2009

Interrater Reliability Of Psychomotor Skill Assessment In Athletic Training

Jason Craddock
University of Central Florida

INTERRATER RELIABILITY OF PSYCHOMOTOR SKILL ASSESSMENT IN ATHLETIC TRAINING

by

JASON C. CRADDOCK
B.S. Free Will Baptist Bible College, 1996
M.S. Middle Tennessee State University, 1998

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Education
in the Department of Educational Studies
in the College of Education
at the University of Central Florida
Orlando, Florida

Spring Term
2009

Major Professors:
David Boote
Debbie Hahs-Vaughn

ABSTRACT

Assessment in athletic training education is an evolutionary process that is determined by each individual Athletic Training Education Program. The autonomy authorized by national accreditation standards allows academic programs to determine the appropriate assessment practices that facilitate the meeting of student learning outcomes. Even with autonomy, formative and summative techniques are to be employed in both the didactic and clinical arenas of athletic training education programs.

The major objective of athletic training education is to prepare students for entry-level practice in athletic training. The purpose of this study was to assess interrater reliability of athletic training faculty and approved clinical instructors in their rating of athletic training student performance on four psychomotor skills. A total of 115 individuals participated in this study. Thirty two faculty and 83 approved clinical instructors completed the online survey

The results of this study indicate that the overall reliability was high for the entire population as well as the subgroups analyzed. Even though the overall reliability was high, three specific criteria out of a total of 29 criteria had lower reliability scores. These findings may indicate that there may be a high degree of agreement between academic faculty and approved clinical instructors in the rating of athletic training students in their performance of psychomotor skills.

To my family; my lovely wife Miranda, and my three incredible boys; Ethan, Bryce, and Lucas, without your inspiration, prayers, dedication, and motivation this project would not have been completed.

ACKNOWLEDGMENTS

I would like to express my utmost gratitude to the members of my dissertation committee: Dr David Boote, Dr. Debbie Hahs-Vaughn, Dr. Dennis Hunt, Dr. James Lynch, Dr. Nancy Cummings, and Dr. Tammy Boyd. Their expertise and guidance were essential in the completion of this dissertation.

I wish to extend a special note to Dr. Boote and Dr. Hahs-Vaughn for their words of encouragement and advice throughout my course of study and in the development of this dissertation.

Special words of appreciation go to Dr. Hunt and Dr. Cassani for their many words of encouragement as well as brainstorming sessions throughout this project. The time and effort was invaluable. My many thanks.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER ONE
INTRODUCTION

Studies in athletic training education have examined different facets of the educational processes and practices; however there has not been as much information related to the evaluation methods used in athletic training education (Gould & Caswell, 2006). Little analytic attention has been paid to the reliability of assessment of athletic training students (ATS) (Butterwick, Paskevich, Lagumen, Vallevand & Lafave, 2006, Gould & Caswell, 2006). Athletic training education encompasses two main processes: didactic and clinical education (NATA, 2006a). The focus of this research was to examine the interrater reliability of the assessment practices in athletic training clinical education, specifically; psychomotor skill assessment conducted by approved clinical instructors in clinical education and athletic training academic faculty.

The growth and development of athletic training education is closely related to the evolution of the National Athletic Trainers' Association (NATA) (Delforge & Behnke, 1999). Early athletic training education curriculum prepared the way for the development and implementation of specific educational competencies and accreditation standards. The influence of accreditation standards on assessment of ATS and the impact of these standards on athletic training faculty and athletic training approved clinical instructors are examined in this chapter. Also found in this chapter are the goals/scope of the study and the definitions regarding athletic training terms. This chapter concludes with an overview of the remaining chapters.

Background

The first model in athletic training education originated in the 1950s with efforts to establish professional standards (Delforge & Behnke, 1999). The goal of the curriculum development at that time was to prepare ATS for employment in the secondary school setting, preferably in the subject areas of physical education or health. Athletic training students were also prepared to meet entry requirements for preparation into physical therapy schools. The first formal athletic training education program (ATEP) was approved by the NATA in 1969. Thus, the formal evaluative process for assessing an ATEP for approval was also initialized at this time (Delforge & Behnke).

Changes in the athletic training curricular process occurred in the 1970s and 1980s. Athletic training students were no longer required to obtain a teaching certificate in physical education or health, thus allowing the students to determine their own academic progress within athletic training education at that time. Delforge and Behnke stated that the revision included progress toward specific core content for athletic training curriculum. They also assert that within the core content, clinical experiences and learning outcomes were developed in order to enhance the overall learning of ATS (1999).

In 1983, athletic training education was again modified with the approval of the parameters for the development of undergraduate athletic training education programs. These guidelines assisted academic programs in the development of specific subject matter rather than specific courses (NATA, 1983). In addition to these guidelines, the first educational

competencies were formulated. The *Competencies in Athletic Training* (NATA, 1983) were a collection of measureable learning objectives to be used in the development of specific didactic and clinical education. This collection of educational competencies was the first published set of educational requirements for athletic training education programs. Since then, the educational competencies have been modified to better reflect the contemporary athletic training scope of practice. These educational competencies are the knowledge and skills that are necessary for the entry-level practice as a certified athletic trainer (NATA, 2006b).

The next milestone in the development of athletic training education was the recognition of athletic training as an allied health profession from the American Medical Association (AMA). Once athletic training was acknowledged as an allied health profession by the AMA, efforts to create a mechanism for accreditation of entry-level programs by the Committee on Allied Health Education and Accreditation (CAHEA) were initiated. In 1991, the Joint Review Committee on Educational Programs in Athletic Training (JRC-AT) was created. The JRC-AT utilized partners from various other allied health professions for this initiative (CAATE, 2008). Previous approval documents for athletic training education programs were used to create new accreditation standards entitled, *Essentials and Guidelines for an Accredited Educational Program for the Athletic Trainer.* The Commission on Accreditation of Allied Health Education Programs (CAAHEP) was established in 1992 after CAHEA was dissolved. All the programs that were approved under the old criteria were then required to be reviewed utilizing the new criteria. In 2006, the JRC-AT became independent from CAAHEP and changed its name to what is now the Committee on Accreditation of Athletic Training Education (CAATE) (CAATE,

2008). The CAATE is the current accrediting body for all entry-level athletic training education programs.

The CAATE, under the sponsorship of the NATA and other healthcare organizations, developed the accreditation guidelines for ATEPs to follow. These standards ensure that all academic programs adhere to policy and procedures that enable ATS to become entry-level practitioners (CAATE, 2005).

Concurrent with the development of standards and competencies in athletic training education, a certification process was initiated to certify graduating students as athletic trainers. Athletic training students who successfully completed an accredited academic program became eligible for a certifying examination. The National Athletic Trainers' Association Board of Certification (BOC) was established in 1970 to develop the certification process. Two routes to certification were possible: an internship route and graduation from an accredited entry-level athletic training education program. In the internship route, a student had to fulfill "1500 clinical hours under the direct supervision of a NATA – certified athletic trainer" (Prentice, 2000 p. 29). At the time of application for the certification exam, the student also had to present an official transcript indicating evidence of the completion of coursework in the following areas: health, human anatomy, kinesiology, human physiology, exercise physiology, basic, and advanced athletic training. Upon graduation and the completion of the internship hours requirement, the student was eligible to take the BOC certification exam. Alternately, a student graduating from an accredited entry-level athletic training program had to complete more credit hours in an athletic training curriculum with a reduced amount of clinical hours (800 minimum in a two-year

period). The coursework for an accredited program incorporated the same courses that an internship program contained, and also included the following: prevention and evaluation of athletic injury/illness, first aid and emergency care, therapeutic modalities, therapeutic exercise, nutrition, and athletic training administration (Delforge & Behnke, 1999).

In 2004, the BOC no longer recognized internship candidates as eligible for the certification examination, and the minimum clinical hour requirement was removed for students in accredited programs. Students who graduated from an accredited athletic training program were the only ATS eligible for the certification examination. These changes reduced the amount of contact/practice time for ATS, thus placing more emphasis on assessment of psychomotor skill development in laboratory settings, and on the abilities of ATEP faculty and approved clinical instructors (ACIs) to assess student skills (Prentice, 2006).

Prior to 2007, the BOC examination consisted of three sections: written multiple choice, written simulation, and a practical skill completion section. In May/June of 2007, the three-part exam was replaced by a computer-based examination (CBE), utilizing a multiple choice section along with a computer simulation hybrid section (combination of practical skills and written simulation sections) to assess clinical competency and proficiency (Castle Worldwide, 2008). This change eliminated the face-to-face standardized practical skills assessment and placed further emphasis on the assessment of student's psychomotor abilities by ATEP faculty and ACIs.

As a result of growth in athletic training education and the change in the certification

examination, educational programs are now required to measure outcomes related to program effectiveness. The concept of outcome measures is not novel in higher education (Peer & Rakich, 2000). Through outcome measures, ATEPs demonstrate the quality of the academic program and the areas in need of improvement. The nature of measuring the outcomes is determined autonomously by each academic program. One of outcomes requiring measurement is psychomotor skill assessment; therefore, there is now a greater impetus on the psychomotor skill assessment of students within the didactic and clinical education components of the curriculum (Wilmer, 2005).

Statement of the Problem

Assessment in athletic training education is a continuous process molded by accreditation standards and faculty expertise. According to Vendrely (2002) and Sexton (2003), assessment is used to determine the effectiveness of the educational process for achieving intended student learning outcomes. Sexton states "good educational process does not always lead to good educational outcomes, nor do good educational outcomes lead to proper educational process" (p. 6). If this is true, then having assessment mechanisms within the curriculum does not guarantee the development of competent entry-level practitioners. To obtain an accurate evaluation of ATS, accurate and reliable assessment must occur (NATA, 2006a).

Appropriate clinical assessment of ATS is paramount for both formative and summative assessment of students in athletic training education. Formative assessment involves the provision of feedback to the students during their educational process. Vendrely (2002) stated that formative assessment emphasizes correct performance, determines areas of weakness, and

establishes competency and mastery within the construct of the knowledge being evaluated. She indicated that summative assessment techniques provide the student with an overall outcome or grade for the individual course enrolled. The goal of athletic training education is to develop competent and proficient entry-level practitioners. It is therefore essential to have quality formative assessment strategies and properly trained individuals utilizing those strategies.

With the continued development and advancement of athletic training education, the role of the certified athletic trainer (ATC) continues to expand. Certified athletic trainers are moving from the traditional work environments, such as providing health care to athletes and the physically active, into the academy. Athletic trainers are becoming employed as academic program directors, faculty members in ATEPs, and in clinical settings as ACIs. Along with this change, accreditation standards are requiring ACIs employed in the community to supervise and assess ATS (Brummels & Beach, 2008, Craig, 2006).

The requirements placed upon the ACI to instruct and evaluate ATS are crucial to the success of not only the athletic training education program, but also to the ATS. The psychomotor skill development of athletic training-specific skills by ATS is greatly facilitated by the instruction of the ACI. These individuals, through their formative assessment, mold and develop the ATS into the entry-level practitioner. This evolving role of ATCs from primarily clinicians into ACI and faculty members places a strong demand on these individuals to understand and practice effective assessment techniques (Weidner & August, 1997). Currently, "the clinical education format within athletic training education has placed more clinical teaching and evaluation responsibilities on certified athletic trainers who may not have had the

pedagogical training in their professional preparation" (Weidner & Henning, 2004 p. 335). Additionally, ATCs employed as faculty in higher education or employed as clinicians as ACIs must be able to assess ATS effectively to uphold the quality of the ATEP.

Laurent and Weidner (2001, believe that improving athletic training professional services is dependent upon quality ACIs instructing and assessing ATS accurately and effectively. With the perception that a majority of an ATS education comes from clinical education (Weidner & Henning, 2005), quality and accurate feedback from clinical staff (both ACIs and faculty) is essential. Therefore, poor or inconsistent feedback can affect the standard of care in athletic training professional practice. Most of the feedback relates to the performance of psychomotor skills by the ATS.

Standardized assessment of psychomotor skills in athletic training has not been developed (Butterwick, Paskevich, Lugamen, Vallevand, & Lafave, 2006). Based on the aforementioned developments and changes in the training and certification of ATS, the need for congruency in assessment practice between faculty and ACIs is essential. The purpose of this study was to investigate the interrater reliability of ACIs and ATEP faculty when rating ATS on psychomotor skill performance.

Research Questions

This present study evaluated the within group and between group agreement of psychomotor skill assessment by athletic training academic faculty and ACIs. This study was guided by the following research questions:

1. What is the interrater reliability of psychomotor skill assessment within Athletic Training Education Program faculty and within approved clinical instructors?
2. What is the interrater reliability of psychomotor skill assessment between Athletic Training Education Program faculty and approved clinical instructors?

Ancillary Questions

1. What is the interrater reliability of psychomotor skill assessment between ACIs with a bachelors or masters degree?
2. What is the interrater reliability of psychomotor skill assessment between male and female ACIs?
3. What is the interrater reliability of psychomotor skill assessment between ACIs who attended different numbers of approved clinical instructor workshops?
4. What is the interrater reliability of psychomotor skill assessment between academic faculty who hold a masters or doctoral degree?
5. What is the interrater reliability of psychomotor skill assessment between academic faculty with varying lengths of time spent in an ATEP?

Significance of the Study

The pragmatic significance of this research is based on the need for consistency between athletic training academic faculty and ACIs in rating ATS. Disagreement in the scoring of a student's ability is possible when there are several different individuals assessing the competency

of each ATS. The findings of this study will provide insight into the ability of ACIs to reliably assess ATS. It will also allow ATEPs to determine if student assessment by ACIs can be used to verify student readiness to take the BOC examination or if alternate forms of student assessment are needed. It also assumes that an ATEP faculty member and/or ACI will assess a skill the same way if they are in the clinic/lab setting as they would in this study if they were given the same format.

Assumptions

This study assumed that ACI and athletic training faculty would complete the study assessment as though they were evaluating an actual student. It is further assumed that the rating of each criterion on the psychomotor skills test used in this study is indicative of the actual rating ability of the ACIs and athletic training faculty.

Limitations

The results of this study cannot be used to evaluate the athletic training education program at any institution nor can they be used to evaluate individual faculty members on their assessment ability. The assessment instrument used in this study cannot be used in its present form for the assessment of ATS.

Definition of Terms

For the purpose of clarification, the following definition of terms will be used throughout this manuscript. All the terms listed below with their following descriptions have been taken from the Clinical Instructor Educator Seminar Handbook (NATA, 2006a).

- *American Medical Association* (AMA): promotes the art and science of medicine and the betterment of public health. Recognized the athletic training profession as an allied health profession in 1990.
- *Approved clinical instructor* (ACI): an appropriately credentialed professional identified and trained by the CIE to provide instruction and evaluation of the Athletic Training Educational Competencies and/or Clinical Proficiencies (Clinical Instructor Educator Seminar Handbook, 2006).
- *Athletic training education program* (ATEP): an entry-level professional education program accredited by the Committee on Accreditation of Athletic Training Education Programs.
- *Athletic training education program faculty:* Board of Certification Certified Athletic Trainers and other faculty who are responsible for classroom and/or clinical instruction within the athletic training major.
- *Athletic training student* (ATS): students enrolled in a CAATE accredited athletic training education program.
- *Board of Certification* (BOC): sets the standards for the practice of athletic training and is the only certifying body for athletic trainers in the United States.
- *Certified Athletic Trainer* (ATC): certified health care provider through the BOC who collaborate with physicians to optimize activity and participation of patients and clients.
- *Clinical education*: the application of knowledge and skills, learned in classroom and laboratory settings, to the actual practice on patients under the supervision of an approved clinical instructor or clinical instructor.

