sixth, it assists physicians, trainers or coaches in concussion management and making careful RTP decisions. Seventh, ImPACT 2.0 offers alternate test forms to reduce practice effect. In addition, test takers can choose the language in which they want to complete the test; the test is available in 12 different languages. Detailed instructions are described before each and every test module. Completing the test requires approximately 40 minutes.

The structure of ImPACT (based on Lovell, 2006). ImPACT consists of seven sections including the six test modules (i.e., the actual test). The seven sections are:

(1) Demographic data (date of birth, age, gender, handedness, height, weight, native country, native language, second language, years of speaking);

(2) Education (years of education completed; if they have ever repeated one or more years of school; received speech-language therapy; were diagnosed with learning disability or hyperactivity and attention deficit disorder (ADD); attended special education);

(3) Current sport history (type of sport, level of participation, position, years of experience at this level);

(4) Concussion history (number of times diagnosed with a concussion, concussion resulted in confusion and/or loss of consciousness, difficulty in remembering events that occurred before and/or after the injury, total games missed);

(5) Health history (treatment for headaches, migraines, seizures, substance abuse, psychiatric condition by a physician, history of brain surgery, and meningitis);

(6) Symptom inventory; this section includes 22 concussion-related symptoms. Athletes have to self-report and rank the intensity of the range of symptoms they experience on the test date from 0 to 6: with 0 meaning not experiencing the symptom through 6, indicating the severity of the symptom. The post-concussion symptom scale was adapted from Lovell & Collins (1998);

(7) The actual neurocognitive test battery that includes 6 test modules. The six cognitive modules (i.e., the actual test) consist of the following designs: (1) Word Memory; (2) Design Memory; (3) the X's and O's; (4) Symbol Match; (5) Color Match; (6) Three Letter Memory. The next section introduces these modules with further details.

Word Memory. This module measures attention processes, learning, memory and verbal recognition. The task includes 12 target and 12 non-target words. Each non-target word is semantically related to a target word. At first, the twelve target words are displayed on the screen for 750 milliseconds each. Subsequently, the words are displayed once again. After that a discrimination task follows, in which test takers have to decide from a 24-word list (including both target and non-target words) whether or not the word presented was on the original list of target words by clicking on the ”yes” or ”no” button.

Design Memory. This module measures and evaluates attention processes, learning, memory and visual (abstract shape) recognition. The task includes 12 target and 12 non-target designs. At first, the twelve target abstract designs are displayed two times for 750 milliseconds each – just like in the Word Memory module – which is followed by a discrimination test. The test taker is required to decide if the target design was among the stimuli or not by clicking on the ”yes” or ”no” button with the mouse while both target and non-target designs are displayed sequentially.

The X's and O's. This module measures and evaluates (spatial) working memory and reaction time (Echemendia, 2006). The test taker is required to recall where the yellow-highlighted X's and O's are placed on the display immediately after a reaction time task. At first, the program displays X's and O's randomly, in which three are highlighted in yellow. This is immediately followed by a distracter task where the test taker has to use the buttons “p” and “q” on the keyboard and push them as fast as they can when seeing the stimuli paired with the buttons. Immediately after the distracter task, the test taker has to recall the three highlighted X's and O's and click on the stimuli where s/he thinks the items appeared. This subtest consists of four tasks.

Symbol Match. This module measures and evaluates visual processing speed, learning, and memory. The test taker is asked to learn and recall nine shapes that belong to their paired numbers (from 1 to 9). After multiple trials of the task, all the shapes disappear from the screen, and they need to be paired with their number counterpart one-by-one. The shapes are presented randomly.

Color Match. This module measures and evaluates reaction time. The words “red”, “green”, and “blue” are presented on the screen with their matched colors, and the test taker is asked to click on the right stimulus if the colors match their labels. The words with the three colors are presented variably and randomly.

Three Letter Memory. This module measures and evaluates the processing of verbal working memory and processing speed. The test taker has to learn and recall three random consonants that are presented on the screen next to each other. After presenting the 3 letters another distracter task follows: the test taker has to count backwards from 25 to 1 by clicking on the correct numbers on the screen as fast as they can. Immediately after the distracter task, the test taker has to recall and type the three consonants into the three letter-box. This subtest consists of four tasks.

At the end of the test, the test taker is required to recall the words and abstract shapes again (i.e., the stimuli of the Verbal Memory and the Design Memory subtests).

After completing the six test modules, the program software measures the performance based on the subtests and calculates five composite scores: (1) Verbal Memory Composite (VeMC); (2) Visual Memory Composite (ViMC); (3) Processing Speed Composite (PSC); (4) Reaction Time Composite (RTC); (5) Impulse Control Composite (ICC), and Cognitive Efficiency Index Composite (CEI).

Verbal Memory composite score includes the average percent correct for the Word Memory (module 1), Three Letter Memory (module 6), and Symbol Match (module 4) designs. Scores are also represented in age-referenced percentile. The higher score indicates better performance on verbal memory.

The Visual Memory composite score consists of the average percent correct for the Design Memory (module 2), the X’s and O’s (module 3) tests. It measures visual memory and attention. The higher score indicate better performance on visual memory.

The Processing Speed composite score is calculated by the average of correctly recalled X's and O's (module 3) and the number of correct backwards clicking on the numbers (i.e., distracter task that is part of the Three Letter Memory test). The higher score indicates higher processing speed, and consequently better performance on the battery.

The Reaction Time composite score is composed of the average of reaction time scores for the Symbol Match (module 4), the Color Match (module 5), and the distracter task of the X's and O's (module 3) test (pushing the buttons "p" and "q" displayed between the X's and O's test). The lower score indicates less time; therefore, better performance.

The Impulse Control composite score indicates the number of errors made while taking the ImPACT test. This composite also helps to identify test takers who do not put adequate effort into completing the test or are simply confused by test instructions (Iverson, Lovell, & Collins, 2003). It is useful in determining test validity; scores above 20 should be considered as invalid.

The Cognitive Efficiency Index composite score has been recently (in early 2011) added to the ImPACT test measures by the test developer of ImPACT, Inc. According to ImPACT, the CEI score as an additional measure shows the interaction between accuracy and reaction time on the Symbol Match Test. The score ranges from 0 to 0.70 with a mean of 0.34 and indicates the extent to which the test taker performed on both speed and accuracy. For example, a low score (i.e., below 0.20) demonstrates that the test taker performed low on accuracy with slow response time, whereas a high score marks good accuracy with fast response time. This measure demonstrates how well participants performed on the Symbol Match Test at each test session in accuracy and reaction time. Therefore, CEI may also indicate the stability of test effort within participants.

Psychometric properties of ImPACT

Test-retest reliability. According to Anastasi (1968), the term reliability always means the consistency of scores obtained by the same person when retested by the same test or with an equivalent form of the test. Although there are different types of reliability in general, in this paper, test-retest reliability is the essential reliability value that literature took into account in terms of checking stability of test results with ImPACT. The reason behind that is that test-retest reliability is the most often used method for measuring how consistent the results are when the identical test is repeated (Anastasi, 1968). Accordingly, the value that demonstrates the stability of a test is called the reliability coefficient. This coefficient is the correlation between the scores obtained by the same person on at least two administrations of the same test (Anastasi, 1968). In addition, retest reliability shows "the extent to which scores on a test can be generalized over different occasions: the higher the reliability, the less susceptible the scores are to the random daily changes in the condition of the subject or of the testing environment" (Anastasi, 1968, p. 78).

Retest reliability also includes the exploration of practice effect. Practice effect is a cumulative change in performance over time when a test is administered to the same person more than once (Echemendia & Julian, 2001). Tasks that involve measuring speed, learning, and memory are susceptible for demonstrating certain amount of practice effect. Consequently, the magnitude of practice effect depends on the time between test administrations and the number of test administrations (Echemendia & Julian, 2001). Testing occurs serially in sports-related concussion assessment until the athlete becomes symptom-free and reaches his/her baseline performance. If the effect of practice is suspected during these testing sessions, the clinician or professional managing concussion cannot be sure whether the improvement in performance is due to recovery or the effect of practice.

There is, however, reason to believe that practice effect on the ImPACT battery may not be the major issue that imposes risks to proper sports-related concussion management. First, when designing the battery the test developers at Pittsburgh were focusing intensely on reducing practice effects with this new computerized battery (i.e., the current version; Version 2.0). They randomized and designed as many stimuli in the battery, so that those taking the test would not come across with the same stimuli the next time they come in for another post-concussion evaluation. In this sense, the stimuli keep changing in each and every testing session. Second, research on ImPACT claimed that practice effects were not present when testing non-concussed amateur athletes four times on the battery showing stable performance across testing sessions (Lovell et al., 2003). This study will be presented in much detail on page 55.

Test-retest reliability of ImPACT. The first studies on ImPACT test-retest reliability are unclear and difficult to follow. With the absence of peer-reviewed, published examinations, there are only a couple of reports from past presentations and review articles that briefly discuss test-retest reliability of ImPACT without reference to the primary source. Therefore, it is difficult to review those reliability studies on ImPACT that were either presented (Lovell, 2001) but not published, or reported but not referred to the primary source for further details. For that reason, this section on initial studies could only rely on these secondary sources.

For example, Lovell (2006) briefly summarizes an initial study in which ImPACT was administered 4 times to 24 high school athletes; however, no reference was provided. The time period between the testing sessions varied from 2 to 8 days. Test-retest reliability coefficients for memory ranged between 0.66-0.85 between test sessions 1-2, 2-3, and 3-4. Using the same assessment comparisons, test-retest correlation coefficients for visual processing speed ranged between 0.75-0.88. Reaction time test-retest reliability values fell between 0.62-0.66. Lovell adds further on that although reaction time correlations were consistent across all testing sessions, results on memory and processing speed show "slight" variability in which the correlation between times 1-2 was somewhat weaker than between time 2-3 and time 3-4. However, these "slight" changes, as the author noted, indicated that the performance on these indices improved after the first testing sessions, with small practice effect after additional administrations (Lovell, 2006). However the author did not provide specific information about how slight these changes were, and how much practice effect was present.

Another initial study presented test-retest reliability based on ImPACT Version 1.0 (Iverson et al., 2002), but only the abstract got published. Although Version 2.0 is very similar to the original Version 1.0, there are some crucial changes. Test developers added an extra module to Version 2.0 (i.e., Design Memory which measures visual memory), whereas one of the working memory tasks (i.e., X's and O's) was modified to be more difficult than the one in Version 1.0. Therefore, Version 2.0 provides two memory composite scores (Verbal and Visual Memory), as opposed to Version 1.0 that only contains 1 (verbal) memory composite score (Iverson et al., 2003a). However, the results appeared in another review without clear reference (Randolph et al., 2005). According to the review, Iverson et al. (2002) posted reliability data on the ImPACT website. The sample included 49 healthy, non-concussed high school and collegiate athletes, and reliability was assessed using short retest intervals with a maximum of two weeks between testing sessions. The reliability coefficient was the following: for the memory composite 0.54; for the processing speed 0.76, and for the reaction time 0.63. According to Randolph et al. (2005), these reliability coefficients were below suggested levels (i.e., $r = 0.9$), though generally equivalent to other reliability coefficients for individual paper-pencil tests. The authors (Randolph et al., 2005) argued that test-retest reliability should be close to 0.90 for individual decision making about an athlete's cognitive status after a concussion because if the test-retest reliability is only 0.70 (which is typical for many NC tests) an individual would have to demonstrate a decline of 20 IQ points to move out of the same confidence interval. As a result, this can affect the calculation of reliable change scores, which is crucial for individual decision making about an athlete's cognitive status.

The problem with these brief summaries and reviews is that they do not introduce the study adequately; they lack important details about 1) the participants, 2) the version of ImPACT, 3) the time intervals between testing sessions, 4) the methodology for calculating reliability. There are, however, three peer-reviewed primary sources on stability measures of ImPACT 2.0., of which the review is the following.

Iverson, Lovell, & Collins (2003a) reported test-retest reliability for ImPACT Version 2.0. The purpose of the study was to measure the stability of ImPACT and to calculate the reliable change confidence intervals for the various test-retest scores. The rationale behind the study was that test scores can be influenced by many factors, such as practice effects, regression to the mean, and other unpredictable forms of measurement error. To interpret test results properly, it is crucial to study the probable range of measurement error first, because measurement error is present in test-retest difference scores. Therefore, if there is an initial understanding of measurement error within ImPACT 2.0, it will allow interpreting deterioration, improvement, and recovery more precisely (Iverson et al., 2003a).

For the study the authors selected 56 (29 males, 27 females, age range: 15-22) healthy, non-concussed amateur athletes who completed ImPACT 2.0 battery on two occasions with varying time intervals between the testing sessions. The time interval between the 2 testing sessions was not the same for all participants: retesting sessions varied between 3-11 days. Sixty-four percent of their sample was high school athletes, while 34% were college athletes. The first analysis included a within subject design in which the level of performance within subjects was examined. Test-retest reliability coefficients were calculated using Pearson *r* correlation. Reliable change parameters were estimated by the methodology proposed by Jacobson and Truax (1991) that corrects for practice when practice effects are present (Iverson et al., 2003a). According to the authors, this reliable change methodology has been applied in clinical neuropsychology extensively, because it allows the clinician to estimate measurement error of test-retest difference scores. The dependent measures for the study were the four composite scores and the score on concussion symptoms from ImPACT 2.0: verbal memory, visual memory, reaction time, visual processing speed, and the total symptom score. After the analysis, the test-retest correlation coefficients were the following: for Verbal Memory = 0.70, Visual Memory = 0.67, Reaction Time = 0.79, Processing Speed = 0.86, total symptom score = 0.65. As a summary of the results the authors argued that these test-retest coefficients – though seemingly modest – are higher than in many other NC tests, including the Wechsler Memory Scale 3rd edition, the California Verbal Learning Test 2nd edition, or the Delis-Kaplan Executive Function System Trail Making Test and Color Word Test. The second finding showed that practice effects were only present in processing speed, which means that although participants were faster at retest, they did not perform better on other measures of ImPACT (e.g., Verbal Memory, Visual Memory).

The second part of the study examined the sensitivity of ImPACT 2.0 to the acute effects of concussion. For that reason the authors tested 41 concussed amateur athletes who completed the test battery within 74 hours of post injury. These athletes also underwent baseline testing prior to injury. The majority of concussed athletes were males (90%) and mostly football players (88%); approximately 71% were in high school, and 29% were in college. As a group, they demonstrated a “large” decline in Verbal Memory, Reaction Time; a large increase in the total symptom score, and a “medium-to-large” change in Visual Memory and Processing Speed. Moreover, when the authors applied the reliable change index scores to the concussed athletes’ performance, 44%-54% showed statistically reliable declines across all five dependent measures.

In conclusion, ImPACT 2.0 proved to be a sensitive and reliable measure for concussion assessment. According to the authors, ImPACT 2.0 battery was designed to reduce practice effects by randomizing stimuli. This design feature was an essential factor when developing the test since it is intended to be administered repeatedly, over short time intervals.

The reliability of ImPACT test results was also supported by another study from Lovell, Collins, Iverson, Field, Maroon et al. (2003) that measured recovery from concussion in high school athletes. The purpose of the study was to evaluate memory deficits along with self-reported symptoms in concussed high school athletes. Although the main objective of the study was not specifically designed to assess test-retest reliability, the authors provided practice effect calculations, which are important and bear relevance in terms of stability and reliability of the test itself.

Lovell and colleagues (2003) used a healthy, non-concussed control group (N = 24) and assigned them to the same test conditions as the concussed athletes (N = 64). The purpose of using a control group was to examine whether practice effects due to prior exposures to the test affect the performance on the test battery. Non-concussed athletes were tested on the ImPACT 4 times (baseline and 3 follow-up tests) together with a group of concussed athletes. The four testing sessions were based on 1) baseline, 2) 36 hours, 3) 4 days, and 4) 7 days post injury. The results revealed that normal controls performed relatively the same across the testing sessions with no significant practice effect. However, both the methods and the results indicating the findings lacked specific statistical details. The findings on practice effect were crucial for the study because these results provided the basis for comparing the changes on post-injury performance in the concussed group. The results on the concussed group of athletes showed that duration of on-field mental status changes (amnesia and confusion) showed a relationship with the length of recovery. High school athletes who suffered amnesia, confusion, and disorientation longer than 5 minutes did not fully recover in a week. They demonstrated longer-lasting post concussion symptoms and a decline in memory performance. On the other hand, athletes with on-field mental status changes lasting less than 5 minutes returned to their baseline level of performance by day 7 post-injury.

In a different study, Broglio, Ferrara, Macciocchi, Baumgartner, Elliott (2007a) measured test-retest-reliability of ImPACT. The purpose of their study was to measure test-retest reliability of three computerized concussion assessment programs. The three programs were ImPACT (ImPACT Applications, Pittsburgh, PA, 1998), Concussion Sentinel (CogState Ltd, Victoria, Australia), and Headminder Concussion Resolution Index (Headminder Inc, New York, NY, 1999). The authors argue that previous studies assessing test-retest reliability used arbitrary time intervals rather than test-retest intervals commonly observed in a clinical setting (i.e., when managing concussion). Therefore, the rationale behind the study was to determine test-retest reliability in these NC concussion assessment programs because recent literature lacks such reliability studies on the currently available test programs. The study took a number of confounding factors into consideration for assessing test-retest reliability including lack of effort that could reduce reliability. To control for suboptimal effort during testing, the authors attempted to identify participants who may not demonstrate adequate effort in order to examine valid test-retest reliability. Therefore, to control for effort, the computerized version of the Memory and Concentration Test for Windows (MACT, Green's Publishing Inc., Edmonton, Alberta, 2004) was administered.

118 healthy college students volunteered to take part in the study, but only 73 remained after completing all 3 testing sessions with all 3 computerized concussion assessment batteries. The reason for subject attrition was mainly due to lack of effort that led to invalid baseline performance. If participants had invalid performance on any of the computerized batteries (ImPACT, Concussion Sentinel, CRI, and MACT), they were excluded from the study. Among the remaining participants, 12 reported having a history of concussion ranging from 1 to 5

injuries. Apart from that, no participant reported sustaining a concussion during the testing process or 6 months prior the study. Participants completed the three concussion assessment programs three times: baseline, 45 days after the first testing session and then five days after the second testing session (i.e., 50 days after baseline). No concussed participants took part in the study.

The statistical design differed from Iverson et al (2003a). The authors calculated a two-way random effect analysis of variance intra-class correlation (ICC) to establish test-retest reliability for each assessment programs. The calculation values are predicted to be between 0 and 1: values closer to one would indicate stronger reliability and less error variance. Broglio and colleagues argue that ICC is often used with different ANOVA models especially when a data set has various groups in contrast to Pearson correlation in which each variable calculation is considered separate from one another; ICC calculates the correlation of one variable between two members within a group. Since this study involves a repeated measure design with multiple groups, ICC estimations appear to be appropriate for test-retest reliability. The minimum ICC level for this study was 0.60, which the authors chose based on the recommendation of Anastasi (1968). After statistical analysis the results varied between tests. As it was reported by the authors, test-retest reliabilities on all the three computer assessment programs fell below the levels commonly recommended (i.e., 0.75) for making clinical decisions. According to the results, the ICC coefficient estimates from baseline to day 45 assessments ranged between 0.15 and 0.39 on the ImPACT, 0.23 and 0.65 on the Concussion Sentinel, and 0.15 and 0.66 on the Concussion Resolution Index. The ICC coefficient estimates from day 45 to day 50 ranged from 0.39 to .61 on the ImPACT, 0.39 to 0.66 on the Concussion Sentinel, and 0.03 to 0.66 on the Concussion Resolution Index. Further interpretation of test-retest reliability revealed that Concussion Sentinel and Concussion Resolution Index had similar test-retest values from baseline to day 45, while Concussion Sentinel provided the highest reliability value between the days 45 and 50. In general, test-retest reliability was the highest in all 3 batteries between day 45 and 50. The majority of indices exceeded the minimum test-retest reliability criterion (i.e., 0.60) over a 5-day period. However, the ICC estimates on several test indices met or just exceeded the 0.60 level. Furthermore, the ICC values for all 3 test programs were lower than those reported in the literature. For example, test-retest reliability measures for ImPACT, ranged among 0.54 (memory), 0.63 (reaction time) and 0.76 (processing speed). But the most crucial result of the study of Broglio et al (2007) was that none of the 3 computerized test batteries reached ICC levels greater than 0.75 test-retest reliability, which is typically considered as "good" test stability.

Summary. To summarize the results, studies of test-retest reliability on ImpACT demonstrate inconsistency with one another. First, there is inconsistency about the statistical methodology with regards to test-retest reliability. Broglio et al. (2007) applied ICC as the "right" methodology for calculating stability coefficients, while Iverson et al. (2003a) used the traditional Pearson *r* statistics for estimating reliability values. Therefore, the statistical analysis for test-retest reliability raises issues, on which literature is divided.

Second, reliability coefficients lack consistency between studies. Based on the different statistical estimations, reliability in the study of Iverson et al. (2003a) indicated that ImPACT 2.0 is a stable and consistent measure, even across multiple administrations; whereas Broglio et al. (2007a) found that all three computerized assessment tools (ImPACT, Concussion Sentinel, Headminder Concussion Resultion Index) showed low to moderate reliability coefficients.

Third, it is necessary to control for lack of effort not only when testing athletes for baseline, but when testing participants for psychometric purposes. Broglio et al. (2007a) clearly demonstrated that subject attrition was mainly due to insufficient effort during baseline assessment, which further on resulted in invalid performance. Iverson et al. (2003a), however, did not control for lack of effort during reliability testing while testing their healthy participants.

Fourth, time between testing sessions lacked consistency. Broglio et al (2007a) argued that they applied clinically relevant time intervals between testing sessions (i.e., 45 and 50 days after baseline, whereas Iverson et al (2003a) tested the participants on two occasions in which the time interval between the first and second testing sessions varied between 3-11 days. For testing practice effects on ImPACT Lovell et al (2003) tested HS athletes with and without a concussion on four occasions: 1) baseline, 2) within 36 hours, 3) 4 days, and 4) 7 days after baseline assessment. The population for assessing reliability remained the same; they were all amateur student athletes.

After reviewing test-retest reliability literature on ImPACT, it can be concluded that computer-based assessment provided only moderately stable measures in a normal population. As a consequence, inconsistent performance on a test would lead clinicians to inaccurate decisions on NC functions; therefore, clinicians should be careful with relying only on one concussion test results and should use multiple evaluation approaches for concussion management (Broglio et al., 2007a).

Other than these three studies, there is no more peer-reviewed research published on ImPACT specifically related to test-retest reliability and practice effect measures in a normal population. Although the test developers of ImPACT stated that practice effect was not present when testing athletes across testing sessions, it must be emphasized that the findings have not been replicated yet, and the reliability measures of ImPACT differ from one study to the other. Therefore, one of the aims of this dissertation is to measure test-retest reliability and stability of NC performance on ImPACT across testing sessions (i.e., 7-8 days between testing sessions).

Validity of ImPACT. The basic concept of validity is to establish that a test measures what it is purported to measure and how well it does so (Freeman, 1964, Anastasi, 1968, Franzen, 1989, Nunnally & Bernstein, 1994, Randolph et al., 2005, Schlegel & Gilliland, 2007). The validity of today's computerized tests relies on the same type of validity measures as other traditional tests. In NC assessment, exploring validity involves the demonstration that scores of a test can accurately separate brain-injured people from the non-brain injured (Franzen, 1989). This approach shows that a test is sensitive to impairment in brain-injured populations.

This statement is relevant in terms of ImPACT validation research. According to Lovell (2006), the validation of any NC test is a cumulative process that involves the systematic measurement of different aspects of validity (e.g., face content, criterion, construct). The most often referred aspect of validity is the so called criterion-related validity, which measures the ability of a test or subtests to differentiate between impaired subjects from the non-impaired (Lovell, 2006). However, the term criterion validity refers more specifically to the correlation between a test score and some external variable that can be another test which measures the same characteristic of interest (Franzen, 1989). Moreover, differentiating between impaired and non-impaired subjects is more related to the definition of sensitivity. Studies that are not specifically designed to measure criterion validity have clearly demonstrated that ImPACT is capable of separating those with deficits from non-injured (Iverson et al., 2003a; Iverson et al., 2004; Lovell et al., 2003; McClincy et al., 2006). All these studies reported that concussed amateur athletes demonstrated significant declines in memory processes, increase of reaction time and self-reported symptoms with a moderate change in processing speed of information, when compared to non-injured athletes. Although these studies cited above refer to ImPACT as a sensitive measure to change in NC performance, they were not specifically designed to measure any type of validity. So far, only one peer-reviewed study has been published that measured the validity of ImPACT specifically:

Iverson, Lovell, & Collins (2005) tested validity on ImPACT for measuring attention and processing speed following a concussion. In this study, validity refers to “the dimensions or traits that the test was designed to measure” (Franzen, 1989, p. 45). One aspect of the validation process is to compare ImPACT as a computerized test for measuring attention and processing speed to a traditional neurocognitive measure for the same trait, such as the Symbol Digit Modalities Test (SDMT, Smith 1982). According to the authors, traditional SDMT has been used many times as a part of concussion assessment since it measures scanning and tracking aspects of attention and processing speed. Therefore, the hypothesis of the study was that traditional SDMT would be more related to the Processing Speed and Reaction Time composite scores than to Verbal and Visual Memory composites.

The study involved 72 amateur student athletes with a concussion who were evaluated using ImPACT 2.0 and SDMT within 21 days post-injury. The participants were mostly males (83.8%) with different sports background (e.g., football, hockey, soccer, basketball, wrestling). The dependent measures for the study were the following: Verbal and Visual Memory, Reaction Time, and Processing Speed composite scores. To complete the SDMT, the participants were required to substitute a certain number for a randomized series of geometric shapes. The process for this task was the following: each number is paired with a geometric figure. After presenting a geometric figure, the examinee had to write the target number that goes with that figure as quickly as possible.

The results revealed that Pearson correlations were higher between SDMT and Reaction Time and between SDMT and Processing Speed than with the memory composite scores. Therefore, these outcomes suggest that SDMT and the measures for attention and processing speed in ImPACT assess a similar underlying concept. However, the assumption made was based on comparing only one traditional test for measuring attention and processing speed. Even though the validation process of ImPACT is ongoing, more psychometric research is needed to reveal the different aspects of validity.

Sensitivity of ImPACT. Validity and sensitivity are strongly related to one another. Although validity refers to the ability of a test to measure what it has purported to measure, sensitivity refers to the ability of a test to differentiate between clinical patients and normal population (Randolph et al., 2005). Therefore, sensitivity measures of a test are crucial for clinical purposes in order to properly diagnose and manage injuries and deficits. To establish sensitivity in concussion assessment, studies usually measure differences in NC performance between athletes with a concussion and their matched normal controls.

Furthermore, sensitivity is also related to specificity which refers to the detecting subjects who may be classified as impaired based on their NC performance when they are actually normal (Randolph et al., 2005). This phenomenon is also known as “false positives”. In baseline assessments these false positives should be controlled. On measuring sensitivity and specificity of the ImPACT battery, one peer-reviewed research has been published.

Schatz, Pardini, Lovell, Collins, Podell (2005) explored the diagnostic utility of the different composite scores and post-concussion scale of ImPACT. This involved the comparison of a group of high school athletes with a concussion to a group of athletes with no concussion. Inclusion criteria for concussed athletes required that participants got tested baseline and within 72 hours post-injury, whereas for non-concussed athletes only those could participate who had no history of concussion plus underwent baseline assessment. As a result, 72 concussed athletes and 66 normal control athletes took part in the study. The dependent measures were Total Symptoms Scores and the five composite scores based on the cognitive abilities ImPACT measures: Verbal and Visual Memory, Reaction Time, Visual Processing Speed, and Impulse Control composite scores; the independent variable was athletes' concussion status (i.e., concussed vs. non-concussed).

