Development and Validation of the Sense of Competence Scale – Revised

Cara W. McFadden

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Educational Leadership and Policy Studies

Steven Janosik, Co-Chair
Gary Skaggs, Co-Chair
Nancy Bodenhorn
Laura Welfare

May 27, 2010
Blacksburg, VA

Keywords: student development theory, instrument development, instrument validation

Development and Validation of the Sense of Competence Scale - Revised

By

Cara W. McFadden

Abstract

The purpose of this study was to develop an instrument to measure the sense of competence of traditional age college students across the dimensions that define the construct. The Sense of Competence Scale-Revised (SCS-R) was developed to provide a measure of Chickering's (1969) first vector, an important psychosocial construct. Administrators can use data from the instrument to modify an institution's academic and social environment to enhance the development of the intellectual, physical, and interpersonal competencies of college students.

During the development and validation, various aspects of the SCS-R were examined in accordance with the validity framework outlined by Messick (1995). According to Messick (1995), the validity of measures can be examined in terms of six forms of evidence; content, substantive, structural, generalizability, external and consequential. The six forms of evidence function as general standards for all educational measurement (Messick, 1995). During the study the content aspect of validity was addressed through the creation of concept maps and test blueprints. In addition, the content aspect of validity was addressed by creating and selecting items by reviewing the literature and hosting brainstorming sessions, items were then reviewed by student development theory experts, pilot tested, field tested and then items with high technical quality were selected for the final instrument. The substantive aspect of validity was addressed through an analysis of item rating scale functioning, person fit to the measurement scales, and item difficulty. The structural aspect of validity was addressed by evidence of the instrument's dimensionality. The generalizability aspect of validity was addressed through an analysis of item/person reliability.

The evidence generated from the study suggested that the chosen items for the SCS-R provide reliable and valid estimates of a student's personal assessment of their intellectual, physical and interpersonal abilities.

Dedication

Dedicated to my husband, Shay

and son, Caelan

Acknowledgements

First I would like to thank my committee members for their direction, support and guidance during the dissertation process. Thank you to Dr. Janosik for always believing in me and giving me the opportunity to succeed and grow in the Higher Education program. Thank you for your time, guidance, listening ear and encouragement during this process. Thank you to Dr. Skaggs for you willingness to teach, consult and share your expertise with me. Thank you to Dr. Bodenhorn for sharing your knowledge as well as continuous support. And thank you to Dr. Welfare for your support and constant encouragement during the dissertation process.

To Dr. Chickering and Dr. Reisser, it was a pleasure to have the opportunity to consult with you during the instrument development phase of the SCS-R. Thank you for your insight and encouraging words. To my cohort, for the countless hours of conversation and encouragement that assisted me in accomplishing my goals in the Higher Education program. Martha, Tracey, Nick and Fulya, I could not have done it without you.

To my family, thank you to my mother who instilled in me that you can do anything you set your mind too. Thanks for always being there to listen. Your patience and support has meant the world to me during my studies. You are always the one who reminds me that I can do anything I want and that the only way we grow is by challenging ourselves to learn new things. To Wayne, you continuously demonstrate what it means to be dedicated, from family to work to activities. Thanks for always encouraging me.

To my husband Shay, my best friend, I love you with all my heart. I know you have made great sacrifices for me. I would not be the person I am today without your continued support and love. My hopes are that I can do the same for you today, tomorrow and forever. And my son Caelan, thank you for your wonderful smile and laugh that always brings sunshine to my day.

Table of Contents

List of Tables

List of Figures

Chapter One

Introduction

Historically, college and university administrators were at liberty to use state funds however they chose. Today college and university administrators face external demands from the state government and are expected to show how they are meeting state expectations while also supporting values of society (Zumeta, 2001). Some state legislators reduce financial, personnel and other procedural controls while mandating that universities abide by accountability, quality assessment and performance budgeting processes (Schmidtlein & Berdahl, 2005). Now, college administrators maintain institutional accountability by creating assessment methods that show progress. In addition, these methods must be deliberate and must illustrate accomplishments to all constituents (the state and federal government, peer institutions, and the public) (Schmidtlein & Berdahl, 2005). Faculty members and student affairs professionals are asked to provide not only data on defined outcomes but also evidence that results are gained by realistic expenditures (Schmidtlein & Berdahl, 2005).

To achieve academic quality, faculty must make assessment and evaluation procedures a "part of every intentionally designed developmental intervention" (Evans, Forney, Guido, Patton & Renn, 2010, p. 369). In addition, assessment data must reflect the needs of the student community (Evans, et al., 2010). To achieve this goal faculty members and student affairs professionals need to improve outcome assessment efforts of students' development.

The manner in which a student develops varies based on the individuality of each student. Effective faculty members and student affairs professionals provide psychometrically sound measures of a student's development to accurately portray the development of students during his or her time with the university. Yet there are only a few reliable or valid instruments available to measure student development and those that do exist are hard to administer (Evans, Forney, & DiBrito, 1998).

Researchers in higher education need to dedicate time and effort in the development of psychometrically sound instruments (Smart, 2005). Smart contends that lack of attention to detail creates a deficiency in higher education research. Researchers know what constitutes a problem or issue facing American higher education; however, they tend to rush their research to provide

quick answers for policymakers and university administrators. Facilitating research in this manner causes potential problems with the quality of data researchers collect (Smart, 2005).

To develop psychometrically sound instruments and accurately explain inferences made from collected student data, researchers must understand student development theory. Instruments supported by theory encourage collaborative efforts between faculty members and student affairs professionals in enhancing student learning and maximizing viable outcomes in higher education settings (Evans, et al., 1998). Student development is defined as the means for how students grow and progress their developmental capacities as a result of enrollment in a college or university (Rodgers, 1990). Theory provides the foundation for the practice of student development in higher education (Evans, et al., 1998).

For the purpose of this study, Chickering's (1969) student development theory and the revised version (Chickering & Reisser, 1993) of the theory provide the foundation for the development of a psychometrically sound instrument to measure students' development. Chickering (1969) defines seven developmental vectors specific to college students. Each vector is described by its own developmental tasks and outcomes. The vectors are not always sequential in nature as students move through the different vectors at different rates. However, the first three (developing competence, managing emotions and moving through autonomy toward interdependence) are viewed as the foundation for a student's development and future success in the remaining vectors. In addition, the two major works (Chickering 1969, Chickering & Reisser, 1993) include things such as establishing identity, a student's ability to be comfortable with oneself and developing integrity, a student's personal beliefs are more concrete and provide a guide for their behavior (Chickering, 1969).

Chickering's (1969) theoretical framework offers one theoretical model to describe traditional age college students' development of competence. Sense of competence is depicted by three skills, intellectual, physical and manual, and interpersonal. These skills provide a way to understand variables that describe a student's sense of competence. Achieving an increased level of sense of competence not only strengthens assurance brought to adult tasks, but also affects the remaining six vectors of development.

Chickering and Reisser (1993) later expanded upon the original student development theory (Chickering, 1969) to include research from a more diverse population (Chafin, 2006).