- *Clinical instructor* (CI): an individual identified to provide supervision of athletic training students during their clinical experience
- *Clinical instructor educator* (CIE): Board of Certification Certified Athletic Trainer recognized by the institution as the individual responsible for approved clinical instructor training.
- *Commission on Accreditation of Allied Health Education Programs* (CAAHEP): established in 1992 to replace CAHEA as the accrediting body for athletic training education
- *Committee for Accreditation of Athletic Training Education* (CAATE): established in 2006 to replace both JRC-AT and CAAHEP as the evaluators and accrediting agency for athletic training education programs respectively.
- *Committee on Accreditation of Health Education Programs* (CAHEA): organization within the AMA that was given the task of developing requirements for structure and function of entry-level athletic training education programs.
- *Joint Review Committee on Athletic Training* (JRC-AT): entity that was to evaluate athletic training education programs seeking accreditation and make recommendations to CAHEA for accreditation of these programs.
- *National Athletic Trainers' Association* (NATA): is the professional membership association for certified athletic trainers and others who support the athletic training profession.
- *Rater*: A person who evaluates or judges student performance on an assessment against specific criteria.

Organization of the Dissertation

The purpose of this study was to test the interrater reliability between the ATEP faculty and the ACI who evaluates students for the academic program. In chapter two, a review of research pertaining to assessment, interrater reliability, athletic training pedagogical practice in didactic and clinical education is discussed. In chapter three, the research design is described along with sample selection, administration, and evaluation of the development and instrumentation of the research study. Chapter four presents the results of the study, descriptive statistics and the results of the interrater reliability for the sample in regards to the psychomotor skills criteria. Chapter five discusses the practical application of the research study and offers suggestions for continued research in this area.

CHAPTER TWO
REVIEW OF THE LITERATURE

This chapter presents a summary of the literature related to assessment in athletic training education. The first section of this chapter evaluates the research pertaining to clinical education in allied health education programs. This is followed by a discussion of the function of assessment in clinical education, especially as it relates to the education of athletic training students. Included in this section is an evaluation of the role that accreditation plays in the development of assessment practices. The chapter concludes with a discussion of reliability and the concept of interrater agreement in clinical assessment.

Clinical Education

Clinical education is an essential part of the educational process of all allied health education programs (Giles, Wetherbee, & Johnson, 2003, Lauber et. al, 2003, Weidner & Henning, 2004). Research on clinical education in the related fields of nursing, physical therapy, and occupational therapy includes topics pertaining to the development of professional behaviors, critical thinking, peer mentoring, instructional practices and qualifications of clinical instructors (Ladyshewsky, Baer, Jones & Nelson, 2000; Riola, 1997). Research on clinical education in athletic training has not achieved the volume of literature in comparison. The majority of current research pertaining to clinical instruction within athletic training clinical education is geared toward the effective characteristics of clinical instructors who are educating ATS.

Clinical education is the setting under which the competency of skill acquisition and practical educational components are integrated into the every day workings within a health care facility (Stickley, 2005). Many health care professions also utilize clinical education to reinforce theoretical information presented in the didactic format (Laurent & Weidner, 2001). The goal of clinical education is to challenge students to integrate academic knowledge, professional development, and necessary skills to achieve effective patient/client management in diverse settings (Buccieri et al., 2006). Clinical education in athletic training is the setting in which ATS have the opportunity to utilize their psychomotor skill sets developed within laboratory settings in a practical environment. The academic program utilizes clinical education in order to allow the student to gain the necessary skills to be minimally competent and proficient through the application of a thorough base of knowledge. Ideally, students are to perform athletic training clinical skills on "real" patients (Knight, 2008). This teaching method allows ATS a more realistic opportunity to enhance critical thinking skills. The integration of a theoretical knowledge base into of clinical decision-making skills is paramount in the development of ATS (Knight).

Accreditation Standards and AT Education

Athletic training education, both clinical and didactic, has changed over time in conjunction with the development of new accreditation standards. New standards, reflecting a singular educational paradigm, have resulted in a more systematic approach to the educational process (Craig, 2003). Standardization replaces the prior multifaceted approach to athletic training educational practice in the preparation of students to become eligible for BOC

certification. The result of this revision is an increased accountability and an increased teaching responsibility for the ATC serving as a clinical instructor (CI) (NATA, 2006a). These individuals are now responsible for the integration of foundational knowledge within a clinical decision-making environment. Lauber, Toth, Leary, Martin, and Killian (2003) stated that the development, training, and evaluation of quality CIs contribute highly to the successful implementation of the clinical education component in athletic training education. Quality clinical education includes nurturing student development within the professional constructs of the overall education of the student (Seegmiller, 2003).

Clinical Settings in Athletic Training

The role of the ATC is defined through professional practice. Clinical education takes place in a variety of professional practice settings including intercollegiate athletics, outpatient rehabilitation centers, local high schools, area businesses, professional athletic teams, and local hospitals. In these settings the BOC has defined the primary task of an ATC within six domains: prevention; clinical evaluation and diagnosis; immediate care; treatment, rehabilitation, and reconditioning; organization and administration; and professional liability. Each professional setting may have a greater or lesser emphasis of the six domains. For example, an outpatient rehabilitation center focuses on treatment, rehabilitation, and reconditioning, while an ATC at a local high school would focus on clinical evaluation, diagnosis and immediate care. Because of this, the clinical education of the ATS provided by the ACI may reflect a particular domain emphasis within these different professional practice settings.

The quality of the education of ATS stems from the continuity between the ACI and academic faculty (Stroohschein, Hagler, & May, 2002). Consistency in theory and practice allows the student to understand the relationship between knowledge in the didactic and clinical settings. Therefore, a more formalized connection between academia and clinical instructors allows the goals of clinical education to supplement the overall educational process of the ATS (Knight, 2006).

Assessment

The National Athletic Trainers' Association Education Council has developed the essential learning outcomes identified as educational competencies, which are necessary for entry-level practice as an ATC. These educational competencies are required content in ATEP curricula for both didactic and clinical education and are minimal requirements for mastery in order to sit for the Board of Certification (BOC) examination. Assessing these learning outcomes is based on sound assessment procedures as well as the philosophy of the individual institution (CAATE, 2005). However, assessment in the didactic setting differs from that of the clinical education setting. Assessment in the classroom occurs in many different forms, from general questioning during instruction to formal assessment to determine a course grade. Assessment in clinical education usually occurs at two levels: specific psychomotor skills and the ability to integrate knowledge, psychomotor skills, and attitudes into the general work milieu of the athletic training setting (Clinical Instructor Educator Seminar Handbook, 2006, Knight, 2006, Knight, 2008).

Clinical Education Assessment

Assessment in clinical education is a crucial component in the education of ATS. Formative and summative assessment procedures utilized by the ATEP allow for the collection, interpretation, and application of the information in order to make programmatic adjustments to the areas needed (Vendrely, 2002). Formative use of assessment guides the student's learning while it is in progress. Information taken from either this formal or informal process contributes to the overall ATS learning through the provision of feedback to the student (Yorke, 2003). According to Nitko (2004), formative assessment is used for four basic purposes: determining the abilities of the students, diagnosing individual learning needs, diagnosing the group's learning needs, and planning instruction based on the results. The main goal of the formative assessment process is to either reinforce or to correct knowledge and/or performance relating to given material.

Athletic training students may receive feedback from multiple sources on their performance. Peers, clinical instructors, and self-evaluation provide criticism of their performance through the observation of formative evaluations. Informal formative assessment within athletic training psychomotor skill development is necessary for providing immediate feedback for skill acquisition. Informal formative assessments can motivate a student to increase their efforts to achieve a higher standard (Epstein, 2007). Yorke (2003) stated that the effectiveness of formative assessment is the quality of the feedback given to the student. With the many required educational competencies and skills to be mastered, formative feedback contributes to student learning by creating and enhancing a sense of motivation in their development (Epstein, 2007).

Summative assessments occur at the conclusion of the educational process and are used to determine the extent of the achievement of curricular objectives (Yorke, 2003). Summative assessments aid the instructors in evaluating the learning of the student as well as evaluating their own teaching (Nitko, 2004). Summative assessments are used to assign grades for a particular course, to determine a student's readiness to progress, and to determine the effectiveness of the overall course (Alexander, 1996). However, summative assessments provide little feedback other than the overall outcome on a particular assessment or course (Sexton, 2003). Summative assessment in athletic training is accomplished by both didactic and clinical instructors. The use of this form of assessment determines the entry-level competence of the student (Weidner & Henning, 2004).

Vendrely (2002) describes two alternative strategies for both formative and summative assessment: restricted performance and extended performance. Athletic training education utilizes both restricted and extended performance strategies in assessing student competence. Restricted performance assessment is used when a student is asked to perform very specific criteria within a given framework. A practical-based examination is an example of this type of assessment procedure. The amount of time to complete it is limited. The content of the assessment procedures is directed toward specific learning objectives and educational competencies particular to the task to be learned. The use of checklists is common in this type of assessment strategy and may be found in both the clinical and didactic setting.

In contrast, extended performance assessment allows the student to demonstrate an increased depth and breadth of knowledge and skill related to the given situation. Vendrely

(2002) defined extended performance assessments as high realism with extensive time requiring skilled judgment in scoring. This type of assessment generally occurs during clinical education and is the responsibility of the ACI supervising ATS. Extended performance assessment requires increased time and skilled judgment of the assessor to provide professional judgment of ATS performance. The ATEP is relying on the judgment of the ACI to assess the level of performance of ATS while in clinical education.

The ability to assess each student correctly is vital in giving the student appropriate feedback as well as ensuring the integrity of the athletic training profession with competent and proficient professionals (Vendrely & Carter, 2004). Assessment allows for students to understand their progress and provides meaningful feedback in an effort to motivate and correct the students learning processes. Assessment serves as the gatekeeper in accountability as well as a protective barrier in determining further practice or training for entry or continuance in a profession (Epstein, 2007). Assessment of ATS is perceived to be a significant part of athletic training clinical education (Lauber, Toth, Leary, Martin, & Killian, 2003). However, in contrast to other health fields, research related to assessment methods used in athletic training education is limited (Gould and Caswell, 2006).

Assessment by Clinical Instructors/Approved Clinical Instructor

Athletic Training accreditation standards require the use of CIs to educate ATS in developing entry-level athletic training behaviors and skills; however, it is not uncommon for allied health CIs to be selected to mentor students because of their professional aptitudes rather than their teaching and assessment abilities (Weidner and Henning, 2004). There is an adequate

amount of research in the allied health and medical literature to support the belief that clinical education faculty play an influential role in the overall development of the student (Giles, Wetherbee, & Johnson, 2003). Research in the field indicates that teaching and assessment are important characteristics of effective clinical instructors (Laurent & Weidner, 2001).

Clinical teaching differs from the traditional classroom in that clinical instructors must provide quality patient care as well as assess the clinical skills and reasoning of the student (Bowen, 2006). The effectiveness of these individuals serving the healthcare education programs is dependent on multiple variables. Sellheim (2006), states that teaching methodology of the educators stems from their beliefs on teaching and learning as well as their experiences going through their own educational process.

Certified athletic trainers serve an important role as clinical educators in the integration and facilitation of athletic training knowledge and skills to ATS (Laurent & Weidner, 2001). These individuals supervise, instruct, assess, and provide optimum learning experiences for ATS. With this realization, standards and guidelines were developed by the NATA Education Council to equip athletic training education programs with resources to train the ATC's who mentor ATS (Weidner & Henning, 2004). This information was used to develop Approved Clinical Education workshops by the academic institution to train the ATC as an Approved Clinical Instructor (ACI). An emphasis of the workshops is to educate the ATC in instructional methodologies as well as formative and summative assessment techniques utilized by the academic program. Upon completion of the workshop, the ATC is recognized as an ACI for that individual ATEP (NATA, 2006a).

Teaching within clinical education requires the ATC to instruct the student in athletic training specific competencies as well as provide critical feedback on the implementation of those skills. An integral part of teaching is the provision of constructive feedback during the learning process. Buccieri et al (2006) stated that students feel that an integral component of their learning is the ability of their clinical instructor to provide constructive feedback through their teaching. The evaluation of the students facilitates the progression toward development of entry-level competence (Weidner & Henning, 2004). The ability of the ATC to teach and assess students in athletic training clinical education has become more apparent (Craig, 2006).

Certified athletic trainers serving as ACIs have the opportunity to mold and influence the professional development of ATS through education by demonstrating and teaching professional skills and behaviors (Buccieri et al, 2006). Laurent and Weidner (2001) further clarify the important role of the ATC serving as an ACI in the educational process through the ability to integrate and model athletic training skills and behaviors. The modeling of behaviors is described as one of the most important characteristics of practicing professionals whether in athletic training, medicine, nursing, or physical therapy (Curtis, Helion & Dohmsohn, 1998; Laurent & Weidner, 2001; Buccieri et al, 2006).

Athletic training students believe, according to Weidner and Henning (2005), that 53% of their athletic training specific education comes from clinical education. With a majority of a student's athletic training education perceived to come from clinical education, having quality ATCs serving as ACIs is vital.

Weidner and Henning, through a series of research studies, (2004 & 2005), determined

seven important characteristics that an ACI should possess. These characteristics are an important facet of the ACI in the implementation of athletic training clinical education. Of these characteristics, the ability of the clinical instructor to teach and evaluate ATS effectively and correctly in their clinical education was recognized as a significant characteristic (Lauber et al, 2003; Weidner & Henning, 2005). The ability of the ACI to correct ATS in a manner that will elicit professional growth is a key component in the overall effectiveness of the ACI. According to Giles, Wetherbee, and Johnson (2003), clinical instructors must be able to accurately compare the performance of the student to the criteria determined for entry-level into the profession. Again, adequate and appropriate feedback concerning performance was shown as necessary to enhance a student's professional growth.

"Clinical Education should also help students to learn the necessary skills and apply the appropriate theoretical knowledge in the correct circumstances; therefore, continuous improvement of athletic training services is dependent upon building and maintaining quality clinical education (Laurent & Weidner, 2001, p. 58). The quality of the ACI is an important aspect in how effective the overall clinical education of ATS will be. Quality is being defined as characteristics that positively support the growth and development of the ATS. One key component in the overall effectiveness of this process is the type of assessment the ATS is receiving by the ACI during the course of their clinical education (Weidner & Henning, 2004).

The evaluation of the clinical performance is a necessary and critical component in athletic training clinical education (Weidner and Henning, 2004). The evaluation of a student's competence within clinical education aids in the determination of their overall academic

progress. Both ATS and athletic training program directors believe that quality and timely feedback through evaluative processes are vital in the entry-level development of the student (Lauber et al, 2003; Weidner & Henning, 2004).

Outcome Measures in Athletic Training

The assessment of ATS is an important aspect in the outcomes measurement plan of the ATEP. One form of outcome measures are student learning outcomes. Athletic training accreditation standards require that academic programs must routinely collect data pertaining to the effectiveness of the ATEP based on the student learning outcomes (Commission on Accreditation of Athletic Training Education, 2005). A part of the outcomes assessment plan is the clinical skills assessment performed by the ACI in the clinical setting. Assessments performed by the ACI are vital in the collection of data that allows the ATEP to determine if ATS meet the standards for entry-level practice (Weidner & Henning, 2004).

Accreditation standards dictate the minimal requirements for establishing an outcomes assessment plan. Currently, the use of standardized instruments that produce data that are both reliable and valid and that are used to measure technical skills has not been reported in athletic training (Butterwick, Paskevich, Lugamen, Vallevand, & Lafave, 2006). Accreditation standards allow ATEPs to determine how best to assess students in psychomotor skills in either the classroom setting or in clinical education (Commission on Accreditation of Athletic Training Education, 2005). This approach offers academic programs the freedom to develop instruments that meet their current needs. With the possibility of a variety of different assessment tools,

questions are raised concerning the comparative reliability among ACIs as well as academic faculty who are assessing ATS within athletic training education.