The analysis included a one-way ANOVA to identify any group differences in age and education, whereas between group chi-square analysis was used to identify further differences in gender, handedness, learning disability, and history of special education. The results revealed no significant differences on the variables. MANOVA was conducted to estimate between group differences on the dependent variables. According to the analysis, MANOVA revealed multivariate effect of concussion group. Further ANOVA identified these effects confirming that the group with concussion performed significantly worse on all the cognitive measures. Impulse control did not differ between groups. Further stepwise discriminant analysis showed that all composite scores and total symptom score were significant factors in classifying concussion; as a result 85.5% of cases were classified correctly. After calculating the positive and negative predictive value (i.e., the probability that a concussion is present or not when the test is positive or negative), the sensitivity of ImPACT was 81.9%, while the specificity was 89.4%. As a results, it can be concluded that ImPACT computerized battery is a sensitive and specific measure for the assessment and management of concussion.

Concussion assessment and management today

Concussion management relies heavily on computerized NC testing. There are several assessment tools offered, ImPACT is the most widely used program. Today's concussion assessment actually developed from NC assessment; it measures cognitive functions such as reaction time, memory, learning, attention, processing speed, and it also involves a questionnaire of current symptoms. Testing these cognitive functions forms an essential part of concussion management. However, NC assessment in concussion is also complementary to the evaluation and management process; together with other forms of assessment (i.e., testing auditory comprehension, verbal fluency, hearing, balance), NC tests help to build accurate diagnosis and better prognosis.

Concussion management and assessment has many components. Its purpose is 1) to monitor signs and reported symptoms over time to track recovery; 2) to help avoid premature return-to-play (RTP) decisions together with other assessment methods (i.e., postural stability, verbal fluency, health history); 3) to measure cognitive functions such as memory, attention, learning, visual processing speed of information, and reaction time over time to monitor recovery of these functions independently; 4) to provide measures for individualized assessment since it offers baseline testing and a number of follow-up examinations; 5) to identify cumulative effects of concussion. With the help of computerized assessment technique such as ImPACT, this is all possible. In addition, the contribution of concussion assessment and management to research is exceptional and genuine. What is known today of concussion has to do a lot with the development of NC assessment instruments.

By now many forms of assessment techniques have been developed for concussion assessment and management. According to the Second International Conference on Concussion in Sport (2005), there is general agreement that concussion assessment should have common principles. In concussion assessment and management this is done in 3 ways: 1) sideline assessment, 2) baseline testing, and 3) post-injury examinations. In sideline assessment, the purpose is to measure the severity of head injury during the acute phase and to make RTP decisions on spot in the same game or practice session. The most acclaimed and validated sideline assessment tool is the Standardized Assessment of Concussion (SAC) and the Sideline ImPACT. These on-field cognitive testing tools require only five minutes administration, and they provide results immediately after injury. They measure the signs first, then orientation, immediate memory, and delayed recall. If the athlete fails to give the right answer on any of the short questions and/or tests, he or she is no longer allowed to return to play, hence further assessment is necessary (McCrory et al., 2005).

Post-injury assessment is serial and interdisciplinary in its nature. Athletes with a concussion are tested and assessed, until they no longer report concussion-related symptoms and their cognitive performance on the ImPACT battery either returns to baseline-level or returns to normal (in case the athlete was not tested at baseline, his/her performance is compared to normative data).

Issues in today's clinical management of concussion. Even today concussion assessment and management presents a difficult diagnostic and management situation. First, not all concussions can be detected by neuroimaging techniques (Aubry et al., 2002). Second, patients suffering a concussion present a wide variety of symptoms that can include somatic, psychological, and cognitive symptoms all together (Hall et al., 2005). Third, symptoms and the length of symptoms arising after a head injury may vary individually (Covassin et al., 2006). These symptoms of acute cerebral concussion most often include disorientation, word-finding difficulties, slurred speech, amnesia, confusion, attention and concentration deficits, reduction in processing speed of information and reaction time, impairment of verbal and non-verbal memory, learning difficulties, fatigue, headache, dizziness, trouble with sleeping, sensitivity to light and noise, tinnitus, nausea and/or vomiting, irritability, and depression (Aubry et al., 2002). As the literature reports, these symptoms and deficits of neurological function, however, tend to be short-lived and resolve spontaneously usually within 14 days post-injury (McCrory et al., 2009). Fourth, since LOC may not necessarily accompany a concussion, it is difficult do detect who may be at risk when sustaining one.

To recognize this risk is even more problematic since there have been more reports on athletes not reporting or underreporting their symptoms for the sake of returning to competition (Barth et al., 1989, Echemendia & Julian, 2001; Echemendia & Cantu, 2003; Broglio, Macchiocchi, & Ferrara, 2007b). Further research suggests that athletes may be even unaware of the signs and symptoms that accompany a minor head injury (McCrea et al., 2004; LaBotz et al., 2005; Valovich McLeod et al., 2008).

The reasons why athletes do not report or minimize their symptoms and their injury/injuries may vary individually. Athletes may feel that concussions are "part of the game" (Echemendia et al., 2001, p. 23) and admitting a "small head injury" would be the sign of weakness, and would make them feel that they let their team mates and coach down. Furthermore, athletes indicated that reporting their injuries would jeopardize their scholarship, would lose their position in the team, would miss the opportunity to demonstrate their skills in front of a professional agent, would see themselves as a failure even by their family and team mates, and would risk the opportunity to continue to play as well as removing the athletes from the upcoming competition (Barth et al., 1989, Echemendia et al., 2001; Hunt et al., 2007).

These reasons tend to match with our clinical experience at the UTEP Concussion Management Clinic. When asking our patient-athletes here in the El Paso area, some of them express their concern that they would lose their scholarship, or would miss the opportunity to demonstrate their skills in order to receive a scholarship. Furthermore, most of the athletes asked often report that they do not want to miss the opportunity to play at an upcoming competition.

There are many reasons why "not reporting" a concussion poses many risks. First, if a head-injured athlete returns to play before his/her symptoms have resolved completely, s/he may be at an increased risk to suffer a disabling or even fatal injury. Research and clinical practice names such incidences as Second Impact Syndrome (SIS, Cantu & Voy, 1995, Echemendia &

Julian, 2001). Therefore, management of first concussion is crucial in order to avoid the devastating consequences of SIS. Second, recurring concussions may lead to chronic, long-lasting cognitive, physical, and psychological deficits (Iverson et al., 2004). As a result there is a risk of developing Post-Concussion Syndrome. Third, athletes with a history of multiple concussions may increase the risk of experiencing late-life cognitive impairment (Guskiewicz et al., 2005).

There is a growing amount of evidence suggesting that recurrent head injuries may result in longer-lasting and severe post-concussion symptoms (Collins et al., 1999; Iverson et al., 2004; Covassin et al., 2008), and late-life cognitive impairment (Guskiewicz et al., 2006). Furthermore, if athletes do not report their injuries and symptoms in time, they may increase the risk of cumulative and catastrophic effects of recurrent head injuries. Therefore, raising awareness about concussion in general and about the consequences of unreported head injury in young athletes are extremely crucial (McCrea et al., 2004). In addition, clinical observation of the concussed athlete along with well-structured patient interview and a sensitive and reliable concussion assessment battery may account for these challenges in the diagnosis and management (i.e., tracking recovery, deciding when to start RTP protocol) of sports-related concussion.

Confounding issues in NC assessment: Effort

Whenever measuring cognitive functions, there are a number of confounding factors that could affect NC performance. These factors include 1) lack of effort during baseline testing, 2) demonstrating maximum effort during post-concussion assessment, 3) regression to the mean, 4) and practice effects. Since NC tests are standardized and believed to yield objective measures of cognitive abilities, it is hard to control for factors that rely on mainly the test taker's co-operation and attitude towards testing. These factors not only confound the true results of one's actual test performance, but they may influence concussion management and RTP decisions as well.

Lack of effort. Standardized NC tests require adequate effort from test takers to get valid results of cognitive performance. In concussion assessment and management, baseline testing of athletes is crucial for post-concussion evaluation and making safe RTP decisions. If test takers are not cooperative and perform insufficient effort during baseline assessment, test results are no longer considered to be valid, and clinicians can no longer interpret changes of test performance following a concussion. Furthermore, poor test performance due to insufficient effort may yield to poor concussion management and RTP decision that may be harmful for the athlete. Consequently, sufficient effort and being cooperative during test taking are crucial for valid assessment and rehabilitation (Hunt et al., 2007). The reason behind this lies in the nature of standardized assessment. NC tests usually classify one's performance as superior, high average, average, low average, borderline, and impaired based on normative samples. Although effort is not usually assessed in normative samples, but subjects often participate voluntarily and/or may even get monetary compensation after test taking. Moreover, they have an interest in performing to the best of their ability (Stevens, Friedel, Mehren, & Merten, 2008).

It is difficult to measure effort based on subjective examiner impressions. In addition, such impressions may also be erroneous (Hunt et al., 2007). Therefore there is a growing tendency to use standardized effort assessment methods to measure sufficient effort during test taking. The most prominent methods include the computerized Word Memory Test (WMT, Green, 2003) or the Medical Symptom Validity Test (MSVT, Green, 2004). Both tests can be used with children and adults. The MSVT is a derived form of the WMT (i.e., is comprised of 20 word pairs) that contains 10 word pairs. This list of word pairs appears twice on the screen following a word discrimination task. Effort is assessed in immediate recognition (IR) and delayed recognition (DR) in the following way: the computer presents word pairs again, one that was among the target list of words and one that did not appear among the list. The test taker has to identify which word appeared previously on the screen. Furthermore, the test also measures response consistency based on the first two recognition trials. Based on the WMT and MSVT, second grade students (age 7-9), children with 3rd grade reading level, children with a verbal IQ less than 70, children with learning disability, and children with an IQ of 65 could pass both tests. The rationale behind using these tests in a clinical-medical setting is that if children with such abilities and impairments could pass the test successfully, then adults with mild head injuries should pass them as well. If an adult with mild neurological injury performs lower than the established cut-offs, he/she clearly demonstrates poor effort on the NC test (Carone, 2008).

According to Stevens et al. (2008), uncooperativeness can have many factors: lack of interest in taking the test, fatigue, other intentions such as fraud or malingering. Although malingering is not expected among athletes, athletes may demonstrate insufficient effort during baseline testing. When athletes with no self-reported learning disability and attention deficit disorder reach below average or borderline performance level, this may not necessarily mean cognitive impairment but rather lack of effort during completion of the test. However, knowing the health history of the athlete is essential when drawing such assumptions.

Hunt et al. (2007) measured the effect of effort on baseline NC performance in high school football athletes. They administered a short NC battery together with the Dot Counting Test and the Rey-15 Item that measure effort. Among the 199 high school athletes, 177 demonstrated sufficient effort during NC testing, whereas 22 showed poor effort. The difference between these effort groups based on their Dot Counting Test was statistically significant as well. Although only a small percentage of athletes demonstrated poor effort during NC testing (approximately 1 from every 10 athlete), examiners and test developers have to account for lack of effort from test takers. There are several reasons that may explain lack of effort during testing. The list includes distractions (e.g., especially in group testing), fatigue, poor comprehension of task instructions (Hunt et al., 2007). Here at the UTEP Concussion Management Clinic we observed that inadequate effort was often demonstrated by athletes often talking to one another during baseline assessment, using their cell phones, watching each others' screen, not always reading or following the instructions carefully, comparing each other's test stimulus and performance, and sacrificing speed for accuracy. For example, invalid baseline test results were often spotted among athletes in those test sessions when athletes were tested in large groups. However, adequate effort may be uncertain for examiners; therefore, it would be useful to include standardized effort assessments in NC batteries. Because obtaining valid results require adequate effort and cooperativeness from athletes, it would be also advantageous to explain athletes why baseline assessment is important and draw their attention to poor effort and the necessity to perform the best of their abilities. This may help enhance cooperativeness for baseline assessment with athletes until more is discovered about the effect of effort in NC testing.

Effort at post-concussion testing. There is little research about the role of effort in sports-related NC assessment and how it confounds performance. With regards to non-athlete populations there is a considerable amount of literature that studies and debates malevolent malingering (i.e., receiving compensation for injuries, avoid going back to work, etc.). More specifically, studies found that malingering post-concussion symptoms in non-athlete population who receive financial compensation after their injury demonstrate prolonged and intense symptoms (Youngjohn, Burrows, Erdal, 1995, Binder et al., 1996). This is not generally expected within the athlete population. Based on the study of McCrea et al. (2004), LaBotz et al. (2005), and Valovich McLeod et al. (2008), amateur athletes tend *not* to report previous injury for a number of reasons: 1) thinking that the injury was not serious enough for further medical attention; 2) do not want to be withdrawn from competition due to professional aspirations; and 3) lack of awareness of probable concussion.

In recent concussion literature, Echemendia & Julian (2001) and Echemendia & Cantu (2003) assumed that effort may impact NC testing of athletes who have sustained concussion suspecting that these athletes would be “highly motivated” to return to play. For that reason, they would minimize their symptoms in the hope of getting back to game as soon as possible, or on the contrary, in some cases they might exaggerate their symptoms to avoid practice.

As for NC assessment, effort demonstrated during baseline and post-concussion NC testing could affect NC performance on a test. According to Echemendia & Julian (2001), an athlete may be “unmotivated” during baseline testing because they may feel bored or a bit threatened. The same athlete’s attitude towards testing could change when the post-injury test has to be completed, especially during the season. Therefore, the drive to minimize their symptoms in order to return to play may vary among athletes. Research suggests that reasons behind underreporting symptoms might be money matters or fulfilling contractual agreements, their love for the game, pressure from the coach or trainer, or the pressure to improve their status in the team, and the decrease of chance that not playing at a certain game would impact their future career (Echemendia & Julian, 2001, Bailey & Arnett, 2006). Moreover, these differing levels of effort demonstrated during NC testing may cause further problems in concussion management if improved post-injury scores due to the drive to return to competition were interpreted as “recovery”. Consequently, the role of the neuropsychologist, trainer, and physician is crucial in the assessment and concussion management process (Echemendia & Julian, 2001). For example, athletes who do not demonstrate optimal effort at baseline testing and increase their effort greatly during post-injury assessment, their baseline performance will fall below their true cognitive abilities. Therefore, after a concussion occurs, these athletes may appear to return to their baseline by one week post-injury due to a change in their effort when in fact they may be still recovering. Such diagnostic issues may place athletes to further risks if they are permitted to return to play prematurely based on inaccurate baseline cognitive abilities and underreported concussion-related symptoms. More specifically, such a situation may have dangerous implications on RTP protocols, and therefore could result in premature RTP and in increased likelihood of getting injured again (Guskiewicz et al., 2000). Therefore, it is suggested that confounds of serial assessment and factors that motivate athletes such as love of the game, return to play/practice, endangering scholarship, role in the team, professional aspirations could be related. Though investigating such "motivational factors" is not the purpose of this dissertation study, the effort that is demonstrated during testing, however, could answer whether such "drive" affects NC performance or not. This is another aim of this dissertation study.

How does effort affect NC performance and how can it be measured in athletes? Since athletes are administered baseline and post-injury testing, clinicians can compare the performance gained from the two testing sessions. At baseline level, examiners can identify the strengths and weaknesses of athletes. However, after post-injury assessment, the results can be compared to the results of pervious testing and identified how concussion affected cognitive functioning. Because the whole NC assessment process has a comparative nature, it is essential to get accurate results both times (Bailey & Arnett, 2006). Consequently, if one of the testing sessions provides inaccurate results, it would obscure the actual effect of concussion. As it was reviewed previously, NC assessment requires adequate effort and cooperativeness from test takers. If they lack the effort, the test would not reflect athletes' accurate performance on the baseline test. During baseline testing, however, the above mentioned factors (e.g., not to be withdrawn from competition) are not present; furthermore, this general disinterest may obscure the impact of a concussion on the cognitive composite scores (Bailey & Arnett, 2006).

Bailey, Echemendia, & Arnett (2006) were the only research group who examined the impact of effort on baseline and post-concussion NC performance within athletes. The purpose of this study was to determine the impact that effort has on sports-related concussion testing. They wanted to separate athletes who did not put optimal effort at their baseline testing from those athletes who provided optimal effort during NC assessment. The assumption behind that group differentiation was that athletes with "high motivation" probably put adequate effort during baseline assessment, while those athletes who performed well below the mean (i.e., one standard deviation below) would lack the effort due to "suboptimal motivation". Accordingly, the authors selected athletes that sustained a concussion, completed a post-injury assessment one week after the incident, and underwent baseline assessment previously. At first, the authors were focusing on grouping the athletes based on their baseline performance. Hence, those athletes who performed one standard deviation or more below the mean were grouped as athletes with suspect or "suboptimal motivation" at baseline. Those athletes that performed one or more standard deviation above the mean were grouped as athletes with "high motivation" at baseline. In other words, suboptimal and high motivation is reflected as lack of adequate effort and maximum effort during NC testing. The reason behind having athletes with both baseline and post-injury assessment was to demonstrate that if athletes really show suboptimal motivation as opposed to athletes with optimal motivation then 1) athletes with suboptimal motivation at baseline would demonstrate larger changes between their baseline and one week post-injury testing; and 2) the changes between baseline and post-injury assessment would be an improvement over their baseline scores despite sustaining a concussion. Since athletes with a concussion are often not allowed back to play after one week post-injury assessment; therefore, the performance on NC test batteries could be crucial for identifying the relationship between motivational factors and NC assessment in concussed athletes. After statistical analysis (using True Score Adjustment for every baseline test, Analysis of Covariance, Reliable Change Index scores for each participant) it turned out that the results support the hypotheses but only with certain NC tests: on the Trail

Making Test A and B, Stroop – Color portion, Digit Span Test, and the Vigil Computerized Performance Test, athletes with suboptimal motivation showed greater improvement than athletes with high motivation. However, the authors also noted that some tests may be relatively resistant to change in motivation such as the Stroop – Word portion. Therefore, future research is necessary to further determine the impact of motivation on NC assessment. In this sense, it is necessary to define which tests are sensitive to change in motivation and to find empirical evidence which factors may further determine shifts in motivation during testing. For possible causative factors Bailey & Arnett (2006) hypothesize that personality style, lack of knowledge, and misconception about the impact of concussion. In addition, active misrepresentation (i.e., malingering) may stand behind representing lack of effort at baseline testing. According to the authors, there is some literature indicating a strong correlation between personality and cognitive testing. Second, what coaches or trainers know and how they think about concussion may influence athletes' attitude towards NC testing. Third, athletes may be aware that if they misrepresent their cognitive abilities and symptoms at baseline they may gain similar results at post-injury assessment and accordingly, they may have better chances to return to play sooner despite the actual existence of a concussion and its cognitive deficits.

Although there is no empirical evidence supporting these hypothesized factors, there is a possibility to identify lack of adequate effort during NC assessment. Objective measures of lack of effort or malingering are getting more and more attention nowadays since there is a growing need for more objective NC assessment. The application and integration of the previously mentioned Word Memory Test (WMT), Medical Symptom Validity Test (MSVT), or Test of Memory Malingering (TOMM) into NC assessment could provide for more accurate measures of cognitive abilities. If a subject scores below chance level on one of these tests, then the results suggest that the person is either malingering or lacks the adequate effort necessary for obtaining valid test results. Because there is evidence that performance can change due to changes in effort during testing, emphasis should be put on studying such factors of effort in athletes and examine what parts of cognitive testing are further affected by effort. Furthermore, there is reason to assume that athletes may show different levels of effort (i.e., lack of effort at baseline testing vs. effort demonstrated at post-concussion testing in order to return to play as soon as possible) indicating that effort may have an effect on NC test performance, especially in those athletes who show lack of effort at baseline testing and the effort demonstrated at post-concussion testing to return to competition (Bailey et al., 2006; Bailey & Arnett, 2006). Such a situation may have dangerous implications on RTP protocols, and therefore could result in premature RTP and in increased likelihood of getting injured again (Guskiewicz et al., 2000) and suffering long-term effects of mild traumatic brain injury. Therefore, it is suggested that effort and cognitive performance could be related. Consequently such factors may also affect the reliability and stability of concussion screening test results.

Regression to the mean. Regression to the mean (RTM) is a confounding factor in NP assessment because it can make natural variation in repeated assessment look like real change when in fact it is not (Barnett, van der Pols, Dobson, 2004). Therefore, RTM means a reasonable threat to internal validity that can occur in any measurement study and in any variable that is subject to random error (Barnett et al., 2004). Once RTM is ignored, it can result in wrongly concluding that an effect is due to treatment when it is due to chance (Morton & Torgerson, 2003).

According to statistics, RTM usually occurs in the following situations: 1) when there is a non-random sample of people with extreme values from the normal population; 2) when this non-random sample is tested on two consecutive independent variables; as a result, the mean of this sample on the first measurement will get closer to the mean of the second measurement; 3) when the follow-up measurements are only examined on a selected sub-sample using a baseline value

(Barnett et al., 2004); 4) if the value from the population mean is extreme, the more likely the results will regress to the mean (Morton & Torgerson, 2003); 5) if the correlation between the two variables is weak, the larger is the effect of RTM (Bland & Altman, 1994a), and 6) if the degree of measurement error is high due to random variation.

In health care RTM can have potentially serious implications. As Morton & Torgerson (2003) describes it, RTM can confound performance on diagnostic tests because of random chance. Therefore, any intervention to a group that is extremely different from the normal population will seem to be successful due to RTM. In clinical practice this may lead to false interpretation of test results concluding that treatment is affective when in fact this may not be true. Since RTM effects could occur in pretest-posttest designs, they may be present in the study sample. Therefore, such effects have to be assessed before the main data analysis. RTM effects will be assessed using a method described in Rocconi & Ethington (2009).

Summary. There is limited information available regarding the impact of effort and incentives on the neurocognitive assessment of sports-related concussion diagnosis and management. Because serial testing of athletes involve various confounding issues (such as practice effect, lack of effort at baseline testing, demonstrating more effort at post-concussion testing, reliability of self-report about actual symptoms), research is needed to investigate whether incentives and effort (e.g., raising interest in performing well on a test) could affect NC performance on the ImPACT battery. If so, we need to know how reliable the results are. Although there are numerous studies demonstrating that ImPACT is a sensitive measure of detecting cognitive decrements due to concussion (Iverson et al., 2003a; Iverson, Franzen, Lovell, Collins, 2003b; Iverson et al., 2005; Iverson et al., 2004; Lovell et al., 2003; McClincy et al.,, 2006; Schatz et al., 2006), so far no studies have been conducted that address changes of effort in serial testing when using the

ImPACT battery and how it may affect the diagnostic and management process prior to and after a concussion.

It is known from previous studies that using monetary incentives, feedback, and other forms of incentives (i.e., receiving tokens/points, various privileges, and prizes) result in improving outcomes and/or modifying certain behaviors, especially in rehabilitation programs (Drebing et al., 2005, Geller et al., 1982, Hagenzieker, 1991, Hermann et al., 1973, Petry et al., 2005, Phillips, 1968, Phillips et al., 1971). This provides reason to presume that consequenting test taking behavior to increase participants' effort with extra monetary incentives (e.g., raising interest in performing well on the test by providing opportunity to receive monetary incentives) could result in better performance on neurocognitive tests; therefore, may affect the diagnostic process.

Clinical observations. The assumption that change in effort could affect NC performance is also based on our clinical observation here, at the UTEP Concussion Management Clinic. We have been assessing and treating athletes since 2007. During these assessments we have observed that athletes do not necessarily demonstrate adequate effort when tested at baseline, especially when those testing sessions were organized in groups. This inadequate effort was often demonstrated in athletes 1) talking to one another during testing; 2) watching others completing the test and not focusing on their individual performance; 3) using mobile phones during testing; 4) not reading the instructions carefully.

Furthermore, when sustaining a concussion we noticed that some athletes tended to minimize and underreport their symptoms, especially when their injury occurred during the season with an upcoming sports event the following week. We also noticed that by the time these athletes were about to start the RTP protocol, they tended to exceed their baseline performance on the ImPACT battery in at least one of the four cognitive composite scores (i.e., Verbal

Memory, Visual Memory, Visual Processing Speed, and Reaction Time). The following pilot studies were conducted in order to examine whether these observations could be substantiated statistically.

Pilot Studies

The main focus of this dissertation is to examine whether change in effort affects NC performance. Over the years, the UTEP Concussion Management Clinic has been following the "baseline assessment approach". Since 2007 our Clinic has assessed more than 900 athletes at baseline including young amateur athletes from 9 major high schools in El Paso, UTEP, and El Paso Community College. We have been assessing and evaluating athletes with a concussion serially using the best computerized assessment tools, and working together with the athletic trainers and a neurologist to provide the best care for our athletes in this community. During these assessments we had the opportunity to observe the behavior and interview our athlete-patients. We have noticed that athletes do not always demonstrate adequate effort when tested at baseline, especially when athletes were organized in groups to complete ImPACT. In addition, when athletes came back to our Clinic with a concussion, we also observed that some of them seemed to minimize and underreport their symptoms, especially when the injury occurred during the season, and there was an upcoming sports event the following week. Furthermore, by the time these athletes became symptom-free and their performance on the ImPACT battery returned to their baseline, we have noticed that some of them exceeded their baselines in at least one of the four cognitive composite scores (i.e., Verbal Memory, Visual Memory, Visual Processing Speed, and Reaction Time).

Examining whether the last post-concussion test performance differs significantly from the baseline performance of athletes with a concussion explains whether athletes demonstrate a change in effort on the ImPACT battery across multiple administrations. In addition, it also suggests that effort as a factor may affect concussion diagnosis and management. In these pilot studies, the last post-concussion test performance is always the performance that is compared to the athlete's baseline performance indicating that the athlete has recovered from the concussion, and after which the athlete is recommended to start the RTP protocol. Such recommendation is based on the following criteria: 1) athletes are symptom-free; and 2) athletes' performance returned to the level of baseline performance on the ImPACT battery. Since these athletes were tested at baseline and serially at post-concussion levels, improved performance on the last test – when compared to baseline – could indicate the presence of some practice effect and/or that there is a difference in effort between the two levels. Consequently, such difference in performance could also indicate that athletes may be at risk to return-to-play prematurely while they are still under recovery due to demonstrating less effort at their baseline performance. Therefore, this difference in performance would also call for a more sophisticated and elaborate concussion diagnosis and management in order to make sure our athletes can truly return to competition safely. However, no significant findings between the last post-concussion and the baseline test would indicate that the test battery is sensitive to changes in effort and practice; therefore, it provides reliable and stable test results on cognitive functioning.

Complementary to this analysis, it is necessary to examine how sensitive our concussion assessment tool, ImPACT, is with our athlete-patients for concussion management and which cognitive functions are impaired following a concussion. For that purpose, the second exploratory analysis compares baseline performance to the first post-concussion test results using ImPACT. The first post-concussion performance is always the performance that measures the athlete's cognitive functions within 3-4 days after sustaining a concussion. Such analysis explains which cognitive composites are the most likely and sensitive to identify a true deficit due to a concussion. Based on previous studies from the test developers of ImPACT, all four cognitive measures (i.e., Verbal and Visual Memory, Visual Processing Speed, and Reaction Time composite scores) are expected to indicate a significant decline in performance when compared to baseline test results (Iverson et al., 2003a, Lovell et al., 2003b). In the case of "Total Symptom Scores", self-reported symptoms are expected to increase significantly at the first post-concussion assessment when compared to baseline self-reported symptom scores.