Currently, there are three instruments that measure a number of Chickering's vectors, (Student Development Task and Lifestyle Assessment (SDTLI), 1987, Student Development Task and Lifestyle Assessment (SDTLA) updated version of the SDTLI, 1999, Iowa Student Development Inventories (ISDI), 1986 and the Sense of Competence Scale (SCS), 1987); however, the majority of the measures are based on the 1969 concepts of the seven vectors. The SDTLI (Winston, Miller, & Prince, 1987), for example, is based on Chickering's constructs in a general manner but does not conform specifically to the vectors proposed in 1969 (Winston, 1990). The SDTLA on the other hand measures some of Chickering's vectors however does not measure sense of competence. With this in mind, researchers should use caution when using the SDTLI to assess Chickering's theoretical constructs (Evans, et al., 1998). Faculty, staff, and students at the University of Iowa developed the ISDI (Hood, 1986), which is a compellation of six instruments that measure a specific vector of development. The instruments are based on Chickering's original definitions of the vectors (Evans, et al., 1998). IDCI (Hood & Jackson, 1983a) is one of the six instruments that measures competency, but is limited to measuring specific competencies such as self-confidence, competency in writing and math competencies. The Sense of Competence Scale (SCS) (Janosik, 1987) was developed to measure students' competence, however, the instrument centers on only two (intellectual and interpersonal) of the constructs in Chickering's (1969) developing sense of competence vector. In addition, no validity studies are reported on the SDTLI, IDCI (Evans, et al., 1998) or the SCS.

Statement of the Problem

Better methods to assess student development must be constructed. "Too many studies use 'home grown' assessment tools that may have face validity but have not been thoroughly examined for reliability or validity in a systematic way" (Evans, et al., 2010, p. 368). Using appropriate measures makes the development of reliable and valid instruments important since existing instruments are becoming outdated (Evans, et al., 1998).

The absence of reliable and valid measures of student development in the higher education literature provides support for the development and validation of the Sense of Competence Scale – Revised (SCS-R). Administering accurate analyses supports the validity of measures stressing the development of psychometrically sound instruments (Smart, 2005).

For purpose of this study, developing sense of competence, the first vector of Chickering's (1969) theory was the primary focus. Competence was defined as intellectual, physical and manual, and interpersonal skills (Chickering, 1969; Chickering & Reisser, 1993). There are instruments (SDTLI, SDTLA & IDCI) that measure areas of Chickering's (1969) model; however, there is no instrument that measures solely sense of competence as defined by Chickering and Reisser (1993).

A variety of instruments measure student development, however, few instruments with the exception of the SCS (Janosik, 1987) have developed and validated an instrument that solely assesses sense of competence as its primary focus. However, the SCS (Janosik, 1987) does not include the third skill, physical and manual as described in Chickering and Reisser's (1993) vector of achieving sense of competence.

Purpose of the Research

The purpose of this measurement study was to develop additional items for an existing instrument and validate all items that measure sense of competence of traditional age college students. The SCS (Appendix A) is an existing instrument that was developed to measure a student's level of interpersonal and intellectual competence (Janosik, 1987). The goal of the Sense of Competence Scale –Revised (SCS-R) was to measure a student's sense of competence (intellectual, physical and manual, and interpersonal skills) and was based on Chickering's theory (1969; Chickering & Reisser, 1993). The SCS-R provides a measure of Chickering's (1969) first vector, an important psychosocial construct. Administrators could use data from the instrument to modify an institution's academic and social environment to enhance the development of the intellectual, physical, and interpersonal competencies of college students.

The main focus of this study was to develop items for the absent construct (physical and manual) along with additional items for the other two dimensions (intellectual and interpersonal) and examine the validity of all three dimensions of the SCS-R instrument. The SCS-R may be used to enhance understanding and evaluation of traditional age college students' sense of competence. Validity was examined through Messick's (1995) definition of validity as a detailed analytical framework for understanding the technical components of a particular measure or measures within an instrument. Messick (1995) distinguishes six forms of evidence used to illustrate the validity of measures obtained from an instrument. These are content,

substantive, structural, generalizability, external, and consequential aspects of construct validity. The six forms of evidence function as general standards for all educational measurement (Messick, 1995).

Research Questions

The study explored five research questions concerning the development and validation of the SCS-R:

1. What is the item technical quality of the items? (Content)
2. Is there evidence that the theoretical model explains the observed stability among item responses (person item fit)? (Substantive)
3. What dimensional structure best depicts traditional age college students' sense of competence? (Structural)
4. Which measurement model and rating scale configuration best depicts the rating scale structure of these data? (Substantive)
5. Do the three dimensions (intellectual, physical and manual, and interpersonal) of the revised SCS-R establish a reliable measurement to measure a student's personal assessment of their intellectual, interpersonal and physical and manual competence? (Generalizability)

Definitions

The following terms were defined specifically for the purpose of this study:

1. *Sense of Competence Scale (SCS):* Original instrument measuring a student's personal assessment of their intellectual and interpersonal capabilities (Janosik, 1987).
2. *Sense of Competence Scale-Revised (SCS-R):* Revised version of the SCS, measures a student's personal assessment of their intellectual, physical and manual, and interpersonal capabilities.
3. *Student Development Theory:* "the ways that a student grows, progresses, or increases his or her developmental capabilities as a result of enrollment in an institution of higher education" (Rodgers, 1990, p. 27).

4. *Sense of competence*: "increases as they (students) learn to trust their abilities, receive accurate feedback from others, and integrate their skills into a stable self-assurance," (Chickering & Reisser, 1993. p. 46).
5. *Intellectual competence*: "involves mastering content, gaining intellectual and aesthetic sophistication, and, most important, building a repertoire of skills to comprehend, analyze and synthesize" (Chickering & Reisser, 1993, p. 45).
6. *Physical and manual competence*: "can involve athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline" (Chickering & Reisser, 1993, p. 46).
7. *Interpersonal competence*: "entails not only the skills of listening, cooperating, and communicating effectively, but also the more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of the group, and to choose from a variety of strategies to help a relationship flourish or a group function" (Chickering & Reisser, 1993, p. 46).

Significance of the Study

In summary, there are several reasons for the development and validation of the SCS-R, they include the preference for using psychometrically sound instruments, the expansion of analyses to support the validity of the SCS-R and the need for further research concerning students' development of competence within university and college environment.

The present study was significant for future practice, research, and policy. In terms of practice, several campus constituencies might benefit from the results. One group includes student affairs professionals. The results of this study provided professionals with data about the validity of an instrument to assess intellectual, physical and manual competencies and interpersonal skills of traditional age college students. Student affairs professionals might use the SCS-R to assess participants in environments within the campus such as residence halls, recreation facilities or student activity programs and events.

Faculty members are a second group that might benefit from the results of the study. The results of this study provided faculty members with data about the validity of an instrument to assess intellectual competencies of traditional age college students. Faculty members might use the SCS-R to assess intellectual competency of students enrolled in their classrooms.