Reliability

Reliability is the extent to which scores from a measurement are reproducible and dependable, have very little error, and are consistent (Arnold, Gansneder, & Perrin, 2005). Linn and Miller (2005) state that unless the measurement can be shown to be reasonably consistent over different occasions, with separate raters, and/or different samples of the same construct, confidence in the results may be suspect. In addition, Thompson (2003) emphasizes that reliability is a matter of degree due to random fluctuations within scores. "Reliability is fundamental to all aspects of measurement for the reason of having confidence in the data that is collected as well as drawing accurate conclusions" (Portney & Watkins, p. 61). According to Thompson (2003) and Vacha-Haas, Kogan, Tani and Woodall (2001), reliability of the data should be disclosed within every research study. They emphasize that the reliability of the data in a given study aids in the determination of the results.

Reliability is measured by a combination of the true score and measurement error. Measurement error is further delineated into two categories: systematic and random error. Systematic errors are predictable errors that are "constant and do not present a problem for reliability" (Portney & Watkins, 2000, p. 62). This bias can be accounted for through the correction of the assessment process. Random error is unpredictable influence that occurs by chance. An example of random error is the concept of subjectivity within the individual's ability to perform the assessment. There are many factors during a testing/assessment procedure that can

influence the responses of raters performing the assessment, thus influencing the reliability of the score obtained. An important aspect of random error is that it does not have a consistent effect across the entire sample and the scores obtained are randomly dispersed. This variability within the data does not adversely affect the observed score in relationship to the true score.

Subjectivity is introduced in many circumstances involving the assessment of students. Subjectivity in assessment is defined as the personal qualities of an investigator that can affect the outcome of the research project (Portney & Watkins, 2000). These personal qualities include personal views, experience, and/or background that can affect the manner in which an individual views the performance of the student.

There are many different areas that influence an individual's ability to perform an assessment. According to Elder, Knoch, Barkhuizen, & von Randow (2005), some raters perceive certain criteria as being more important than others within an assessment process. These authors emphasized that raters are more critical towards the areas or criteria they deem more important; therefore, accurate scoring of the student may not occur. For example, an ACI may be more critical of certain criteria within the overall assessment procedure because of a self-perceived understanding of the importance of those criteria. In the performance of an assessment, the raters draw from their own personal knowledge, feelings, tendencies, and education in evaluating the student completing the task. These personal tendencies of the rater may unconsciously introduce subjectivity into results of the assessment. Students may also influence the rater in the assessment based on multiple factors that can impact the congruency of personality of the student and instructor (Alexander, 1996). Student's attitudes toward their

overall education, learning environment, clinical placement criteria, and personality may effect how an ACI and faculty view the students during an assessment period. Students who meet or fail to meet the subjective, unwritten requirements of the ACI and faculty may also influence the outcome of the assessment procedure (Alexander).

Ultimately, professional judgment in the assessment of a student's abilities is necessary to determine aptitude (Vendrely & Carter, 2004). Alexander (1996), and Vendrely and Carter, believe that individuals assigned the task of assessing students must be aware of how subjectivity can affect the accuracy of assessment. The ability of ACIs and faculty to manage personal influences is necessary. This raises the question of whether one can truly remove the subjectivity within any observational assessment when the assessor has to make a determination (Alexander 1996). A common practice to address this issue, according to Portney & Watkins (2000) and Elder et al, (2005), is through standardization of the assessment criteria within grading tools in order to reduce subjectivity. A method to address the issue of subjectivity is through the use of grading rubrics with explicit criteria. Standardizing criteria provides the assessor with the important norms that are to be evaluated. Providing instruments that provide reliable and valid scores may offer the context in which the assessor can assess the student with increased accuracy and efficiency.

Reliability analysis specific to athletic training has been used in several different ways including determining the effectiveness of certain assessment tools. Thompson concurs stating that the investigated data must be established as reliable before it can be analyzed for clinical significance (2003). Eechaute, Vaes, Duquet, and Van Gheluwe (2007) believe that acceptable

test-retest reliability must be determined before a testing procedure is implemented. As an example, Clapper and Harris (2008) modified the Maslach Burnout Inventory to determine the factors that contribute to burnout of collegiate certified athletic trainers. In this study, Cronbach's alpha was used to analyze each construct of the tool to establish reliability. In two other studies, a test-retest procedure was used to determine whether the results of specific assessment tools were reliable. Broglio et al. (2007) analyzed the reliability of a new instrument for screening functional ankle instability. Docherty, Gansneder, Arnold, and Hurwitz (2006) examined several computerized concussion assessment programs for test-retest reliability.

Interrater Reliability

Athletic training education utilizes multiple ACI and faculty to assess students in didactic and clinical education. Through the course of assessment procedures, not all raters will agree that students perform each criterion in the manner they see as sufficient. There is variability in the perception by the rater as to what is deemed important within the assessment (Gamaroff, 2000). According to Tinsley and Weiss (1975) this variability in rating is a concern when trying to determine the influence of idiosyncratic judgments on the rater's results. Athletic training education programs rely upon ACI and academic faculty to accurately determine the proficiency of ATS. Within the practice of athletic training, there are multiple ways to perform a skill that meets the standard of proficiency. Therefore, it is necessary for the raters to have a consensus in the overall agreement of skill acquisition made by the student.

Interrater reliability is concerned with the variation between two or more raters and the scores that are made among them (Portney & Watkins, 2000). Interrater reliability is the ability

of two or more raters (e.g. ACI and ATEP faculty) to agree on an assessment of specific criterion. Agreement refers to the ability of two or more raters to arrive at an identical judgment. The main purpose of interrater reliability is the reduction of measurement error between multiple raters. In order to accomplish that task, a process that allows the raters to assess the student under the exact same conditions is essential. Portney& Watkins (2000) hold that simultaneous evaluation of the subject by raters who are isolated from each other is the preferred circumstance in which to measure interrater reliability. This reduces variability in subject performance, thus reducing measurement error. While desirable, simultaneous evaluation may not be practical in athletic training education, however, Vendrely & Carter (2004) observe that video can be used to aid in the process of establishing interrater reliability.

Recent literature in athletic training pertaining to interrater reliability has been in the development of assessment tools as well as evaluating the ability of clinicians to perform specific assessment practices. A research study by Lafave, Katz, Donnan, and Butterwick (2008) analyzed rater agreement using a text-based content-validated assessment tool. Raters attended a three hour orientation and training session to become familiar with the assessment tool. The tool contained between 200 and 250 specific criteria in four individual categories. Raters evaluated students in simulated orthopedic setting with standardized patients as subjects. The authors noted that the raters were required to use their professional judgment in assessing the student's performance even though the rubric was scored at the dichotomous level. This was due to the multiple approaches that could be taken by the student to fulfill the testing requirements. The authors found high levels of interrater reliability among the raters using the assessment tools for

the shoulder (α = .82), knee (α = .83), and ankle (α = .91). The authors also suggest that the training and orientation session given to the participants might have played a role in the reliability scores obtained.

Research performed by Peeler and Anderson (2008) evaluated the interrater reliability of clinicians performing a clinical assessment skill. This study analyzed the reliability between clinicians who assessed patient performance on a standardized pass/fail instrument. The clinicians attended an orientation session before performing the assessment procedure. Peeler and Anderson reported low to moderate levels of interrater reliability, using a chance-corrected kappa statistic (K = .26 and K = .41). The authors attribute the low to moderate levels of reliability to differences in the administration of the assessment procedure by each clinician even though a training session was implemented before the research was performed.

Shirk, Sandrey and Erikson (2006) analyzed interrater reliability of a clinical assessment protocol for which reliability data had not been reported. The goal of the study was to determine whether there was a difference in the outcome of the assessment, based on whether the raters were classified as novice or experienced. The authors reported poor to fair levels of interrater reliability (K = .02 to .26) between and among the two sets of raters, using percent agreement and chance-corrected kappa. As with the study by Peeler and Anderson, an orientation session was conducted for the raters; however, the low reliability scores in this analysis were attributed by the authors to the lack of standardization of the assessment protocol.

Percent Agreement

One method for assessing interrater reliability is by calculating percent agreement. Percent agreement is the most often utilized descriptive method when the data is nominal or categorical (Haley & Osberg, 1989). Percent agreement is defined as the frequency that raters produce the same score on the same criteria. In the case of nominal dichotomous variables, raters determine whether or not a particular criterion is met. Therefore, a percentage can be calculated for the overall agreement or disagreement among the raters performing the assessment.

According to Hunt (1986), Portney &Watkins (2000), and Sim and Wright (2005), there are limitations in the interpretation of the obtained score using percent agreement. While this form of statistical measure is somewhat easy to calculate, the literature suggests a lack of consensus in the appropriateness of its use. A common argument from the literature is that percent agreement does not take into account the possibility that the observed agreement between raters can occur by chance. The "chance" occurrence of scores does not give an accurate interpretation of true reliability. According to Haley and Osberg (1989) "chance agreement is estimated by the proportion of agreements that would be expected if the rater's judgments were perfectly random" (1989, p.972). When agreement occurs as a matter of chance, the question of true agreement arises (Cohen, 1960). As a result, the kappa statistic was developed to determine a percent agreement that accounts for the extent of agreement between raters that can occur by chance (Fleiss, 1971).

Kappa Statistic

The kappa statistic (*K)*, introduced by Jacob Cohen in 1960, was developed to measure nominal/categorical scale agreement between a fixed pair of raters (Cohen, 1960). Research literature suggests that Kappa is the preferred statistical measure for measuring reliability with nominal scale data (Cohen, 1960; Fleiss, 1971: Hunt, 1986; Donner & Klar, 1996; Portney & Watikins, 2000; Sim & Wright, 2005). Kappa values range from 0.00 to 1.00. The kappa statistic is the quantification of agreement beyond chance between two raters for dichotomous judgments (Hunt, 1986). Guggenmoos-Holzmann (1996) suggests that correcting for chance agreement is related to the supposition that the inconsistency in rating is due to the indecisiveness of raters. This indecision could be based upon the inability of the rater to decipher the rating criteria or understand the performance of the rating protocol. A consideration when using Cohen's kappa is that it is only a comparison between two raters and may not be generalizable to a larger number of raters (Crewson, 2005).

Several considerations exist in the interpretation of kappa in the determination of interrater reliability. One consideration is the homogeneity of multiple raters. Similarities among the raters may increase the overall percent agreement of the ratings (Feinstein & Cicchetti, 1990). Another key issue is sample size. Since kappa is based upon proportions, the total number of participants as well as the number of categories to be assessed may affect the kappa value. The kappa statistic developed by Fleiss is affected by this kind of variation (Fleiss, 1971; Randolph, 2005).

Fleiss' kappa allows for more than two raters to rate any number of items and accounts for agreement expected by chance. Fleiss' kappa is often used for fixed-marginal studies (Randolph, 2005). A fixed-marginal study is one in which participants know in advance that cases will be distributed proportionally. Prevalence exists in a fixed marginal study when participants know how many cases of a specific kind are expected in each category (Randolph, 2005; Feinstein & Cicchetti, 1990). Randolph asserts that the concept of prevalence can influence the fixed-marginal kappa statistic when the numbers of cases in each category are not balanced across categories.

An alternative to fixed-marginal versions of kappa is a free-marginal kappa statistic proposed by Randolph (2005). According to Randolph (2005), a free-marginal kappa is recommended when the raters do not know in advance the number of cases each category should have. The multi-rater free-marginal kappa adjusts for the concept of prevalence when the distribution of cases are allowed to vary. Randolph's free-marginal kappa is a version of a bi-rater kappa statistic that is not influenced by prevalence. The present study does not allocate a pre-determined number of cases to specific categories (is free-margin and not prevalent), and uses multiple raters. Therefore the Randolph kappa has been chosen for use in determining interrater reliability in this study.

Documenting Interrater Reliability in Athletic Training

Percent agreement and kappa scores have been calculated in several interrater reliability studies in athletic training. Research by Boyle, Witt, and Krugh (2003) tested interrater reliability

"using a commonly employed" (p. 281) joint mobility index. They analyzed percent agreement for interrater reliability of the categorized data. Results indicated high percent agreement (89%) scores. The authors attribute the high agreement scores to the standardized and common usage of the index. They suggest that the assessment tool contributed to consistency of the ratings which led to high percent agreement scores.

In contrast to the research of Boyle, Witt, and Krugh (2003), Shirk, Sandrey and Erickson (2006) reported no validity with their assessment tool. The data were analyzed using both percent agreement scores and the kappa statistic due to the nominal character of the data. The authors reported low interrater reliability scores (42%, $K = .26$) and attribute this to the lack of a validated assessment tool. These studies suggest the need for the use of validated assessment instruments which are both specific in content and straightforward in implementation, in order to produce reliable assessment results between and among raters using the instrument.

Chapter Summary

This dissertation describes research on the evaluation of interrater reliability in the assessment of ATS between and among ACIs and ATEP faculty. A review of the literature on assessment in the context of athletic training education was presented in chapter two. A description of the methods of the study are presented in chapter three and the results of the study are in chapter four. Chapter five includes an evaluation of the results, their application to athletic training program assessment and clinical practice, and suggestions for further research.

CHAPTER THREE
METHODS

Assessment of ATS' psychomotor skill competency is an integral aspect of athletic training education and is an ever-increasing component in determining the effectiveness of an ATEP. The purpose of this study was to determine interrater reliability within and between athletic training program faculty and ACIs on assessment of the psychomotor skills of ATS. Chapter three includes the following sections: development of the instrument, study population, administration of the instrument, and data analysis.

Instrument

Athletic training students must master both theoretical content and content-specific psychomotor skills. The psychomotor skills used in this study were chosen based on the Board of Certification Role Delineation Study (2004) for athletic training. The role delineation study identified four domains of psychomotor skills that were then included in the practical exam portion of the certification examination. These four domains are: prevention, clinical evaluation and diagnosis, immediate care and treatment, and rehabilitation and reconditioning. For this study, one skill was chosen from each of the four domains: closed basketweave ankle taping (prevention domain), unmodified anterior Lachman test (clinical evaluation and diagnosis domain), ankle horseshoe pad with compression wrap (immediate care and treatment domain), and axillary crutch fitting (rehabilitation and reconditioning domain).

Psychomotor Skills Assessment Instrument

The Psychomotor Skills Assessment Instrument (PSAI) used in this study consisted of two parts: a video demonstration of the psychomotor skill being assessed, and a criteria rubric for determining whether the skill was performed correctly. In the video demonstration, one actor (senior athletic training student) portrayed the ATS while another student portrayed the injured athlete. The same actors demonstrated all the skills used in the video, in the same studio (a university athletic training laboratory), and all videos were recorded on the same day. All demonstrations were performed in silence. Table 1 lists the duration of each psychomotor skill demonstration video.

Table 1

Duration of Psychomotor Skills

	Duration of Skill Demonstration (minutes)
Closed Basketweave Ankle Taping	1:57
Unmodified Anterior Lachman Test	0:16
Ankle Horseshoe Pad with Compression Wrap	0:56
Axillary Crutch Fitting	1:43

In the video, the ATS either followed or deviated from the rubric described for each skill. The unmodified anterior Lachman test and the axillary crutch fitting were performed correctly

according the criteria listed in the rubrics. Conversely, one criterion in the closed basketweave ankle taping procedure and one in the ankle horseshoe pad with compression wrap were not performed correctly. Specifically, in the closed basketweave ankle taping procedure, more than two anchors were used (criteria four). In the ankle horseshoe pad with compression wrap, the distal pulse check was omitted (criteria seven). The rubric and the criteria are described in the following section. Copies of the video may be obtained from the author.