The third analysis between the first and last post-concussion performance explains whether athletes demonstrate recovery and whether athletes show the same performance pattern when comparing the first post-concussion performance to their baseline performance. If the test battery is less sensitive to identify deficits (i.e., baseline vs. first post-concussion performance) but more sensitive to identify improvement (first post-concussion vs. last post-concussion performance), such result would indicate two important aspects besides recovery form a concussion: 1) there could be some practice effect present due to the serial nature of assessment in concussion management; 2) athletes may demonstrate change in their effort when completing the test across multiple administrations. There is, however, reason to believe that practice effect on the ImPACT battery may not be the major issue that imposes risks to proper sports-related concussion management. First, when designing the battery the test developers at Pittsburgh were focusing intensely on reducing practice effects with this new computerized battery (i.e., our current version at the Clinic: Version 2.0). They randomized and designed as many stimuli in the battery, so that those taking the test would not come across with the same stimuli the next time they come in for another post-concussion evaluation. In this sense, the stimuli keep changing in each and every testing session. Second, research on ImPACT claimed that practice effects were not present when testing healthy, non-concussed amateur athletes four times on the battery (Lovell et al., 2003). In another ImPACT study, practice effects were only present in Visual Processing Speed, but not in Verbal Memory, Visual Memory, and Reaction Time. Because practice effects were found in one cognitive measure, reliable change index calculations were adjusted within the battery. Therefore, when the battery calculates the composite scores at the end of the test, the implemented reliable change calculations ensure that the performance shows a "real" change and is not due to chance or random variation. However, more issues may arise since reliable change calculations were adjusted for only one cognitive measure but not for the other measures (i.e., Verbal Memory, Visual Memory, and Reaction Time). Even though, the test developers stated that there was no need for such adjustment due to the absence of practice effects in these cognitive measures, it must be emphasized that the findings have not been replicated yet, and the reliability measures of ImPACT differ from one study to the other. In addition, test developers never really considered that effort may impact performance on the battery, although recent literature has been emphasizing to conduct more research in this area (Bailey et al., 2006). Since the suggestion is that athletes want to achieve a performance good enough to get permission to return to play, effort seems to be the variable that should be investigated further on.

In this third analysis, the main aim is to measure whether the difference between the first and last PC test would show the same pattern as in the difference between the baseline and the first PC test. Furthermore, it would also demonstrate whether athletes recover from their concussion. The reason for such an investigation is to examine if the test is as sensitive to detect recovery after a concussion as it is to detect a deficit due to a concussion. If ImPACT is more sensitive to identify recovery than it is to detect a concussion in a sample of athletes with probable concussion than it would also support the assumption that effort may affect NC performance on the test. Furthermore, the more significant differences the analysis shows between the first and the last post-concussion assessment than between the baseline and first post-concussion performance, the more it supports the assumption that changes in effort could have an effect on NC performance. As a result, it could impact concussion diagnosis and management. Since athletes often express their desire to return to practice/competition, these pilot investigations would demonstrate which cognitive composites are more sensitive for impairment and which test composites are more susceptible to improve due to recovery and effort. In addition, these pilot analyses would also provide some ground for further study in the impact of effort on NC performance in concussion management.

To examine such an impact, the purpose of these exploratory analyses is to investigate 1) which cognitive functions are impaired on the ImPACT battery due to a concussion when comparing the baseline and the first post-concussion performance of our patient-athletes at the UTEP Concussion Management Clinic; 2) whether performance on the last post-concussion test (i.e., when we recommend to start the RTP protocol with the athlete) shows recovery and/or exceeds the baseline performance; and 3) whether there is a tendency to improve performance on those composite scores that were not impaired on the ImPACT battery. In this way, we are able to analyze whether effort has an impact on NC performance in athletes with a concussion, and whether ImPACT is sensitive enough to identify deficits due to mild traumatic brain injuries.

Study 1

Purpose

The purpose of the study is to measure the sensitivity of the ImPACT battery on 22 collegiate athletes who recently sustained a concussion and were tested at the UTEP Concussion Management Clinic both at pre- and post-concussion levels. The concussed athletes' performance in cognitive functioning at the first post injury evaluation is compared to their previous baseline performance.

Participants

There were 22 concussed collegiate athletes taking part in the study. Athletes who reported having been diagnosed with attention deficit disorder (ADD), learning disability (LD), and receiving special education at the baseline testing were excluded from the study (1 male athlete with self-reported ADD and a history of 2 concussions was excluded). In addition, only those athletes were included for further analysis whose performance at baseline was valid; therefore, 4 athletes with invalid baseline test results were excluded previously.

From the 22 athletes 16 were males (73%) and 6 were females (27%). 14 male athletes were football players, 2 were ice-hockey players. Among the female athletes 3 were softball, 2 were basketball, and 1 was a soccer player. All 22 participants underwent preseason baseline testing at the UTEP Concussion Management Clinic. At the time of the preseason (i.e., baseline) testing the mean age of the 22 athletes was 18.91 (SD = 0.97, age range 17-21). For the women the mean age was 19 (SD = 0.89, age range 18-20). The mean age of the men was 18.87 (SD = 1.03, age range 17-21). At the time of the first post-concussion testing, the mean age of the 22 athletes was 19.53 (SD = 0.9; age range: 8-22). The mean age for women at PC1 was 19.50 (SD = 0.84), whereas the mean age for the men at PC1 was 19.56 (SD = 1.37). Years of education for the whole group ranged between 12 and 15 years (M = 12.5, SD = 0.86). See Table 2 for descriptive statistics of this sample in Appendix A.

From the 22 athletes 13 reported having no history of concussion (59%, 10 males, 3 females), 7 reported having a history of one concussion (32%, 2 females, 5 males), 1 male athlete reported having a history of 2 concussions (4.5%), and 1 female athlete reported having a history of 5 concussions (4.5%).

All 22 athletes suffered a concussion and got tested at the UTEP CMC approximately 3-4 days post injury (M = 3.55, SD = 3.05). Time post between baseline and the first post-concussion (PC) testing ranged from 2 weeks to 21 months (M = 8 months, SD = 6.5 months).

Procedure

Athletes were tested with the ImPACT concussion battery Version 2.0 at the UTEP Concussion Management Clinic at baseline and post-concussion levels. Version 2.0 of ImPACT is a computerized neurocognitive test battery designed for concussion assessment and management. ImPACT consists of six individual test modules that measure aspects of cognitive functioning such as attention, memory, reaction time, and visual processing speed before and after a head injury. The six cognitive modules (i.e., the actual test) consist of the following designs: (1) Word Memory; (2) Design Memory; (3) the X's and O's; (4) Symbol Match; (5) Color Match; (6) Three Letter Memory.

After reading our information letter about concussion management at UTEP and signing the consent forms, the athletes filled out the demographic, sports and health history forms. Having filled out the 22 symptom-inventory and rated their current symptoms, the athletes proceeded to the test. Participants were told before they started the test that they have to read the instructions carefully before they proceed to the test, and they have to make sure they understand what they have to do in the task. If they have any questions about the task, the examiner would help the participants. Approximately, the test required 40 minutes to complete. Detailed, written instructions were provided prior to each and every test module. After completing the six test modules, the program software measured the performance based on the subtests and calculates five composite scores: (1) Verbal Memory Composite (VeMC); (2) Visual Memory Composite (ViMC); (3) Processing Speed Composite (PSC); (4) Reaction Time Composite (RTC); and (5) Total Symptom Scores (TSS).

Once the athletes returned for follow-up assessments after sustaining a concussion, the examiner(s) interviewed the athletes about the circumstances of their injuries as well as their current symptoms. Following that the athlete was administered to a series of neurocognitive tests including ImPACT. After the assessment, the examiner(s) evaluated the athletes' self-reported symptoms, the signs of the concussion in the athlete and their current performance on the tests. Athletes' PC performance on the ImPACT battery was evaluated by comparing their test results to their baseline performance. If the athlete reported post-concussion symptoms that was considered unusual (i.e., the subsequent category after 'normal'; applying the normative data for ImPACT Version 2.0, (2003): TSS $\geq$ 7 for university men; TSS $\geq$ 10 for university women) and their cognitive performance on the ImPACT battery showed any deficits compared to their baseline performance, the athlete was scheduled for further follow-up evaluations. These assessments and evaluations of acute concussion lasted until the athlete's symptoms returned to normal (i.e., in most cases symptom-free) and their last PC performance returned to their level of baseline. In addition, athletes were provided with further clinical recommendations in order to manage their concussion and advance their recovery. Recently the Concussion Management Clinic started to recruit and evaluate athletes 3, 6, and 12 months post-concussion to track their cognitive functioning and symptom-reporting. For this analysis, athletes' baseline (BL) and first PC (PC1) performance is compared to examine which measures are sensitive to detect deficits due to concussion.

Design & Analysis

This set of analysis is a within-subject design. The dependent measures for this study are the performance (i.e., composite scores) on Verbal Memory, Visual Memory, Visual Processing Speed, Reaction Time, and Total Symptom Scores. The independent variable is sensitivity (i.e., the presence or absence of a deficit due to concussion). The independent variable is dichotomous, whereas all five dependent variables are continuous. Since preliminary data analysis did not meet the assumptions for applying parametric statistical methods in this sample, the equivalent non-parametric methods were applied (i.e., the Friedman related samples test and the Wilcoxon related-samples test) to analyze within-subject differences. Prior to the main analysis factors that may affect the outcomes were examined. These factors were namely ‘gender’ and ‘history of concussion’. To measure whether a) gender and b) history of concussion affect performance on the ImPACT battery in this sample, the Mann-Whitney *U* Test was applied additionally. Further analysis showed that there were no missing data across the repeated measures and values were added accurately. Since this set of analyses compares performance across testing sessions (i.e., BL, PC1, and last PC) within the same sample of athletes, the non-parametric Friedman related samples test was applied first to measure differences in the five individual dependent measures (e.g., BL verbal memory, PC1 verbal memory, and last PC verbal memory was compared). To measure differences within the individual performance (e.g., in this analysis between BL and PC1 performance), the Wilcoxon related-samples test was used with each of the five dependent measures. P-values were adjusted for multiple comparisons to control alpha for inflation of Type I error.

Results

Differences across the repeated test administrations were measured in a sample of 22 athletes who were tested before and serially after sustaining a concussion. Time post BL and PC1 was approximately 8.2 months (SD = 6.5 months). Time post last concussion and the last PC evaluation was approximately 2 weeks (SD = 8.6 days) and ranged between 0 to 43 days (i.e., 6 weeks; see Table 2.). Time post PC1 and last PC was approximately 8.6 days (SD = 9 days). Time post BL and the last PC evaluation was approximately 8.5 months (SD = 6.5 months; see Table 3 and 4 in Appendix A).

Prior to the main analysis, gender as a factor was investigated (6 collegiate female (f) athletes and 16 male collegiate male (m) athletes). Mann-Whitney *U* Test estimations showed no gender effects to be present in this sample when comparing baseline performance between the two groups (two-tailed, α = 0.05; BL Verbal Memory: Median(f) = 79.50, Median(m) = 86.00, p = 0.375; BL Visual Memory: Median(f) = 77.50, Median(m) = 74.50, p = 0.51; BL Visual Processing Speed: Median(f) = 39.53, Median(m) = 37.98, p = 0.58; BL Reaction Time: Median(f) = 0.56, Median(m) = 0.57, p = 0.317). No differences were measured between male and female athletes on reporting symptoms (BL Total Symptom Scores: Median(f) = 0, Median(m) = 4.5, p = 0.408).

History of concussion (HOC) was also examined as a factor that may affect results within the analysis. The ImPACT performance of 13 athletes (10 males, 3 females) with a (HOC) was compared to 9 athletes (6 males, 3 females) with no HOC. Mann-Whitney *U* Test measures indicated no differences in cognitive performance between athletes with and without a self-reported history of concussion (two-tailed, α = 0.05; BL Verbal Memory: Median (HOC) = 86.00, Median (no HOC) = 85.00, p = 0.62; BL Visual Memory: Median (HOC) = 74.00, Median (no HOC) = 75.00, p = 0.947; BL Visual Processing Speed: Median (HOC) = 36.55, Median (no HOC) = 38.25, p = 0.48; BL Reaction Time: Median (HOC) = 0.57, Median (no HOC) = 0.55, p = 0.056). Furthermore, no differences were detected between the two groups in relation to symptom-reporting (BL Total Symptom Scores: Median (HOC) = 8.00, Median (no HOC) = 0, p = 0.101).

When applying the Friedman related samples test (two-tailed with α = 0.05) to measure differences in all the five individual composite scores across the multiple administration of the ImPACT battery, it estimated that three of the five measures showed significant differences across testing sessions.[1] These measures were Visual Memory (BL Median = 74.50, PC1 Median = 69.00, last PC Median = 81.50, df = 2, $p < 0.008$*), Reaction Time (BL Median = 0.5600, PC1

[1] For more accurate descriptive purposes the mean (M) and standard deviation (SD) values were included additionally to the median values of each composite score across testing sessions.

Median = 0.6000, last PC Median = 0.5300, df = 2, $p < 0.005$*), and Total Symptom Score (BL Median = 2.00, PC1 Median = 16.50, last PC Median = 3.00, df = 2, $p < 0.001$*). In Visual Motor Speed (BL Median = 37.975, PC1 Median = 39.000, last PC Median = 40.440, df = 2, $p = 0.48$) and Verbal Memory (BL Median = 85.00, PC1 Median = 87.00, last PC Median = 89.50, df = 2, $p = 0.61$) the Friedman related samples test did not detect any significant differences across the repeated test administrations, even though the median values for these two composite scores show the tendency to increase across testing sessions. These results are illustrated in Figure 1 & 2 in Appendix B.

The sensitivity of the composite scores to the acute effects of concussion was estimated in the sample of 22 amateur athletes who were tested preseason and within 3-4 days post-injury. When using the Wilcoxon related/paired samples method (two-tailed with $\alpha = 0.05$), the athletes demonstrated a significant decline in 2 of the 5 composite scores: Reaction Time (baseline M = 0.5650, SD = 0.0506, Median = 0.5600, PC1 M = 0.7105, SD = 0.3573, Median = 0.6000, $p < 0.021$*, $dz = 0.44$, moderate effect size) and Total Symptom Score (baseline M = 4.86, SD = 6.97, Median = 2.00, PC1 M = 22.45, SD = 21.65, Median = 16.50, $p < 0.003$*, $dz = 0.81$, large effect size). The athletes, however, did not demonstrate any significant decline in Verbal Memory (baseline M = 84.00, SD = 10.85, Median = 85.00, PC1 M = 81.73, SD = 16.25, Median = 87.00, $p = 0.66$); Visual Memory (baseline M = 72.23, SD = 13.84, Median = 74.50, PC1 M = 67.55, SD = 15.64, Median = 69.00, $p = 0.19$); and Visual Processing Speed (baseline M = 38.42, SD = 5.21, Median = 37.97, PC1 M = 35.89, SD = 9.97, Median = 39.00, $p = 0.66$). These results are illustrated in Table 5 – Appendix A.

Discussion

When comparing athletes' first post-concussions test results to their baseline performance on the ImPACT battery, the results indicated that the most sensitive measures to concussion in the ImAPCT battery are Reaction Time and Total Symptoms Scores. When comparing the first post-concussion test results to athletes' baseline performance, athletes demonstrated a significant decline in reaction time after sustaining a concussion. This sample of athletes was much slower in making decisions and completing the test in general than before their concussion occurred.

This group of athletes' symptom scores increased significantly from their baseline symptom scores. This finding indicated that our athletes after sustaining a concussion report and rate more symptoms than before sustaining a concussion. In this sense, the ImPACT battery seemed to be a sensitive measure in two of the five measures in this sample of collegiate athletes with a concussion. Therefore the findings of this analysis show that our participants not necessarily demonstrate cognitive deficits in all the cognitive areas ImPACT measures based on the median values. On the other hand, since athletes' performance at their PC1 assessment did not indicate significant improvement when compared to their baseline performance, it is assumed that the battery is sensitive to the effects of concussion. Furthermore, it is designed in order to reduce the chance of probable practice effect.

Comparing the results of this study to the study of Iverson et al (2003a), this sample of athletes did not demonstrate significant decline in the measures of Verbal Memory, Visual Memory, and Visual Processing Speed. There may be several reasons why this study could not replicate the results of Iverson and his colleagues. First, our sample size was smaller than the study of Iverson's; therefore, not necessarily representing the distribution of the cognitive measures of the population with acute sports-related concussion. Second, there is a chance that athletes' effort at their preseason baseline was not adequate or simply decreased; therefore, not necessarily representing their accurate performance on the ImPACT battery. Consequently, when assessing them at post-concussion, the effort of the athletes completing the battery might have increased than their effort at baseline, and as a result ImPACT could not detect deficits due to such discrepancy in effort. This is a factor that has not been investigated when interpreting change on the ImPACT battery, and as such waits for further testing.

Study 2

Purpose

The purpose of this analysis is to examine whether this group of athletes recovered from their concussion and whether they exceeded their baseline performance significantly compared to their last PC assessment on ImPACT. Since it is assumed that athletes with a concussion are more keen to return to play, this comparison would illustrate whether athletes' effort has an effect on NC performance using the ImPACT battery or whether there is some practice effect present.

This analysis also presumes that athletes will not demonstrate any significant difference between their self-reported symptoms at BL and at the last PC test session, indicating that they no longer experience concussion symptoms, therefore, they are ready to start the RTP protocol. The analysis expects this group of athletes to return to their baseline level of performance indicating that they have recovered from the concussion. However, this analysis also presumes that athletes will exceed their baseline performance on at least one composite score when compared to their last PC test results.

Participants & Procedure

Participants and test procedure were the same as in Study 1.

Design & Statistical analysis

This set of analysis is a within-subject design. The dependent measures for this study are the performance (i.e., composite scores) on Verbal Memory, Visual Memory, Visual Processing Speed, Reaction Time, and Total Symptom Scores. To measure whether there is a difference between BL and the last PC performance on the individual composite scores, the Wilcoxon related-samples test was applied.

Results

Change in cognitive performance was examined in a sample of 22 athletes who were tested at baseline and got tested serially after a concussion. To estimate whether athletes demonstrate recovery, their BL and last PC performance was compared. Using the Wilcoxon related/paired samples test (two-tailed with $\alpha = 0.05$) for the individual composite scores, the athletes demonstrated significantly better performance after having tested serially in Visual Memory (baseline M = 72.23, SD = 13.84, Median = 74.50, last PC M = 79.23, SD = 12.39, Median = 81.50, $p < 0.021$*, $dz = 0.61$, medium effect size). Athletes, however, did not demonstrate any significant differences in their performance on Verbal Memory (baseline M = 84.00, SD = 10.85, Median = 85.00, last PC M = 87.59, SD = 9.07, Median = 89.50, $p = 0.22$), Visual Processing Speed (baseline M = 38.42, SD = 5.21, Median = 37.975, last PC M = 40.17, SD = 5.94, Median = 40.44, $p = 0.11$), and Reaction Time (baseline M = 0.5650, SD = 0.0506, Median = 0.5600, last PC M = 0.5614, SD = 0.0886, Median = 0.5300, $p = 0.36$). There was no significant difference within athletes in relation to symptom-reporting (baseline M = 4.86, SD = 6.97, Median = 2.00, last PC M = 3.73, SD = 5.13, Median = 3.00, $p = 0.57$). These values are demonstrated in Table 6 – Appendix A.

Discussion

The analysis indicated that our group of athletes returned to their baseline performance indicating that they recovered from the effects of a concussion. However, athletes' performance was significantly exceeding their baseline performance on one composite score: Visual Memory. This finding is surprising for a few reasons. First, the study of Iverson et al. (2003a) found significant improvement on Visual Processing Speed but not on Visual Memory test results when conducting test-retest reliability on ImPACT. Second, Lovell et al. (2003) found no practice effects and reported stable performance on all the composite scores when testing healthy, non-concussed amateur athletes four times. However, by examining the mean and median values on each individual score across testing sessions, our findings indicate that there is improvement on all the composite scores except for Reaction Time and Total Symptom Scores. These two measures returned to baseline and did not exceed that level. Such pattern not only indicates that athletes show recovery, but they may also demonstrate maximum effort during PC testing. Consequently, effort may affect cognitive performance on the test, especially when athletes with a concussion are tested during season.

The reasons behind exceeding baseline performance and showing significant improvement on one composite score could be the assumption that there may be a difference in effort in those athletes who did not sustain a concussion and those athletes who sustained one. Once athletes come back to the Clinic for a concussion assessment, they may recognize that this time they have to do “well” on the test in order to consider the permission to RTP. Therefore, there may be a difference in the effort of an athlete with a possible concussion who may feel there is much at stake for him/her than a non-injured athlete who does not have to experience this kind of pressure. In this sense it is possible that non-concussed athletes may not demonstrate change in their effort to complete the test a few times, and as a result they may not show change in their performance across the multiple administrations. This assumption would explain why previous studies did not find improvement of performance on the composite scores when using non-concussed athletes across the repeated screenings (e.g., in the study of Lovell et al., 2003). However, those athletes who want to return to play may demonstrate a change in effort and as a result may show significant improvement in their performance on the test. Such pattern, nonetheless, may pose many risks to reliable and accurate concussion diagnosis and management. One of them is the concern that effort may obscure post-concussion evaluations. Although we do not assume athletes actively misrepresenting themselves at baseline by demonstrating poor performance on purpose; however, disinterest, apathy, and lack of knowledge can lead to demonstrating lack of effort at the first testing (Bailey & Arnett, 2006). Consequently, the second risk is showing not significant difference in performance when tested after a concussion, despite the actual existence of cognitive deficits. The findings of this analysis confirm that further study is necessary to examine whether effort impacts NC performance and whether it may impact concussion diagnosis and management.

Study 3

Purpose

The aim of this analysis is to examine whether this sample of athletes demonstrates recovery and whether athletes show the same performance-pattern when comparing the results of the first comparison (i.e., baseline vs. the first PC performance) to the performance-pattern of their recovery (i.e., PC1 and last PC performance). In addition, this analysis also shows whether there is a tendency to improve performance on those composite scores that were not impaired on the ImPACT battery when comparing each composite score across testing sessions (i.e., BL, PC1, last PC). In this case we already know that our sample of concussed athletes showed deficits in reaction time when completing ImPACT after their injury. In addition, they indicated more concussion-related symptoms with more intensity. However, since athletes' performance did not show significant decline in their verbal and visual memory, as well as their processing speed, it is not expected that athletes would show significant improvement when comparing their first and last PC performance Accordingly, in this analysis it is presumed that athletes would show significant improvement in their reaction time and their self-reported symptom scores when comparing their first and last PC performance. More specifically, this would indicate that athletes have become just as fast as they were before their concussion, and they no longer experience any concussion-related symptoms.

Participants and Procedure

Participants and test procedure were the same as in Study 1.

Design & Statistical analysis

This set of analysis is a within-subject design. The dependent measures for this study are the performance (i.e., composite scores) on Verbal Memory, Visual Memory, Visual Processing Speed, Reaction Time, and Total Symptom Scores. To measure whether there is a difference between PC1 and the last PC performance on each composite score, Wilcoxon related-samples test was applied.

Results

Using the Wilcoxon related/paired samples test (two-tailed with $\alpha = 0.05$) for the individual composite scores, the comparison of PC1 and the last PC performance showed that athletes recover in their Reaction Time (PC1 M = 0.7105, SD = 0.3573, Median = 0.6000, last PC M = 0.5614, SD = 0.0886, Median = 0.5300, $p = 0.006$*), and they return to the level of BL (BL Median = 0.5600, last PC = 0.5300). Athletes also reported significantly less symptoms on the Concussion Symptom Inventory (PC1 M = 22.45, SD = 21.65, Median = 16.50, last PC M = 3.73, SD = 5.13, Median = 3.00, $p < 0.001$*) indicating that their TSS returned to the 'normal' level (BL Median = 2.00; last PC = 3.00). These results measuring recovery correspond to the findings of the first study, in which athletes showed significant decrease in their Reaction Time as well as significant increase of their self-reported symptom scores. However, further results marked significant improvement on Visual Memory (PC1 M = 67.55, SD = 15.64, Median = 69.00, last PC M = 79.23, SD = 12.39, Median = 81.50, $p < 0.003$*), a result that was estimated significantly not different in this sample of athletes when comparing their BL to their first PC performance. Athletes did not demonstrate any significant differences in their performance on Verbal Memory (PC1 M = 81.73, SD = 16.25, Median = 87.00, last PC M = 87.59, SD = 9.07, Median = 89.50, $p = 0.13$), and Visual Processing Speed (PC1 M = 35.89, SD = 9.97, Median = 39.00, last PC M = 40.17, SD = 5.94, Median = 40.44, $p = 0.055$). These findings correspond to the findings of the first study, in which our athletes did not demonstrate any deficit in Verbal Memory and Visual Processing Speed when testing them at PC1. See Table 7 for descriptive statistics in Appendix A.

Discussion

The results of this analysis support the hypothesis that effort could affect NC performance on the ImPACT battery even if athletes are recovering from a concussion. Participants in this study showed that they returned to their baseline performance in Reaction Time. In addition, the participants indicated that they no longer experience concussion-related symptoms. See Figure 3 & 4 in Appendix B.

The analysis also demonstrated four additional improvements on cognitive performance. In this sample, athletes enhanced their results on their verbal memory, visual memory, reaction time, and visual processing speed of information. These findings are illustrated in Figure 5 & 6 in Appendix B.

Comparing this pattern to the results of the first analysis (i.e., BL vs. PC1), it becomes more apparent that the battery is sensitive to detect recovery and sensitive to the effects of concussion within the participants. However, when examining baseline and the last PC test results on the cognitive functions, it is questionable whether the battery accounts for confounds such as effort in Visual Memory. The reasons behind such assumptions lie in the different performance-patterns this sample of athletes demonstrated. When comparing our athletes' average baseline test results to their first PC test results, we found that athletes showed an overall decline in all cognitive abilities tested approximately 4 days post-concussion. However, only one function showed a significant deficit; our sample of athletes declined in how fast they respond to an item. In addition, they also reported more concussion-related symptoms on the Concussion Symptom Inventory. Because ImPACT identified these two composite scores to be most affected by a concussion in our sample of athletes, we presumed that these two measures should be the ones that would show a significant difference when comparing the first PC test results to the last PC performance within our sample. However, this last comparison indicated more measures to be improving along with our athletes' Reaction Time and Total Symptom Score: Visual Memory. In addition, this group of athletes demonstrated slight gradual improvement on Verbal Memory and Visual Processing Speed scores across testing sessions despite the concussion. Present findings correspond to previous findings of Iverson et al (2003a), who already indicated improved performance on Visual Processing Speed. On the other hand, these results contradict some of the previous findings in other research. For example, Lovell et al (2003) – when testing participants four times – did not find any significant improvement on any of the cognitive composite scores. Furthermore, our second comparison already showed that this sample of athletes exceeded their baseline performance on Visual Memory. Since there are more composite scores at the last PC test that exceed the baseline level of performance, there is reason to presume that change in effort may affect performance on the ImpACT battery.

To clarify the impact on concussion diagnosis and management, those athletes who do not demonstrate optimal effort at baseline testing and increase their effort greatly during post-injury assessment, their baseline performance will fall below their true cognitive abilities (Bailey et al., 2006, Bailey & Arnett, 2006). However, after a concussion occurs, these athletes are most likely to decrease their cognitive performance at 2 hours (Bailey et al., 2006), or in this study 4 days post-injury when looking at the mean performance, even though they may appear to return to their baseline by 1 week post-injury due to a change in their effort. Situations such as these may have dangerous consequence on the clinical decision-making that is based on inaccurate baseline cognitive abilities and underreported concussion-related symptoms. Therefore, there is a risk that the athlete is sent back to play prematurely before s/he resolved from the concussion. However, returning to play before recovering from the previous head injury increases the likelihood of sustaining another concussion (Guskiewicz et al., 2000). Hence suffering prolonged somatic-psychological-cognitive symptoms or a more dangerous Second Impact Syndrome on rare occasions (Cantu & Voy, 1995). Since there is only limited research that has been conducted within this area, it only emphasizes the importance of examining effort on NC performance in sports-related concussion assessment. The overall findings of this study may seem to underlie previous assumptions and call for a more targeted examination on the impact of effort on NC testing and on concussion diagnosis and management. Considering that previous studies already demonstrated the positive effect of extra incentives on behavior, this dissertation project proposes to apply extra monetary incentives as a method to enhance and maintain maximum effort in cognitive performance across the repeated administrations of the ImPACT battery.

Conclusion

Using our current data on collegiate athletes with a concussion, three pilot studies were conducted to examine 1) which cognitive functions are impaired on the ImPACT battery due to a concussion; 2) whether our concussed athletes recover and exceed their baseline performance; and 3) whether there is a tendency to improve performance on those composite scores that were not impaired on the ImPACT battery. Findings tend to support the hypothesis that this sample of athletes even with a concussion may demonstrate higher test results due to change in effort across the testing sessions.