The present study also had significance for future research. For example, this study revised an existing instrument to measure sense of competence of traditional age college students. Future studies might develop and validate new instruments for the remaining six vectors in Chickering's student development theory to assess students.

This study explored the development and validation of the SCS-R. Future studies might examine group differences among students along the three dimensions. Such a study would expand on the information available about influences on traditional age college students' competencies.

Finally, the study was significant for future policy. The results of the study offered insight into the importance of accurate assessment of traditional age college students' sense of competence. These data may be used to expand existing accountability policies to include information about the development and validation of psychometrically sound instruments that measure a student's developmental process while studying at a university or college.

In addition, the results provided policymakers with information about the importance of developing valid measures to assess student growth and progression during their time with the university. Policymakers might use this information when considering appropriate instruments to measure college student learning by selecting those that are psychometrically sound in nature.

Delimitations

The primary delimitation dealt with sample. All the participants in the study were students at the same university who had sophomore class status. It is possible that students designated as freshmen, junior or senior status differed in some important ways from students who were classified as sophomores at the institution. In addition, it is possible that students at this institution differed in some way from students at other colleges and universities. If so, this might have influenced the results in some unforeseen manner.

Organization of the Study

The present study is organized around five chapters. Chapter One introduced the topic of study, the research questions and the significance of the study. The second chapter reviews the literature relevant to the study. Chapter Three offers a description of the methodology of the study, including the sampling techniques and the procedures used to collect and analyze the data. The fourth chapter and fifth chapter include two manuscripts each with a distinct focus. These

manuscripts are designed to stand alone. Each will include its own introduction, literature review, methodology, and discussion. Each will highlight the most compelling findings relative to its respective purpose. That being the case, Chapters Four and Five may deviate from what is proposed in the first three chapters. The first manuscript centers on the development of items for the SCS-R instrument. The second manuscript focuses on the validity and the reliability of the dimensions (intellectual, physical and manual, interpersonal). Each article has a different target audience. The development article is targeted toward audiences in the higher education field. The validation article is targeted toward audiences in the measurement field.

Chapter Two

Literature Review

This study was designed to develop and validate the Sense of Competence Scale – Revised (SCS-R). This chapter contains a review of the literature relevant to three major areas of this study. The first concerns student development theory. Relevant literature in this area includes discussion of student development theory and an overview of Chickering's (1969) Vector 1 (developing competence). The chapter describes current measures of competency based on Chickering's (1969) Vector 1 and the rational for the SCS-R. The Student Development and Lifestyle Inventory (SDTLI), Student Development and Lifestyle Assessment (SDTLA), Iowa Developing Competency Inventory (IDCI), and the Sense of Competence Scale (SCS) are discussed. In addition, the chapter illustrates the rational and the background of the SCS. A review of instrument development and validation practices was important for understanding the design of the SCS (Janosik, 1987) and the development and validation of the SCS-R.

Student Development Theory

Since the 1960's, the literature (Erikson, 1968, Chickering, 1969, Chickering & Reisser, 1993, & Pascarella & Terenzini, 2001) surrounding student development theory continues to grow and is researched by a number of individuals in higher education as well as other disciplines.

The literature often refers to a number of developmental theories with one being psychosocial theory. Psychosocial theory examines the personal and interpersonal lives of individuals (Evans, 1996). Following psychosocial development research conducted by Erikson (1950, 1968), many scholars began developing theories (Chickering, 1969, Chickering & Reisser, 1993) and using identity models (Marcia, 1966, Josselson, 1987, Cross, 1991, Phinney, 1990, & D'Augelli, 1994) to assess college student development. Chickering's vector theory was one of these (Chickering, 1969).

Chickering's Student Development Theory

Chickering's (1969) student development theory is comprised of seven vectors that illustrate how traditional age college students progress in finding a sense of identity during their time at college. Chickering (1969) defines the first vector as developing intellectual, physical

and manual and interpersonal skills. Managing emotions, the second vector describes a student's ability to have self-control (Chickering, 1969). The third and final foundational vector for a student's development is moving through autonomy toward interdependence. Completion of this third vector is achieved once a student has both emotional and instrumental independence as well as acquiring the ability to acknowledge one's interdependencies (Chickering, 1969). Chickering (1969) describes the fourth vector as establishing identity, a student's ability to be comfortable with oneself. Achieving the ability to be comfortable with oneself fosters change in the latter three vectors. Once students establish identity, the fifth vector, they are able to develop interpersonal relationships by tolerating individuals who are different from them (Chickering, 1969). Development of purpose (sixth vector) is achieved once a student has formulated preparations and priorities that integrate avocational, vocational and life-style plans (Chickering, 1969). In the final vector, developing integrity students' personal beliefs are more concrete and provide a guide for their behavior (Chickering, 1969). Chickering's (1969) theory was revised (Chickering & Reisser, 1993) for purposes of researching a more diverse population (Chafin, 2006). For purposes of this study Chickering's (1969) earlier writings along with his collaborative work (Chickering & Reisser, 1993) for Vector 1 (developing competence) provided insight for the development of the newly generated items for the SCS-R.

Developing Competence

In developing the SCS-R, Vector 1 (developing competence) was the primary focus. "Competence is a three-tined pitchfork. Intellectual competence, physical and manual skills, and interpersonal competence are the tines. But the handle is most important. Without it, no work can be done, no matter how sharp and sturdy the tines. A sense of competence stems from the confidence that one can cope with what comes and achieve goals successfully," (Chickering & Reisser, 1993, p. 53). The focus for the SCS-R was on three interrelated dimensions (intellectual, physical and manual, and interpersonal) of developing competence. The SCS-R planned to measure college students' personal assessment of their sense of competence.

Intellectual. Intellectual dimension involves, "mastering content, gaining intellectual and aesthetic sophistication, and, most important, building a repertoire of skills to comprehend, analyze, and synthesize," (Chickering & Reisser, p. 45). Most importantly, intellectual skills comprise the ability to reason, solve problems and participate in active learning opportunities

(Chickering & Reisser, 1993). Chickering and Reisser (1993) urge faculty members and college administrators not to strictly define intellectual competence as "skills at passing tests or mastery of some 'essential' knowledge" (p. 63). It should also include the ability to listen, question, reflect, and communicate (Chickering & Reisser, 1993). In addition, students should be an active participant in searching for knowledge rather than using a more passive approach (Chickering & Reisser, 1993).

Physical and Manual. Physical and manual abilities contribute to a student's sense of competence. These skills come from a variety of activities being either athletic or artistic in nature. These skills are derived from participation in athletics and recreational sports, attention to wellness, and involvement in performing arts, tangible creations, and hands-on learning (Evans, et al., 1998, Chickering & Reisser, 1993). For a few students, participation in such activities become a vocation while for others the skills become an avocation (Chickering & Reisser, 1993). Vocation is defined as a career pursuit or routine, while avocational activities are described as a hobby or leisure pursuit. Physical skill development seems obvious when one learns to kick a soccer ball, take photographs, dance or sculpt, however little research exists illustrating the development of these skills while in college (Chickering & Reisser, 1993).