Psychomotor Skills Rubric

Scoring rubrics from the text *Practical Exam Preparation Guide of Clinical Skills for Athletic Training* (Amato, Hawkins, & Cole, 2002) were used for the rubric for the PSAI. Permission was obtained from the publisher for use of the rubrics (Appendix C). Rubric criteria for each of the four skills in the PSAI were worded exactly as in the *Practical Exam Preparation Guide.* Response choices for each of the criteria were "Yes" (acceptable performance) or "No" (unacceptable or not performed). Four skills from the *Practical Exam Preparation Guide* were selected for use in the PSAI: closed basketweave ankle taping procedure, unmodified anterior Lachman test, ankle horseshoe pad with compression wrap, and axillary crutch fitting.

The closed basketweave ankle taping procedure is primarily used for newly sprained or chronically weak ankles (Prentice, 2009). The corresponding rubric is designed to assess 11 critical aspects of the procedure, including preparation of ankle to be taped, placement of pads and anchors, and stabilization of joint with a series of interlocking procedures (Table 2).

Table 2

Closed Basketweave Ankle Taping Procedure

Criteria:	Yes	No
Seated with ankle dorsi-flexed to 0 deg		
Places heel and lace pads		
Secures pads with layer of pre-wrap		
Two anchors around base of gastrocnemius muscle		
One anchor around instep, proximal to base of fifth metatarsal		
Applies a stirrup		
Places a horseshoe around the foot just below malleoli		
Repeats above 2 steps two additional times		
Applies a figure eight starting at the dorsal aspect of ankle followed by heel-locks.		
Closes taping with horseshoes, overlapping by half the tape from distal to proximal		
Places final anchor around the instep, proximal to the base of the fifth metatarsal		

The Unmodified Anterior Lachman Test is used to test the integrity of the anterior cruciate ligament in the knee. The corresponding rubric is designed to assess four critical aspects including hand placement and testing procedure (Table 3).

Table 3

Unmodified Anterior Lachman Test

Criteria	Yes	No
Athlete is supine with knee in 10 – 25 degrees of flexion		
Stabilizes posterior calf with one hand		
Stabilizes anteriorly on distal femur with other hand		
Attempts to displace tibia on femur		

The Ankle Horseshoe Pad with Compression Wrap is designed to do reduce swelling and effusion. The corresponding rubric is designed to assess seven critical components including ankle placement, application of elastic wrap, and the security of the wrap (Table 4).

Table 4

Ankle Horseshoe Pad with Compression Wrap

Criteria	Yes	No
Seated with ankle dorsiflexed to 90 degrees		
Applies horseshoe pad to lateral malleolus		
Applies compression wrap just proximal to toes		
Applies compression wrap proximally to lower third of calf		
Overlaps compression wrap by half		
Secures compression wrap with tape		
Checks distal pulse/sensation/capillary refill		

The Axillary Crutch Fitting procedure is used to aid the patient with ambulation when weight bearing on the effected limb is contraindicated. The corresponding rubric is designed to assess seven important criteria including the height and alignment of the crutch and patient safety (Table 5).

Table 5

Axillary Crutch Fitting

Criteria	Yes	No
Feet shoulder width apart		
Wearing low-heeled/flat shoes		
Places crutch 2-4 inches in front of involved limb		
Places crutch 4-6 inches to the side of the involved limb		
Places axillary pad 1-1.5 inches below axilla		
Places elbow in approx. 30 deg of flexion		
Tightens all fasteners		

Assembly of the PSAI

The video tape demonstrations and the rubrics were placed on the web-based *Checkbox* (Prezza, 2007) platform for administration. A demographics survey was added to the PSAI as part of this study. The survey included questions regarding age, gender, credentials and experience. The demographics survey is described later. A copy of the survey questions is located in Appendix A.

Content Validity

The goal of content validity is to show that the items on the test are representative of the intended content domain. Sireci (1998b) views content validity as test quality. Test quality pertains to the quality of the material specific to the construct of interest. Haynes, Richard, Kubany (1995) defined construct as the target or the intended concept of the assessment instrument. They hold that content validity is an important part of the validation process in that it provides confirmation regarding the elements of the instrument as they relate to and characterize the intended target. Therefore, content validity can affect interpretation of data. Validity also includes any aspect of the measurement process that can affect the data obtained with an assessment instrument. This includes measurement procedure, format, instructions, and time of assessment. The two components of content validity are essential in the instrument development phase and are addressed by content experts.

One way to establish content validity is for a panel of experts to review the instrument in order to assure relativity to content domain. To ensure content validity in the present study, criteria for each psychomotor skill shown in the video were taken from a peer reviewed text entitled *Practical Exam Preparation Guide of Clinical Skills for Athletic Training* (Amato, Hawkins, & Cole, 2002). This text was developed as an exam preparation guide for the BOC examination and was authored by three content experts who represent areas of practice within the field of athletic training. The authors are recognized leaders within the athletic training profession through their research, public speaking, and authorship. The *Practical Exam Preparation Guide* was peer reviewed by educators and clinicians who provided feedback on the content published in the text.

Pilot Studies

Two pilot studies were conducted to evaluate the PSAI. The first pilot study was conducted to evaluate the overall instrument content and process. Two individuals who met the population criteria (one faculty and one ACI) were personally solicited by phone to evaluate the instrument in order to provide feedback concerning the following: (a) overall reactions to the process; (b) clarity of instructions; (c) the length of the process; (d) suggestions for revisions. They were asked to provide feedback within two week either by phone or email. After thorough evaluation of the feedback from these two reviewers, changes were made to the instructions for each grading rubric to reduce verbiage and improve clarity.

A second pilot study was conducted using content experts for the purpose of establishing content validity for the entire instrument. Four certified athletic trainers were contacted by phone to evaluate the PSAI as experts in athletic training education. These individuals were chosen based on their professional credentials and meeting population criteria. Two were athletic trainers who had academic doctorates, had been teaching for at least 10 years in an ATEP, and were involved in the assessment of psychomotor skills of ATS. The remaining two were ACIs for an ATEP with master's degrees, had been an ACI for at least 5 years, and were involved in the assessment of psychomotor skills. The four content experts were asked to respond within two week by either phone or email. The content experts were asked to complete the PSAI (complete the nine demographic questions, view the videos of the psychomotor skills and complete the rubrics) and provide feedback on both the process and content of the instrument. All four indicated no changes were necessary.

Study Population

The target population for this study consisted of athletic trainers who teach and supervise students in CAATE-accredited ATEPs. The population was divided into two groups. The first group was comprised of athletic trainers who serve as academic faculty teaching the didactic component of athletic training curriculum. This group was labeled as faculty. The second group consisted of athletic trainers who serve as ACIs and teach in the clinical education component of the curriculum. This group was labeled as ACI for this study.

Study participants were selected by convenience consecutive sampling. A list of CAATE-accredited ATEPs was acquired from the CAATE web-site (http://www.caate.net/). Using the Accredited Programs link on the CAATE homepage, programs in the states of Florida (13), Georgia (2), North Carolina (1), Ohio (1), West Virginia (1), Tennessee (1), and Texas (2) were selected from the list generated, based on personal knowledge of individuals in the programs. For the selected programs, contact information for the program directors in the form of email addresses as well as telephone numbers was available. Also available on the CAATE website is a link to the ATEP home web site for the programs contacted for this study. The homepage for each ATEP website was reviewed for faculty contact information. The faculty identified on the program website were contacted by either telephone or email to participate in the study. If contact information was not available, the program director was contacted via phone and/or email to retrieve this information. A majority of the contact information for faculty contacted for this was found by locating and reviewing the ATEPs web site.

All potential participants in this study were asked for ACI contact information. Only one

ATEP sent ACI contact information to the investigator. This group was emailed the link to PSAI. The ATEPs that did not send ACI contact information were then contacted by either telephone or email or both to forward the study participation request to the ACIs.

Emails requesting study participation explained the research study, included the informed consent agreement, and contained a link to the PSAI. A copy of the email may be found in Appendix A**.** Subject responses on the instrument were compiled and stored on a mainframe computer and supplied to the author in spreadsheet form. Respondent anonymity was maintained through the lack of unique identifiers in the demographic portion of the instrument. Demographic information included: (a) age of the participant, (b) sex, (c) classification as faculty or ACI, (d) number of years in this classification, (e) highest degree earned, (f) if an ACI, how many ACI training workshops had they attended, and (g) if an ACI, what was their primary practice setting.

Sample Size

Percent agreement and kappa can be considered forms of a correlation since this study is looking at the association between two groups. A power of .80 is considered the conventional standard for research studies (Portney & Watkins, 2000). Due to the descriptive nature of this study, a power analysis not applicable and was not performed.

Administration of the Instrument

The PSAI was set up to be completed in a linear fashion. Participants read the informed consent and clicked on the “I agree” to continue or “I do not agree” to exit. Participants were then asked to complete the demographics portion of the PSAI. All of the questions appear

simultaneously to the participant. Once the demographic questions were completed, the participant clicked on the "next" icon on the bottom of the page to move onto the psychomotor skill assessment portion of the PSAI. For this portion of the PSAI, the rubric was reviewed first, followed by a video demonstration of the skill, and finally the rubric was again presented for completion by the reviewer. The closed basketweave ankle taping procedure was the first psychomotor in the sequence. The rubric criteria were first presented and participants were allowed to view the criteria for as long as desired. After review, the participants clicked on the "next" icon on the bottom of the page to move to the psychomotor skill video. The psychomotor skill video played immediately without pause. No repeat of the video was allowed. Upon completion of the video, participants clicked on the "next" icon and the entire scoring rubric appeared with the exact same criteria as seen prior to the psychomotor skill video. An unlimited amount of time was allowed to complete the scoring rubric with the opportunity to change selections if needed. This process was then repeated for the remaining three psychomotor skills. A final page was generated thanking them for their participation.

Data Analysis

Dichotomous criteria choices on the PSAI were YES and NO, requiring dummy coding in order to generate percent agreement scores and the kappa statistic. YES responses were coded as zero, and NO responses were coded as one. Demographic data was coded as found in Appendix D.

Research Question 1

What is the interrater reliability of psychomotor skill assessment within Athletic Training Education Program faculty and within approved clinical instructors?

This question was analyzed by determining a percent agreement score for each psychomotor skill and an overall percent agreement score for all skills combined for the entire sample. A kappa statistic was also calculated for each psychomotor skill and for all skills combined for the entire sample.

Research Question 2

What is the interrater reliability of psychomotor skill assessment between Athletic Training Education Program faculty and approved clinical instructors?

This question was analyzed by determining a percent agreement score for each psychomotor skill for each group and an overall percent agreement score for all skills combined for each group. A kappa statistic was also calculated for each psychomotor skill for each group and for all skills combined for each group.

Ancillary Questions

1. What is the interrater reliability of psychomotor skill assessment between ACIs with a bachelors or masters degree?
2. What is the interrater reliability of psychomotor skill assessment between male and female ACIs?

3. What is the interrater reliability of psychomotor skill assessment between ACIs who attended different numbers of approved clinical instructor workshops?
4. What is the interrater reliability of psychomotor skill assessment between academic faculty who hold a masters or doctoral degree?
5. What is the interrater reliability of psychomotor skill assessment between academic faculty with varying lengths of time spent in an ATEP?

Each of these questions was analyzed by dividing the sample into appropriate groups. A percent agreement score for each psychomotor skill and an overall percent agreement score for all skills combined was determined for each group. A kappa statistic was also calculated for each psychomotor skill and for all skills combined for each group.

Post Hoc Analysis

Criteria four and five of the closed basketweave ankle taping procedure and criterion seven of the ankle horseshoe compression wrap were analyzed. A percent agreement and kappa scores were calculated for each individual criterion.

Statistical Formulas/Calculations

All individual psychomotor skill percent agreement scores, composite percent agreement scores, and the kappa statistic were calculated using an online Kappa calculator developed by Justus Randolf (Randolph, 2008). This online calculator was designed to calculate the chance adjusted measure of agreement for any number of cases and raters. Descriptive analyses using SPSS v.16.0 were used to determine the characteristics of the sample.

Chapter Summary

This chapter described the methods used in this research. The following chapters provide a description of the results of this study (chapter 4) and an evaluation of the results, including the importance of the study to the field of athletic training education, and suggestions for further research (chapter 5).

CHAPTER FOUR
RESULTS

This chapter describes the results of this study beginning with a description of the characteristics of the sample. The results of the analysis of each of the research and ancillary questions are then presented. The chapter concludes with a brief summary.

Characteristics of the Population

Demographics

Age was distributed across four categories with a majority of the participants in the 31–40 range (n = 54, 47%; Table 6). Approximately one-half the respondents were female (n = 57, 50%). Thirty-two (28%) participants were full-time faculty members and 83 (72%) were full-time ACIs (Table 6). A majority of the ACIs (n = 35, 30%) indicated they had attended at least three ACI workshops. Forty-seven ACIs (41%) specified their primary work setting as the high school, with the remainder divided among intercollegiate, professional, clinical, industrial, and other settings.

Table 6

Distribution of Age, Sex and Faculty Status of Participants

	Age				
Position& Gender	20-30	31-40	41-50	51 & older	Total
Faculty					
Male Faculty	0	2	12	4	18
Female Faculty	1	2	11	0	14
ACI					
Male ACI	5	24	11	0	40
Female ACI	12	26	4	1	43
Total	18	54	38	5	115

The number of years certified as an athletic trainer were broken into six categories (Table 7) with a majority of the participants (n = 38, 33%) having been certified at least 11 – 15 years.

Table 7

Years Certified as an Athletic Trainer

	Years certified						
Faculty Status	1-5	6-10	11-15	16-20	21-25	25 or more	Total
Faculty	1	1	2	13	9	6	32
ACI	8	20	36	12	7	0	83
Total	9	21	38	25	16	6	115

Most of the participants held a master's degree (n = 66, 57%) with the remainder holding degrees at the doctoral and bachelors level (Table 8).

Table 8

Distribution of Sex, Faculty Status, and Degree Earned of Participants

		Degree Earned				
Faculty Status		BS/BA	MS/MA/MEd	EdD/PhD	Other	Total
Faculty						
	Male	0	9	8	1	18
	Female	0	12	2	0	14
ACI						
	Male	16	24	0	0	40
	Female I	20	21	1	1	43
	Total	36	66	11	2	115

A majority of the participants (n = 79. 69%) had 6 – 10 years of experience as a faculty member or ACIs (Table 9).

Table 9

Distribution of Years as Faculty or ACI of Participants

	Years as faculty or ACI			
Faculty Status	1-5	6-10	11-15	Total
ATEP Faculty	3	13	14	30*
ACI	31	50	2	83
Total	34	63	16	113

* Two faculty members did not indicate number of years served as faculty.

Interpretation of Kappa

According to Portney and Watkins (2000), several factors need to be taken into consideration when interpreting the kappa scores obtained in reliability analysis. One factor that was discussed earlier in this report was variance among raters. When a group has similar characteristics, agreement scores tend to be high. Secondly, kappa "represents an average rate of agreement for an entire set of scores" (Portney & Watkins, 2000, p. 574). Kappa scores are also influenced by the number of categories used within the reliability analysis. As the number of categories increase, the extent of the agreement can possibly increase or decrease due to the larger combination of possibilities in the assignment of scores. However, the present study only uses dichotomous outcomes. Because kappa was designed for use with dichotomous nominal level data (Cohen,1960; Fleiss, 1971), interpretation of kappa in this study can be expected to reflect actual agreement.

Landis and Koch (1977) proposed the following guidelines for strength of agreement: ≤ 0 = poor; .01 - .20 = slight; .21 - .40 = fair; .41 - .60 = moderate; .61 - 80 = substantial; and .81 - 1.0 = almost perfect. These guidelines are used in the literature to interpret the strength of kappa.

Research Questions

Two research questions were formulated to evaluate interrater reliability among ATEP faculty and ACIs. Each of the research questions are listed below with the results of the statistical analyses.