Expected findings

Based on previous studies (e.g., Iverson et al., 2003a), performance was expected to decrease significantly on all the cognitive measures of ImPACT when comparing pre- to the first post-concussion test results. In addition, the composite scores measuring memory, reaction time, and processing speed were expected to increase and return to the level of baseline indicating that our patient-athletes are recovering from their concussion.

Self-reported symptom scores were expected to be significantly higher at the first post-concussion assessment than at baseline. Total symptom scores were also expected to return to baseline or in the normal range when comparing the last post-concussion to baseline test results indicating that our patient-athletes are recovering from their concussion.

Comparisons

Valid baseline NC performance was compared to the first and to the last post-concussion performance of 22 athletes to investigate whether athletes display any deficits and recovery from their concussion. Comparing the baseline and last post-concussion test results would also indicate whether this sample of athletes exceed their baseline performance on the ImPACT battery. In addition, the first and last post-concussion performance was compared for reliability purposes (i.e., to examine whether pattern of cognitive recovery is consistent with the cognitive deficits marked on the first post-concussion evaluation). Verbal Memory, Visual Memory, Reaction Time, Visual Processing Speed, and Total Symptom Scores were measured as the dependent variables for this study.

Findings

Non-parametric related-samples Friedman and Wilcoxon tests revealed that our sample of concussed athletes displayed significant deficits only in Reaction Time and reported significantly more concussion-related symptoms when comparing baseline to the first post-injury test results. In addition, our athletes' performance tended to decline on Visual Memory as a result of the concussion. These findings showed that the ImPACT battery is a sensitive measure for detecting cognitive deficits due to a concussion.

When looking at findings in comparing the baseline to the last post-concussion performance on the ImPACT battery, we found that this group of concussed athletes performed significantly better on Visual Memory. In addition, they also displayed the tendency to improve on Verbal Memory, Visual Processing Speed despite the concussion, when comparing their *median* cognitive performance across the testing sessions (i.e., baseline, first, and last post-concussion assessment). There was a slight (i.e., not significant) improvement on Reaction Time score as well when comparing athletes' baseline and last post-concussion performance indicating that by the time athletes recover and they are ready to go back to practice/competition, they could become faster than pre-concussion. These findings may support the hypothesis that effort may affect NC performance in sports-related concussion assessment.

Discrepancies in findings

Further comparison between the performance of the first and last post-concussion indicated that ImPACT may detect improvement on more composite scores than detecting deficits between the baseline and first post-concussion performance. This finding suggests that the test is sensitive to detect athletes recovering from their concussion. The conclusion on ImPACT "detecting more improvement than deficits" should be interpreted carefully since mean values showed an overall decline in cognitive functioning on the battery from baseline to PC1. However, performance increased in our athletes on almost all cognitive measures (except for

Reaction Time) even when examining the mean values between baseline and the last post-concussion performance. See Figures 7-10 in Appendix B.

When looking at the mean and the median *cognitive* performance, the results show a different pattern of performance change. The median performance shows that performance gradually increases in Verbal Memory and Visual Processing Speed despite the fact that our athletes sustained a concussion. In Verbal Memory, the figures with median values demonstrate that our sample of athletes became better on recalling words, whereas in Visual Processing Speed the findings indicate that our athletes tended to process information faster across test sessions. However, the mean performance shows exactly the opposite pattern in which performance has decreased in all the four cognitive composite scores. Although this decline was significant *only* in Reaction Time, still these results call for a careful interpretation of the findings. First, the mean and the median performance across testing sessions provide a different slope of change in cognitive functioning. In Reaction Time the higher the score means lower performance; the athlete has become slower in responding to a test stimulus (See previous Figures 5-6 and Figure 7-10 in Appendix B).

There are several reasons why the median and mean performance shows different patterns and slopes when indicating change in cognitive performance across testing sessions. First, in a normal distribution the mean and the median values are the same. In this study, difference between the mean and median performance shows that our distribution did not follow the normal distribution curve. Second, our sample is not large enough for representing the athlete population. Finally, this difference between the two values also projects that there may be outliers in our sample that affect the mean values in the composite scores. However, since we have been examining an amateur athlete population that is exposed to the risks and consequences of mild traumatic brain injury, we strongly believe that all head injuries are unique and as such contribute to a better understanding of the phenomenon. Consequently, it is important to take all cases of concussive head injuries into serious consideration for clinical purposes in order to develop more refined and accurate diagnosis and management for our athlete-patients. Therefore, in this study both the mean and median performance was used for a more refined interpretation of the results such as change in the athletes' performance across testing sessions. For statistical analysis, however, non-parametric methods were applied since our sample did not hold and support the assumptions for applying parametric methods. As a result, differences measured within athletes' performance on the battery are based on the median values.

Another crucial factor that has to be considered in interpreting the results is time and the number of times athletes were tested. First, the number of times athletes got screened differed individually. Some athletes got tested twice, others four or five times depending on the recovery of the athlete. This could affect the findings indicating that concussions varied among the athletes in terms of time of recovery and exposure to serial assessment. Second, time between the onset of concussion and the last PC assessment varied among participants ranging up to 6 weeks indicating that athletes' recovery pattern differs individually (See Tables 3 & 4 in Appendix A.). In addition time post onset and the first PC assessment also ranged on a broad scale. Although athletes took their first PC test approximately 3-4 days post-injury, some of this group took their first PC assessment up to 12 days post-concussion. Therefore, time and the number of times participants are tested have to be controlled in order to make sure these two factors do not confound the actual findings.

Summary

The overall findings of these pilot studies suggest that this sample of athletes show cognitive deficits due to a concussion and demonstrate recovery in their cognitive functioning as well. Furthermore, this group of athletes also demonstrated that they can exceed their baseline performance by the time they recover and ready to start the RTP protocol. These results could indicate that athletes may demonstrate change in effort and improve their cognitive performance on at least one cognitive measure when tested serially (pre- and last post-concussion) using the ImPACT battery. Furthermore, the data may also underlie the hypothesis that effort – due to the expressed desire to get back to play as soon as possible – may affect cognitive performance, and may impact the clinical evaluation and management of sports-related concussion. Therefore, future study is necessary to further examine the impact of change in effort on NC performance and its impact on concussion diagnosis and management.

Statement of the Problem

The issues addressed in this dissertation are two-fold. First, it addresses the controversies surrounding the reliability and stability of the ImPACT battery. Previous studies reported a lack of consistent stability and reliability measures when using the ImPACT battery (Iverson et al., 2003a, Broglio et al., 2007a). Second, it is related to research indicating that athletes whose NC performance falls one standard deviation below average performance at baseline show larger change when evaluated at 1 week post-injury. Furthermore, the findings also indicate that these athletes tend to exceed their baseline performance 1 week post-concussion. Such change between test results at baseline and at 1 week post-concussion would not be expected since athletes usually are in the process of recovering from a concussion. Therefore, this interesting finding may pose difficulties for the clinician to interpret changes in test performance accurately (Bailey et al., 2006, Hunt et al., 2007).

When examining our pilot data with collegiate athletes, we found that athletes with "average" baseline performance showing deficits at approximately 3-4 days post-concussion. Furthermore, their performance recovered following a concussion as would be expected approximately 12 days post-injury, though time of recovery varied across athletes. However, our group of athletes also showed that the last post-concussion performance (i.e., approximately 12 days post-concussion) exceeded the baseline performance on certain composite scores. It remains unclear whether such unexpected improvement from baseline to post-concussion is due to 1) random variation due to lower than expected test-retest reliability or 2) lack of effort at baseline and maximum effort at post-injury assessment, or 3) learning from previous exposure to the battery.

Statement of the Purpose

As discussed in previous sections, ImPACT 2.0 battery was designed to assess memory, speed of information processing, reaction time, visuo-spatial abilities, visuo-motor abilities, and attention. To measure the psychometric properties of this widely used assessment instrument, test developers conducted reliability, sensitivity, and validity measures mainly on athletes following a concussion. These studies compared (a) the test results of normal non-concussed athletes with the test results of concussed athletes and/or (b) the baseline performance with the post-injury performance of concussed athletes. These studies while using the ImPACT battery showed inconsistent findings of test-retest coefficients and practice effects present in visual processing speed. However, test-retest studies did not take one crucial factor into consideration: whether effort may affect NC performance on the battery and how this may affect the stability of the scores obtained on the battery. Based on the study of Bailey et al. (2006) and our clinical observations with the athletes during baseline and post-concussion testing, we used the complete post-concussion collegiate dataset we have at the UTEP Concussion Management Clinic to investigate in which cognitive abilities athletes demonstrate deficits when administrating the ImPACT battery after sustaining a concussion, and whether these athletes exceed their baseline performance across the repeated NC assessments while they recover from the injury.

These investigations demonstrated that athletes show cognitive deficits due to a concussion. On the one hand, these deficits were memory-related indicating that this sample of athletes had problems with recalling visual-abstract shapes accurately. On the other hand, the deficits the 22 athletes displayed affected the speed of responding to a specific stimulus. Therefore, due to the effect of concussion our athletes became significantly slower than prior to their concussion.

When looking at the recovery pattern of the 22 athletes, performance between baseline and last post-concussion assessment significantly increased in Visual Memory. Furthermore, we saw a tendency of increasing post-concussion performance on Verbal Memory and Visual Processing Speed composite scores of the ImPACT battery. Therefore, this sample of collegiate athletes demonstrated that they could exceed their baseline performance on certain cognitive measures following a concussion. Although the improvement of their composite scores demonstrated recovery of cognitive functions from a concussion, the tendency to show higher composite scores on the last post-concussion test than at baseline has raised further questions: "Why did performance on the last post-concussion test exceed baseline?" "Does effort have a role in NC performance when assessing athletes?" "Does learning play a role and are these scores truly reliable?"

Consequently, it is important to examine the basis of the peculiar test performance patterns exhibited by athletes pre-to-post concussion. Therefore, this dissertation addresses factors, other than brain injury and recovery that may influence performance on the ImPACT battery. Specifically, it examines the roles of random variation, effort, and learning across repeated test administrations of the test battery.

In order to operationalize these objectives, the effects of concussion were eliminated by examining test performance of non-concussed college students. Participants were tested on four occasions with 7-8 days between testing sessions. Additional monetary incentive for these non-athlete participants on the 3rd test session (i.e., treatment session) was used as a reasonable analog for athletes wanting to return to play. Test-retest reliability was measured across all 4 test sessions. Performance between the 2nd and 3rd, as well as the 3rd and 4th test session was compared to measure the effects of extra incentive on NC performance. Performance between the 1st and 2nd test sessions was compared to measure practice effects (i.e., learning).

Research Questions and Hypotheses

To address the issues of random variation, effort, and learning on NC performance when using the ImPACT battery, the following research questions were asked:

Research question #1: Does ImPACT provide stable and high test-retest reliability measures of cognitive performance within a group of non-concussed college students across the repeated administration of the ImPACT battery?

Operationally defined: What are the test-retest reliability coefficients of each composite score across the four test sessions when a group of non-concussed collegiate students are administered to the alternate test forms of the ImPACT battery?

Null hypothesis for statistical analysis:

> Test-retest coefficients on the ImPACT battey will be statistically unrelated and will not vary across test sessions for all five individual composite scores.

Research question #2: Does extra monetary incentive affect cognitive performance on the ImPACT battery *within* a group of non-concussed collegiate students?

Operationally defined: Is there a statistically significant difference within subjects' composite scores before (i.e., 2nd test session) and after (i.e., 4th test session) applying extra monetary incentives on the 3rd test session?

Null hypothesis for statistical analysis:

There will be no statistically significant difference within subjects' composite scores on the five measures of verbal and visual memory, visual motor speed, reaction time, and cognitive efficiency index before and after extra monetary incentives are applied.

*Research question #**3***: Are there any practice effects present within the cognitive performance of non-concussed college students when administering the alternate test forms of the ImPACT battery?

Operationally defined: Is there a statistically significant difference within subjects' composite scores achieved on the 1st and 2nd test session?

Null hypothesis for statistical analysis:

There will be no statistically significant difference within subjects' composite scores between the 1st and 2nd test session.

CHAPTER THREE

Methods & Design

Overview

This study investigates 1) the stability of test-retest reliability coefficients across all four test sessions; 2) whether there is a difference within non-concussed college students' NC performance before and after applying extra monetary incentive; and 3) whether there are practice effects present within subjects' cognitive performance in the first two test sessions. Participants were asked to complete the ImPACT battery a total of four times. Before all test sessions start, participants were offered a one-time-only compensation of $30 in gift card for participating in the experiment. At the 3rd test session, participants were told that they could earn extra money if they do well on the individual test modules. This extra incentive of $30 was offered only in the 3rd (i.e., treatment) session. A posttest was conducted. The participants were not offered any extra incentives on the 4th test session.

Setting

The testing sessions were conducted at the Concussion Management Clinic (CMC) at University of Texas at El Paso (UTEP) located at the College of Health Sciences (CHS). The Clinic and the CHS provided the necessary computers and software programs that were essential for the testing. Because the ImPACT battery Version 2.0 is computerized and Internet-based, participants could complete the tests online using laboratory computers at a prescheduled appointment with no interruptions. Completing the test required attention, focus, and concentration from the test-takers; therefore, the Concussion Management Clinic also offered a quiet, secure, and comfortable environment for the participants during their testing.

Recruitment

33 non-concussed college students were recruited from the main Campus of the University of Texas at El Paso together with the College of Health Sciences and School of Nursing. An IRB approved research flyer was used to recruit men and women from the UTEP student population. An electronic flyer was also posted on the monitors of "UTEP Today" at the College of Health Sciences. In addition, participants were recruited through classroom visits as well as via contacting faculty requesting for assistance in recruiting UTEP college students.

Participants were tested in the Fall Semester 2010. After the first testing session eligible participants were invited to participate in the study and come back three more times to complete the test again approximately 7-8 days between testing periods. Time between test sessions was based on our clinical practice with athletes who sustain a concussion. We usually schedule our athletes for follow-up assessment 7-8 days between test sessions.

The principal investigator (PI) kept contact with the participants via telephone and/or e-mail to follow-up on scheduled appointments with them. For example, at the UTEP Concussion Management Clinic, we usually call and/or leave a message via phone or e-mail to confirm a test appointment with the athletes and/or athletic trainer. Based on this approach, the PI contacted the participants one day prior to testing to reconfirm a scheduled appointment for testing. The participants were free to reschedule an appointment for the same day of the scheduled testing (e.g., changing appointment from morning to the afternoon) or move the appointment for testing the next day if necessary. The PI, however, had to make sure that time between testing sessions did not exceed 7-8 days for all participants.

Consent Form

Upon selection, the purpose of the investigation, all procedures, benefits and/or risks associated with the research were thoroughly explained to each participant by the principal investigator. They also had the opportunity to read the written consent for themselves and ask questions about the research and their participation. The Consent Form provided a brief overview of the procedure explaining the use of the ImPACT battery, how long the testing is going to take and how many times they are asked to complete ImPACT. Furthermore, an explanation of the participant's rights about not participating and/or withdrawing from the study was provided. Participants were assured that participation is voluntary and withdrawal from participation is allowed at any time during the study without any penalties. A self-report on health history, medical condition was required from the participants for study eligibility. In the Consent Form it was stated the compensation of $30 will be offered for participating in all four test sessions. The compensation will be provided at the end of the 4th testing session. Only those participants who completed all four test sessions and keep the time interval of 7-8 days between test sessions were compensated. The consent form did not provide any information about receiving extra (extra) incentives for the 3rd test session. Consent form had to be signed by the participant to take part in the experiment.

Participants

Inclusion criteria. Those participants with a history of attention deficit disorder, learning disability, as well as with a history of one concussion within the past 12 months, with a history of two or more concussions, brain surgery, psychiatric condition, epilepsy/seizure, meningitis, and substance abuse were excluded from taking part in the study. In addition, participants had to be enrolled to college. In the total sample there were 33 non-concussed college students who completed the ImPACT battery 4 times.

Gender, age, and years of education. There were 18 (54.5%) female and 15 male (45.5%) college students taking part in the study. The average age of the total sample was 21.36 (SD = 1.66), with an average of 14.09 (SD = 1.36) years of education. Age ranged from 18 to 25; education ranged from 12-17 years. The sample included 29 undergraduate (16 females, 13 males) and 4 graduate students (2 females, 2 males).

In terms of age and years of education, females and males had similar attributes (see Table 8 in Appendix A). The average age for college women was 21.11 (SD = 1.18, range = 19-24) with an average of 14.33 (SD = 1.19. range = 13-17) years of education, whereas the average age for college men was 21.67 (SD = 2.09, range = 18-25) with an average of 13.80 (SD = 1.52, range = 12-17) years of education. No participant reported having diagnosed with attention deficit disorder, learning disability, or receiving any special education.

Language. Of the total sample there were 26 (79%) who reported speaking a second a language (i.e., Spanish) other than English. Fifteen participants (45.5%) selected Spanish as their first language. However, only four of these 15 students selected Spanish as their test/exam language on the ImPACT battery.

Health and sports history. There were two participants (1 male and 1 female) who indicated having a history of one concussion which was more than 6 years post onset. Participants were not excluded from the study if they reported having a history of one concussion at more than 18 months post onset. No participant reported any attention deficit disorder, learning disability, or a history of receiving special education services. In addition, none of the participants indicated receiving any treatment for substance abuse, psychiatric condition, meningitis, epilepsy/seizure, or brain surgery. During the data collection (i.e., after the 2nd testing) one participant indicated being treated with migraine headaches. However, on the day of the last two test sessions the student reported not experiencing any migraine headaches in the past 24 hours. From the total sample none of the participants sustained any head injuries during the time of the data collection.

Participants ranged on a wide scale in their sport- and recreational activities: football, basketball, swimming, volleyball, softball, racquetball, cheerleading, cross-country, soccer, tennis, ballet, and cycling. Eighty-eight percent (n = 29) of the total sample reported pursuing some sports or other recreational activities in their life, mainly during their high school years. Seventy-nine percent (n = 26) of the students indicated taking part in high school sports such as football, volleyball, swimming, soccer, softball, or basketball. Twenty-one percent (n = 6) still pursues some sport activity on a daily or weekly basis such as cycling, swimming, baseball, or racquetball. From the total sample, 4 participants (2 women and 2 men) reported never doing any sport or recreational activity.

Symptom-reporting. On average, both college men and women reported symptoms that were in the normal range when compared to ImPACT Normative Data (Iverson et al., 2003, see Table 9 – Appendix A). ImPACT Normative Data reported university men having an average of 4.5 (SD = 7.5, 0 - 5 normal) total symptom score, whereas university women had an average of 8.0 (SD = 10.3, 0 - 10 normal) total symptom score. See Table 2 for descriptive analysis in total symptom scores by gender. In this sample total symptom scores ranged from 0-14 at test session 1 (M = 4.27, SD = 4.79, normal), 0-18 at test session 2 (M = 3.55, SD = 5.03, normal), 0-13 at test session 3 (M = 3.36, SD = 4.66, normal), and 0-14 at test session 4 (M = 2.39, SD = 3.82, normal). Preliminary analysis showed no significant gender differences and within-group effects of self-reported symptom scores on ImPACT.

Materials and Measures

The ImPACT Version 2.0 Internet-based test battery is a computerized test designed for concussion management by measuring cognitive skills before and after head injury. It offers five tests altogether: a baseline and four consecutive tests. All tests use the same structure of tasks but with different stimuli. Test instructions of the test modules are the same across test sessions. Approximately, the test requires 30-40 minutes to complete. ImPACT consists of seven sections including the actual test. Participants have to fill out and complete all seven sections each test session. The seven sections are the following:

(1) Demographic data (i.e., date of birth, age, gender, handedness, height, weight, native country, native language, second language, years of speaking the second language);

(2) Education (years of education completed; if they have ever repeated one or more years of school; received speech-language therapy; were diagnosed with learning disability or hyperactivity and attention deficit disorder (ADD); attended special education);

(3) Current sport history (type of sport, level of participation, position, years of experience at this level);

(4) Concussion history (number of times diagnosed with a concussion, concussion resulted in confusion and/or loss of consciousness, difficulty in remembering events that occurred before and/or after the injury, total games missed);

(5) Health history (treatment for headaches, migraines, seizures, substance abuse, psychiatric condition by a physician, history of brain surgery, and meningitis);

(6) Symptom inventory; this section includes 22 concussion-related symptoms. Athletes have to self-report and rank the intensity of the range of symptoms they experience on the test date from 0 to 6: with 0 meaning not experiencing the symptom through 6, indicating the severity of the symptom. The post-concussion symptom scale was adapted from Lovell & Collins (1998);

(7) The actual neurocognitive test battery that includes 6 test modules. The six cognitive modules (i.e., the actual test) consist of the following designs: (1) Word Memory; (2) Design Memory; (3) the X’s and O’s; (4) Symbol Match; (5) Color Match; (6) Three Letter Memory. The next section introduces these modules with further details.

Word Memory. This module measures attention processes, learning, memory and verbal recognition. The task includes 12 target and 12 non-target words. Each non-target word is semantically related to a target word. At first, the twelve target words are displayed on the screen for 750 milliseconds each. Subsequently, the words are displayed once again. After that a discrimination task follows, in which test takers have to decide from a 24-word list (including both target and non-target words) whether or not the word presented was on the original list of target words by clicking on the ”yes” or ”no” button.

Design Memory. This module measures and evaluates attention processes, learning, memory and visual (abstract shape) recognition. The task includes 12 target and 12 non-target designs. At first, the twelve target abstract designs are displayed two times for 750 milliseconds each – just like in the Word Memory module – which is followed by a discrimination test. The test taker is required to decide if the target design was among the stimuli or not by clicking on the ”yes” or ”no” button with the mouse while both target and non-target designs are displayed sequentially.

The X’s and O’s. This module measures and evaluates (spatial) working memory and reaction time (Echemendia, 2006). The test taker is required to recall where the yellow-highlighted X’s and O’s are placed on the display immediately after a reaction time task. At first, the program displays X’s and O’s randomly, in which three are highlighted in yellow. This is immediately followed by a distracter task where the test taker has to use the buttons “p” and “q” on the keyboard and push them as fast as they can when seeing the stimuli paired with the buttons. Immediately after the distracter task, the test taker has to recall the three highlighted X’s and O’s and click on the stimuli where s/he thinks the items appeared. This subtest consists of four tasks.

Symbol Match. This module measures and evaluates visual processing speed, learning, and memory. The test taker is asked to learn and recall nine shapes that belong to their paired numbers (from 1 to 9). After multiple trials of the task, all the shapes disappear from the screen, and they need to be paired with their number counterpart one-by-one. The shapes are presented randomly. The Cognitive Efficiency Index score is calculated based on the performance on this task.

Color Match. This module measures and evaluates reaction time. The words “red”, “green”, and “blue” are presented on the screen with their matched colors, and the test taker is asked to click on the right stimulus if the colors match their labels. The words with the three colors are presented variably and randomly.

Three Letter Memory. This module measures and evaluates the processing of verbal working memory and processing speed. The test taker has to learn and recall three random consonants that are presented on the screen next to each other. After presenting the 3 letters another distracter task follows: the test taker has to count backwards from 25 to 1 by clicking on the correct numbers on the screen as fast as they can. Immediately after the distracter task, the test taker has to recall and type the three consonants into the three-letter-box. This subtest consists of four tasks.

At the end of the test, the test taker is required to recall the words and abstract shapes again (i.e., the stimuli of the Verbal Memory and the Design Memory subtests).

After completing the six test modules, the program software measures the performance based on the subtests and calculates five composite scores: (1) Verbal Memory Composite (VeMC); (2) Visual Memory Composite (ViMC); (3) Processing Speed Composite (PSC); (4) Reaction Time Composite (RTC); (5) Impulse Control Composite (ICC).

Verbal Memory Composite includes the average percent correct for the Word Memory (module 1), Three Letter Memory (module 6), and Symbol Match (module 4) designs. Scores are also represented in age-referenced percentile.

The Visual Memory Composite consists of the average percent correct for the Design Memory (module 2), the X’s and O’s (module 3) tests. It measures visual memory and attention.

The Processing Speed Composite is calculated by the average of correctly recalled X's and O's (module 3) and the number of correct backwards clicking on the numbers (i.e., distracter task that is part of the Three Letter Memory test). The higher score indicates higher processing speed; therefore, better performance.

The Reaction Time Composite is composed of the average of reaction time scores for the Symbol Match (module 4), the Color Match (module 5), and the distracter task of the X's and O's (module 3) test (pushing the buttons "p" and "q" displayed between the X's and O's test). The lower score indicates better performance.

The Impulse Control Composite indicates the number of errors made while taking the ImPACT test. This composite also helps to identify test takers who do not put adequate effort into completing the test or are simply confused by test instructions (Iverson, Lovell, & Collins, 2003). It is useful in determining test validity; scores above 20 should be considered as invalid.

The Cognitive Efficiency Index Composite score has been recently (in early 2011) added to the ImPACT test measures by the test developer of ImPACT, Inc. According to ImPACT, the CEI score as an additional measure shows the interaction between accuracy and reaction time on the Symbol Match Test. The score ranges from 0 to 0.70 with a mean of 0.34 and indicates the extent to which the test taker performed on both speed and accuracy. For example, a low score (i.e., below 0.20) demonstrates that the test taker performed low on accuracy with slow response time, whereas a high score marks good accuracy with fast response time. This measure demonstrates how well participants performed on the Symbol Match Test at each test session in accuracy and reaction time. Therefore, CEI may also indicate the stability of test effort within participants.

Design

This study is a time series, two pretest-one posttest (A-A-B-A, in which A = baseline, B = treatment) quasi experimental design with equally-spaced time intervals (i.e., 7-8 days) between test sessions. Participants were seen on four occasions. They were tested either individually in the Concussion Lab.

Procedure

Participants had the opportunity to choose the language they prefer to complete the test. All four test sessions started with the same instructions:

> "You are about to start the test. Make sure that you read the instructions very carefully and understand what you have to do before you proceed to the task. From this point on, you are on your own. However, if you have any questions about the instructions, I am here to assist you to understand them. You can proceed now."

First, participants had to complete the demographic, health history forms, and fill out the symptom inventory. The principal investigator (PI) helped the participants to make sure they provide all the necessary demographic, education, sports and health history, and current symptoms data during these portions of the battery. Following the completion of the demographic, health history, and symptom inventory, the actual test started. Each ImPACT test consists of 6 test modules.

However, in the ***3rd test session*** (i.e., treatment), each participant were interrupted before a test module starts. Before participants proceeded to the first test module, the following information was presented verbatim to every participant:

> "You are about to start the first test module. If you do well on the test, you can earn additional money. You can proceed now."

In case participants asked questions about how much they can earn, the PI provided the following verbatim information to those participants: "It depends how well you perform today". No other information was provided for the participants.

After participants completed the first test module, they were asked to stop. Following a verbal praise ("Good Job!"), the PI handed out a $5 token saying "Here is a $5 token". After the verbal praise and dispersing the $5 token, the PI gave the following instruction:

> "You are about to start the next test module. If you do well on the test, you can earn additional money. You can proceed now."

This procedure was repeated for all 6 test modules. Therefore, after each of the 6 module the PI said the following to each participant verbatim:

> "Good job! Here is a $5 token. You are about to start the next test module. If you do well on this test module, you can earn additional money. You can proceed now."

Extra incentives were offered to all 33 participants regardless of how they performed on the battery. The PI made a record about the $5 tokens dispersed for each participant while completing the test modules. Participants could keep all six $5 tokens and trade it in at the end of the 4th test.

No results were presented to the participants any time in the experiment. If participants required additional information on their test results across test sessions, the PI scheduled a new appointment with them once the experiment was over (i.e., after completing the 4th test session). Furthermore, PI did not provide any information about how much they would earn during the experiment. For example, if any participant asked whether they receive extra money during the 4th test session, the PI repeated the test instructions again to the participant.