Interpersonal. Along with physical ability as a component of developing competence comes the facility to interact with others. Student interaction with others contributes to the level of interpersonal competence. Interpersonal skills include things like listening, self-disclosing, and participating in dialogue that brings insight and satisfaction (Chickering & Reisser, 1993). These skills are, "a prerequisite for building successful friendships and intimate relationships," (Chickering & Reisser, p. 77) and "for playing one's role as a citizen" (Chickering & Reisser, p. 77). Intellectual, physical and manual, and interpersonal abilities are all components of developing competence.

However, it is important to remember that students' overall sense of competence is subjective; sense of competence stems from how individual students feel about their achievements and can trust their own abilities. Some students may take their level of competence for granted by having strong interpersonal skills, while other students may think no matter what they achieve it is never enough leaving them unsure of their abilities (Chickering & Reisser, 1993). University administrators who strive to provide opportunities for students such

as electives and extracurricular activities will create a foundation for students to build upon during their time in college. Through these "increments of growing mastery and assuredness…the development of competence occurs" (Chickering & Reisser, p. 82, 1993). The challenge is for university administrators to identify methods for measuring sense of competence to discern student's development and self-confidence.

Rational for SCS-R

Since psychosocial development is complex, its assessment is not an easy task (Miller & Winston, 1990). An individual's development is a continuous pattern occurring in many different facets and assessment provides a limited evaluation of a particular aspect of development at a specific point in time (Evans, et al., 1998). A few instruments (SDTLI, SDTLA, IDCI & SCS) are based on Chickering's constructs in a more general way. Before reviewing the SCS it is important to briefly review the SDTLI and IDCI.

Student Developmental Task and Lifestyle Inventory (SDTLI)

The SDTLI (Winston, et al., 1987) was derived from Chickering's (1969) student development constructs in a general manner but does not follow the seven vector structure as defined by Chickering (Winston, 1990, Evans, et al., 1998). Establishing and Clarifying Purpose (ECP), Developing Mature Interpersonal Relationships (DMIR) and Academic Autonomy (AA) are the three developmental tasks assessed by the SDTLI (Evans, et al., 1998).

The validity of the SDTLI was tested by examining its factor structure and intercorrelations. Reliability data of the SDTLI illustrated internal consistency based on a generated sample of 1200 respondents (college students) enrolled in 20 colleges and universities in the United States and Canada. A Cronbach Alpha (α) score above .70 for the three dimensions (ECP, DMIR, and AA) was reported (Winston, 1990). However, faculty members interested in assessing Chickering's theoretical constructs should use caution when using the SDTLI since a series of revisions have taken place based on the data collected from the respondents (Evans, et al., 1998). The changes created the development of the Student Development Task and Lifestyle Assessment (SDTLA).

Iowa Developing Competency Inventory (IDCI)

IDCI is a second instrument based on Chickering's (1969) original definitions of the vectors. The instrument is made up of three subscales: Self-Confidence, Competency in Writing,

and Competency in Math. Evans, Forney, and Guido-DiBrito (1998) stated that no validity studies have been reported as of 10 years ago. In relation to Chickering and Reisser's (1993) explanation of Vector 1 the IDCI primarily focuses on interpersonal skills and intellectual skills as they relate to knowledge of writing and math.

Several of the IDCI measures were developed from Chickering's (1969) student development theory. Hood and Jackson (1983) used two of the three types of competencies to create a pool of items, intellectual and interpersonal. However, Hood and Jackson (1983) chose not to use the third competency, physical and manual, due to problems they encountered. Hood and Jackson (1983) questioned the desire for students to grow in the physical and manual area during college and stated that physical ability may or may not be present for many students during this time in their life. The IDCI is made up of 70 items. Potential issues for using IDCI are the instrument uses negatively scored items (see Wolfe & Smith, 2007a for further explanation) and the assessment of physical and manual competency is absent.

Background of SCS

The development of the SCS took place more than 20 years ago after finding that there was no instrument at the time that dealt satisfactorily with Chickering's (1969) notion of sense of competence. With the absence of an existing instrument, the SCS was developed and consisted of 23 items to measure intellectual and interpersonal competence of college students (Janosik, 1987). Items were developed after a review of a number of instruments that measured self-esteem and other personal characteristics of college students. Two instruments in particular proved to be of assistance in the development process. Janosik (1987) designed the instrument by reviewing the Personal Orientation Inventory (POI) developed by Shostrom (1963) and the Student Information Form developed by the ACE/UCLA Cooperative Institutional Research Program (1982), now known as CIRP.

The SCS (Janosik, 1987) focused on the intellectual and interpersonal dimensions for developing competence. Based on factor analysis, thirteen of the items were assigned to the interpersonal subscale and remaining items to the intellectual subscale. A four-point Likert Scale was used to score the items. Janosik (1987) conducted a pilot study since there was no statistical information available, which assisted in determining the reliability of the instrument.

The pilot test was distributed to 100 male and 100 female undergraduate students. A total of 97 (48.5%) of the students responded to the instrument. Using Cronbach's Alpha (α) the reliability coefficient for the 23-item SCS was .78. For each subscale, intellectual .79 and interpersonal .76, the inter-item correlation was .14 and the correlation between the two subscales was .11. The results confirmed that the SCS was acceptably reliable and each item tended to represent a component of the respondent's sense of competence (Janosik, 1987).

For the purpose of this study it is equally important to review current measures of sense of competence as well as review the background of the existing instrument being modified. The SCS (Janosik, 1987) focused on measuring intellectual and interpersonal competence of traditionally aged college students. The SCS-R was designed to assess all three dimensions of developing competence as defined by Chickering (1969) and Chickering & Reisser (1993). For the SCS-R to accurately measure sense of competence, measurement theory and psychometric analyses cannot be ignored.

Measurement Theory

In measurement theory items are developed or selected and specific procedures are followed to obtain and quantify the observations of each dimension being measured. In addition, there is a validation process that tests the accuracy and sensitivity of the instruments and procedures being used (Crocker & Algina, 1986).

In the behavioral sciences the process for measurement consists of three steps. The first step is instrument administration where researchers observe or record behaviors of an object of measurement. Second is the coding process, observations become standardized and are placed into categories that represent a set of observations that provide similar information about the object of measurement. Third is the scaling, the codes are divided into one or more numerical descriptors of the object of measurement (Wolfe & Smith, 2007a). These descriptors are used to determine differences between those objects of measurement on a variable. The variables represent a theoretical framework and "represent important indicators or outcomes that have implications for decisions that are made about individuals," (Wolfe & Smith, 2007a, p. 2). Those variables represent the theoretical construct in question, which serve as primary conceptual ideas within scientific theories. For measures to be useful they must be reliable and

"provide information that allows researchers and practitioners to make valid inferences and decisions," (Wolfe & Smith, 2007a, p.2).