Question 1: What is the interrater reliability of psychomotor skill assessment among Athletic Training Education Program faculty and approved clinical instructors?

The results for research question one are presented in Table 10. Percent agreement scores were calculated for each psychomotor skill as well as a combined percent agreement for the entire sample. A kappa statistic was also calculated for each psychomotor skill and for all skills combined for the entire sample, using the Randolph online calculator (Randolph, 2008). The psychomotor skills percent agreement scores ranged from 89.8% (K = .80; substantial) for the closed basketweave ankle taping procedure to 98.3% (K = .97; almost perfect) for the unmodified Lachman test, with an overall percent agreement of 93.1% (K = .86; almost perfect).

Table 10

Percent Agreement and Free-Marginal Kappa Statistic for the Aggregate Group (Faculty & ACIs)

	Percent Agreement	Kappa
Closed Basketweave Ankle Taping	89.8%	.80
Unmodified Anterior Lachman Test	98.3%	.97
Ankle Horseshoe Compression Wrap	92.6%	.85
Axillary Crutch Fitting	95.7%	.91
Overall	93.1%	.86

Question 2: What is the interrater reliability of psychomotor skill assessment between Athletic Training Education Program faculty and approved clinical instructors?

The results for research question two are presented in Table 11. Percent agreement scores were calculated for each psychomotor skill for each group as well as a combined percent agreement for each group. A kappa statistic was also calculated for each psychomotor skill for each group and for all skills combined for each group.

The percent agreement scores of the psychomotor skills for the faculty group ranged from 87.1% (K = .74; substantial) for the closed basktweave ankle taping procedure to 96.9% (K = .94; almost perfect) for the unmodified Lachman test, with an overall percent agreement of 90.8% (K = .82; almost perfect). The percent agreement scores of the psychomotor skills for the

ACI group ranged from 90.2% (K = .80; substantial) for the closed basketweave ankle taping procedure to 98.8% (K = .98; almost perfect) for the unmodified Lachman test, with an overall percent agreement of 93.2% (K = .86; almost perfect).

The ACI group had the highest percent agreement and kappa scores on three out of four skills with the ankle horseshoe compression wrap having the lowest scores. Percent agreement scores were within five percentage points or less between groups. Kappa scores were within .09 or less for all four psychomotor skills.

Table 11

Percent Agreement and Free-Marginal Kappa for the Disaggregated Groups (Faculty & ACIs)

	Faculty		ACI	
	Percent Agreement	*K*	Percent Agreement	*K*
Closed Basketweave Ankle Taping	87.1	.74	90.2	.80
Unmodified Anterior Lachman Test	96.9	.94	98.8	.98
Ankle Horseshoe Compression Wrap	92.5	.85	91.8	.84
Axillary Crutch Fitting	91.4	.83	96.1	.92
Overall	90.8	.82	93.2	.86

Ancillary Questions

Ancillary Question 1: *What is the interrater reliability of psychomotor skill assessment between ACIs with a bachelor's or master's degree?*

The results for ancillary question one are presented in Table 12. Two degree levels were reported for ACIs: bachelor's and master's degree. Percent agreement scores were calculated for each psychomotor skill within each group (bachelor's and master's) as well as a combined percent agreement score for all skills within each group. A kappa statistic was also calculated for each psychomotor skill within each group as well as a combined kappa score for all skills within each group.

Within the ACIs with bachelor's degrees, percent agreement scores ranged from 94.1% (K = .88; almost perfect) for the closed basketweave ankle taping procedure to 100% (K = 1.0; perfect) for the unmodified Lachman test, with an overall percent agreement of 93.2% (K = .93; almost perfect). Percent agreement scores for the ACI group with a master's degree ranged from 90.1% (K = .81; almost perfect) for the closed basketweave ankle taping procedure to 100% (K = 1.0; perfect) for the unmodified Lachman test, with an overall percent agreement of 93.4% (K = .88; almost perfect).

The ACIs with a bachelor's degree had the higher percent agreement and kappa scores on three out of four skills as well as a higher overall percent agreement and kappa score. The unmodified Lachman test was the only skill with perfect agreement between the two groups. Percent agreement scores were within four percentage points or less between groups. Kappa scores were within .07 or less for all four psychomotor skills.

Table 12

Percent Agreement and Free-Marginal Kappa Scores of Degree Earned for ACIs Disaggregated by Degree Earned

	Degree Earned			
	BS/BA		MS/MA/MEd	
	Percent Agreement	*K*	Percent Agreement	*K*
Closed Basketweave Ankle Taping	94.1	.88	90.1	.81
Unmodified Anterior Lachman Test	100	1.0	100	1.0
Ankle Horseshoe Compression Wrap	94.5	.89	93.7	.87
Axillary Crutch Fitting	99.2	.98	95.4	.91
Overall	96.2	.93	93.4	.88

(n = 82) The terminal degree was not specified for one ACI. Scores for this individual were not included.

Ancillary Question 2: What is the interrater reliability of psychomotor skill assessment between male and female ACIs?

The results for ancillary question two are presented in Table 13. Examining only ACIs, percent agreement scores were calculated for each psychomotor skill for males and for females. A kappa score was also calculated for each psychomotor skill for males and for females as well as a combined kappa score for both groups.

Male percent agreement scores ranged from 93.1% (*K* = .86; almost perfect) for the ankle horseshoe compression wrap to 100% (*K* = .99, almost perfect) for the unmodified Lachman test, with an overall percent agreement of 95.2% (*K* = .90; almost perfect). Female percent agreement scores ranged from 89.2% (*K* = .78; substantial) for the closed basketweave ankle taping procedure to 98.8% (*K* = .98; almost perfect) for the unmodified Lachman test, with an overall percent agreement of 92.5% (*K* = .85; almost perfect).

Overall, males had higher percent agreement and kappa scores for each psychomotor skill as well as a higher overall percent agreement and kappa. Percent agreement scores for female were no less than three percentage points lower than male scores. Kappa scores for females were within .07 of the male scores for all four psychomotor skills.

Table 13

Percent Agreement and Free-Marginal Kappa of Sex for the ACI

	Male		Female	
	%	K	%	K
Closed Basketweave Ankle Taping	93.2	.86	89.2	.78
Unmodified Anterior Lachman Test	100	.99	98.8	.98
Ankle Horseshoe Compression Wrap	93.1	.86	91.8	.84
Axillary Crutch Fitting	97.9	.96	94.5	.89
Overall	95.2	.90	92.5	.85

(*n* = 83)

Ancillary Question 3: What is the interrater reliability of psychomotor skill assessment between ACIs who attended different numbers of approved clinical instructor workshops?

The results for ancillary question three are presented in Table 14. Approved clinical instructor workshop attendance is required only once every three years. Workshop attendance among ACIs was divided into six categories: one through five, and more than five workshops. No ACI attended only five workshops therefore this category is not presented in Table 14. Percent agreement scores were calculated for each psychomotor skill for each group, as well as a combined percent agreement score for each group. A kappa statistic was also calculated for each psychomotor skill for each group as well as a combined kappa score for each group.

The ACIs who attended only one ACI workshop had the lowest percent agreement and kappa scores for all four psychomotor skills and for the overall percent agreement and kappa. The percent agreement scores for ACIs who attended one workshop ranged from 78.1% (K = .56; moderate) for the closed basketweave ankle taping procedure to 90.0% (K = .80; substantial) for the unmodified lachman test, with an overall percent agreement of 83.4% (K = .67; substantial). The percent agreement score for the closed basketweave ankle taping procedure was 16.5 percentage points lower than the score for same skill in the ACI group with the highest score (two ACI workshops). The kappa score was also .33 lower.

The remaining groups (3, 4, and > 5) had very similar percent agreement and kappa scores across all four psychomotor skills. The scores for these three groups were lower than the two workshop group but kappa scores were still in the substantial and almost perfect categories. The procedure with the widest range in kappa scores was the closed basketweave ankle taping

procedure. Participants with the highest kappa (.86; almost perfect) while participants with >5 workshops had kappa score of .78 (substantial). The percent agreement scores for the ankle horseshoe compression wrap varied less than four percentage points within this group. Kappa scores varied .08 or less. Percent agreement scores for the axillary crutch fitting varied less than four percentage points also. Kappa scores varied .07 or less for this group. The overall percent agreement and kappa scores were similar in this group.

Table 14

Percent Agreement and Free-Marginal Kappa for ACI Workshops Attended

	# of Workshops									
	1 (n = 5)		2 (n = 30)		3 (n = 35)		4 (n = 7)		> 5* (n = 5)	
	%	*K*	%	*K*	%	*K*	%	*K*	%	*K*
Closed Basketweave Ankle Taping	78.1	.56	94.6	.89	91.0	.82	93.1	.86	89.1	.78
Unmodified Anterior Lachman Test	90.0	.80	99.9	.99	100	1.0	100	1.0	100	1.0
Ankle Horseshoe Compression Wrap	85.7	.71	95.3	.91	92.4	.85	91.8	.84	91.4	.83
Axillary Crutch Fitting	85.7	.71	96.6	.93	96.8	.94	100	1.0	100	1.0
Overall	83.4	.67	96.0	.92	94.0	.88	95.4	.91	93.8	.88

* No member of the ACI group indicated they attended only five ACI workshops. (n = 82)

Ancillary Question 4: What is the interrater reliability of psychomotor skill assessment between academic faculty who hold a masters or doctoral degree?

The results for ancillary question four are presented in Table 15. Faculty were divided into two groups based on the highest degree held: master's and doctorate. Percent agreement scores were calculated for each psychomotor skill for each group as well as a combined percent agreement score for each group. A kappa statistic was also calculated for each psychomotor skill for each group as well as a combined kappa score for each group.

Faculty members possessing a master's degree had higher percent agreement scores on both the closed basketweave ankle taping procedure and the axillary crutch fitting than the faculty possessing a doctorate degree. Percent agreement scores for the faculty group with a masters degree ranged from 90.9% (K = .82; almost perfect) for the closed basketweave ankle taping procedure, compared to 88.9% and K = .78 (substantial) for the doctoral degreed faculty. Axillary crutch scores in the master's group were 96.3% and K = .93 (almost perfect) while the doctoral degree faculty produced scores of 93.3% and K = .87 (substantial). Scores were nearly identical in the two groups for the remaining two procedures. Overall percent agreement and kappa scores were similar between the two groups (93.3%, K = .87 and 91.6%; K = .83 respectively).

Table 15

Percent Agreement and Free-Marginal Kappa for Faculty Academic Degree Earned

	Academic degree			
	MS/MA/MEd		EdD/PhD	
	%	*K*	%	*K*
Closed Basketweave Ankle Taping	90.9	.82	88.9	.78
Unmodified Anterior Lachman Test	100	1.0	100	1.0
Ankle Horseshoe Compression Wrap	91.6	.83	92.1	.84
Axillary Crutch Fitting	96.3	.93	93.3	.87
Overall	93.3	.87	91.6	.83

Ancillary Question 5: What is the interrater reliability of psychomotor skill assessment between academic faculty with varying lengths of time spent in an ATEP?

The results for ancillary question five are presented in Table 16. Numbers of years as ATEP faculty were divided into three categories: 1 - 5, 6 - 10, and 11 or more. Percent agreement scores were calculated for each psychomotor skill in each category, as well as a combined percent agreement score for each category. A kappa statistic was also calculated for each psychomotor skill in each category as well as a combined kappa score for each category.

The faculty group with 1 – 5 years of experience had both the lowest and the highest percent agreement and kappa scores for the four psychomotor skills. The percent agreement on the ankle horseshoe compression wrap Procedure was 81.0% (K = .62; substantial), compared to faculty with 6 – 10 years of experience (94.5%, K = .89; almost perfect), a 14.5 percentage point difference and a kappa difference of .27. Faculty with 11 or more years had slightly lower scores (92.3%, K = .85; almost perfect).

Less than 10 percentage point separated all three groups for both the Closed Basketweave Ankle Taping Procedure and the Axillary Crutch Fitting Procedure. All groups scored perfect agreement on the Unmodified Anterior Lachman Test.

There was less than a 10 percentage point difference in the overall percent agreement scores between the three groups. Faculty with 6 – 10 years of experience had the highest scores (94%, *K* = .88; almost perfect) among the three groups. Faculty with 11 or more years of experience had similar overall scores, while faculty with 1 – 5 years of experience had the lowest scores (88.5%, *K* = .77; substantial). There was more consistency of scores between procedures in the faculty group with 11 or more years of experience than in the other two groups, with faculty having 1 – 5 years of experience as the most inconsistent.

Table 16

Percent Agreement and Free-Marginal Kappa of Years Serving as Faculty

	Years as faculty					
	1-5		6-10		11 or more	
	%	K	%	K	%	K
Closed Basketweave Ankle Taping	81.8	.64	86.5	.73	91.3	.83
Unmodified Anterior Lachman Test	100	1.0	100	1.0	100	1.0
Ankle Horseshoe Compression Wrap	81.0	.62	94.5	.89	92.3	.85
Axillary Crutch Fitting	100	1.0	94.5	.89	94.8	.90
Overall	88.5	.77	94.0	.88	93.6	.87

Post Hoc Analysis

A post hoc analysis was performed to analyze the percent agreement and kappa scores for criteria four and five of the closed basketweave ankle taping procedure and criterion seven ankle horseshoe compression wrap. A review of the raw data revealed that there was an increased number of "No" ratings for these three criteria. Criterion four of the closed basketweave ankle taping Procedure was performed incorrectly and criterion seven of the ankle horseshoe compression wrap was omitted on the video

according to the criteria described in the *Practical Exam Preparation Guide of Clinical Skills for Athletic Training* (Amato, Hawkins, and Cole, 2002). Criterion five of the closed basketweave ankle taping procedure was also analyzed due to the increased number of raters who selected "No" for this criterion. Nineteen out of 115 raters chose "No" on criterion four and 25 out of 115 chose "No" on criterion five of the closed basketweave ankle taping procedure. Thirty-six raters out of 115 chose "No" for criterion seven of the ankle horseshoe compression wrap. The results of the post hoc analysis are presented in Table 17. Percent agreement and kappa scores were calculated for each criterion.

Table 17

Analysis of Individual Psychomotor Skill Criteria

Criteria	Percent Agreement	*K*
Closed Basketweave Ankle Taping Procedure		
Criterion 4	72.2%	.44
Criterion 5	65.7%	.31
Ankle Horseshoe Compression Wrap		
Criterion 7	56.6%	.13

Chapter Summary

Results of procedures conducted to review interrater reliability were reported in this chapter. Analysis and report of results included descriptive statistics for the entire sample as well as a break down into smaller groups. These results are discussed in chapter 5, along with implications for future research and practical application of this study.

CHAPTER FIVE
DISCUSSION

Assessment of psychomotor skills is a necessary and important component of athletic training education in determining the competency and proficiency of ATS. The responsibility for making this determination includes ATEP faculty and ACIs. In 2007, the BOC modified the certification examination procedure from a three part examination to a complete computer based examination (CBE). The modification uses multiple choice questions, but replaces the face-to-face practical and written simulation sections with a hybrid section. This change from a face-to-face practical skill assessment to an online assessment has placed greater emphasis on the psychomotor skill assessment by ATEP faculty and ACIs. The goal of this study was to review the interrater reliability of athletic training faculty and ACIs in the assessment of psychomotor skills specific to athletic training.