At the end of the 4th test session all participants were paid the compensation of $30 and the extra monetary incentive of $30. Therefore, at the end of the experiment each participant received a total of \$60. To earn the compensation and the extra monetary incentive, participants had to complete all four test sessions and keep the time intervals between test sessions (i.e., 7-8 days). Incentives were handed out after the participants at the end of the experiment after the 4^{th} test session.

Short survey. After participants completed the 4^{th} test, they were asked to fill out a 7 point Likert-scale short survey to get basic and initial feedback on the incentive condition. This short survey was filled out once participants had received the incentive condition and completed all four test sessions. The survey included the following two instructions:

1) On a scale of 0-6 rate/indicate the degree of how well you tried to perform the 3^{rd} time on the ImPACT battery because of the extra money offered.

0 – Not at all.

1 2 3 4 5 6

2) On a scale of 0-6 rate/indicate the degree of how well you tried to perform NOT because of the extra \$30 but because of other reasons (e.g., out of curiosity, see if you can perform better this time).

0 – Not at all.

1 2 3 4 5 6

Power analysis

Using the G*POWER software Version 3.0.10 (Faul, Erdfelder, Lang, & Buchner, Heinrich Heine Universität, Düsseldorf, Germany, 2007), a power analysis was calculated to determine sample size based on a t-test (means: difference between matched pairs) with the following input parameters: two tails, $\alpha = 0.05$, Power $(1 - \beta) = 0.80$. The reason behind choosing a matched samples t-test as a main method for deriving the sample size for this study was based on the clinical significance this study mainly focuses on; and that is whether extra monetary incentive has an effect on NC performance. Since a small effect size may not bear specific clinical significance (Jacobson & Truax, 1991) in the diagnostic and recommendation process of concussion management, effort has to indicate a medium-large effect size to be taken as a threat and risk in our concussion diagnostic evaluation. Based on within-subject effect size calculation (on GPower it is Cohen's *dz*, which adjusts the value of Cohen's *d* and converts it suitable for repeated measures design, GPower User's Guide), the effect sizes (Cohen's *dz* for a matched-pairs t-test) of .2 are considered as small, .5 as medium, and .8 as large. To determine the minimal increase in performance that would be considered reliable increase of performance is based on the "Interpreting change" study on the ImPACT battery of Iverson et al. (2003a). For all the four dependent variables the effect size (*dz*) ranged between 0.66-0.99. Based on the rationale that a large effect size is necessary for detecting differences in NC performance as a result of change in effort, a moderate and the most conservative effect size calculated (dz = 0.66) was applied to estimate the sample size for this study. In addition, choosing medium-large effect size is also necessary to be considered as a risk to accurate concussion diagnosis and evaluation The program computed a sample size of approximately 21 for a given effect size of 0.66 with 0.80 power, and an alpha level of 0.05. Assuming that 40% - 43% of individuals are lost to attrition (approximately 8- 9 individuals), this would imply recruitment a total of 33 participants. Rate of attrition is based on Broglio et al. (2007a) who reported a rate of 38% for loss of participants in their test-retest study (i.e., from the 118 only 73 was included in the analysis, 45 was excluded).

Statistical Analysis

Variables

The dependent measures for this study were the performance (i.e., the five composite scores) on Verbal Memory, Visual Memory, Visual Processing Speed, Reaction Time, and

Cognitive Efficiency Index. The independent variables were time and the presence/absence of extra monetary incentives.

Data screening & checking for outliers

SPSS version 17.0 statistical program was used for all data screening procedures. Before the main statistical analysis, careful examination of the data was carried out. Accuracy of data entry was checked and necessary corrections were made. There were no missing data. In order to carry out the main statistical analyses, a fit between data and assumptions of the main statistical analysis were assessed. This included the examination of normality, outliers, regression to the mean, and homogeneity of covariance matrix for each dependent measure.

Outliers. Descriptive statistics, skewness, kurtosis, and graphic representation (i.e., histograms following normal distribution curve) of the collected data was carried out to examine if all continuous variables were within range, and if means and standard deviations were plausible. For detecting outliers, standardized *z* scores were computed for all five continuous dependent variables. A *z* score larger than 3.29 and -3.29 indicated an outlier. In this study there were no outliers.

Regression to the mean (RTM). Regression to the mean is described as a common threat to internal validity in pretest-posttest designs. It is generally defined as “the tendency to score closer to the mean the second time a test is administered” (Rocconi & Ethington, 2009, p. 369). The current study includes a pretest-posttest design; therefore, a possible threat to internal validity needs to be investigated. For example, RTM effects may be present if some of the participants with a high pretest score would tend to shift slightly closer to the mean in the consecutive test session. Vice versa, if some of the participants scored low on the test the first time, their performance may shift slightly around the mean on the consecutive test session; therefore, RTM effects may be present.

RTM effects are detected if there is a negative Pearson correlation between change (i.e., the difference between pretest and posttest score) and pretest score (i.e., the scores obtained from the first test session). In case RTM effects are present on any of the dependent measure(s), pretest scores will be adjusted for RTM using a statistical method developed by Roberts (1980). Description of the method was obtained from Rocconi & Ethington (2009). The formula for adjusting pretest scores is the following:

Adjusted pretest score = pretest score + (1 - $r_{\square y}$) (grand mean of pretest score – pretest score), in which $r_{\square y}$ = test-retest reliability value between pre- and posttest.

After initial calculations, RTM effects were present for each dependent variable. Therefore, all composite scores obtained on the first test session were adjusted for RTM (see Table 10 – Appendix A). These adjusted pretest scores were used for the main statistical analysis. However, it has to be noted that the adjusted and non-adjusted scores did not change the pattern of statistical outcomes when calculating main effects. Still, to make sure the outcomes are not affected by RTM, the adjusted scores were used to calculate further statistical effects.

Main statistical analysis

Test-retest reliability. To answer the first research question (i.e., test-retest reliability of the ImPACT battery across test sessions), Pearson correlation was used to calculate test-retest coefficients for each composite score across the repeated administration of the ImPACT battery.

Repeated measures ANOVA. Repeated measures Analysis of Variance (ANOVA) model was used to answer the second research question (i.e., whether extra incentive affects NC performance when performance on the 3rd test session is compared to the performance on the 2nd and 4th test sessions), and the third research question (i.e., whether practice effects are present between the 1st and 2nd test session). More specifically, repeated measures ANOVA was employed to evaluate the overall significance of time with regards to the participants' performance on the ImPACT battery. In addition to the repeated measures univariate analyses, the Bonferroni test was employed for post hoc to determine any significant differences in cognitive performance among the 4 test sessions. The Bonferroni test of pairwise comparisons provides a strict and conservative adjustment of alpha level for each dependent variable (e.g., verbal memory between test session one and two, one and three, one and four, two and three, etc.) to avoid experimental error.

For all repeated measure analyses Mauchly's Test of Sphericity was conducted to control for homogeneity of covariance. The sphericity assumption was not violated if significance value was greater than 0.05 (i.e., *p* □ 0.05). In case the sample data failed the sphericity test (i.e., p value < .05), the more conservative (i.e., less likely to reject the null hypothesis) Greenhouse-Geisser adjustment was employed to determine within-group differences in cognitive performance. The Greenhouse-Geisser test modifies the degrees of freedom and increases the value of *F* to test statistical significance within the group and hence to avoid experimental error.

ImPACT Normative Data (2003) was used to compare our participants' performance to ImPACT norms and see whether there is a change in severity range of performance. Severity categories range from "impaired" to "very superior ". Composite score ranges, however, differ for gender and age range for all five measure. Since ImPACT Normative Data is based on gender and age range; therefore, the ImPACT norm-categories of University Men and Women were applied for this sample.

Trend Analyses. TA offers a method to identify stages where changes are significant across test sessions. In addition to repeated measures ANOVA, Trend Analysis (TA) was conducted to identify whether extra incentives affect the sequence of observations. Furthermore, TA was applied to examine whether individual composite scores show a significant trend of moving upwards or downwards across test sessions and to identify practice effects besides the effect of extra incentives.

CHAPTER FOUR

Research Results

The purpose of the study was 1) to examine the test-retest reliability of the ImPACT battery across test sessions, 2) to determine whether extra incentive affects neurocognitive performance on the 3^{rd} test session, and 3) to examine whether practice effects are present across the repeated administration of the ImPACT battery in a group of non-concussed college students. The results of the study are reported in this chapter.

Test-retest reliability.

The test-retest reliability analyses were based on healthy, non-concussed college students tested 4 times. This was a within-subject design. Similarly to the Iverson et al. (2003a)study reporting-test-retest correlations for all dependent measures, relative position across the test-retest distributions were examined with a Pearson correlation in this sample as well (see Table 13 – Appendix A). Reaction time showed the highest test-retest reliability coefficient, and visual processing speed showed the smallest test-retest coefficient value. Overall, test-retest coefficients were low and variable across test sessions.

For verbal memory test-retest reliability ranged from 0.19 to 0.65. These values also indicated that test-retest reliability coefficients gradually improved with the number of times the test was administered to the participants. A different phenomenon was observed in visual memory.

In visual memory, test-retest reliability ranged from 0.05 to 0.45 across test sessions. Overall, reliability remained low and did not show any gradual improvement with the number of times the test was administered. As it is seen in Appendix A – Table 13 test-retest reliability improves from Test 1 to Test 3, but drops drastically from Test 3 to Test 4 indicating variable performance levels within participants across time.

For visual processing speed, test-retest reliability ranged from 0.60 to 0.79, showing a gradual decrease of test-retest reliability. The highest test-retest reliability measure appeared between Test 1 and Test 2, and as the test administration increased, the reliability measure dropped to 0.60 by the fourth test session.

In reaction time, test-retest reliability ranged from 0.60 to 0.88. In this measure reliability gradually improved with the number of times the test was administered. Reaction time seemed to be the most reliable measure in the ImPACT battery. For the CEI score, test-retest reliability coefficients have never been reported. In this sample, test-retest reliability gradually increased ranging from 0.33 to 0.62.

Repeated measures ANOVA

ImPACT yielded composite scores in verbal memory, visual memory, visual processing speed, reaction time, and a recently added cognitive efficiency index. A high score indicated better performance in all measures except for reaction time. In ImPACT the scores in verbal and visual memory composite scores range from 0-100, and they are presented as percentages correct of 100. Scores in visual processing speed range from 0 to 60.

Table 11 presents the detailed descriptive statistics in Appendix A (i.e., means and standard deviations) for each dependent variable: verbal and visual memory, visual processing speed, and reaction time. Test session 3* represents the treatment session in which participants were offered extra incentive to boost performance on the battery. Pairwise comparisons across test sessions are illustrated in Figures 11-15 – Appendix B.

Verbal memory. The assumption of sphericity in the data was violated ($\chi^2 = 12.25$, $p = 0.038$); therefore, degrees of freedom were corrected using the more conservative Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.809$). For the verbal memory composite score, there was a significant within-subject effect in performance across test sessions ($F(2.43, 77.66) = 12.41$, $p <$

0.001, $\eta^2 = 0.279$, large effect size). Pairwise comparisons using Bonferroni adjustment revealed significant increase in verbal memory performance from the first to the second test session ($p < 0.001$), the first to the third test session ($p < 0.001$), and from the first to the last test session ($p = 0.002$). No other significant differences were found between test sessions in verbal memory performance. In the total sample, 66% (n = 22) of participants increased their performance in verbal memory from Test 1 to Test 2.

When using ImPACT Normative Data University Women (2003), the performance of our female participants did not shift in its severity category. Therefore, performance remained in the average range across all four test sessions (see Table 12 in Appendix A). However, when comparing our college men data to ImPACT Normative Data University Men, performance varied from the average range to high average range across the four test sessions.

Visual memory. The assumption of sphericity in the data was violated ($\chi^2 = 11.83$, $p = 0.037$); therefore, degrees of freedom were corrected using the more conservative Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.791$). In terms of visual memory, a marginal significant within-subjects effect was found ($F(2.37.89) = 2.67$, $p = 0.067$, $\eta^2 = 0.077$, small effect size). Pairwise comparisons using Bonferroni adjustments showed a significant increase in performance between the first and the last test session ($p = 0.048$). No other significant difference was found in performance among the test sessions. When offering extra incentive on the third test session, participants showed a small and statistically not significant increase. However, a significant linear trend was found across test sessions, indicating steady increase of visual memory performance within-subjects across time ($F(1,32) = 8.54$, $p = 0.006$, $\eta^2 = 0.211$, large effect size). No statistically significant quadratic ($F(1,32) = .415$, $p = 0.470$) or cubic ($F(1,32) = 1.23$, $p = 0.276$) trends of performance were shown across test sessions. Although change of performance was not statistically significant from Test 1 to Test 2, 61% (n = 20) of participants received higher composite scores at Test 2 than at Test 1.

When using ImPACT Normative Data University Men and Women, our college male and female participants did not shift in performance category. Therefore, performance remained in the average range for both genders across all four test sessions (see Table 12 in Appendix A).

Visual processing speed. The assumption of sphericity in the sample was not violated ($\chi^2 = 8.95$, $p = 0.11$). With sphericity assumed, significant within-subjects effects (time) were found on visual processing speed performance ($F(3, 96) = 6.145$, $p < 0.001$, $\eta^2 = 0.161$, large effect size). Pairwise comparisons using Bonferroni adjustments indicated significant differences in visual processing speed between the first and second test sessions ($p < 0.001$), the first and third test sessions ($p = 0.029$), and the first and last test sessions ($p = 0.009$) indicating increase in visual processing speed performance in all comparisons. No other significant differences were found within-subjects when comparing performance across test sessions in visual processing speed. In the total sample, 73% (n = 24) of participants received better composite scores in visual processing speed at Test 2 than at Test 1.

When using ImPACT Normative Data University Men and Women, our college male and female participants did not shift in severity category. Therefore, performance remained in the average range for both genders across all four test sessions (see Table 12 in Appendix A).

Reaction time. Mauchly's test indicated that the assumption of homogeneity of covariance in the data was violated ($\chi^2 = 12.53$, $p = 0.028$); therefore, degrees of freedom were corrected using the more conservative Greenhouse-Geisser estimates of sphericity ($\varepsilon = 0.807$). For the reaction time composite score, there was a significant within-subject effect in performance across test sessions ($F(2.42, 77.51) = 12.83$, $p < 0.001$, $\eta^2 = 0.286$, large effect size). Pairwise comparisons using Bonferroni adjustments revealed significant differences in reaction time performance between the first and second test session (*p* = 0.014) the first and the third test session (*p* = 0.026), and the first and the last test session ($p < 0.001$). Further increase in performance was found between the second and the last test session (*p* = 0.044), and between the third and last test session (*p* = 0.009). No other significant differences were found between test sessions in reaction time performance. In the total sample, 64% (n = 21) of participants received better composite scores in reaction time at Test 2 than at Test 1. These fin dings indicate that participants became faster on reaction time measures across test sessions.

When using ImPACT Normative Data University Men (2003), the performance of our college male participants did not shift in its severity category. Therefore, performance remained in the average range across all four test sessions (see Table 12 in Appendix A). However, when comparing our college women data to ImPACT Normative Data University Women, performance changed from the low average range to average range across between the first and second test sessions.

Cognitive Efficiency Index (CEI). Mauchly's Test of Sphericity was not violated ($\chi^2 = 3.34$, *p* = 0.65); therefore, no adjustment of degrees of freedom was necessary. For the CEI score, there was a significant within-subject effect in performance across test sessions ($F(3, 96) = 6.65$, $p < 0.001$, $\eta^2 = 0.172$, large effect size). Pairwise comparisons using Bonferroni adjustments showed that CEI improves when the participants' performance is compared between the first and second test session (*p* = 0.002), the first and the third test session (*p* = 0.004), and the first and the last test session (*p* = 0.026). However, change in CEI was not significant between Test 2 and Test 3, and Test 3 and Test 4. In the total sample, 76% (n = 25) of participants received better CEI scores at Test 2 than at Test 1. These fin dings also indicate that participants became faster and more accurate on the Symbol Match subtest from Test 1 to Test 2. ImPACT Normaitve Data was not available for CEI.

Trend analyses

In addition to the MANOVA model, a trend analysis was employed to identify underlying patterns of performance in time series (i.e., across test sessions) on each dependent variable. Figures 16-20 presents trend lines in each dependent measure. Test 3* represents the treatment session in which participants were offered extra incentive to boost performance on the battery. Figures 16-20 are presented in Appendix B.

Verbal memory. In verbal memory performance the participants showed a significant linear trend indicating a pattern in verbal memory performance that increases from Test 1 to Test session 2 ($F(1, 32) = 14.32$, $p < 0.001$, $\eta^2 = 0.309$, large effect size). However, a significant quadratic trend indicates a fraction in the linear trend line and showing stable performance on the verbal memory measure from Test session 2, 3, and 4 ($F(1, 32) = 18.25$, $p < 0.001$, $\eta^2 = 0.363$, large effect size). No significant cubic trend was found in this measure ($F(1, 32) = 2.29$, $p = 0.067$). Trend line of verbal memory performance is shown in Appendix B – Figure 16.

Findings of this trend analysis support the results found in the univariate pairwise comparisons between test sessions. When comparing all consecutive test sessions to test 1, participants' performance significantly increased, however, no change in performance was found within Test session 2, 3, and 4 as seen in Figure 16 (in Appendix B).

Visual memory. In terms of visual memory, there was a significant linear trend across test sessions indicating increasing performance from test 1 across test 4 ($F(1, 32) = 8.54$, $p = 0.006$, $\eta^2 = 0.211$, medium effect size). No significant quadratic ($F(1, 32) = 0.683$, $p = 0.415$) or cubic trend ($F(1, 32) = 1.23$, $p = 0.276$) was found in this measure within the group indicating only a linear trend of performance change in the four test sessions. See Figure 17 for trend line in visual memory performance within the group in Appendix B. Although there was an increasing linear pattern in visual memory performance, no significant change was detected between test sessions within the subjects.

Visual processing speed. For the visual processing speed composite score, there was a significant linear (F(1, 32 = 9.95, $p = 0.003$, $\eta^2 = 0.237$, large effect size) and cubic trend (F(1, 32 = 4.62, $p = 0.039$, $\eta^2 = 0.126$, medium effect size) across test sessions (see Figure 18 – Appendix B). The linear trend line shows that participants increased in their visual processing speed performance from test 1 across test 4, while the cubic trend indicates a fracture within the linear increase of processing speed performance. As it is seen in Appendix B – Figure 18, the trend line shows no change from test 2 to test 3, but a further increase of visual processing speed performance follows at test session 4. Findings of trend analysis support the univariate comparisons between test sessions; showing significant increase of performance from test 1 to test 2, from test 1 to test 3, and from test 1 to test 4, but no significant change from test 2 to test 3, and from test 3 to test 4.

Reaction time. With regards to reaction time, there was a significant linear trend across test sessions (F(1, 32 = 25.59, $p < 0.001$, $\eta^2 = 0.444$, large effect size). No quadratic (F(1, 32 = 0.516, $p = 0.48$) trend was found in reaction time performance; however, there was a marginal significant cubic trend line ($F(1, 32) = 3.88$, $p = 0.057$, $\eta^2 = 0.108$, small effect size) indicating faster response time overall when performance is compared to the first test, but a non-linear improvement of performance within Test 2 and 3. As it is seen in Appendix B – Figure 19, performance is almost identical between the second and third test sessions. However, decrease in response time indicates that participants became faster in performance from Test 1 across Test 4. Appendix B – Figure 19 shows trend line in reaction time performance across time. The post hoc comparisons support the results of this trend analysis indicating significantly faster response time from Test 1 to Test 2, from Test 1 to Test 3, and from Test 1 to Test 4. In addition, performance improved from Test 2 to Test 4, and from Test 3 to Test 4. However, no change in reaction time performance was found between Test 2 and Test 3.

Cognitive efficiency index (CEI). For the CEI there was a marginal linear and a significant quadratic trend of performance found across test sessions. In cognitive efficiency the participants showed a marginal linear trend of performance indicating a pattern that is increasing from test 1 to test session 2 ($F(1, 32) = 7.02$, $p = 0.012$, $\eta^2 = 0.180$, large effect size). However, a significant quadratic trend indicates a fraction in the linear trend line and showing stable performance on the CEI measure from Test session 2 across Test session 4 ($F(1, 32) = 9.45$, $p = 0.004$, $\eta^2 = 0.228$, large effect size). No significant cubic trend was found in this measure ($F(1, 32) = 2.82$, $p = 0.103$). Trend line of CEI is shown in Appendix B – Figure 20. Within-group trend line in CEI supports the findings of the univariate pairwise comparisons in this measure; participants showed a significant increase in their cognitive efficiency score from Test 1 to Test 2, Test 1 and Test 3, and Test 1 and Test 4, but no significant change from Test 2 to Test 3, and from Test 3 to Test 4.

Additional findings

Gender differences. Table 12 (see in Appendix A) shows the average performance on the ImPACT battery in college male and female participants. To evaluate gender differences, MANOVA was computed for time, gender and subtest. There was no significant interaction between time and gender for any of the composite scores ($F(3,29) = 1.07$, $p = 0.38$). However, significant interaction was found between gender and subtest ($F(4,28) = 5.44$, $p = 0.002$, $\eta^2 = 0.437$, large effect size). Significant gender differences were found in one of the dependent variables: visual memory ($F(1,31) = 11.75$, $p = 0.002$, $\eta^2 = 0.275$, large effect size). Not significant gender differences are marked as n.s. in Table 12. Pairwise comparisons revealed that college men performed significantly better in visual memory at test session 1 ($p = 0.044$) and test session 2 ($p = 0.004$) than college women. However, no significant gender differences were found in test sessions 3 and 4. Furthermore, no gender differences were found in any other dependent measure. Consequently, there were no differences between college men and women regarding their verbal memory, reaction time and visual processing speed performance. Neither did college males and females showed any significant gender difference in CEI.

Relationship between test effort and neurocognitive performance. After asking participants whether they tried to perform better on the third test sessions because of the extra monetary incentive, their answer ranged from 0-6 on a Likert-scale, 0 meaning “not at all”, 1 “very little”, 2 “a little”, 3 “somewhat”, 4 “moderately”, 5 “a lot” and 6 indicating “very much”. From the total sample 42.4% (n = 14) reported 5 and 6, indicating being affected by the extra incentive, while 15.2% (n = 5) reported not affected by the extra incentive at all. 12.2% (n = 4) of the total sample said the extra incentive played very little in performing better on the battery because of the extra incentive, and 30.4% (n = 10) indicated 3 or 4 on the Likert-scale (i.e., trying to perform somewhat better on the test).

The second question asked the participant whether they tried to perform better not because of the extra incentive but because of some other factor such as self-competitiveness, curiosity, or interested in becoming a research participant. Similarly to the first question, answers ranged on a 0-6 point Likert-scale, 0 meaning “not at all” and 6 indicating “very much”. Of the total sample, 69.7% (n = 23) scored 5 and 6 on the scale indicating to put high effort into performing better on the battery not because of the monetary incentive, but because other factors. 18.2% (n = 6) indicated being somewhat affected to perform better, and 4 (12.1%) participants reported not affected or affected only a little by other motivational factors than the extra money to improve their performance. Except for four, almost every participant indicated that they wanted to outperform their previous test performance. After the experiment, however, approximately one-third of participants came back to see what scores they achieved on the battery over time.

When correlating the answers (i.e., the points indicating on the 7-point Likert-scale) to the scores achieved on the third test session (i.e., the dependent variables: verbal and visual memory, visual processing speed, reaction time, and CEI), no significant relationship was found between the scores reported on the questionnaire and the performance achieved on the treatment (i.e., third) session. More specifically, extra incentive did not show to have any statistically significant effect on ImPACT performance. In addition to this finding, no significant correlation was present between scores achieved on the second questionnaire and the actual ImPACT performance.

CHAPTER FIVE

Discussion

To date standardized neurocognitive testing is a commonly used assessment method to diagnose and manage sports-related concussion. The results of such assessment play a crucial part in treating concussion and performing accurate and safe return-to-play decisions. Today's sports-related concussion management involves testing athletes at baseline (i.e., preseason, before a concussion may occur), and re-testing them (serially) if a concussion occurs. Athletes usually return to competition once they are symptom-free and their cognitive performance post concussion returns to their baseline cognitive performance.

Retesting athletes with a concussion now commonly involves using an alternate form of the same neurocognitive test they received at baseline/preseason. This way the clinician can determine whether the athlete's cognitive performance has declined when compared to his/her baseline. The ImPACT battery, that was specifically designed to assess sports-related concussion, offers such alternate test forms. Although this approach in sports-related concussion management seems specific and clear-cut, there are a number of factors that can interfere with the accuracy of assessment. This study addressed three such factors: test-retest reliability, the effect of extra monetary incentive on neurocognitive performance, and practice effects.

In terms of test-retest reliability (i.e., the 1st research question), this study presents important aspects of the psychometric properties of ImPACT version 2.0. The test-retest (i.e., Test 1 and Test 2) coefficients for the five cognitive composite scores ranged from 0.19 to 0.76 between the first and second test sessions. These stability coefficients are lower and different from the stability coefficients reported for ImPACT in the study reported by Iverson et al. (2003a). After testing 56 high school and collegiate participants twice on the ImPACT battery, the authors reported test-retest reliability coefficients ranging from 0.65-0.86. The days between test sessions ranged from 3 to 11 days. Just like the present study, their sample included both men (n = 29) and women (n = 27). Since time between test sessions varied in the study, it may have also affected test-retest coefficients. According to Kaplan & Saccuzzo (2009) higher test-retest reliabilities can be obtained when participants are retested within a short time interval. On the other hand Broglio et al. (2007a) place more emphasis on empirically determined, clinically relevant time intervals (i.e., 50 days between the first and second test sessions) between test sessions. Long time intervals between baseline and first post-concussion assessment, however, could lead to low reliability values (Kaplan & Saccuzzo, 2009). In this study, 7-8 days between test sessions were chosen, because we usually evaluate athletes with a concussion 7 days between assessments.

When comparing the test-retest reliability coefficients of this study, it became apparent that in verbal and visual memory as well as in reaction time our measures were much smaller than the ones reported by Iverson et al. (2003a). In the present study only the visual processing speed composite reliability value (i.e., 0.76) was similar to the value in the Iverson et al. study (i.e., 0.79). However, in this experiment Pearson *r* coefficients varied between indices and across test sessions as well. The reliability coefficients on several measures met or just exceeded the 0.60 level. The minimal acceptable level of reliability is usually defined between 0.65-0.70. Nunnally (1978) regards a reliability value of 0.70 as a minimum level, whereas DeVellis (2003, p. 95) suggests a reliability value between 0.65-0.70 as a minimally acceptable level. Both authors regard a 0.65 level or below as unacceptable and undesirable. Based on these criteria, only the visual processing speed composite reached a value that is “respectable”, as defined by DeVellis (2003). As reliability improves across test sessions, stability coefficient for verbal memory was met and reaction time exceeded the minimum acceptable level. Still, most test-retest coefficients in this study demonstrated low and unacceptable test-retest reliability values on the

ImPACT composite scores. Since ImPACT is the most widely used computerized concussion assessment tool, diagnosis and RTP decisions should not depend on using only this assessment measure.

Another observation was the change of test-retest reliability coefficients as the number of test administrations increased. When comparing the reliability coefficients of the last two test sessions to the Iverson et al. study, our values in verbal memory and reaction time composite scores showed more similarities to the test-retest coefficients reported by the test developers. However, in visual memory and visual processing speed composite scores reliability coefficients dropped across test sessions and remained smaller than those test-retest coefficients from Iverson et al. With such variability present and no concussion affecting neurocognitive performance, these results may suggest that changes of performance across time and test sessions may result from other factors (Broglio et al., 2007a).

According to Kaplan and Saccuzzo (2009), one of the issues with classical test theory is that it presumes that a behavior is constant over time. However, these authors argue that important behavioral characteristics, such as motivation, mental or health status can be expected to vary. In classical test theory, these characteristics are considered as errors. Still, if these behavioral characteristics are assumed to be present across time, then clinicians could account for them by applying multidimensional assessment techniques to diagnose and treat a concussion.