A failure to adhere to sound practices during the development and pretesting phase of instrument development may in part account for the ambiguous results that so often differentiate educational research. The researcher or practitioner who hastily formulates research questions and defining the variables to collect data is likely to not measure what was truly intended (Crocker & Algina, 1986). For measures to be useful they must be reliable and must provide information that allows researchers and practitioners to make valid decisions. Until recently, depictions of validity concentrated solely on content, criterion and construct related forms (Parker, 2007). In the past 15 years, however, the field of measurement theory has adopted more comprehensive approaches to instrument validation. Messick (1995) provides a definition for validity and six forms of evidence for validation.

Validity and Validation Activities

According to Messick (1995), "Validity is an overall evaluative judgment of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment" (p. 741). Messick believes that validity is not to be interpreted as a property of any instrument in question. Rather, validity should be viewed as a functional characteristic of the scores obtained from an instrument, the persons responding and the context of the assessment itself (Messick, 1995). Messick (1995) distinguishes six forms of evidence used to illustrate the validity of measures obtained from an instrument. These are content, substantive, structural, generalizability, external, and consequential aspects of construct validity. The six forms of evidence function as general standards for all educational measurement (Messick, 1995). For validation analyses of this study four (content, substantive, structural, and generalizability) of the six aspects of construct validity were the primary focus.

Content Validity

The content aspect of validity is based on the relevance and representativeness for which the items are developed and the technical quality of those items. Content validity is determining the knowledge, attitudes, skills and other attributes the researcher wishes to reveal from the instrument. Content relevance and representativeness are traditionally evaluated by, "expert

professional judgment, documentation which serves to address the content aspect of construct validity" (Messick, 1995, p. 745). Evidence for this aspect takes form in the documentation of the purpose of the instrument along with the processes in place through which the construct was defined. Expert reviews as well as the technical quality of items should be documented to support and provide additional evidence relating to the development of the instrument (Wolfe & Smith, 2007b). The following sections describe the evidence in more detail by describing the instrument purpose and specification, item development and item technical quality.

Instrument purpose and specifications. The first task in the instrument development phase is to identify the purpose of the instrument. During this phase formal decisions are made based on the measures, constraints and frame of reference are identified, and a population is chosen against which measures will be compared (Wolfe & Smith, 2007a). It is important to identify the purpose as the first step in the instrument development phase to remain consistent throughout the development process. Creating instruments and thinking about the kinds of items or scores that are desired in the beginning of the development process is a common mistake among some test developers (Wolfe & Smith, 2007b). To avoid making this mistake one must identify the domain of inference and the type of inference. The domain may relate to instruction, cognitive theory, knowledge, skills or abilities required for achievement. In the case of this study the domain of inference is sense of competence. The type of inference relates to the description of participants or describes a group. An accurate description of a student's level of sense of competence during their time as a traditional age college student is the type of inference for this particular study. Outlining the domain and type assists the instrument developer with setting clear goals for the duration of the development process (Wolfe & Smith, 2007a).

Item development. Creating a detailed description of the construct is an important task in the development process. Information for the construction of the construct model

> "may be drawn from a number of sources: (a) real-world observations, (b) literature reviews focusing on theoretical models of the construct in question or a related construct, (c) literature reviews focusing on empirical research involving the construct in question or a related construct, (d) reviews of instruments that measure the construct in question or a related construct, (e) expert and lay viewpoints about the construct or related constructs,

> and (f) content and task analyses within the domain in question (Wolfe & Smith, 2007a, p. 7)."

Expert reviews can be used throughout the instrument development process. Experts may include those who have familiarity with the target population, psychometricians, those familiar with the construct in question and for those who plan to use the information provided by the instrument in applied settings.

Item technical quality. For purposes of this section, understanding the development of Likert scale items is important. The guidelines for developing Likert scale items state:

> (a) avoid references to the past, (b) avoid factual statements, (c) avoid absolutes such as 'all', 'none', 'always', and 'never', (d) restrict words such as 'only', 'just' and 'merely' as much as possible, (e) write simple, short and clear statements, (f) avoid double-barreled items, (g) avoid suggesting a particular answer (Do you agree that…), (h) avoid non-distinguishing statements,(i) avoid negatives, (j) write items at a reading level much lower than the target population, (k) avoid irrelevancies, and (l) write at least twice as many items as would like to appear on the final instrument (Wolfe & Smith, 2007a, p. 19).

Also, it is important to mention that Cronbach (1951) and Wolfe and Smith (2007a), recommend using an even number of categories leaving less room for error variance.

Using a four-point Likert scale is supported by Wolfe and Smith (2007a) in that, "fewer rating scale categories (between three and five) tend to provide better fit to the Rasch Rating Scale Model" (p. 13). A four response rating scale was also supported by research conducted by Stone (1998). Stone (1998) suggested that four choices along a rating scale are usually enough and two are often sufficient. Researchers (Stone, 1998, Wright & Stone, 2003) found that using a scale with more than four options for the Wolfe's *Fear Survey Schedule* proved to be inefficient. Three options were found more efficient than a five-point scale after validation analyses.

Researchers (Crocker & Algina, 1986, Wolfe & Smith, 2007) recommend avoiding negatively worded items. "Analyses of negatively worded items may result in less reliable data, may demonstrate more rater bias and response sets, and may result in more homogeneous item

response distributions due to respondents being less willing to use extreme rating categories, particularly negative ones" (Wolfe & Smith, 2007a, p. 20).

Field test responses are also used to define the technical quality of each domain (intellectual, physical and manual, and interpersonal) within an instrument. During the evaluation process the item point-measure correlation and the mean-squared fit indices should be considered. The item point-measure correlation indicates the extent to which scores on an item are consistent with the average score across all other items (Wolfe & Smith, 2007b). The purpose of the fit analysis is to predict potential problematic response patterns as well as flag any misfit items. The mean square statistics can be interpreted as either over-fit, fit, or misfit. Overfit determines all observed responses are equal to the expected responses while misfit is the opposite; all observed responses differ greatly from expected responses. The fit interpretation observed responses will differ from the expected responses based on the amount predicted (Wolfe & Smith, 2007b).

Substantive Validity

The substantive aspect adds to the content aspect of construct validity the need for empirical evidence of response consistencies. Evidence derives how the dimensions relating to item content and processing models explain the observed stability among item responses (Wolfe & Smith, 2007a). The substantive aspect of validity, "addresses the degree to which theoretical rationales relating to both item content and cognitive processing models adequately explain the observed consistencies among item responses," (Wolfe & Smith, 2007b, p. 207). Think aloud protocols, correlation patterns, and consistencies in response times for task segments are examples of how to collect evidence (Messick, 1995). The following sections describe the evidence in more detail by describing rating scale function procedures and person fit analyses.

Rating scale function. The Rasch Rating Scale Model (RSM) was the measurement model used to evaluate Likert-type items for the SCS-R. Response categories setup as Likert scale items represent a respondent's increasing partiality towards the concept being questioned (Wright & Mok, 2004). RSM includes a threshold difficulty parameter that depicts the difficulty of moving from one scoring category to another. The RSM model assumes that the thresholds are to be constant across items (Wright & Mok, 2004).