The assessment instrument used in this study consisted of a video demonstration of four psychomotor skills and a criteria rubric for each skill. The rubric for the four psychomotor skills was used with permission from the publisher and taken from the text *Practical Exam Preparation Guide of Clinical Skills for Athletic Training* (Amato, Hawkins, and Cole, 2002). Athletic training faculty and ACIs were solicited to complete a demographics questionnaire, review criteria for each psychomotor skill, view the video demonstrations, and then complete a scoring rubric pertaining to each skill. Results for each skill was analyzed for interrater reliability for the entire sample in aggregate and then disaggregated by position (faculty and ACI). Data was also analyzed to review the interrater reliability of several sets of subsamples. Analytical

results addressing to primary research questions are discussed first, followed by a discussion of five ancillary questions. A post hoc analysis was also completed and this analysis is discussed after the research questions. The final section of this chapter presents limitations and recommendations for future research. The chapter ends with a brief conclusion.

Research Questions

Question 1: What is the interrater reliability of psychomotor skill assessment among Athletic Training Education Program faculty and approved clinical instructors?

Overall, percent agreement and kappa scores were high for the entire sample. This is an unanticipated result. It was expected that overall interrater reliability would be lower than was demonstrated. Psychomotor skill assessment is an integral part of all ATEP programs and assessment performed by both faculty and ACIs. The high overall interrater reliability found in this study suggests that face-to-face assessment of ATS by faculty and ACIs underscores the professional ability of these individuals to qualify students as competent and proficient in psychomotor skills.

Foster and Leslie (1992) reported that clinical educators with teaching degrees or training in pedagogy placed a higher value on teaching and assessing students. They also reported that these individuals with advanced training treated the clinical setting as an educational opportunity for the student. The authors also reported that greater amounts of experience in teaching methodology (> 6 years) contributed to an increased comfort level in clinical education. Craig (2006) concurs that teaching effectiveness is influenced by experience level. In the current study, a majority of the participants (n = 79; 69%) had 6 or more years of experience and a

master's/doctorate degree (n = 77; 67%). The findings by Foster and Leslie (1992) and Craig (2006) regarding experience level and advanced training may align themselves with the results of this study. A positive relationship may exist between advanced training and experience level in the ability to educate and assess ATS.

Question 2: What is the interrater reliability of psychomotor skill assessment between Athletic Training Education Program faculty and approved clinical instructors?

The overall percent agreement and kappa scores were similar between both groups with the faculty having slightly lower percent agreement and kappa scores. It was expected that the scores would be lower for the ACI group when compared to faculty group due to the faculty's continual involvement in formative and summative assessment techniques. It was also assumed that ACIs would have a lower degree of reliability if the number of participants in this study were from multiple ATEPs. With the autonomy in assessment practices that is allowed in ATEP, there may be an opportunity for multiple different assessment techniques to be used by ACIs in assessing ATS, thus reducing the reliability scores. Based on the results, this expectation was unfounded.

Laurent and Weidner (2001), Weidner and Henning (2005), and Sliwinski et al (2004) all agree that ACIs are critical in the professional development of ATS. Carefully designed clinical education learning opportunities are critical for an ATEP. Also having ACIs that can instruct, mentor and assess ATS is necessary to promote these learning opportunities in clinical education. Laurent and Weidner (2001) feel that ACIs may lack the ability to understand the importance

and the direction that must be taken to instruct ATS. To address this understanding, Weidner and Henning (2005) describe the importance of the ACI workshop as the means to train the ACIs to effectively and accurately teach and assess ATS. With the advent of ACI workshops, any disparity between the two groups (faculty and ACIs) in teaching and assessment of ATS that may have existed in the past may be decreasing.

Conversely, Sliwinski, Schultze, Hansen, Malta, and Babyar (2004) reported that there may be a disconnect between academic faculty and clinical instructors in the assessment of students. The authors believed that a raters implicit criteria relating to the skill rather than the published criteria may influence the assessment ability of the clinical faculty, thereby producing different outcomes than the academic faculty. In the present study, scores between academic faculty and ACIs were similar for this study suggesting that this was not the case.

Ancillary Questions

Ancillary Question 1: *What is the interrater reliability of psychomotor skill assessment between ACIs with a bachelors or masters degree?*

Both groups had high reliability scores with the bachelor's prepared ACIs having higher scores for all four psychomotor skills. Assessment practice is a component in teaching methodology (Vendrely & Carter, 2004), and it was expected that ACIs with a master's degree would have higher scores based on additional education obtained through the advanced degree. Foster and Leslie (1992) found that ATCs with a master's degree demonstrated an increased effectiveness in educating ATS. Craig (2006) also noted that training in teaching methodology may influence how ACIs educate ATS, whether the training occurred in the undergraduate or

graduate coursework. The results of present study contradict the conclusions presented by Foster and Leslie (1992) however, pedagogical education of the ACIs in this study is unknown. Therefore, one must interpret the results for this question with caution in determining whether individuals with a bachelor's degree are more consistent at evaluating psychomotor skills.

Ancillary Question 2: What is the interrater reliability of psychomotor skill assessment between male and female ACIs?

Overall, there was high agreement between the male and female ACI groups with less than a three percentage point difference between the two groups. The overall percent agreement (92.5%) and kappa (.85; almost perfect) for females (n = 43) was slightly lower than the percent agreement (95.2%) and kappa (.90; almost perfect) for males (n = 40) in the ACI group. The sample sizes were nearly equal. Laurent and Weidner (2001) reviewed clinical instructor characteristics that were helpful in the instruction of ATS. They found that sex differences did not contribute in any way in the teaching of ATS. The consistency between genders in scoring is a positive finding and supports Laurent and Weidner's results.

Ancillary Question 3: What is the interrater reliability of psychomotor skill assessment between ACIs who attended different numbers of approved clinical instructor workshops?

Training ATCs in assessment methods is an important process in aligning these individuals with the ATEP. The ACI workshop is intended to provide a mechanism to train ATCs in the instructional and assessment practices of that ATEP. The goal of the workshop is to provide a more "consistent method of teaching and measuring the acquisition of skills used in

athletic training clinical education" (NATA, 2006a, p. xi). Workshop attendance is required every three years.

It could be assumed that attendance at more ACI workshops would lead to an increase in percent agreement and kappa scores. Although all of the reliability scores in the present study were in the substantial to almost perfect categories, scores did not increase as the attendance at workshops increased. While the low number of raters in the single workshop category (n = 5) may have negatively affected reliability scores in this group, the same cannot be said for the four and greater-than-five workshops groups (n = 7 and 5, respectively) where scores were higher. Since greater numbers of workshops may reflect more years of service, it may be that skills gained and/or reinforced through workshops and experience provided the higher reliability in these groups, in spite of the low number of participants. In contrast, Shirk et al. (2006) concluded that low reliability scores obtained in their study could not be attributed to experience. These authors used a test-retest methodology with experienced (6 or more years in clinical practice) and inexperienced (less than 2 years of clinical practice) ATCs performing a clinical evaluation. Shirk et al. attributed the low reliability scores to lack of validity in their assessment instrument.

Ancillary Question 4: *What is the interrater reliability of psychomotor skill assessment between academic faculty who hold a masters or doctoral degree?*

The percent agreement and kappa scores between faculty who have a master's degree (n = 21) and faculty who have a doctoral degree (n = 10) were high and similar across the four psychomotor skills and in overall reliability. According to Seegmiller (2006), athletic training

faculty spend large amounts of time fostering and assessing students' abilities in athletic training in comparison to research and service. This increased time reviewing and analyzing student work may have contributed to the consistency between these two groups.

In the interrater reliability studies reviewed for this research, no study explored the differences between the rating ability of academic faculty with a master's or doctoral degree regarding the assessment psychomotor skills. In a study by Wimer (2005), interrater reliability was reviewed between athletic training accreditation site visitors. Athletic Training accreditation site visitors are ATCs who review ATEPs for compliance to accreditation standards. Accreditation site visitors were to review hypothetical scenarios and determine compliance or noncompliance based on accreditation standards. The author analyzed the results for interrater reliability among between several subgroups. In analyzing the results, the author stated that there was not a difference in the ratings between site visitors who held a master's or doctoral degree. This author suggested that the demographic factor of degree did not influence that rating ability of the participants.

Ancillary Question 5: What is the interrater reliability of psychomotor skill assessment between academic faculty with varying lengths of time spent in an ATEP?

A majority of the faculty group (n = 27) had at least six or more years of experience. It was thought that as the number of years serving as a faculty member increased, reliability would also increase. This was not the case. It was anticipated that faculty with the lowest years of service would have the lowest reliability scores which was true for two of the four skills.

However, the same group had the highest reliability scores on the remaining two skills. There was a wider range of scores between skills in the 1 – 5 years of service group compared to the 11 or more year group. It could be assumed based on these results that academic faculty experience is not a factor in reliability scores of the participants. This is similar to the results of the study by Shirk et al. (2006), wherein the authors found that years of experience was not a factor in determining reliability scores. While it may be tempting to conclude that length of service contributes to the consistently high reliability scores that were found in the 6 years and above groups in this study.

Post Hoc Analysis

Of the 29 skill criteria in the PSAI, one was deliberately performed incorrectly and one was omitted entirely. Using the PSAI in this study, considerably lower interrater reliability scores were found for three of the 29 criteria. Two of these criteria included one that was performed incorrectly (criterion four of the Closed Basketweave Ankle Taping Procedure) and one that was omitted (criterion seven of the Ankle Horseshoe Compression Wrap). A majority of the raters indicated that the incorrect criterion was performed correctly, and that the omitted criterion was in fact performed.

The third criterion having lower reliability scores was criterion five of the Closed Basketweave Ankle Taping Procedure, which was demonstrated correctly according to the published rubric (Amato, Hawkins & Cole, 2002). There are multiple ways to perform the ankle taping procedure that may be considered correct by faculty and ACIs (Beam, 2006). This may

have influenced the rater's choice on the two criteria in the closed Basketweave Ankle Taping Procedure, thus producing the low reliability scores.

Criterion seven of the Ankle Horseshoe Compression Wrap was obviously omitted. This criterion was the last in the series of seven criteria to be performed and the 22nd criteria performed in the assessment sequence. The placement of this omission may have influenced the reliability scores. One factor that may have influenced the raters is that 20 of 21 criteria that were previously rated were all performed according to the published rubrics. Raters may have become accustomed to viewing the psychomotor skills being performed correctly. There is also variability between texts on this criterion in the performance of this skill, which may have contributed to the low reliability score. Peeler and Anderson (2008) found similar rating disparities in their study. The authors found that individual raters varied in their application of assessment criteria, in the evaluation of a clinical testing procedure (Peeler & Anderson, 2008). They attributed their lower reliability scores to this variability.

The results of this research indicated an overall high interrater reliability among and between groups suggesting that faculty and ACIs can reliably assess ATS ability to perform psychomotor skills. Recent literature in athletic training has described both high and low levels of interrater reliability. Research studies that have high reliability attribute this result to rater training, content validated assessment tools, similarities among the raters, and standardized patients (Lafave et al., 2008; Lagumen et al., 2008 & Boyle et al., 2003). This is similar to the current study in that content-validated criteria were used and there was homogeneity among raters and patient consistency. Studies reporting low levels of reliability attributed this to a lack

of validation in the assessment tools used (Peeler & Anderson, 2008; Shirk et al., 2006). While the PSAI was validated for content, post hoc analysis suggests that the construction of the instrument itself has an influence on the reliability scores. Raters had high reliability on the criteria performed correctly and lower reliability on criteria performed incorrectly or omitted. Since most of the criteria (27 of 29) were performed correctly according to content validated standards (Amato, Hawkins & Cole, 2002), the raters may have been swayed to agree with the correct performance of the criteria as opposed to disagreeing with incorrect or omitted performance. This could skew the overall reliability scores.

Limitations

Several limitations are associated with this study. The limitations are discussed to aid in the understanding of the results and implications of this research.

Evaluating the Instrument

The instrument developed for this study included a video demonstration of the four psychomotor skills. Criteria for each skill was presented for review prior to the video and again after the video to be scored as "Yes" for completed correctly or as "No" for not completed correctly. The first psychomotor skill, closed basketweave ankle taping, had 11 criteria, followed by the unmodified anterior Lachman test with 4 criteria. The third skill, ankle horseshoe compression, and the fourth skill, axillary crutch fitting, both had 7 criteria each.

The highest reliability scores were found in the procedure with the lowest number of criteria (Unmodified Anterior Lachman Test) while the lowest scores were found in the

procedure with the largest number of criteria (Closed Basketweave Ankle Taping Procedure). This suggests that the design of the instrument may have influenced the reliability scores regardless of sample characteristics. Eleven criteria may be more than raters can keep in memory during the evaluation process, while four are more easily retained. Both Lafave et al (2008) and Portney and Watkins (2000) suggest that the number of criteria can influence the reliability scores. This study used content-validated criteria (Amoto, Hawkins, & Cole, 2002), however, variation in procedural criteria exists between textbooks used in athletic training education. Low reliability scores as discussed in the post hoc analysis may be a reflection of these variations.

Rater training is a key component in any interrater reliability study (Lagumen et al., 2008 & Portney & Watkins, 2000). Rater understanding of the purpose, design, as well as the implementation of the instrument is necessary for the reduction of measurement error. Peeler and Anderson (2008) hold that rater training is critical in the reduction of subjectivity and variations in the understanding the grading criteria. In the present study, raters were practicing ATCs experienced with the psychomotor skills being evaluated. Instrument training consisted of allowing the raters to review the scoring criteria for an unlimited amount of time before viewing the video. The criteria were not available while the rater viewed the video. Once the video was complete, the rater completed the scoring rubric. If the scoring rubrics were available during the video, the reliability scores may have been higher for all groups in this study.

Practical Significance

This study contributes to the research on interrater reliability in several ways. The study affirms that high reliability exists in the published criteria (Amato, Hawkins & Cole, 2002) for

the four psychomotor skills used in this study. The use of standardized criteria in the assessment of skill performance may aid the rater in reducing subjectivity in the determination of correct performance. The results of this study also provide data on the effect of different experience levels, rater training, and degree on the rating ability of both faculty and ACI. Additionally, the methods used in this study (i.e. video demonstration) may provide a template for training raters in assessment protocols, as well as for research in determining the reliability of assessment instruments used in athletic training education.

Recommendations for Future Research

The results of this research suggest several directions for further research. The overall interrater reliability for this study was high for the entire sample as well within the faculty and ACI groups. This suggests that majority of these individuals agreed in their analysis of the performance for this individual for these four psychomotor skills. The numerous correct criteria may be hiding the effect that this study was intended to review. Low agreement scores were found for the only two incorrect/omitted criteria. An alteration to the instrument having more criteria performed incorrectly, and/or using only clinically significant errors may create a more authentic assessment procedure. Butterwick et al. (2006) feel that accurate assessment of technical is important in established competent entry level professionals. The authors also believe having objective structured assessment tools is essential to this process. In addition, repeating the study with only one psychomotor skill would allow for a more detailed analysis by allowing the rater to concentrate on one skill instead of multiple skills.

Using senior level ATS as the individuals performing the psychomotor skills would create a more realistic evaluation. Additionally, a sequence of multiple students performing the same psychomotor skill but with different errors would enhance the realism of the evaluation. This method was used in the research of Lafave et al. (2008). They found that having multiple raters rating multiple students was a more authentic research design. Having the grading rubric available for reference while viewing the videos would also be more realistic, and may provide a more authentic assessment setting.

Modifying the experimental design by having equal numbers of faculty and ACI in the sample would allow for more balanced comparisons between the two groups. Within these two groups, examining more attributes of the groups such as specific degree earned, teaching methods used in the curriculum, specific pedagogical training, and other types of professional certifications would also provide more specific data for comparison within and between these groups. Craig (2006), discussed the need to have more specific demographic information related to the participants in an effort to determine the relationship that this information may have on outcomes of her study.

Research in interrater reliability in athletic training is limited. Additional research is necessary to develop specific assessment tools for interrater reliability.