For example, one of the factors affecting performance fluctuation and improvement of performance in this non-concussed sample could be due to the ambition to perform better. Although no significant relationship was found between extra monetary incentive and the actual performance on the composite scores, 70% of the participants reported trying to perform better because of other factors. Participants often referred to these factors as being self-competitive, and curious to see if they can exceed their previous performance.

Within the context of this experiment, however, extra monetary incentive did not affect neurocognitive performance on the ImPACT battery. One of the reasons why extra incentive did not show any significant change in performance could be related to the amount of incentive offered in this study. When asking participants how hard they tried to exceed previous performance because of the extra money offered, 27.4% of them reported not being motivated by it at all or only a little. In addition, 30.4% indicated being moderately motivated by extra monetary incentive. Based on previous studies when extra monetary incentive was used in a rehabilitation setting a positive change in behavior was reported (Drebing et al., 2005, Geller et al., 1982, Hagenzieker, 1991, Hermann et al., 1973, Petry et al., 2005, Phillips, 1968, Phillips et al., 1971). Possibly the $5 token per subtest used in this study may not have been sufficient enough to boost cognitive performance on the ImPACT battery.

However, significant practice effects were found across all composite scores (i.e., verbal memory, visual processing speed, reaction time and the newly added cognitive efficiency index score) except for visual memory. Even after eliminating the effect of regression to the mean, a significant linear trend was found in visual memory and reaction time performance as well. These linear trend lines indicated gradual improvement of performance on the ImPACT battery across test sessions.

Practice effect has been a major issue in neurocognitive testing. Previous research found that practice effect shows a nonlinear relationship: those tests with high practice effect measures have poor reliability coefficients, and vice versa; tests with high reliability show small practice effects. In addition, there is evidence that practice effect does improve performance significantly on both traditional “paper – pencil” and computerized neurocognitive tests (Collie et al., 2003b; Collie et al., 2004; Duff et al., 2001; Rosenbaum et al., 2006).

Current findings were interesting since the battery was designed with a Reliable Change Index method being implemented in the battery. Previously it was mentioned that Reliable Change method helps to identify change in performance due to true change (in the statistical sense) and not due to error or random variation (Iverson et al., 2003 a).

Furthermore, when ImPACT's test-retest reliability was established, Iverson et al. (2003a) reported practice effects in only one neurocognitive measure: visual processing speed. In addition, Lovell et al. (2003) never found any significant improvement of performance on the ImPACT battery when testing 24 non-concussed high school athletes on four occasions. In the study of Lovell et al. (2003) test-retest coefficients were not reported. Interestingly, the present study found significant improvement when comparing neurocognitive performance from the first test session to the second test session. Linear trend analyses also supported findings for practice effects across composite scores. However, when comparing the means to ImPACT Normative Data (2003) across all test sessions, performance did not change severity categories as described in Table 12 – Appendix A. More specifically, performance mainly remained in the average range when Normative Data was applied based on the Reliable Change method to examine changes of performance. Such findings suggest that even though performance on the composite scores improved, they did not improve to such an extent that performance moved into another severity category. Normative data, however, varies in its ranges of performance. In addition, there are several issues the Reliable Change theory did not account for such as: 1) practice effect, 2) regression to the mean, and 3) other factors that include age, gender, education, and assessment related factors such as anxiety, fatigue, or lack of effort (Collie et al., 2004, Hunt et al., 2007; Stevens et al., 2008). The combination of these confounding factors may give the false impression that the patient needs to show large changes in cognitive function when even small changes may be just as meaningful and sufficient (Collie et al., 2004). Such findings have clinical implications since athletes get tested serially to evaluate cognitive status necessary to make return-to-play decision. In addition, athletes post-concussion often express their desire to return to competition as quickly as they can. Therefore, clinically one should be cautious when interpreting athletes' post-concussion test results. As an example, with practice effects present, no change in neurocognitive performance between baseline and post-concussion could also indicate that the athlete is still not ready to return-to-play and is under the effects of a mild traumatic brain injury. Since having low test-retest reliability and significant practice effects on the ImPACT battery, performance would be expected to improve across assessment sessions if no concussion has occurred.

Conclusions

Overall, research on sports-related concussion showed that neurocognitive assessment has been sensitive to the consequences of a concussion, and standardized, computer-based tests have many practical advantages that a paper-pencil test cannot offer.

At the UTEP Concussion Management Clinic, we have been applying ImPACT to assess and manage sports-related concussion for nearly 5 years. Our observations and analysis have also supported the necessity and usefulness of standardized neurocognitive testing. ImPACT has shown to be sensitive to the changes of neurocognitive performance due to a concussion and provides multiple domains to be examined over time. In our pilot data, 22 college athletes with a concussion became significantly slower in their reaction time than in their baseline performance. In addition, they reported significantly higher symptom scores post concussion than before they sustained a concussion. This significant increase in reaction time and self-reported symptom scores are all indicators of the effects of a concussion. At the same time these results also indicate that ImPACT was sensitive to detect the effects of a concussion. Still, understanding the psychometric properties of an assessment tool is crucial for a clinician in concussion management. Because of the serial nature of treating mild traumatic brain injuries, accurately identifying cognitive changes in athletes post injury requires a conservative and cautious approach from clinicians when making return-to-play decisions. As Broglio et al. (2007) suggests that greater focus should be placed on those neurocognitive measures that showed the highest test-retest reliability values. In ImPACT these measures were reaction time and visual processing speed. Our analysis in athletes with a concussion also suggested reaction time being the most sensitive measure to the consequences of a concussion on the ImPACT battery. However, because of the limitations of ImPACT, our concussion clinic also strongly supports the recommendation that a concussion assessment method should be based on multiple assessment techniques. For example, in addition to neurocognitive performance, we evaluate signs and self-reported symptoms, observe postural stability and balance, measure verbal fluency, evaluate post-traumatic stress, and assess auditory comprehension across test sessions. Since psychometric issues of concussion assessment in tools such as the ImPACT have not been clarified yet, a single test or method may not provide enough information to diagnose and manage mild traumatic brain injuries accurately and safely. Furthermore, because of the threat of practice effects, unreliable symptom-reports and research demonstrating the effect of effort on neurocognitive performance (Bailey et al., 2006, Hunt et al., 2007), a more refined multidimensional standardized assessment technique could identify neurocognitive deficits more accurately (Broglio et al., 2007). Therefore, future research in sports-related concussion should continue clarifying psychometric properties of neurocognitive tests using clinically relevant assessment methods. In addition, because of the possible threat of behavioral changes in neurocognitive performance, especially in the field of athletics, research should continue to adapt and evolve to new challenges presented by the diagnosis and management of sports-related concussion. . The training of professional in the use of these assessment instruments should include an understanding not only of the advantages but the limitations of a test battery and how to assess those modalities that – as we know from evidence –could reliably be affected by a mild traumatic brain injury.

Furthermore it is imperative to improve concussion assessment batteries to be less prone to practice effects. The ImPACT test battery provides alternative test forms to reduce the possible threat of learning. However, in this experiment it was evident that performance can be increased even if alternate test forms were used. Therefore, a short practice session before each test module may be useful to reduce the effect of learning. In this way test takers could learn and understand what is expected from them and what to pay attention to before the actual test begins. Another possible solution for obtaining test takers true score (i.e., in the statistical sense) preseason is to administer the baseline test twice. By choosing the most optimal test between the two baselines, dual baseline assessment could serve as a strategy to reduce practice effects (Duff et al., 2001).

Limitations of the Study

There are a number of limitations in the study that needs to be taken into consideration. One of these limitations is related to the sample size. The sample size of this study was not as large as the sample size of previous studies examining test-retest reliability. Iverson et al. (2003a) had 56 non-concussed high school and college athletes for the purpose of calculating test-retest reliability, while Broglio et al. (2007a) used 73 non-concussed college student volunteers. Still, initial power analysis and effect size measures for this present study suggested that sample size was sufficient to generalize statistical findings to college-age adults. Due to the increasing popularity of college athletics and the substantial number of college athletes, the results of the present experiment should not be neglected.

Secondly, it was previously noted that the amount of extra monetary incentive might not be sufficient enough to raise interest in exceeding performance on the ImPACT battery.

Therefore, it is presumed that a more substantial incentive could lead to a significant improvement of neurocognitive performance.

References

American Academy of Neurology. (1991). Practice parameter: The management of concussion in sports (summary statement). Report of the Quality Standards of Subcommittee. *Neurology, 48,* 581-585.

Anastasi, A. (1968). *Psychological testing* (3rd ed.). New York: MacMillan.

Armstrong, C. L. (1987). Mild head injuries. *Neuropsychology, 1,* 15-19.

Aubry, M., Cantu R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., Lovell, M., McCrory, P., Meeuwisse, W., Schamasch, P. (2002). Summary and agreement statement of the first International Conference on Concussion in Sport, Vienna 2001. *Clinical Journal of Sport Medicine, 12*(1), 6-11.

Bailey, C. M., & Arnett, P. A. (2006). Motivation and the assessment of sports-related concussion. In S. M. Slobounov & W. J. Sebastianelli (Eds.), *Foundations of Sports-Related Concussion* (pp. 171-195). Norwell, MA: Springer-Verlag.

Bailey, C. M., Echemendia, R. J., & Arnett, P. A. (2006). The impact of motivation on neuropsychological performance in sports-related mild traumatic brain injury. *Journal of the International Neuropsychological Society, 12,* 475-484.

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology, 34*, 215–220.

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

Bender, S. D., Barth, J. T., Irby, J. (2004). Historical perspectives. In M. R. Lovell, R. J. Echemendia, J. T. Barth, M. W. Collins (Eds.) *Traumatic brain injury in sports: An international neuropsychological perspective* (pp. 3-23). Lisse, The Netherlands: Swets & Zeitlinger Publishers.

Berner, O. (1935). Concussion of the head. *American Journal of Surgery, 29*(2), 273-281.

Binder, L. M., & Rohling, M. L. (1996). Money matters: A meta-analytic review of the effects of financial incentives on recovery after closed-head injury. *The American Journal of Psychiatry, 153*(1), 7-10.

Bland, J. M. & Altman, D. G. (1994). Statistic notes: Regression towards the mean. *British Medical Journal, 308*, 1499.

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

Broglio, S. P., Macciocchi, S. N., & Ferrara, M. S. (2007b). Neurocognitive performance of concussed athletes when symptom free. *Journal of Athletic Training, 42*(4), 504-508.

Cantu, R. C. & Voy, R. (1995). Second Impact Syndrome: A risk in any contact sport. *The Physician and Sports Medicine, 23*, 27-34.

Cantu, R. C. (1996). Head Injuries in sport. *British Journal of Sports Medicine, 30*, 289-296.

Cantu, R. C. (2001). Classification and clinical management of concussion. In J. Bailes & A. L. Day (Eds.) *Neurological sports medicine: A guide for physicians and athletic trainers* (pp. 25-35). Rolling Meadows, Ill: American Association of Neurological Surgeons.

Carone, D. A. (2008). Children with moderate/severe brain damage/dysfunction outperform adults with mild-to-moderate brain damage on the Medical Symptom Validity Test. *Brain Injury, 22*(12), 960-971.

CBS Sports News. (2009, November 18). Survey: NFL players hide, fear their concussion. Retrieved March 20, 2010, from http://www.cbsnews.com/stories/2009/11/18/sportsline/main5700174.shtml

Centers for Disease Control (2007). Nonfatal traumatic brain injuries from sports- and recreation activities – United States, 2001-2005. *Morbidity and Mortality Weekly Report, 56*(29), 733-737.

Cernich, A., Reeves, D., Sun, W., & Bleiberg, J. (2007). Automated Neuropsychologcal Assessment Metrics sports medicine battery. *Archives of Clinical Neuropsychology, 22S*, S101-S114.

Cohen, J. (1992). "A power primer". *Psychological Bulletin, 112,* 155–159.

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

Collie, A., Maruff, P., Darby, D., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test–retest intervals. *Journal of the International Neuropsychological Society, 9*, 419-428.

Collie, A., Maruff, P., Makdissi, M., McStephen, M., Darby, D. G., McCrory, P. (2004). Statistical procedures for determining the extent of cognitive change following concussion. *British Journal of Sports Medicine, 38,* 273-278.

Collins, M. W., Lovell, M. R., McKeag, D. B. (1999). Current issues in managing sports-related concussion. *Journal of the American Medical Association, 282*(24), 2285-2283.

Collins, M. W., Stump J, Lovell, M. R. (2004). New developments in the management of sports concussion. *Current Opinion in Orthopaedics, 15*(2), 100-107.

Colorado Medical Society. (1991). Guidelines for the management of concussion in sports (Rev.) Denver (CO): Colorado Medical Society.

Conboy, T. J., Barth, J., Boll, T. J. (1986). Treatment and rehabilitation of mild and moderate head trauma. *Rehabilitation Psychology, 31*(4), 203-215.

Covassin, T., Swanik, C. B., Sachs, M., Kendrick, Z., Schatz, P., Zillmer, E., Kaminaris, C. (2006). Sex differences in baseline neuropsychological function and concussion symptoms of collegiate athletes. *British Journal of Sports Medicine, 40*, 923-927.

Covassin, T., Stearme, D., Elbin, R. (2008). Concussion history and postconcussion neurocognitive performance and symptoms in collegiate athletes. *Journal of Athletic Training, 43*(2), 119-124.

De Kruijk, J. R., Twijnstra, A., Meerhoff, S., & Leffers, P. (2001). Mild traumatic brain injury: lack of consensus in Europe. *Brain Injury, 15*(2), 117-123.

DeVellis, R.F. (2003). *Scale development: Theory and applications*. (2nd ed.) Thousand Oaks, CA: Sage Publications.

Drebing, C. E., Van Ormer, A., & Krebs C. (2005). The impact of enhanced incentives on vocational rehabilitation outcomes for dually diagnosed veterans. *Journal of Applied Behavior, 38*(3), 359-372.

Duff, K., Westervelt, H. J., McCaffrey, R. J., Haase, R. F. (2001). Practice effects, test-retest reliability, and dual baseline assessments with California Verbal Learning Test in an HIV sample. *Archives of Clinical Neuropsychology, 16,* 461-476.

Echemendia, R. J. (2006). *Sports neuropsychology: Assessment and management of traumatic brain injury.* Guilford Press, New York.

Echemendia, R. J., & Cantu, R. C. (2003). Return to play following sports-related mild traumatic brain injury: The role for neuropsychology. *Applied Neuropsychology, 10*(1), 48-55.

Echemendia, R. J., & Julian, L. J. (2001). Mild traumatic brain injury in sports: Neuropsychology's contribution to a developing field. *Neuropsychology Review, 11*(2), 69-88.

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

Erlanger, D. M., Kutner, K. C., Barth, J. T., & Barnes, R. (1999). Neuropsychology of sports-related head injury: Dementia Pugilistica to Post Concussion Syndrome. *The Clinical Neuropsychologist, 13*(2), 193-209.

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

Franzen, M. D. (1989). *Reliability and validity of neuropsychological assessment.* New York: Plenum Press.

Freeman, F. S. (1964). *Theory and practice of psychological testing* (3rd ed.). New York: Holt, Rinehart, and Winston.

Geller, E. S., Paterson, L., Talbott, E. (1982). A behavior analysis of incentive prompts for motivating seat belt use. *Journal of Applied Behavior, 15*(3), 403-415.

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

Goodwin Gerberich, S., Priest, J. D., Boen, J. R., Straub, C. P., & Maxwell, R. E. (1983). Concussion incidencies and severity in secondary school varsity football players. *American Journal of Public Health, 73*(12), 1370-1375.

Grissom, R. J. & Kim, J. J. (2005). *Effect sizes for research: A broad practical approach*. Mahwah, NJ: Lawrence Erlbaum Associates Inc.

Green, P. (2004). Green's Medical Symptom Validity Test (MSVT) for Microsoft Windows. Edmonton, Canada: Green's Publishing.

Gronwall, D. M. (1977). Paced auditory serial-addition task: a measure of recovery from concussion. *Perceptual and Motor Skills, 44*(2), 367–373

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

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

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

Guskiewicz, K. M., Marshall, S. W., Bailes, J., McCrea, M., Harding Jr., H. P., Matthews, A., Mihalik, J. R., & Cantu, R. C. (2007). Recurrent concussions and risk of depression in retired professional football players. *Medicine & Science in Sports & Exercise, 39*(6), 903-909.

Hagenzieker, M. P. (1991). Enforcement or incentives? Promoting safety belt use among military personnel in the Netherlands. *Journal of Applied Behavior, 24*(1), 23-30.

Hall, R. C. W., Hall, R. C. W., & Chapman, M. J. (2005). Definition, diagnosis, and forensic implications of postconcussional syndrome. *Psychosomatics: Journal of Consultation Liaison Psychiatry, 46*(3), 195-202.

Hermann, J. A., de Montes, A. I., Dominguez, B., Montes, F., & Hopkins, B. L. (1973). Effects of bonuses for punctuality on the tardiness of industrial workers. *Journal of Applied Behavior, 6*(4), 563-570.

Hinton-Bayre, A. D., Geffen, G., McFarland, K. (1997). Mild head injury and speed of information of processing: A prospective study of professional rugby league players. *Journal of Clinical and Experimental Neuropsychology, 19*(2), 275-89.

Hootman, J. M., Dick, R., Agel, J. (2007). Epidemiology of collegiate injuries for 15 sports: Summary and recommendations for injury prevention initiatives. *Journal of Athletic Training, 42*(2), 311-319.

Hunt, T. N., Ferrara, M. S., Miller, L. S., Macciocchi, S. (2007). The effect of effort on baseline neuropsychological test scores in high school football athletes. *Archives of Clinical Neuropsychology, 22,* 615–621.

Immediate Postconcussion Assessment and Cognitive Testing (ImPACT) Concussion Management Software version 2.0. ImPACT Applications, Inc., Pittsburgh, PA.

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

Iverson, G. L., Lovell, M. R., & Collins, M. W. (2003b). Reliable change in ImPACT version 2.0 following sports concussion. *Archives of Clinical Neuropsychology,18,* 744.

Iverson, G. L., Lovell, M. R., & Collins, M. W. (2003c). Immediate Post-Concussion Assessment and Cognitive Testing (ImPACT) Normative Data. ImPACT Applications, Inc. Pittsburgh, PA.

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

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

Jacobson, N. S. & Truax, P. (1991). Clinical Significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*(1), 12-19. *British Journal of Sports Medicine, 38,* 273-278.

Kaplan, R. M. & Saccuzzo, D. P. (2009). *Psychological testing: Principles, applications, and issues (7th ed.).* Belmont, CA: Wadsworth, Cengage Learning, Inc.

LaBotz, M., Martin, M. R., Kimura, I. F., Hetzler, R. K., & Nichols, A. W. (2005). A comparison of a preparticipation evaluation history form and a symptom-based concussion survey in the identification of previous head injury in collegiate athletes. *Clinical Journal of Sport Medicine, 15*(2), 73-78.

Lloyd, J. (2009, May 4). High school athletes face serious concussion risks. USA Today. Retrieved April 3, 2010, from http://www.usatoday.com/news/health/2009-05-04-concussions_N.htm

Lovell, M. R. (2006). The ImPACT neuropsychological test battery. In R. J. Echemendia (Ed.), *Sports neuropsychology: Assessment and management of traumatic brain injury* (pp. 193-215). Guilford Press, New York.

Lovell, M. R., Collins, M. W., & Bradley, J. (2004). Return to play following sports-related concussion. *Journal of Clinical Sport Medicine, 23*, 421-441.

Lovell, M. R., Collins, M. W., Iverson, G. L., Field, M., Maroon, J. C., Cantu, R., Podell, K., Powell, J. W., Belza, M., Fu, F. H. (2003). Recovery from mild concussion in high school athletes. *Journal of Neurosurgery, 98*, 295-301.

Lovell, M. R. & Pardini J. E. (2006). New developments n sports concussion management. In S. M. Slobounov & W. J. Sebastianelli (Eds.), *Foundations of Sports-Related Concussion* (pp. 111-136). Springer-Verlag: Norwell, MA.

MacFlynn, G., Montgomery, E. A., Fenton, G. W., Rutherford, W. (1984). Measurement of reaction time following minor head injury. *Journal of Neurology, Neurosurgery, and Psychiatry, 47*, 1326-1331.

Maroon, J. C., Steele, P., Berlin, R. (1980). Football head and neck injuries – an update. *Clinical Neurosurgery, 27*, 414-429.

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

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

McCrory, P. & Berkovic, S. F. (2001). The history of clinical and athophysiological concepts and misconceptions. *Neurology, 57*, 2283-2289.

McCrory, P., Johnston, K., Meeuwisse, W., Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Kelly, J., Lovell, M., Schamasch, P. (2005). Summary and agreement statement of the 2nd International Conference on Concussion in Sport, Prague 2004. *British Journal of Sports Medicine, 39*(Suppl I), i78-i86.

McCrory, P., Meeuwisse, W., Johnston, K., Dvorak, J., Aubry, M., Molloy, M., & Cantu, R. (2009). Consensus statement on concussion in sport: 3rd International Conference on Concussion in Sport, Zurich, 2008. *Clinical Journal of Sports Medicine, 19*(3), 185-200.

Morton, V. & Torgerson, D. J. (2003) Effect of regression to the mean on decision making in health care. *British Medical Journal, 326*, 1083-1084.

Murphy, F. & Simmons, J. C. H. (1959). Initial management of athletic injuries to the head and neck. *American Journal of Surgery*, 98, 379-383.

NBC Sports. (2009, December 29). Concussion warnings too late for N. J. athlete. Retrieved May 9, 2010, from http://nbcsports.msnbc.com/id/34623555/ns/sports-college_football

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C. & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill Inc.

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

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

Petry, N. M., Pierce, J. M., Stitzer, M. L., Blaine, J., Roll, J. M., et al. (2005). Effect of prize-based incentives on outcomes in stimulant abusers in outpatient psychosocial treatment programs. *Archives of Genetic Psychiatry, 62*, 1148-1156.

Phillips, E. L. (1968). Achievement place: Token reinforcement procedures in a home-style rehabilitation setting for “pre-delinquent” boys. *Journal of Applied Behavior, 1*(3), 213-223.

Phillips, E. L., Phillips, E. A., Fixsen, D. L., & Wolf, M. M. (1971). Achievement place: Modification of the pre-delinquent boys within a token economy. *Journal of Applied Behavior, 4*(1), 45-59.

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

Randolph, C., McCrea, M., Barr, W. B. (2005). Is neuropsychological testing useful in the management of sport-related concussion? Journal of Athletic Training, 40(3), 139-154.

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

Rocconi, L. M. & Ethington, C. A. (2009). Assessing longitudinal change: Adjustment for regression to the mean effects. *Research in Higher Education, 50*, 368-376.

Rosenbaum, A. M., Arnett, P. A., Bailey, C. M., & Echemendia, R.J. (2006). Neuropsychological assessment of sports-related concussion: measuring clinically significant change. In S. M. Slobounov & W. J. Sebastianelli (Eds.), *Foundations of Sports-Related Concussion* (pp. 137-169). Springer-Verlag: Norwell, MA.

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

Schlegel, R. E. & Gilliland, K. (2007) Development and quality assurance of computer-based assessment batteries. *Archives of Clinical Neuropsychology, 22S*, S49-S61.

Shaw, N. (2006). Neuropsychology of concussion: Theoretical perspectives. In S. M. Slobounov & W. J. Sebastianelli (Eds.), *Foundations of Sports-Related Concussion* (pp. 19-45). Springer-Verlag: Norwell, MA.

Stevens, A., Friedel, E., Mehren, G., & Merten, T. (2008). Malingering and uncooperativeness in psychiatric and psychological assessment: Prevalence and effects in a German sample of claimants. *Psychiatry Research, 157,* 191-200.

Stuss, D. T., Stethem, L. L., Hugenholtz, H., Picton, P., Pivic, J., Richard, M. D. (1989). Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance. *Journal of Neurology, Neurosurgery, and Psychiatry, 52*, 742-748.

Symonds, C. (1962). Concussion and its sequelae. *The Lancet*, *i*, 1-5.

The Huffington Post. (2010, May 11). Brady Lee Frazier DEAD: 8th grader dies in baseball scrimmage. Retrieved May 11, 2010, from http://www.huffingtonpost.com/2010/05/11/brady-lee-frazier-dead-di_n_571653.html

The Huffington Post. (2010, May 10). Dylan Steigers DEAD: Eastern Oregon football player dies from practice head injury. Retrieved May 10, 2010, from http://www.huffingtonpost.com/2010/05/10/dylan-steigers-dead-dies-_n_570092.html

The San Francisco Examiner. (2009, November 18). Quotes from players interviewed recently by the AP about concussions. Retrieved March 20, 2010, from http://www.sfexaminer.com/nation/quotes-from-players-interviewed-recently-by-the-ap-about-concussions-70402492.html

Valovich McLeod, T. C., Bay, C., Heil, J., and McVeigh, S. D. (2008). Identification of sport and recreational activity concussion history through the preparticipation screening and a symptom survey in young athletes. *Clinical Journal of Sport Medicine, 18*(3), 235-240.

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

Youngjohn, J. R., Burrows, L., Erdal, K. (1995). Brain damage or compensation neurosis? The controversial post-concussion syndrome. *The Clinical Neuropsychologist, 9*(2), 112-123.

Zillmer, E. A., Schneider, J., Tinker, J., & Kaminaris, C. I. (2006). A history of sports-related concussions. In R. J. Echemendia (ed.) *Sports neuropsychology: Assessment and management of traumatic brain injury* (pp. 17-35). Guilford Press, New York.

Appendix A

Table 1. Comparison of grading systems (as cited in Cantu, 2001)

	Grade I	Grade II	Grade III
Cantu grading system (Cantu, R. 1986)	Post-traumatic amnesia less than 30 minutes, no LOC	LOC less than 5 minutes OR amnesia lasting 30 minutes - 24 hours	LOC more than 5 minutes OR amnesia more than 24 hours
AAN grading system (1997)	transient confusion, concussion symptoms resolve in 15 minutes, no LOC	transient confusion, concussion symptoms last more than 15 minutes, no LOC	Any LOC (brief or prolonged)
Colorado Medical Society (1991)	Confusion without amnesia, no LOC	Confusion, post-traumatic amnesia, no LOC	LOC

Table 2. Descriptive information (mean and SD) of age at baseline and at PC1 testing (N = 22)

	Female athletes (n = 6)	Male athletes (n = 16)
BL age range	18-20	17-21
BL Mean	19	18.87
BL SD	0.89	1.03
PC1 age range	19-21	18-22
PC1 Mean	19.50	19.56
PC1 SD	0.84	1.37

Table 3. Time post between date of last concussion and testing sessions

	time post date of last concussion & PC1	time post date of last concussion & last PC
Mean	3.55 days	12.2 days
SD	3.05 days	8.6 days
range	0 – 12 days	0 days – 43 days (6 weeks)

Table 4. Time post between PC1 and last PC assessment

	time post between PC1 & last PC	time post between BL & last PC
Mean	8.6 days	8.45 months
SD	9 days	6.5 months
range	0 days - 42 days (6 weeks)	1 week - 21 months

Table 5. Mean and Median values for each individual composite score at baseline and first post-concussion testing

ImPACT Composite	Median		Mean		*p*
Scores	BL	PC1	BL	PC1	
Verbal Memory	85.00	87.00	84.00	81.73	0.66
Visual Memory	74.50	69.00	72.23	67.55	0.19
Visual Process. Speed	37.975	39.00	38.42	35.80	0.66
Reaction Time	0.5600	0.6000	0.5650	0.7105	0.02*
Total Symptom Scores	2.00	16.50	4.86	22.45	0.001*

Table 6. Mean and Median values for each individual composite score at baseline and last post-concussion testing

ImPACT Composite	Median		Mean		
Scores	BL	last PC	BL	last PC	*p*
Verbal Memory	85.00	89.50	84.00	87.59	0.22
Visual Memory	74.50	81.50	72.23	79.23	0.021*
Visual Process. Speed	37.975	40.44	38.42	40.17	0.11
Reaction Time	0.5600	0.5300	0.5650	0.5614	0.36
Total Symptom Scores	2.00	3.00	4.86	3.73	0.57

Table 7. Mean and Median values for each individual composite score across all testing sessions

ImPACT Composite	Median			Mean		
Scores	BL	PC1	last PC	BL	PC1	last PC
Verbal Memory	85.00	87.00	89.50	84.00	81.73	87.59
Visual Memory	74.50	69.00	81.50	72.23	67.55	79.23
Visual Process. Speed	37.975	39.00	40.44	38.42	35.80	40.17
Reaction Time	0.5600	0.6000	0.5300	0.5650	0.7105	0.5614
Total Symptom Scores	2.00	16.50	3.00	4.86	22.45	3.73

Table 8. Descriptive statistics of the participants: Mean (SD), and range.