The RSM is an item response theory based approach to measurement that was developed by Rasch (1960, 1966, 1980). RSM proponents believe that, "a mental measure is the ratiocination of a discovery process," (Osterlind, 2006, p. 286). A test identifies a construct for consideration and prepares items that require just the relevant aspects of cognition. RSM explains a respondent's trait level by identifying a correct response to difficult items which means more trait (competency) is shown or incorrect responses with less trait (competency) is shown (Osterlind, 2006). In general, the model specifies the probability that an examinee will respond to an item of a particular difficulty is a logistic function of the examinee's ability (θ) and the item's difficulty (δ). The model for constructing measures from responses using the RSM can be defined by the following equation:

$$\log (\pi_{nix} / \pi_{ni(x-1)} = \beta_n - (\delta_i + \tau_x)$$

where :

π_{nix} = the probability that participant *n* will respond in category *x* on item *i*

$\pi_{ni(x-1)}$ = the probability that participant *n* will respond in category *x-1* on item *i*

β_n = the ability of participant *n*

δ_i = the difficulty of item *i*

τ_x = the threshold that corresponds an equal probability for selecting category *x* and *x-1*

Response categories setup as Likert style items represent a respondent's increasing partiality towards the concept being questioned (Wright & Mok, 2004). The equation is described as:

$$\pi_{nix} = \frac{\exp \sum_{j=0}^{x} D\left(\theta_n - \delta_i - \tau_k\right)}{\sum_{k=0}^{m} \exp \sum_{j=0}^{k} D\left(\theta_n - \delta_i - \tau_j\right)}$$

For polytomous items, the RSM includes a threshold difficulty parameter that depicts the difficulty of moving from one scoring category to another. Using the RSM model it is assumed that the thresholds are to be constant across items.

Eight guidelines are typically followed for rating scale analysis (Linacre, 2004). It is important to mention the overarching guideline that all items are oriented with one latent variable. Linacre's (2004) eight guidelines:

> (a) At least 10 observations of each category, (b) Regular observation of distribution, (c) Average measures advance monotonically with category, (d) OUTFIT mean-squares less than 2.0, (e) Step calibrations advance, (f) Ratings imply measures, and measures imply ratings, (g) Step difficulties advance by at least 1.4 logits, and (h) Step difficulties advance by less than 5.0 logits (Linacre, 2004, pp. 263-275).

Person fit. The RSM expectations for a respondent and the actual observed scores of the respondent are used as additional support for the substantive aspect of validity (Messick, 1995) by confirming that the response method used by each respondent is consistent with that predicted by the item developers (Wolfe & Smith, 2007b). Specific requirements are set by RSM and the expected values for each combination of respondent and item score are directed by those requirements. "The standardized mean-squared unweighted person fit indices depict the degree to which the observed responses for a person are in accord with these model-based expectations," (Wolfe & Smith, 2007b, p. 211). Person misfit may occur when examinees guess, if there is a test security breech, or item bias within the instrument itself. The person fit indices are centered on 0.00 and range from $-\infty$ to ∞. Negative values illustrate small residuals and positive values indicate large residuals. Wolfe and Smith (2007b) suggest, "a value of 2.00 as a cut score for flagging persons for misfit" (p. 212).

Structural Validity

The structural aspect of validity evaluates "the fidelity of the scoring structure to the structure of the construct domain," (Wolfe & Smith, 2007b, p. 213). Evidence is illustrated in the form of correlation analyses of response consistencies and indicators. Messick (1995) states for the structural aspect of validity, "the theory of the construct domain should guide not only the selection or construction of relevant assessment tasks but also the rational development of construct-based scoring criteria and rubrics," (p. 746). The following section describes dimensionality analysis as evidence.

Dimensionality analysis. Dimensionality is defined as the number of aspects or dimensions that represent a construct (Wolfe & Smith, 2007a). Dimensionality may be

addressed in several ways. The RSM requires unidimensionality and local independence to assess structural validity. Local independence occurs when all items are independent of one another for examinees of the same ability on a latent tendency (Osterlind, 2006). When analyzing unidimensionality all items of the instrument are scaled equally. A principle component analysis is conducted on the standardized residuals from the previous equal scaling. Eigenvalues from the equalized scaling and principal component analysis of the residuals are placed onto a common scale. The analyses assist in determining if the model is multidimensional rather than unidimensioinal (Wolfe & Smith, 2007b).

Generalizability Validity

The generalizability aspect of validity "addresses the degree to which measures maintain their meaning across measurement contexts…often takes the form of correlations between measures generated in different measurement contexts," (Wolfe & Smith, 2007b, p. 215). The following sections describe the evidence in more detail by describing reliability.

Reliability. Reliability of an instrument refers to "the degree to which test scores are free from measurement error, will be constrained to measures of internal consistency, and which account for error due to content sampling" (Smith, 2004, p. 94). Reliability in terms of rating scale models refers to the degree to which items for a group of test takers are consistent across repeated applications of a measurement procedure. The interpretation determines whether the items are dependable and repeatable. Reliability expresses the degree to which scores are free of measurement errors for a particular group of people by reviewing how the indices depict the ratio of true variance to observed variance while taking into account random measurement error.

Standard error of measurement (SEM) is used, since reliability indices tend to be unhelpful for depicting the stability of individual measures. SEM "depicts the degree to which one can expect fluctuations in the observed measures over repeated administration of the instrument," (Wolfe & Smith, 2007b, p. 218). In addition, reliability is often reported as an index of internal consistency using coefficient alpha (α). In RSM, Person Separation Reliability (Rel_θ) may be seen as an analog to coefficient α, where the degree of internally consistent measures is expressed as one minus the ratio of the average squared errors (MSE_θ) to the variance of observed measures (V_θ):

$$Rel_\theta = 1 - [MSE_\theta / V_\theta]$$

Although Rel_{θ} and coefficient alpha estimate the same theoretical relationship between true variance and observed variance, Rel_{θ} is often noticeably lower, or more conservative, than coefficient α. As such, Rel_{θ} may be viewed as a lower bound estimate of true person reliability.

Within instrument development and validation there are multiple validation activities. Validation procedures for this particular study were shared with emphasis on four out of six of Messick's (1995) forms of validity evidence. Adherence to these principles for instrument construction along with empirical field testing of items before the initial sample is undertaken will only improve the quality of research in education (Smart, 2005, Crocker & Algina, 1986).

External Validity

The external aspect of validity "concerns the degree to which measures are related to external measures of the same construct, similar constructs, and other constructs. Depending on the level of measurement of those external measures, appropriate evidence may be reported as correlations between sets of measures or hypothesis test statistics and effect sizes" (Wolfe & Smith, 2007b, p. 220).

The external evidence provides documentation of theory-based predictions about the relationships between multiple measures of the target construct and other important constructs. The evidence is gathered and displayed in a multitrait-multimethod matrix (MTMM matrix). The matrix illustrates correlations between constructs showing that they are either related or unrelated to one another (Wolfe & Smith, 2007b).