Conclusions

This study explored the interrater reliability of psychomotor skill assessment within athletic training education. Both the faculty group and ACI group had high interrater reliability within each group and between groups for all four psychomotor skills evaluated. Sample

subgroups such as sex, degree held, and years of service also exhibited high interrater reliability within and between each subgroup on all four psychomotor skills. The assessment instrument was found to be useful in assessing interrater reliability; however, caution is necessary in interpreting the result due to the design of the instrument. Changes to both the instrument and experimental design are recommended.

APPENDIX A: PSYCHOMOTOR SKILL ASSESSMENT INSTRUMENT

Title of Research: Interrater Reliability in Psychomotor Skill Assessment in Athletic Training.

Principle Investigator: Jason Craddock, MS, ATC, LAT, CSCS.

Faculty Advisor: Dr. David Boote, Ph.D.

Department: Educational Studies at University of Central Florida

You are being asked to participate in a research project as part of a doctoral studies dissertation, conducted thorough the University of Central Florida's department of Educational Studies. This study titled: Interrater Reliability in Psychomotor Skill Assessment in Athletic Training will take the participant approximately 15-20 minutes to complete. The University requires that you give your agreement to participate in this project. Please read the remainder of this informed consent form and if you agree to participate "click" on the next button on the bottom of the page to see the electronic survey. By "clicking" on the "I Agree" followed by the "Next" button, you will be giving your informed consent to participate and attesting to the fact that you are at least 18 years of age.

The purpose of this study is to examine interrater reliability in psychomotor skill assessment by athletic training faculty and approved clinical instructors. The investigator is evaluating the congruency of rating ability among athletic training educators and approved clinical instructors. Each participant will be asked to complete a few demographic questions and view a grading rubric based on each specific psychomotor skill to be evaluated. After reviewing the grading rubric the participant will view a streamed video clip of the psychomotor skill and then mark the electronic grading rubric. There will be a total of four grading rubrics to complete. Once the electronic grading rubric/survey is completed you will be asked to submit the survey which will complete the participant's involvement in the study.

Anyone who agrees to participate in this study has the right to withdraw from completing the study at any time with no penalty from the investigator. Each participant while taking the survey can decide not to complete any question within the electronic survey if they desire, without penalty. Upon completion and submission of the electronic survey, you will then have the ability to request the results from this study. Refusal to participate in this study will have no effect on any future services you may be entitled to from the University of Central Florida or the researcher.

Benefits

There are no personal benefits to the participant however there are benefits to the athletic training profession in regards to educational practice in assessment within didactic and clinical education.

Risks and Discomforts

No known risks or discomforts.

Confidentiality and Records
In no part of the study will your personal information or results be linked to your name. The investigators will be the only people who have access to the research data. Your results will be collected confidentially by a survey instrument. The researcher will not be able to track the responses back to any individual. The analysis of the data will be in aggregate form and individual answers will not be published.

Compensation
None

Contact Information
If you have any questions regarding this study, please contact
Jason Craddock, MS, ATC, LAT, CSCS (239)590-7535
David Boote, Ph.D. (407)823-4160

This research has been reviewed and approved by the UCF Institutional Review Board. Questions or concerns about research participants' rights may be directed to the UCF IRB office, University of Central Florida, Office of Research & Commercialization, 12201 Research Parkway, Suite 501, Orlando, FL 32826-3246 (407)823-2901.

*

◯ I Agree ◯ I Do Not Agree

Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

This survey is intended for Athletic Training faculty members or approved clinical instructors who serve an Athletic Training Education Program (ATEP).

***Age**

○ 20-30 ○ 31-40 ○ 41-50 ○ 51 & up

***Sex**

○ Male ○ Female

***How many years have you practiced as a certified athletic trainer?**

○ 1-5 ○ 6-10 ○ 11-15 ○ 16-20 ○ 21-25 ○ 25 or more

***Highest Degree Earned:**

○ BS/BA ○ MS/MA/MEd ○ EdD/PhD ○ Other

***Please indicate whether you are full-time Athletic Training faculty member who is a certified athletic trainer or an approved clinical instructor who is a certified athletic trainer for an ATEP.**

○ ATEP Faculty ○ ACI ○ Neither

***How many years have you served as a full-time academic faculty member in an ATEP?**

○ 1-5 ○ 6-10 ○ 11 or more ○ Not academic faculty in ATEP

***How many years have you served as an ACI who is a certified athletic trainer for an ATEP?**

○ 1-5 ○ 6-10 ○ 11 or more ○ Not an ACI

***If ACI, how many ACI workshops have you participated in?**

○ 1 ○ 2 ○ 3 ○ 4 ○ 5 ○ More than 5 ○ Not Applicable

***If an ACI, what is your primary practice setting?**

- ◯ High School
- ◯ Intercollegiate
- ◯ Professional
- ◯ Clinical
- ◯ Industrial
- ◯ Other
- ◯ Not Applicable

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

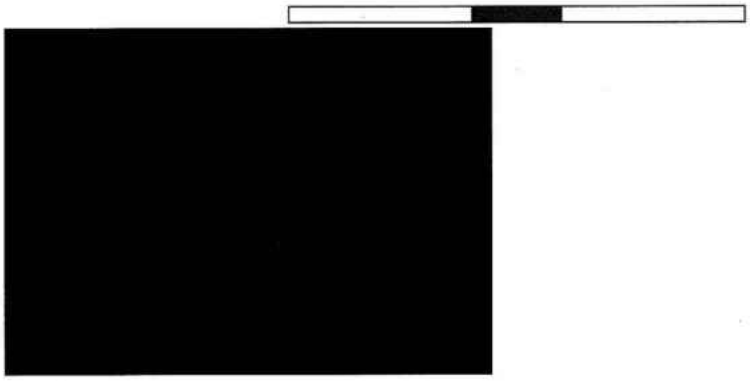
Review the following criteria to prepare to evaluate the ankle taping video:

- Seated with ankle dorsiflexed to 0 deg
- Places heel and lace pads
- Secures pads with layer of prewrap
- Two anchors around base of gastrocnemius muscle
- One anchor around instep, proximal to base of fifth metatarsal
- Applies a stirrup
- Places a horseshoe around the foot just below malleoli
- Repeats above 2 steps two additional times
- Applies a figure eight starting at the dorsal aspect of ankle followed by heel-locks.
- Closes taping with horseshoes, overlapping by half the tape from distal to proximal
- Places final anchor around the instep, proximal to the base of the fifth metatarsal

After reviewing the above criteria, select the next button below to continue to view the ankle taping video.

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance



<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

Select the appropriate box below each criterion to provide your evaluation of the Ankle Taping Procedure.

Seated with ankle dorsiflexed to 0 deg

○ Yes ○ No

Places heel and lace pads

○ Yes ○ No

Secures pads with layer of prewrap

○ Yes ○ No

Two anchors around base of gastrocnemius muscle

○ Yes ○ No

One anchor around instep, proximal to base of fifth metatarsal

○ Yes ○ No

Applies a stirrup

○ Yes ○ No

Repeats above 2 steps two additional times

○ Yes ○ No

Places a horseshoe around the foot just below malleoli

◯ Yes ◯ No

Applies a figure eight starting at the dorsal aspect of ankle followed by heel-locks.

◯ Yes ◯ No

Closes taping with horseshoes, overlapping by half the tape from distal to proximal

◯ Yes ◯ No

Places final anchor around the instep, proximal to the base of the fifth metatarsal

◯ Yes ◯ No

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

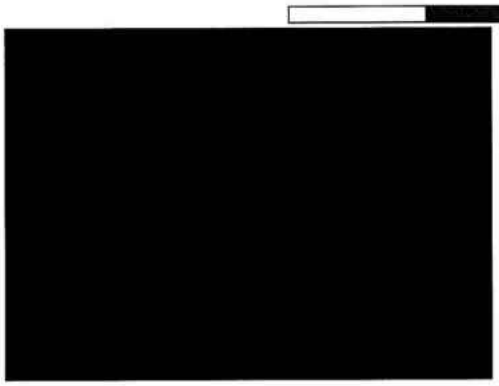
Review the following criteria to prepare to evaluate the Lachman special test video:

- Athlete is supine with knee in 10 – 25 degrees of flexion
- Stabilizes posterior calf with one hand
- Stabilizes anteriorly on distal femur with other hand
- Attempts to displace tibia on femur

After reviewing the above criteria, select the next button below to continue to view the Lachman video.

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance



<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

Select the appropriate box below each criterion to provide your evaluation of the Lachman Test.

***Athlete is supine with knee in 10 – 25 degrees of flexion**

○ Yes ○ No

***Stabilizes posterior calf with one hand**

○ Yes ○ No

***Stabilizes anteriorly on distal femur with other hand**

○ Yes ○ No

***Attempts to displace tibia on femur**

○ Yes ○ No

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

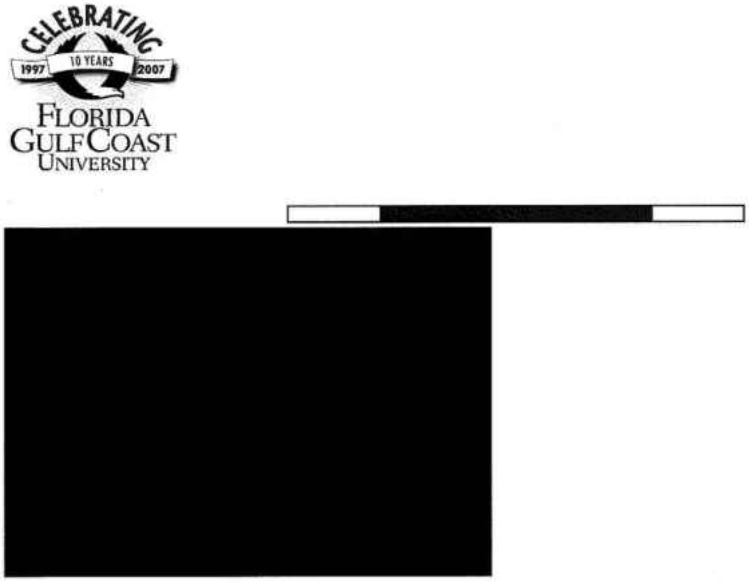
Review the following criteria to prepare to evaluate the Ankle Horseshoe pad video:

- Seated with ankle dorsiflexed to 90 degrees
- Applies horseshoe pad to lateral malleolus
- Applies compression wrap just proximal to toes
- Applies compression wrap proximally to lower third of calf
- Overlaps compression wrap by half
- Secures compression wrap with tape
- Checks distal pulse/sensation/capillary refill

After reviewing the above criteria, select the next button below to continue view the video.

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance



<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

Select the appropriate box below each criterion to provide your evaluation of the Horeshoe Pad technique.

***Seated with ankle dorsiflexed to 90 degrees**

○ Yes ○ No

***Applies horseshoe pad to lateral malleolus**

○ Yes ○ No

***Applies compression wrap just proximal to toes**

○ Yes ○ No

***Applies compression wrap proximally to lower third of calf**

○ Yes ○ No

Overlaps compression wrap by half

○ Yes ○ No

***Secures compression wrap with tape**

○ Yes ○ No

***Checks distal pulse/sensation/capillary refill**

○ Yes ○ No

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

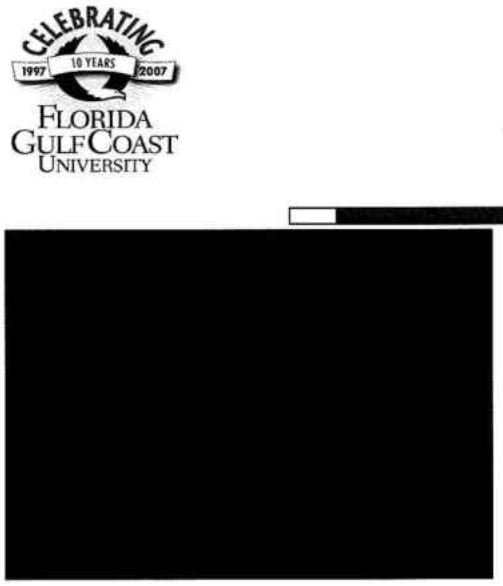
Review the following criteria to prepare to evaluate the Crutch Fitting video:

- Feet shoulder width apart
- Wearing low-heeled/flat shoes
- Places crutch 2-4 inches in front of involved limb
- Places crutch 4-6 inches to the side of the involved limb
- Places axillary pad 1-1.5 inches below axilla
- Places elbow in approx. 30 deg of flexion
- Tightens all fasteners

After reviewing the above criteria, select the next button below to continue to view the video.

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance



<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

Select the appropriate box below each criterion to provide your evaluation of the crutch fitting procedure.

***Feet shoulder width apart**

○ Yes ○ No

***Wearing low-heeled/flat shoes**

○ Yes ○ No

***Places crutch 2-4 inches in front of involved limb**

○ Yes ○ No

***Places crutch 4-6 inches to the side of the involved limb**

○ Yes ○ No

***Places axillary pad 1-1.5 inches below axilla**

○ Yes ○ No

***Places elbow in approx. 30 deg of flexion**

○ Yes ○ No

***Tightens all fasteners**

○ Yes ○ No

<< Back Next >>

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

Thank you for taking the survey.

Official FGCU Survey prepared by the Office of Planning & Institutional Performance

APPENDIX B: IRB APPROVAL LETTER

University of Central Florida Institutional Review Board
Office of Research & Commercialization 12201 Research
Parkway, Suite 501 Orlando, Florida 32826-3246
Telephone: 407-823-2901, 407-882-2012 or 407-882-
2276 www.research.ucf.edu/compliance/irb.ht*ml*

University of Central Florida

Notice of Exempt Review Status

From: ***UCF Institutional Review Board FWA00000351, Exp. 5/07/10, IRB00001138***

To: ***Jason C Craddock***

Date: ***September 20, 2007***

IRB Number: **SBE-07-05181**

Study Title: **INTERRATER RELIABILITY OF PSYCHOMOTORSKAPPAILL ASSESSMENT IN ATHLETIC TRAINING**

Dear Researcher:

Your research protocol was reviewed by the IRB Vice-chair on 9/20/2007. Per federal regulations, 45 CFR 46.101, your study has been determined to be **minimal risk for human subjects and exempt** from further IRB review or renewal unless you later wish to add the use of identifiers or change the protocol procedures in a way that might increase risk to participants. Before making any changes to your study, call the IRB office to discuss the changes.
A change which incorporates the use of identifiers may mean the study is no longer exempt, thus requiring the submission of a new application to change the classification to expedited if the risk is still minimal. Please submit the Termination/Final Report form when the study has been completed. All forms may be completed and submitted online at https://iris.research.ucf.edu.

The category for which exempt status has been determined for this protocol is as follows:

2. Research involving the use of educational tests (cognitive, diagnostic, aptitude, achievement), survey or interview procedures, or the observation of public behavior, so long as confidentiality is maintained.

(i) Information obtained is recorded in such a manner that the subject cannot be identified, directly or through identifiers linked to the subject, **and/or**

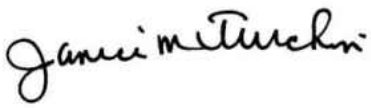
(ii) Subject's responses, if known outside the research would not reasonably place the subject at risk of criminal or civil liability or be damaging to the subject's financial standing or employability or reputation.

A **waiver of documentation of consent** has been approved for all subjects. Participants do not have to sign a consent form, but the IRB requires that you give participants a copy of the IRB-approved consent form, letter, information sheet, or statement of voluntary consent at the top of the survey.

All data, which may include signed consent form documents, must be retained in a locked file cabinet for a minimum of three years (six if HIPAA applies) past the completion of this research. Any links to the identification of participants should be maintained on a password-protected computer if electronic information is used. Additional requirements may be imposed by your funding agency, your department, or other entities. Access to data is limited to authorized individuals listed as key study personnel.