	Male	Female
Age	21.67 (2.09)	21.11 (1.18)
range	18 – 25	19 – 24
Years of education	13.80 (1.52)	14.33 (1.19)
range	12 – 17	13 – 17

Table 9. Descriptive statistics for ImPACT Total Symptom Scores by gender using ImPACT Normative Data (N = 33).

	Male (n = 15)	Normative data for college men (range: 0-5)	Female (n = 18)	Normative data for college women (range: 0-10)
Test 1	4.20 (5.85)	normal	4.33 (3.87)	normal
Test 2	2.80 (4.69)	normal	4.17 (5.35)	normal
Test 3	3.53 (4.87)	normal	3.22 (4.61)	normal
Test 4	1.80 (2.93)	normal	2.89 (4.46)	normal

Table 10. Means (standard deviations) for ImPACT composite scores and for the adjusted scores on test session 1.

Dependent variable	Test 1	Adjusted scores Test 1
Verbal memory composite	88.67 (8.21)	88.67 (1.57)
Visual memory composite	76.82 (11.41)	76.82 (4.21)
Visual processing speed composite	36.63 (4.51)	36.63 (3.43)
Reaction time composite	0.606 (0.069)	0.606 (0.042)
Cognitive efficiency score	0.38 (0.12)	0.38 (0.038)

Table 11. Average ImPACT composite scores with standard deviation (SD) at Test 1, 2, 3*, and 4 (N = 33).

Variable	Test 1 adjusted scores	Test 2	Test 3*	Test 4
Verbal memory composite	88.67 (1.57)	93.42 (5.27)	93.58 (4.22)	93.06 (6.52)
Visual memory composite	76.82 (4.21)	78.48 (10.36)	77.82 (9.89)	82.00 (11.33)
Visual processing speed composite	36.63 (3.43)	38.91 (4.85)	38.35 (4.46)	39.46 (5.55)
Reaction time composite	0.606 (0.042)	0.578 (0.062)	0.575 (0.069)	0.555 (0.064)
Cognitive efficiency score	0.38 (0.038)	0.44 (0.10)	0.44 (0.09)	0.43 (0.10)

Table 12. Means and standard deviations of performance on the ImPACT battery by gender.

ImPACT composite scores	Mean (SD)		*p*	Normative data	
				College men	College women
	Males (n = 15)	Females (n = 18)			
Verbal memory					
Test 1	89.27 (8.37)	88.17 (8.28)	n.s.	average	average
Test 2	94.20 (5.43)	92.78 (5.20)	n.s.	high average	average
Test 3	93.33 (3.85)	93.78 (4.61)	n.s.	average	average
Test 4	94.13 (4.81)	92.17 (7.67)	n.s.	high average	average
Visual memory					
Test 1	81.47 (9.46)	72.94 (11.69)	$p = 0.043$	average	average
Test 2	84.23 (6.70)	73.67 (10.53)	$p = 0.004$	average	average
Test 3	81.07 (8.28)	75.11 (10.52)	n.s.	average	average
Test 4	85.40 (11.19)	79.17 (10.95)	n.s.	average	average
Visual processing speed					
Test 1	36.24 (4.51)	36.95 (4.61)	n.s.	average	average
Test 2	39.28 (5.19)	38.61 (4.68)	n.s.	average	average
Test 3	37.98 (3.96)	38.66 (4.93)	n.s.	average	average
Test 4	40.77 (5.47)	38.37 (5.54)	n.s.	average	average
Reaction time					
Test 1	0.595 (0.041)	0.616 (0.086)	n.s.	average	low average
Test 2	0.588 (0.052)	0.569 (0.069)	n.s.	average	average
Test 3	0.589 (0.072)	0.563 (0.066)	n.s.	average	average
Test 4	0.568 (0.065)	0.544 (0.062)	n.s.	average	average
Cognitive efficiency					
Test 1	0.37 (0.09)	0.38 (0.13)	n.s.	above average	above average
Test 2	0.42 (0.13)	0.46 (0.08)	n.s.	above average	above average
Test 3	0.41 (0.11)	0.46 (0.07)	n.s.	above average	above average
Test 4	0.40 (0.11)	0.46 (0.09)	n.s.	above average	above average

Table 13. Test-retest Pearson correlations for the ImPACT battery.

	Verbal memory	Visual memory	Visual processing	Reaction time	CEI
test 1 & test 2	0.19	0.37	0.76	0.60	0.33
test 2 & test 3	0.49	0.45	0.69	0.75	0.55
test 3 & test 4	0.65	0.05	0.60	0.88	0.62

Table 14. Standard error of measurement across test sessions

	Test 1	Test 2	Test 3	Test 4
Verbal memory	7.39	4.73	3.58	5.61
Visual memory	9.06	8.22	8.68	8.99
Visual processing speed	2.21	2.38	2.50	3.76
Reaction time	0.04	0.04	0.05	0.04
CEI	0.10	0.09	0.08	0.09

Table 15. Comparison of test-retest reliability values on the ImPACT battery.

	Iverson et al. (2003)	present study	
	test 1 & test 2	test 1 & 2	test 3 & 4
Verbal memory	0.70	0.19	0.65
Visual memory	0.67	0.37	0.05
Visual processing speed	0.79	0.76	0.60
Reaction time	0.86	0.60	0.88
Cognitive efficiency index	—	0.33	0.62

Note: Cognitive Efficiency Index at the time of the Iverson et al. study was not developed.

Appendix B

Figure 1. Median performance on Verbal Memory across testing sessions in concussed collegiate athletes

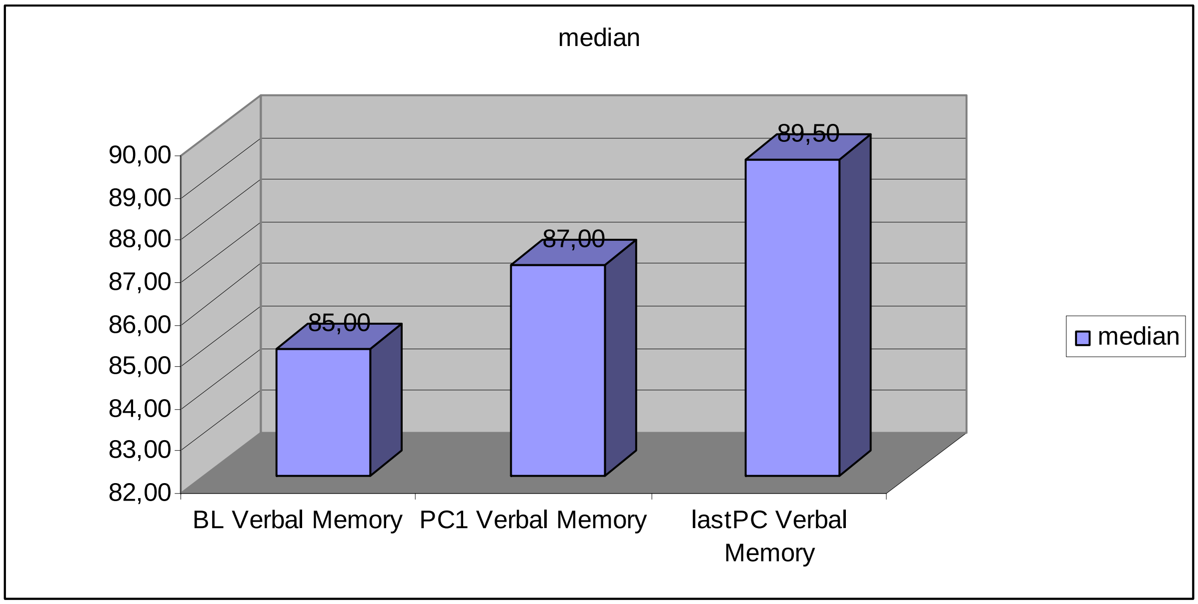


Figure 2. Median performance on Visual Processing Speed in concussed collegiate athletes

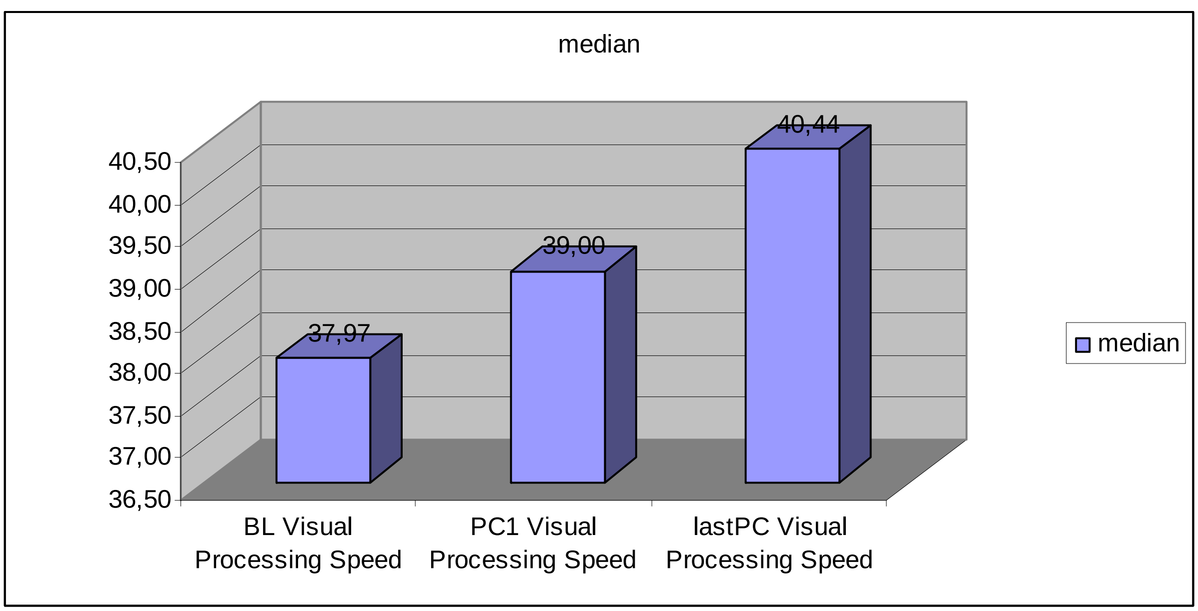


Figure 3. Median performance on Reaction Time across testing sessions in concussed collegiate athletes

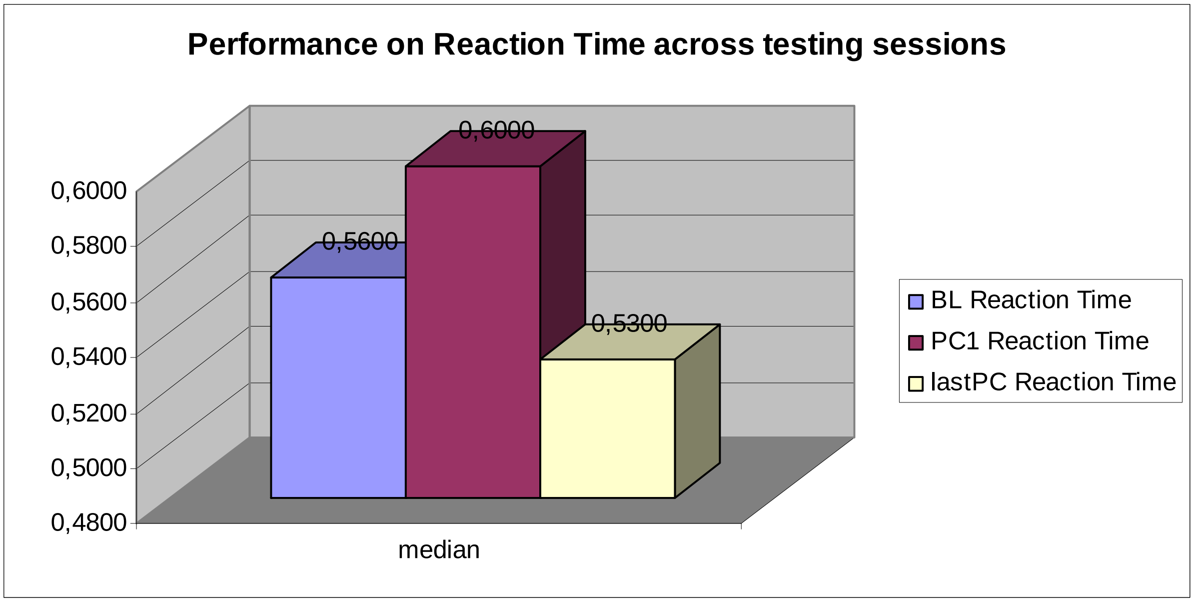


Figure 4. Self-reported Total Symptom Scores across testing sessions in concussed collegiate athletes

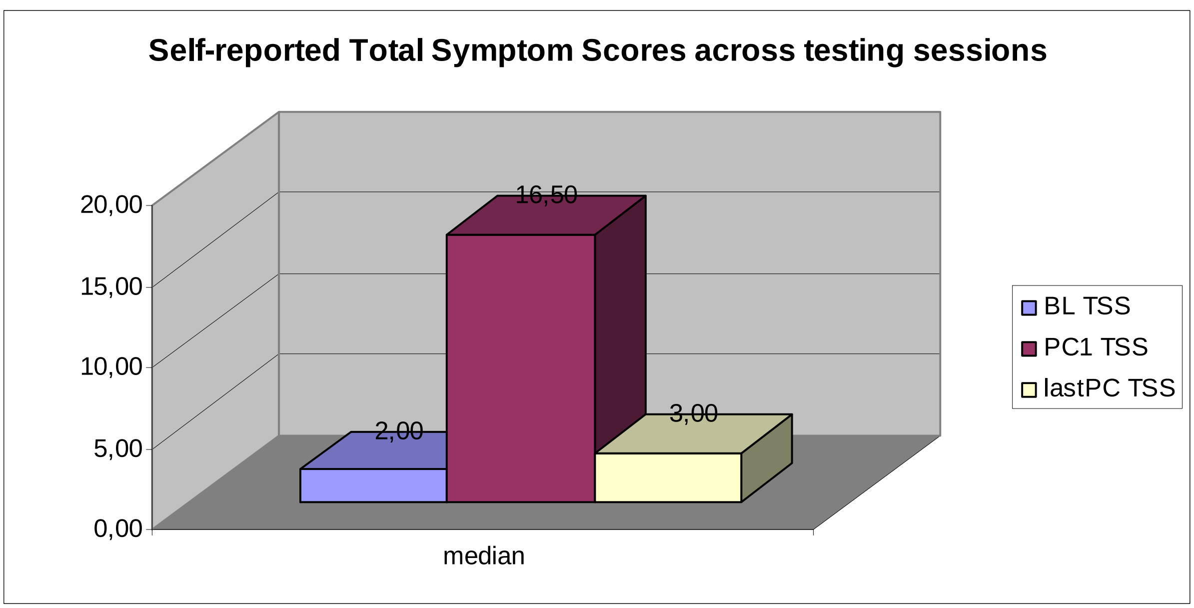


Figure 5. Median performance on Verbal and Visual Memory across testing sessions in concussed collegiate athletes

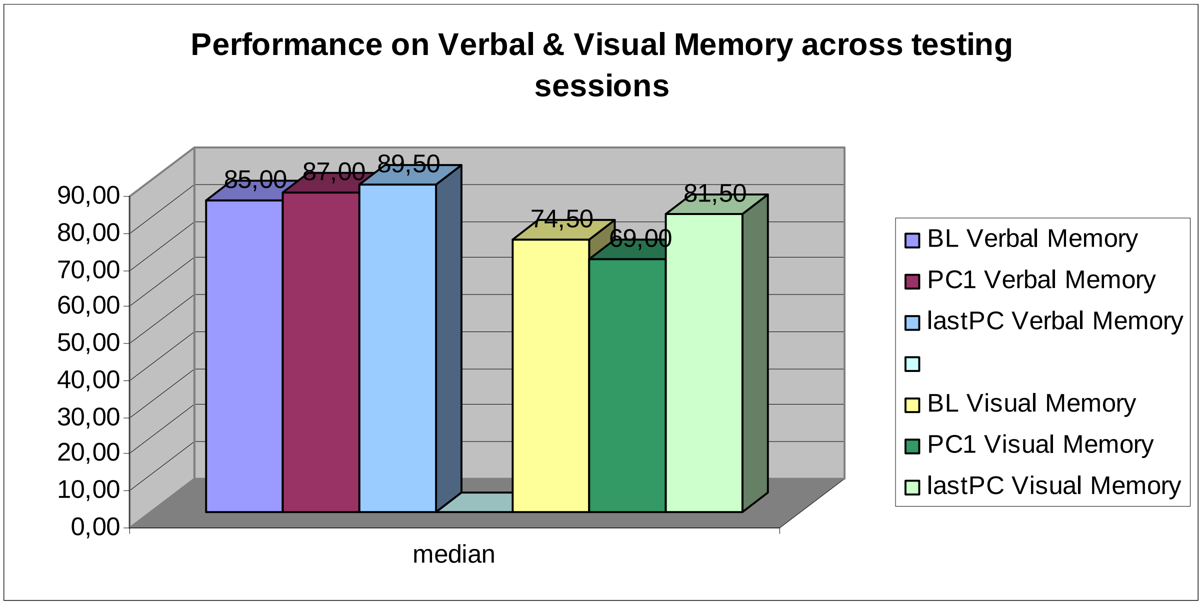
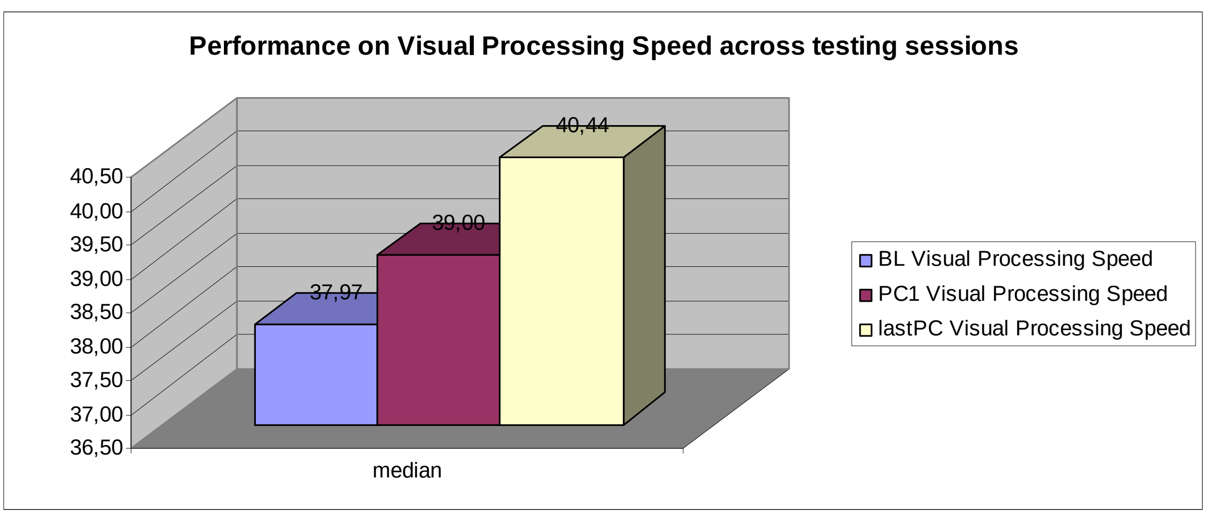
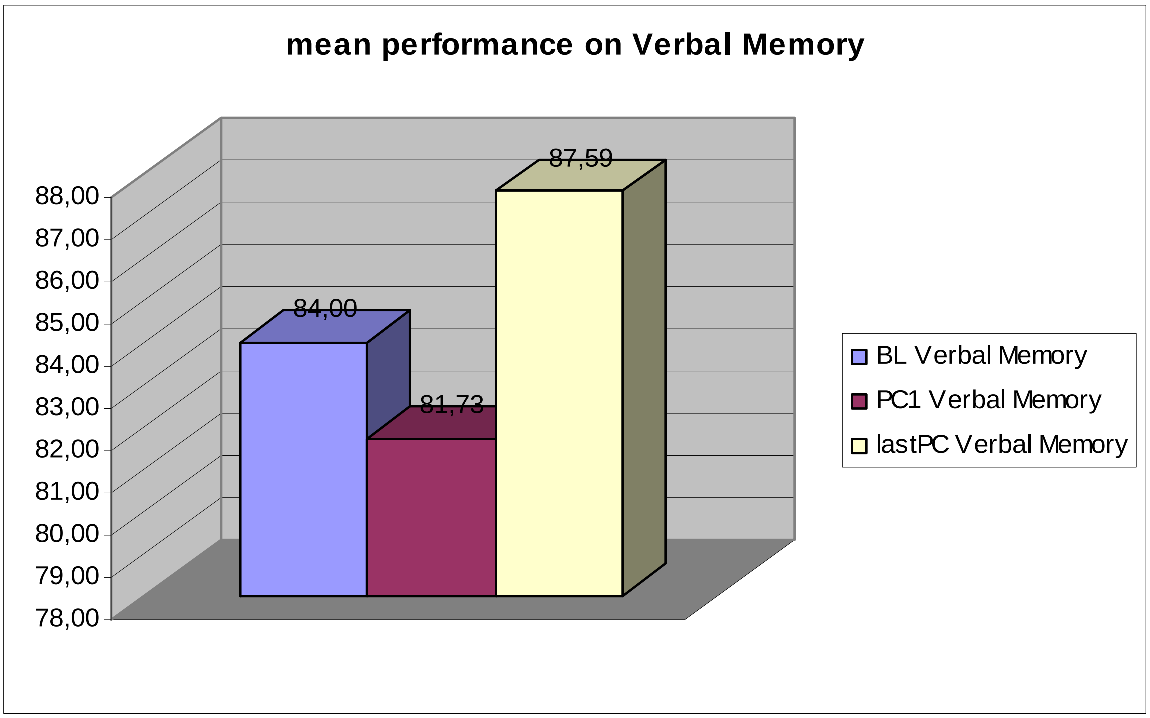
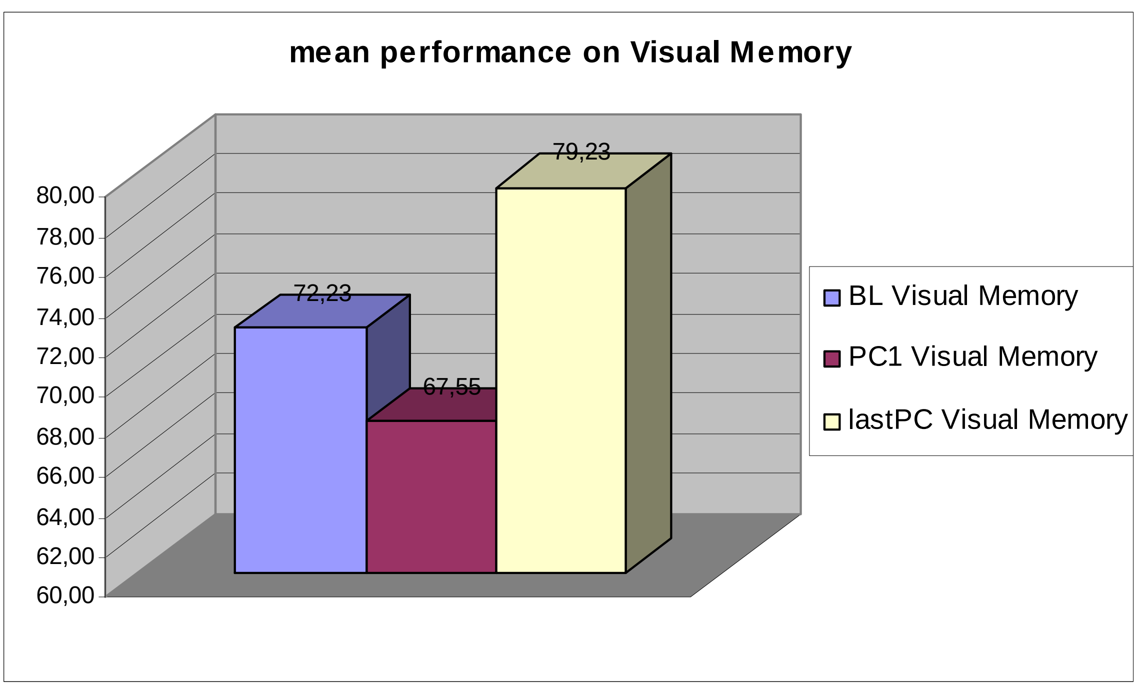


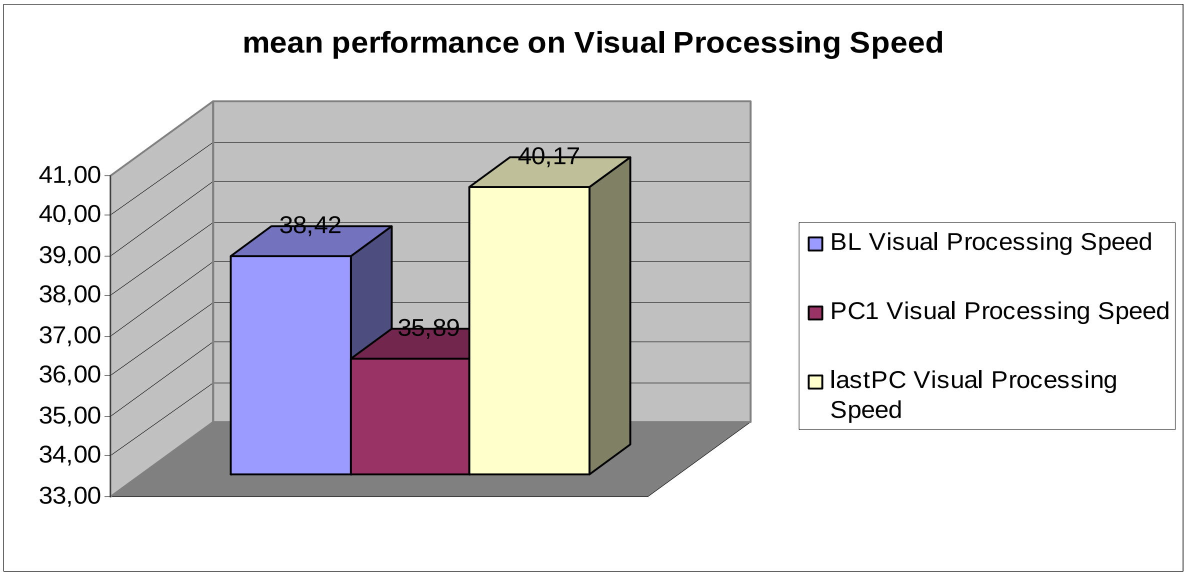
Figure 6. Performance on Visual Processing Speed across testing sessions in concussed collegiate athletes



Figures 7-10. Mean performance on each cognitive measure on the ImPACT battery across testing sessions






mean performance on Visual Processing Speed
41,00
40,00
39,00
38,00
37,00
36,00
35,00
34,00
33,00
38,42
35,89
40,17
BL Visual Processing Speed
PC1 Visual Processing Speed
lastPC Visual Processing Speed

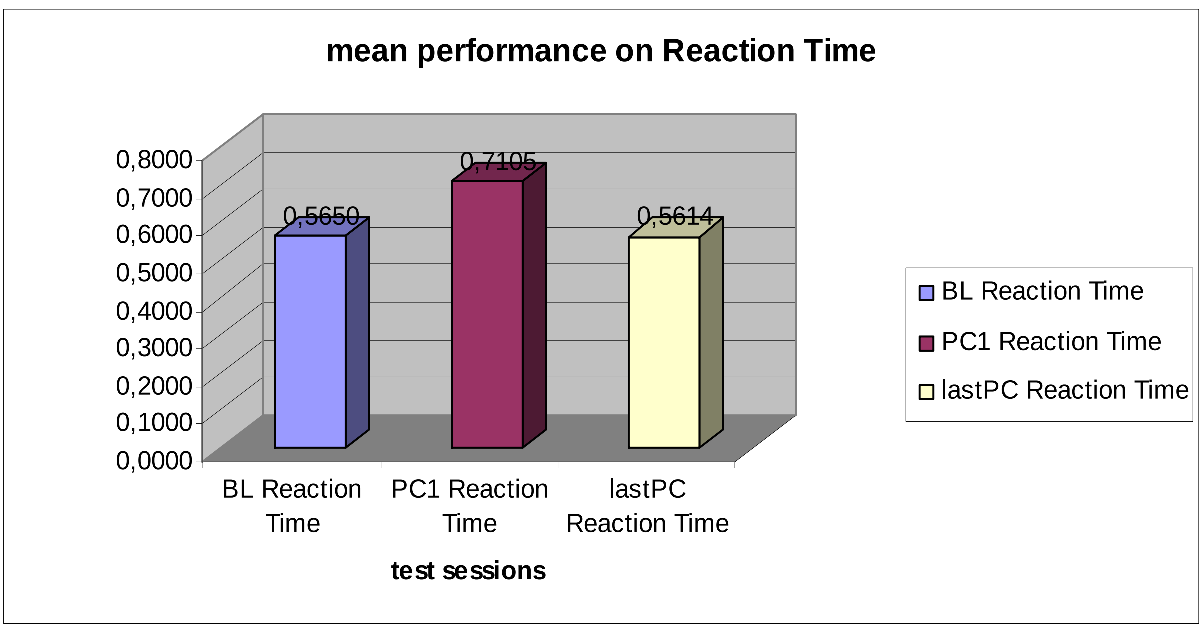

mean performance on Reaction Time
0,8000
0,7000
0,6000
0,5000
0,4000
0,3000
0,2000
0,1000
0,0000
0,5650
0,7105
0,5614
BL Reaction Time
PC1 Reaction Time
lastPC Reaction Time
test sessions
BL Reaction Time
PC1 Reaction Time
lastPC Reaction Time

Figure 11. Mean performance on the ImPACT verbal memory composite score across the four test sessions in non-concussed college students. A higher score represents better performance on the battery.