Consequential Validity

The consequential aspect of validity "focuses on the value implications of score interpretation as a source of action" (Wolfe & Smith, 2007b, p. 224). This aspect creates evidence that addresses actual and potential consequences of how test scores will be used, in regard to sources of invalidity such as fairness or bias (Wolfe & Smith, 2007b).

Consequential evidence is much like substantive evidence in that fair decisions are made about methodology and results. This form of evidence facilitates a process called standard setting. Standard setting is the process of setting passing scores or cut scores to distinguish varying levels of ability. The standard setting identifies different levels of ability and provides a description of what respondents in that level are able to accomplish (Wolfe & Smith, 2007b).

Summary

The preceding chapter provided an overview of student development theory, described a particular psychosocial theory of interest, discussed current measures of competency, rational for the development and validation of the SCS-R, described the background of the SCS, and explained measurement theory. In addition, an explanation of the validation activities required were provided. Six aspects of construct validity were described along with methods used to provide evidence to support each aspect.

Chapter Three

Methods

The purpose of this quantitative study was to develop and validate the Sense of Competence Scale-Revised (SCS-R), a measure of sense of competence of traditional age college students. The Sense of Competence Scale (SCS) is an existing instrument that measures a student's level of interpersonal and intellectual competence (Janosik, 1987). The SCS-R's objective was to measure all three dimensions of college students' sense of competence as previously discussed. Intellectual, physical and manual, and interpersonal skills were defined by Chickering's (1969) theory of student development. The goal for the use of the SCS-R inventory was to measure a student's personal assessment of their sense of competence. The methods and analyses that were used in this process are outlined in this chapter, as well as the content for the two research articles that were written as the last two chapters of this dissertation. The test scores from the SCS-R derive their meaning from the participants' absolute levels of performance on a series of items where each set corresponds to an individual level of proficiency of competence.

Framework for Instrument Development and Validation

The SCS-R's objective was to provide criterion-based scores for an individual student's level of sense of competence. Criterion-referenced instruments classify individuals with similar levels on the trait of interest. Labels are assigned such as pass/fail or meet/below standards (Wolfe & Smith, 2007). For purposes of this study an individual's score was defined by having a high or low assessment of their sense of competence. Various aspects of the SCS-R were examined by applying the validity framework outlined by Messick (1995). Validity is an evaluative judgment of the degree to which data analyses and theory support the appropriateness of interpretations on the basis of an individual's scores on a test. Messick (1995) distinguishes six forms of evidence used to illustrate the validity of measures obtained from an instrument. As discussed in Chapter Two, the forms are content, substantive, structural, generalizability, external, and consequential aspects of construct validity. The six forms of evidence function as general standards for all educational measurement (Messick, 1995).

This chapter describes the means by which content, structural, substantive, and generalizability aspects of Messick's (1989) framework were applied to the SCS-R. Four out of

the six aspects were chosen to answer the research questions listed in Chapter One. The aspects provided evidence of validity for an appropriate interpretation and use of measures obtained from the SCS-R.

Conceptual Framework for the SCS-R

As noted in Chapters One and Two, developing sense of competence consists of three dimensions: intellectual, physical and manual, and interpersonal. Figure 1 diagrams this conceptual model. Each dimension is seen as being a contributor to the construct (sense of competence), no one particular dimension is more important than the other. Answers to the items on the SCS-R reflect a student's assessment of their capabilities among all three dimensions. Figures 2, 3 and 4 illustrate the individual conceptual map for each dimension (intellectual, physical and manual, interpersonal).

Throughout Chapter Three each section discusses the four forms (content, structural, substantive and generalizability) of validity chosen for this study.

Content

During the development phase, each dimension (intellectual, physical and manual, interpersonal) of sense of competence was operationalized by a series of items. It is not feasible to measure all possible indicators of competence. However, through literature and expert review as well as brainstorming sessions an appropriate number of items were developed. A set of 26 original items were created for the physical and manual dimension. The original items from the SCS (Janosik, 1987) were retained and new items for the intellectual and interpersonal dimension developed for the SCS-R. Based on expert reviews and brainstorming sessions held with the target population 35 items represented the intellectual dimension and 31 represented the interpersonal dimension.

Table 1 provides a reference for how items were created and/or adapted in accordance with the cell within the instrument blueprint the item's content intends to represent. The number of items for physical and manual skills listed in Table 1 represents a proposed outcome for the SCS-R, and not necessarily the number of items that were retained for the final instrument. At least twice as many items were created for physical and manual skills. Items for intellectual and interpersonal include original items from the SCS as well as newly generated items. All items of

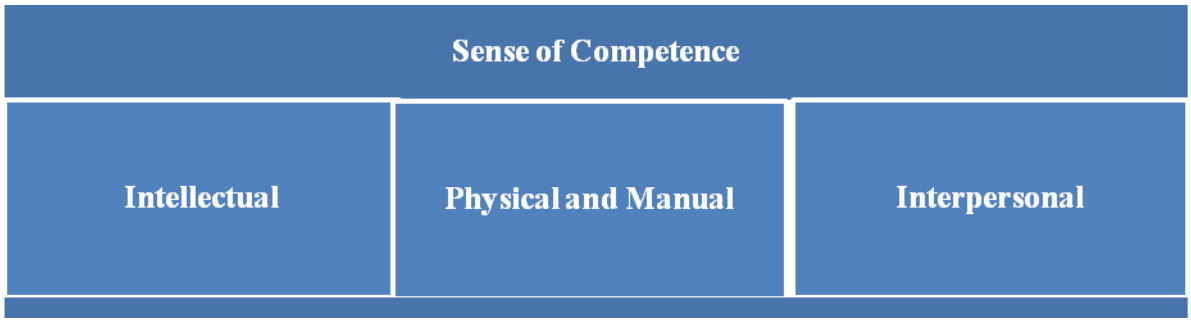

Figure 1. Conceptual Model of Sense of Competence-Revised (SCS-R)

Each dimension is a contributor to the construct sense of competence no one particular dimension is more important than the other.