On behalf of Tracy Dietz, Ph.D., UCF IRB Chair, this letter is signed by:

Signature applied by Janice Turchin on 09/20/2007 02:43:25 PM EDT



IRB Coordinator

APPENDIX C: PUBLISHER APPROVAL

Permission was granted from SLACK Incorporated to use four psychomotor skills from the *Practical Exam Preparation Guide of Clinical Skills for Athletic Training* (Amato, Hawkins, & Cole, 2002). This material is copyrighted, and any further reproduction or distribution is prohibited.

SLACK
INCORPORATED

April 26, 2007

Jason Craddock
Florida Gulf Coast University
jcraddoc@fgcu.edu
1-239-590-7535

Reference #: B000647581
Material Requested: Adaptation of 9 worksheets on pages 27, 38, 44, 96, 234, 245, 278, 308, and 312
Usage Requested: Used for research purposes for a study that will be submitted as a dissertation as well as submitted for publication and presentation; materials requested are NOT to be published or used in any presentation. Materials will be placed on a password protected website for research purposes only.
Citation: Amato HK, Hawkins CD, Cole SL. *Practical Exam Preparation Guide of Clinical Skills for Athletic Training.* Thorofare, NJ: SLACK Incorporated; 2002.

Dear Mr. Craddock:
Permission is granted for the requested materials and usage listed above, subject to the following conditions:

- Permission is granted for one-time use only for one research study.
- Permission is granted for non-exclusive use. Access to the material must only be available through a password protected website to ensure limited access to the materials; at no time may the materials appear on a general website. Materials must be removed after 1 year from the date on this letter.
- The following credit line must be displayed: Adapted with permission from SLACK Incorporated: CITATION. See above for citation information.
- The fee for this use is **$0.00 USD**. This offer is valid for 180 days from the date on this letter. If the requestor does not sign, return, and issue payment during this period, then the permission is rescinded.
- Payment is non-refundable. Payment can be made via credit card or check. Checks are payable to SLACK Incorporated, 6900 Grove Rd, Thorofare, NJ 08086, USA (Fed Tax ID#22301-45602). Fill in credit card information below (we accept AmEx, Visa, or MC):

  Card #: _____ Exp Date: _____

  Name on the card: _____ SVC Code: _____

Please sign and date below, keep a copy for your records, and fax to Attn: Permissions Department. Please include your reference number on all correspondence and payment information. A copy of this form **MUST** accompany payment.

Requestor accepts conditions above:

Signature: _____ **Date:** 5/21/07

Sincerely,
SLACK Incorporated
Permissions Department

APPENDIX D: CODING TABLE

Table 18 *Demographic Coding*

Variable	Response	Code
Age	20 -30	1
	31 – 40	2
	41 – 50	3
	51 & up	4
Sex	Male	1
	Female	2
Years as ATC	1 – 5	1
	6 – 10	2
	11 – 15	3
	16 – 20	4
	21 – 25	5
	25 or more	6
Highest Degree Earned	BS/BA	1
	MS/MA/MEd	2
	EdD/PhD	3
	Other	4
Faculty Status	ATEP Faculty	1
	ACI	2
	Neither	3
Years as Faculty	1 – 5	1
	6 – 10	2
	11 or more	3
	Not Faculty	4
Years as ACI	1 – 5	1
	6 – 10	2
	11 or more	3
	Not ACI	4
ACI Workappashops	1	1
	2	2
	3	3
	4	4
	5	5
	More than 5	6
	N/A	7
ACI Primary Workappa Setting	High School	1
	Intercollegiate	2
	Professional	3
	Clinical	4
	Industrial	5
	Other	6
	N/A	7

REFERENCES

Alexander, H. (1996). Physiotherapy student clinical education: The influence of subjective judgements on observational... *Assessment & Evaluation in Higher Education*, *21*(4), 357. Retrieved July 22, 2008, from Academic Search Complete database.

Amato, H.K., Hawkins, C.V., & Cole, S. (2002). *Practical examination guide of clinical skills for athletic training.* Thorofare, NJ: SLACK Inc.

Arnold, B.L., Gansnedar, B.M., & Perrin, D.H. (2005). *Research methods in athletic training.* Philadelphia, PA: FA Davis.

Beam, J.W. (2006). *Orthopedic, taping, wrapping, bracing & padding.* Philadelphia, PA: FA Davis.

Board of Certification. Inc. (2004). Role Delineation Study. Omaha, NE.

Bowen, J.L., (2006). Medical education: Educational strategies to promote critical diagnostic reasoning. *The New England Journal of Medicine. 355*(21), 2217 – 2226.

Boyle K.L., Witt, P., & Rieger-Krugh, C. (2003). Intrarater and interrater reliability of the Beighton and Horan Joint Mobility Index. *Journal of Athletic Training, 38*(4), 281-285.

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

Brumels, K., & Beach, A. (2008). Role orientation of certified athletic trainers at institutions of higher education. *Athletic Training Education Journal, 1,* 5-12.

Buccieri, K.M, Schultze, K., Dungey, J., Kolodziej, T., Malta, S., Marocco, S., et. al. (2006). Self-reported characteristics of physical therapy clinical instructors: A comparison to the american physical therapy association's guidelines and self-assessments for clinical education. *Journal of Physical Therapy Education, 20,* 1, 47-55.

Butterwick, D.J., Paskevich, D.M., Lagumen, N.G., Vallevand, A.L.C., & Lafave, M.R. (2006). Development of content-valid technical skill assessment instruments for athletic taping skills. *Journal of Allied Health, 35*(3), 147-153.

Castle Worldwide, Inc. (2008). 2007 Annual Summary: Board of certification (BOC) certification examination for athletic trainers. *Board of Certification*. Retrieved October 16, 2008, from http://www.bocatc.org/images/stories/public/2007examreport.pdf

Clapper, D.C., & Harris, L.L. (2008). Reliability and validity of an instrument to describe burnout among collegiate athletic trainers. *Journal of Athletic Training, 43*(1), 62-69.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*, 37-46.

Commission Accreditation of Athletic Training Education. (2005). *Standards for the accreditation of entry-level athletic training education programs.* Round Rock, TX.

Commission on Accreditation of Athletic Training Education (2008). *Overview document.* Retrieved on June 18, 2008 from http://caate.net/documents/OverviewDocument-CAATE.pdf.

Craig, D.I. (2003). Educational reform in athletic training: A policy analysis. *Journal of Athletic Training, 4,* 351-357.

Craig, D.I. (2006). Athletic training instructors: A needs assessment of teaching methodology knowledge and self-perceived competence. *Journal of Athletic Training, 2,* 28-37.

Crewson, P.E. (2005). Fundamentals of clinical research for radiologists. *American Journal of Radiology. 184*, 1391-1397.

Curtis, N., Helion, J.G., & Domsohn, B.S. (1998). Student athletic trainer perceptions of clinical supervisor behaviors: A critical incidence study. *Journal of Athletic Training. 33*(3), 249- 253.

Delforge, G.D., & Behnke, R.S. (1999). The history and evolution of athletic training education in the United States. *Journal of Athletic Training*. *34*(1), 53-61.

Docherty, C.L., Gansneder, B.M., Arnold, B.L., & Hurwitz, S.R. (2006). Development and reliability of the ankle instability instrument. *Journal of Athletic Training, 41*(2), 154-158.

Donner, A., & Klar, N. (1996). The statistical analysis of kappa statistics in multiple samples. *Journal of Clinical Epidemiology, 49*(9), 1053-1058.

Elder, C., Knoch, U., Barkhuizen, G., & von Randow, J. (2005). Individual feedback to enhance rater training: does it work? *Language Assessment Quarterly, 2,* (3), 175-196.

Epstien, R.M. (2007). Assessment in medical education. *New England Journal of Medicine. 356*(4), 387- 396.

Eechaute, C., Vaes, P., Duquet, W., & Van Gheluwe, B. (2007). Test-retest reliability of sudden ankle inversion measurements in subjects with healthy ankle joints. *Journal of Athletic Training, 42*(1), 60-65.

Feinstein, A.R., & Cicchetti, D.V., (1990). High agreement but low kappa. I. The problems of two paradoxes. *Journal of Clinical Epidemiology, 43,* 543-549.

Fleis, J.L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76,* 378 – 382.

Foster, D.T. & Leslie, D.K. (1992). Clinical teaching roles of athletic training. *Journal of Athletic Training. 27,* 298 – 302.

Gamaroff, R. (2000). Rater reliability in language assessment: the bug of all bears. *Elsevier Science Ltd,* 31-53.

Giles, S., Wetherbee, E., & Johnson, S. (2003). Qualifications and credentials of clinical instructors supervising physical therapist students. *Journal of Physical Therapy Education, 17*(2), 50-55.

Gould, T.E., & Caswell, S.V. (2006). Stylistic learning differences between undergraduate athletic training students and educators; Gregoric mind styles. *Journal of Athletic Training, 41*(1), 109-116.

Guggenmoos-Holzman, I. (1996). The meaning of Kappa: Probability concepts of reliability and validity revisited. *Journal of Clinical Epidemiology, 49*(7), 775-782.

Haynes, S.N., Richard, D.C., & Kubany, E.S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological Assessment, 7*(36), 238-247.

Hunt, R.J. (1986). Percent agreement, Pearson's correlation, and Kappa as measures of inter-examiner reliability. *Journal of Dental Research, 65*(2), 128-130.

Knight, K.L. (2006). Educational perceptions vs. reality; classroom and clinical education. *Athletic Training Education Journal, 1,* 15-17.

Knight, K.L. (2008). Progressive skill development and progressive clinical experience responsibility. *Athletic Training Education Journal, 1,* 2-4.

Ladyshewsky, R., Baker, R., Jones, M., & Nelson, L. (2000). Evaluating clinical performance in physical therapy simulated patients. *Journal of Physical Therapy Education, 14*(1), 31-37.

Lafave, M.R., Katz, L., Donnon, T., & Butterwick, D.J., (2008). Initial reliability of the standardized orthopedic assessment tool. *Journal of Athletic Training, 43,* (5), 483-488.

Lagumen, N.G., Butterwick, D.J., Paskevich, D.M., Fung, T.S., & Donnon, T.L. (2008). The intra-rater reliability of nine content-validated technical skill assessment instruments for athletic taping skills. *Athletic Training Education Journal,* 3*, 91-101.*

Landis, J.R. & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33,* 159-174.

Lauber, C.A., Toth,P.E., Leary, P.A., Martin, R.D., & Killian, C.B. (2003). Program directors' and clinical instructors' perceptions of important clinical-instructor behavior categories in the delivery of athletic training clinical instruction. *Journal of Athletic Training, 38*(4), 336-341.

Laurent, T., & Weidner, T.G. (2001). Clinical instructors'and student athletic trainers' perceptions of helpful clinical instructor characteristics. *Journal of Athletic Training, 36*(1), 58-61.

National Athletic Trainers' Association. (1983). *Guidelines for Development and implementation of NATA approved undergraduate athletic training education programs*. Dallas, TX.

National Athletic Trainers' Association. (2006a). *Clinical Instructor Educator Seminar.* Dallas, TX.

National Athletic Trainers' Association. (2006b). *National athletic trainers' association athletic training educational competencies.* (4th ed.).

Nitko, A. J. (2004). *Educational assessment of students.* (4th ed.). Upper Saddle River: Pearson Education Inc.

Peeler, J. D., & Anderson, J. E. (2008). Reliability limits of the modified Thomas test for assessing rectus femoris muscle flexibility about the knee joint. *Journal of Athletic Training, 43*(5), 470-476.

Peer, K. S., & Rakich, J. S. (2000). Accreditation and continuous quality improvement in athletic training education. *Journal of Athletic Training. 35*(2), 188- 193.

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice.* (2nd ed.). Upper Saddle River: Prentice-Hall Inc.

Prentice, W. E. (2009). *Arnheim's Principles of Athletic Training. A competency Based approach.*, New York, NY: Mcgraw-Hill.

Prentice, W. E. (2006). *Arnheim's Principles of Athletic Training. A competency Based approach.*, New York, NY: Mcgraw-Hill.

Prentice. W. E. (2000). *Arnheim's Principles of Athletic Training. A competency Based approach.*, New York, NY: Mcgraw-Hill.

Prezza Technologies (2007). *Checkbox Survey Software.* Prezza Technologies, Cambridge MA.

Randolph, J. J. (2008). *Online kappa calculator.* Retrieved November 11, 2008, from http://justusrandolph.net/kappa/

Randolph, J. J. (2005). Free-marginal multirater kappa: An alternative to Fleiss' fixed marginal multirater kappa. Presentation at the Joensuu learning and Instruction Symposium, Joensuu, Finland.

Riolo, L. (1997). Reliability of assessing psychomotor tasks in physical therapy curricula. *Journal of Physical Therapy Education, 11*(1), 36-39.

Seegmiller, J.G. (2006). Perceptions of quality for graduate athletic training education. *Journal of Athletic Training, 41,* (4). 415-421.

Seegmiller, J. G. (2003). A Model for Clinical Education in Athletic Training. *The Internet Journal of Allied Health Sciences and Practice*, *1*(1). 5/19/2008.

Sexton, P. J. (2003). A perspective on student assessment. *Athletic Therapy Today, 8*(6), 6-10.

Shirk, C., Sandrey, M. A., & Erikson, M. (2006). Reliability of first ray position and mobility measurements in experienced and inexperienced examiners. *Journal of Athletic Training, 41*(1), 93-101.

Sim, J., & Wright, C. C. (2005). The kappa statistic in reliability studies: use, interpretation, and sample size requirements. Physical Therapy, *85*(3), 257-269.

Sireci, S. G. (1998). The construct of content validity. *Social Indicators Research, 45*, 83-117.

Sliwinski, M.M., Schultz, K., Hansen, R.L., Malta, S., & Babyar, S. (2004). Clinical performance expectations: A preliminary study comparing physical therapist students, clinical instructors, and academic faculty. *Journal of Physical Therapy Education, 18,* (1), 50-57.

SPSS Inc. (2007). *SPSS Base 16.0.1 for Windows User's Guide.* SPSS Inc., Chicago IL.

Stickley, L. A. (2005). Content validity of a clinical education performance tool: The physical therapist manual for assessment of clinical skills. *Journal of Allied Health, 34*(1), 24-30.

Strohschein, J., Hagler, P., & May, L. (2002). Assessing the need for change in clinical education practices. *Physical Therapy, 82*(2)*,* 160-176.

Tinsley, H.E., & Weiss, D.J., (1975). Interrater reliability and agreement of subjective judgments. *Journal of Counseling Psychology, 22*(4), 358-376.

Vacha-Haase, T., Kogan, L. R., Tani, C. R., & Woodall, R. A. (2001). Reliability generalization: Exploring variation of reliability coefficients of MMPI clinical scales scores. *Educational and Psychological Measurement, 61*(1), 45-59.

Vendrely, A. (2002). Student assessment in physical therapy education: An overview and literature review. *Journal of Physical Therapy, 16*(2), 64-69.

Vendrely, A., & Carter, R. (2004). The influence of training on the rating of physical therapist student performance in the clinical setting. *Journal of Allied Health, 33*(1), 62-69.

Weidner, T. G., & August, J. A. (1997). The athletic therapist as clinical instructor. *Athletic Therapy Today,* January, 49-52.

Weidner, T. G., & Henning, J. M. (2004). Development of standards and criteria for the selection, training, and evaluation of athletic training approved clinical instructors. *Journal of Athletic Training, 39*(4), 335-343.

Weidner, T. G., & Henning, J.M. (2005). Importance and applicability of approved clinical instructor standards and criteria to certified athletic trainers in different clinical education settings. *Journal of Athletic Training, 40*(4), 326-332.

Wimer, J. W. (2005). An investigation of interrater reliability among athletic training accredited site visitors. *Journal of Allied Health, 34*(2), 65 – 75.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477-501.