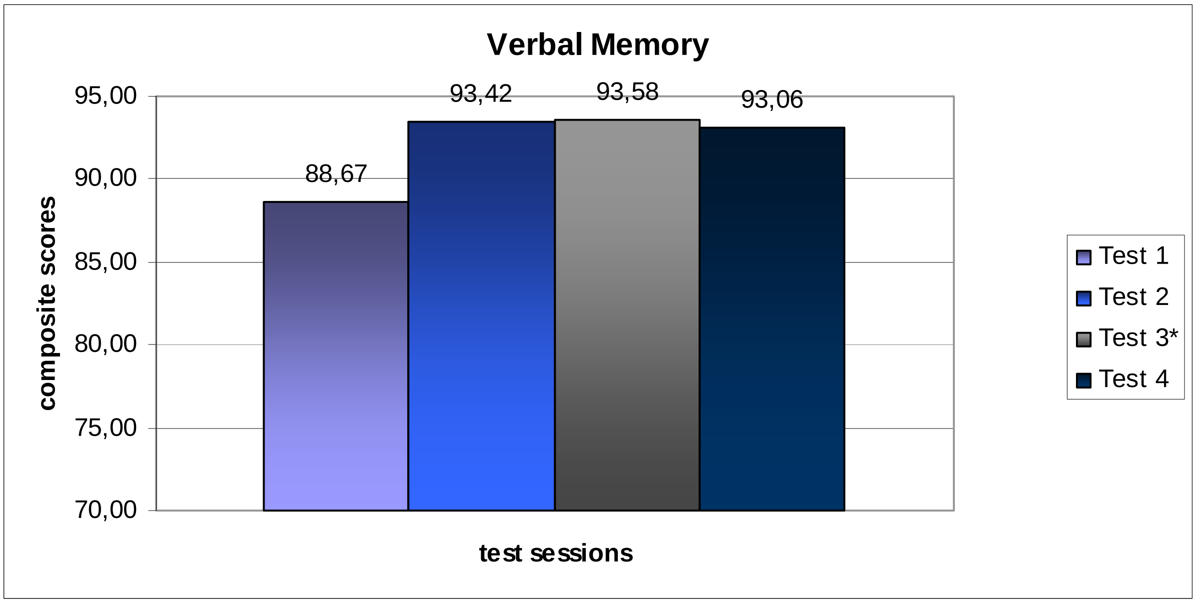


Figure 12. Mean performance on the ImPACT visual memory composite score across the four test sessions in non-concussed college students. A higher score represents better performance on the battery.

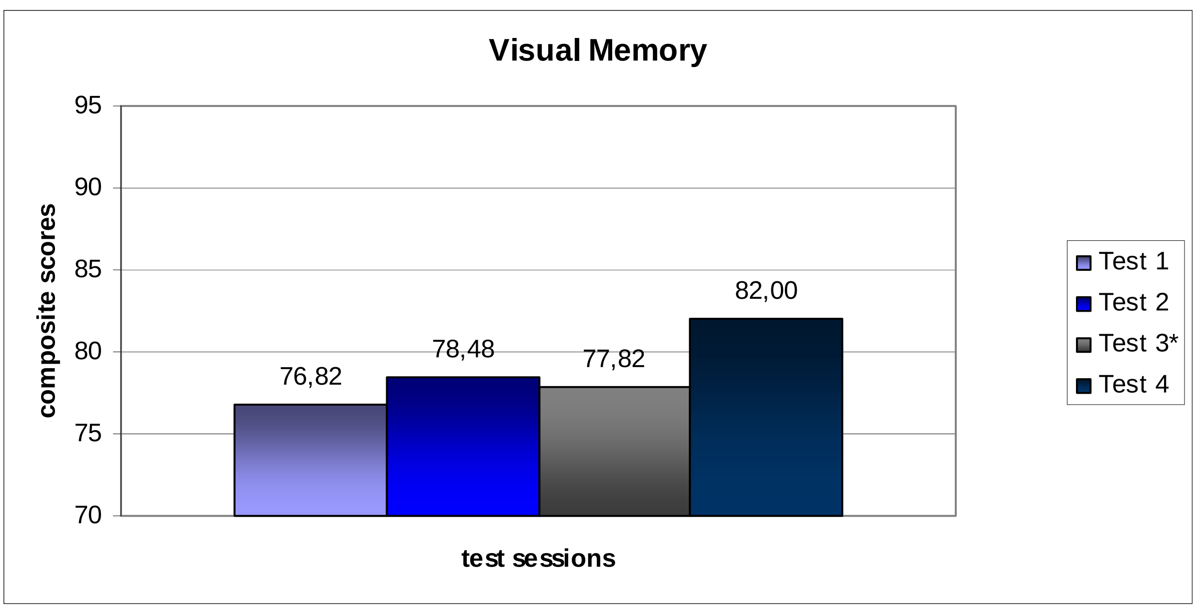


Figure 13. Mean performance on the ImPACT visual processing speed composite score across the four test sessions in non-concussed college students. A higher score represents better performance on the battery.

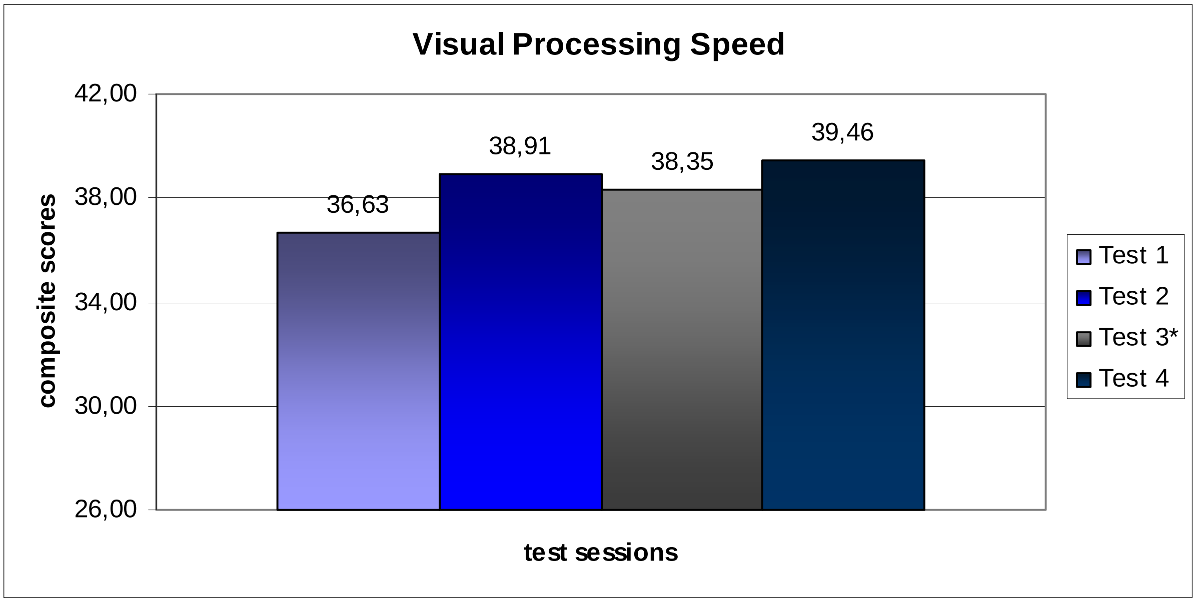


Figure 14. Mean performance on the ImPACT reaction time composite score across the four test sessions in non-concussed college students. A lower score represents better performance on the battery.

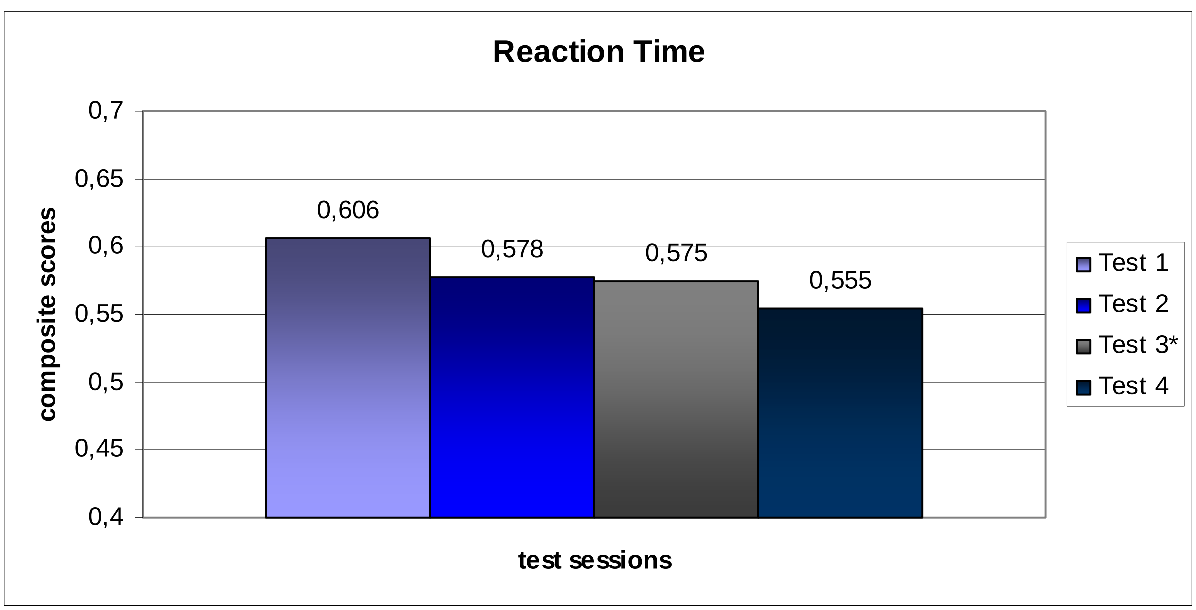


Figure 15. Mean performance on the ImPACT cognitive efficiency index score across the four test sessions in non-concussed college students. A higher score represents better performance on the battery.

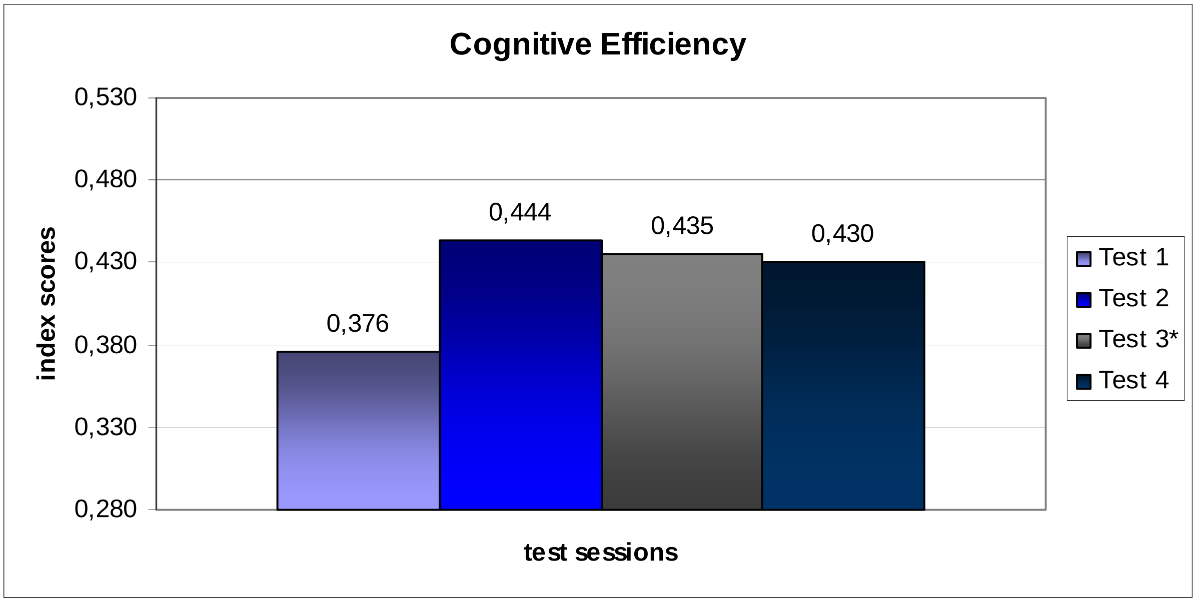


Figure 16. Trend line in verbal memory performance across test sessions in non-concussed college students (N = 33). A higher score indicates better performance.

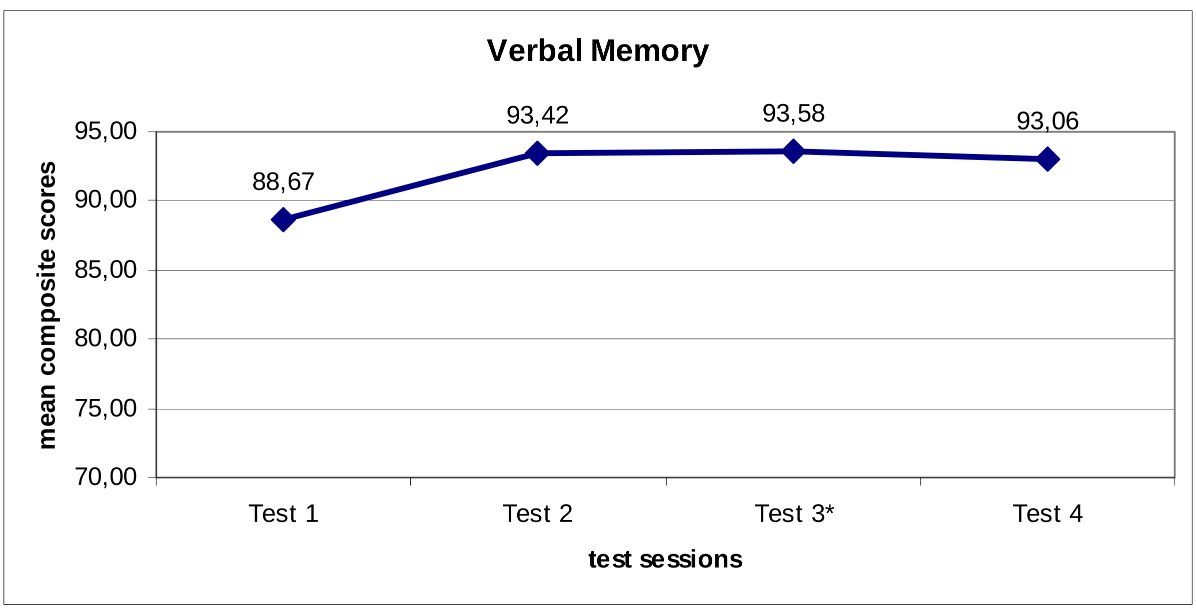


Figure 17. Trend line in visual memory performance across test sessions in non-concussed college students (N = 33). A higher score indicates better performance.

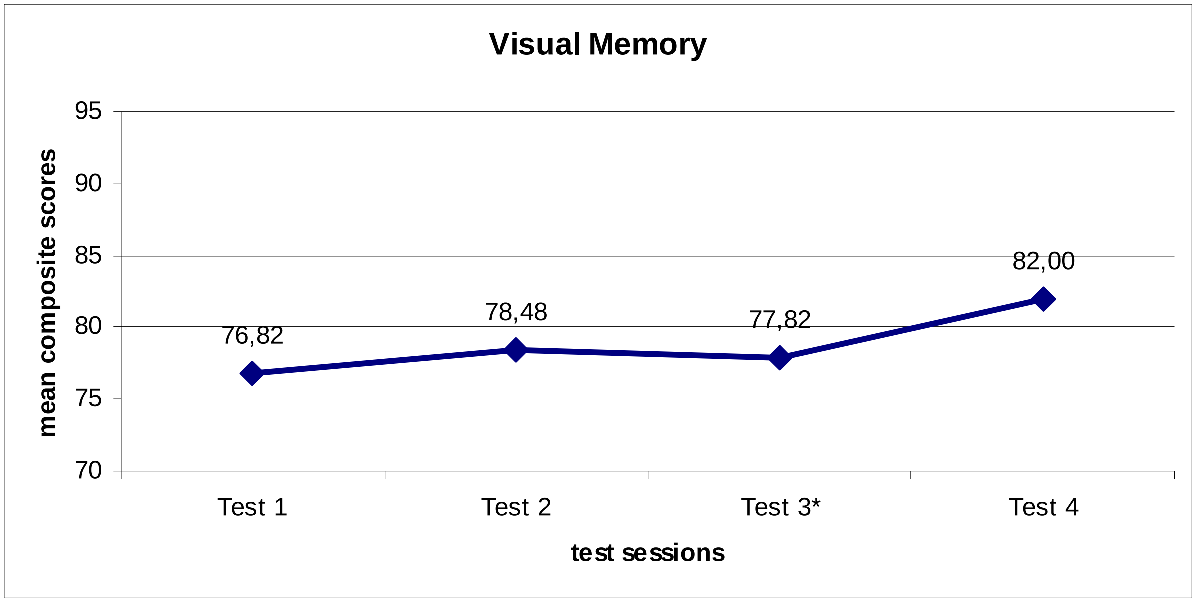


Figure 18. Trend line in visual processing speed performance across test sessions in non-concussed college students (N = 33). A higher score indicates faster processing speed performance.

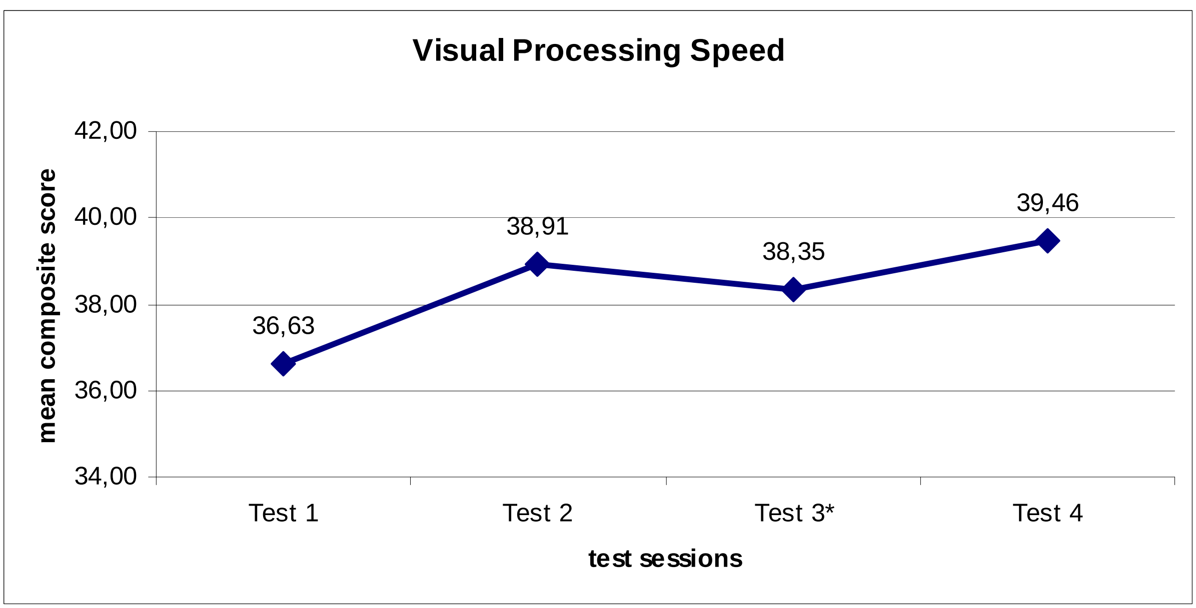


Figure 19. Trend line in reaction time performance across test sessions in non-concussed college students (N = 33). A lower score indicates better (i.e., faster) performance.

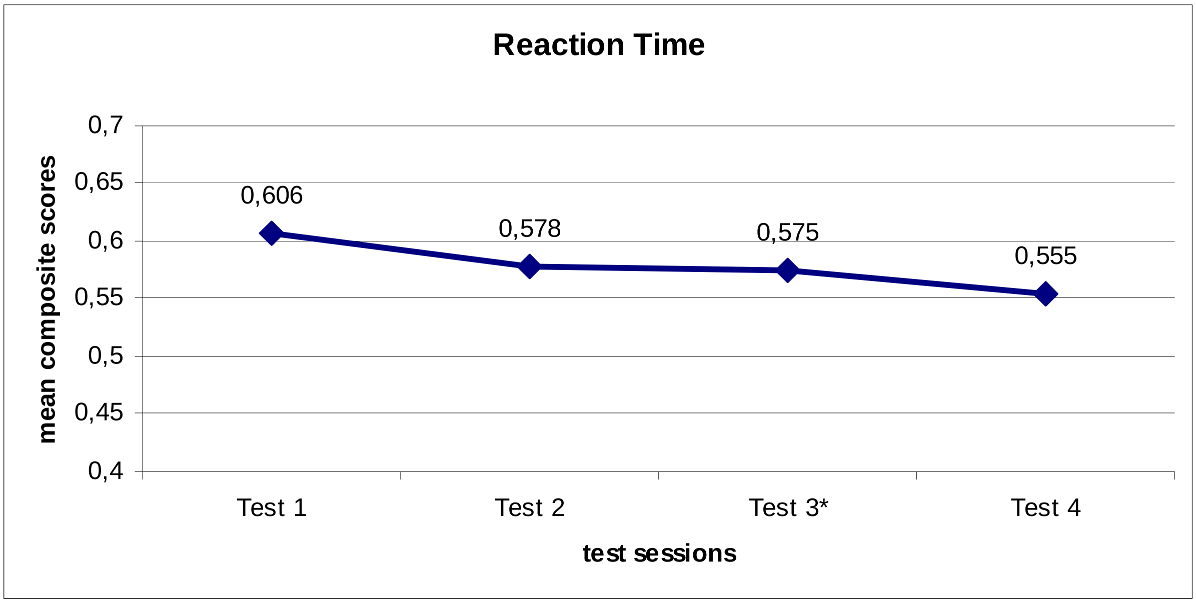
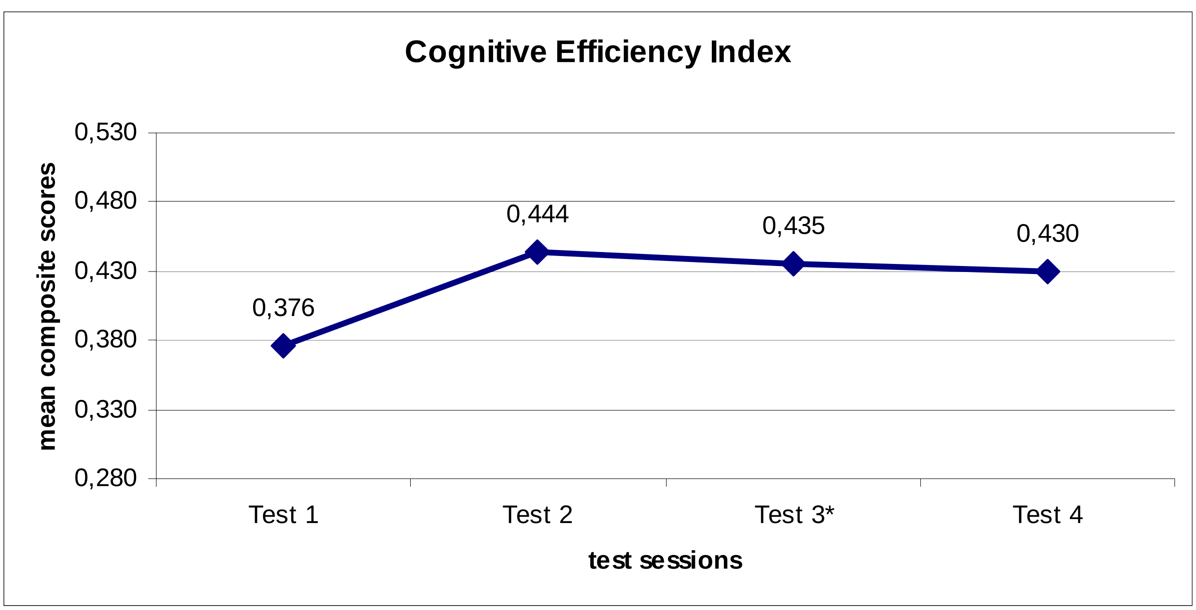


Figure 20. Trend line in cognitive efficiency index score across test sessions in non-concussed college students (N = 33). A higher score indicates better performance.



Edina R. Bene

Interdisciplinary Health Sciences

Edina Bene earned her Master of Arts degrees in English and Hungarian Literature and Linguistics in 2005 at the Pazmany Peter Catholic University in Hungary. In 2007 she joined the Interdisciplinary Health Sciences doctoral program at the University of Texas at El Paso as an international Ph.D. student. Her special field within the doctoral program was speech-language pathology and sports-related concussion research. Her dissertation chair and advisor was Dr. Anthony P. Salvatore.

Dr. Bene has been the recipient of several awards and scholarships, such as the University of Texas at El Paso Graduate School Research Award and the national Donna Fox Scholarship for doctoral students in relation to Speech-Language Pathology. She was also a recipient of the Dodson Dissertation Fellowship in 2011.

While pursuing her doctoral degree, Dr. Bene worked as a research associate for the Concussion Management Clinic within the department of Rehabilitation Science. She also worked as a dissertation fellow at the University of Memphis School of Audiology and Speech-Language Pathology.

Dr. Bene has worked on numerous research projects that were presented at international conferences such as the Annual World Congress on Brain Injury, Guidelines for Management of the Sports Related Concussion Conference, and the International Conference on Motor Speech.

Permanent address: 230 Arbor Commons Circle, Apt. 303
Memphis, TN 38120, USA

This dissertation was typed by Edina R. Bene.