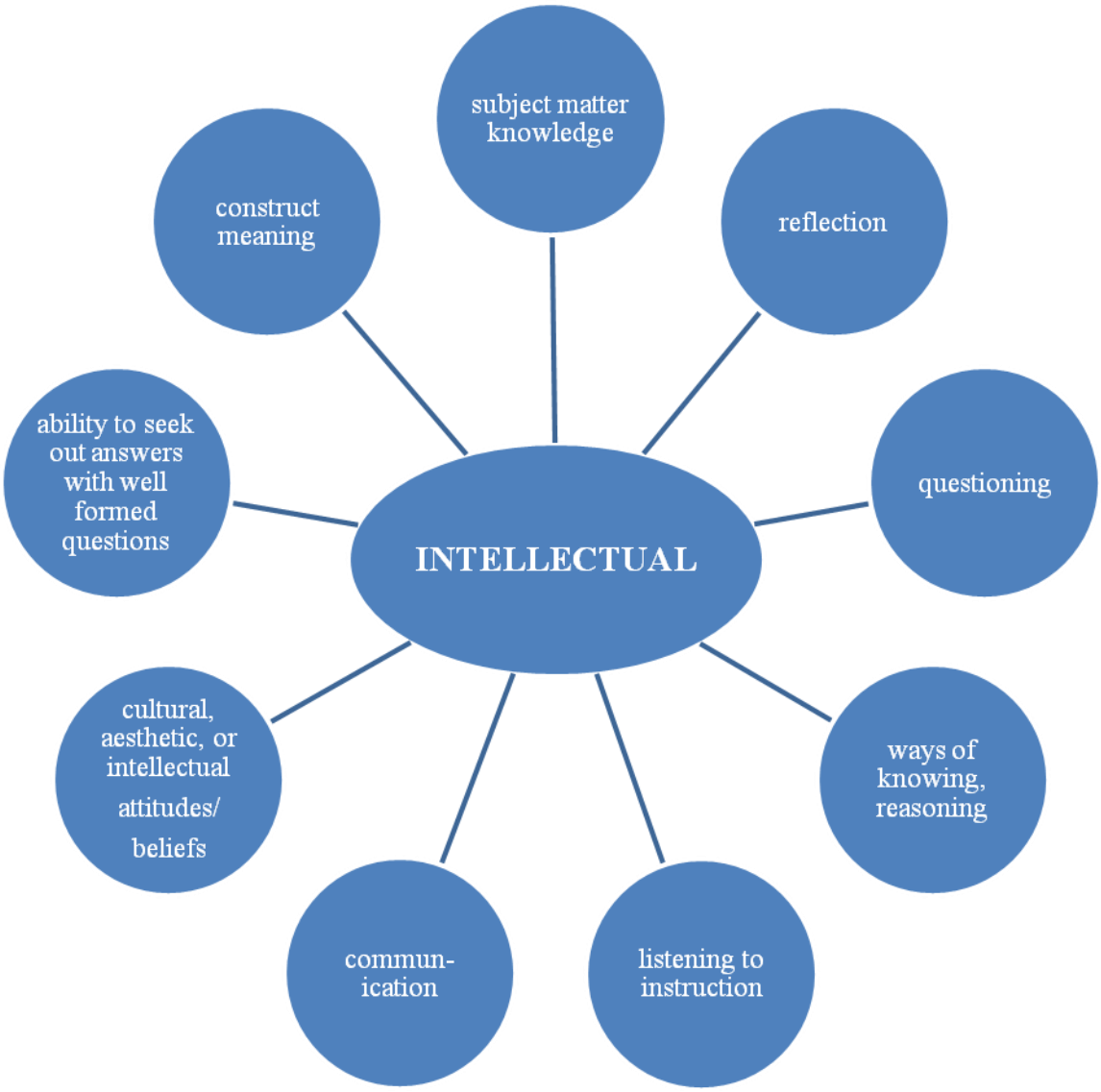

Figure 2. Intellectual Concept Map

Circles depict attributes of the dimension based on Chickering and Reisser (1993).

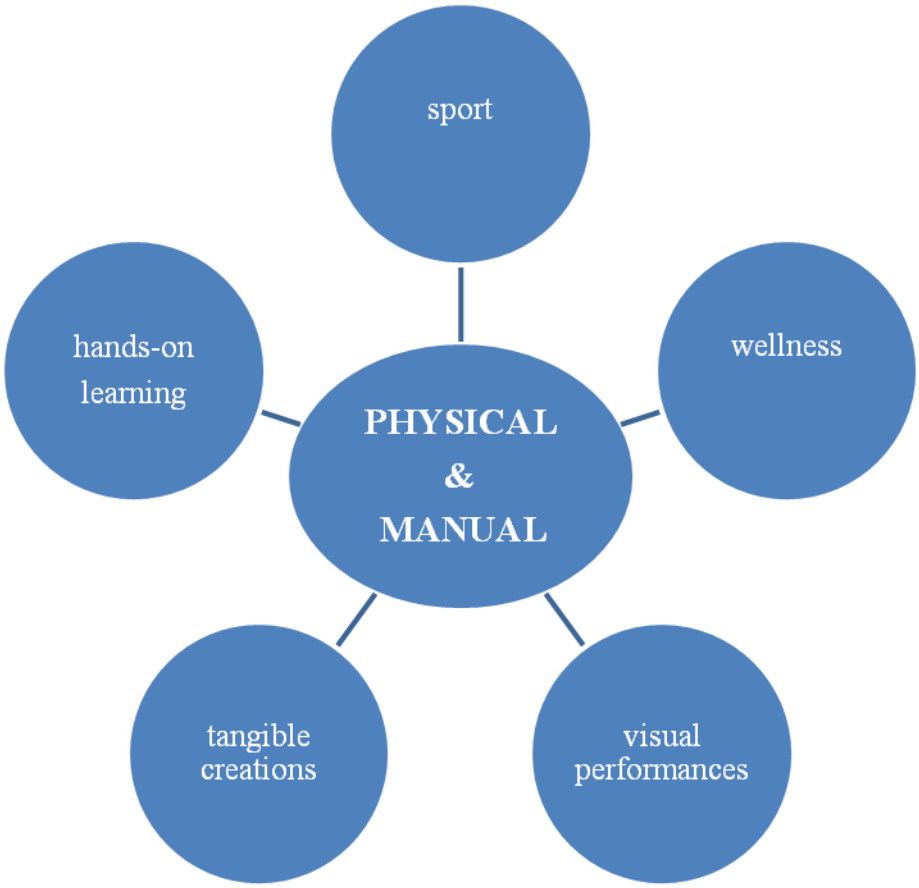

Figure 3. Physical and Manual Concept Map

Circles depict attributes of the dimension based on Chickering and Reisser (1993).

Figure 4. Interpersonal Concept Map

Circles depict attributes of the dimension based on Chickering and Reisser (1993).

Table 1

Sense of Competence-Revised (SCS-R) Instrument Blueprint

Sense of Competence-Revised Dimensions		
		Total
Dimension 1 Intellectual	Involves mastering content, gaining intellectual and aesthetic sophistication, and building a range of skills to comprehend, analyze and synthesize	35 items
Dimension 2 Physical and Manual	Involves athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline	26 items
Dimension 3 Interpersonal	Involves more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of a group, and to choose from a variety of strategies to help a relationship flourish or a group function	31 items
Total		92 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

the SCS-R were refined and field tested leading to the final instrument. Tables 2, 3, and 4 illustrate the test blueprint for each dimension. Each dimension description includes a definition of the skill and an item breakdown for measuring a student's assessment of their capabilities.

After items were drafted, Chickering and Reisser authors of the revised version (Chickering & Reisser, 1993) of Chickering's (1969) student development theory were asked to review the instrument blueprint to assess its relevance to the construct (developing sense of competence). In addition to Chickering and Reisser, a psychometrician and the developer of the SCS (Janosik, 1987) reviewed the items. The intellectual, physical and manual, and interpersonal items were created in an attempt to measure the skills as defined by Chickering and Reisser (1993). All items were reviewed and revised using the general guidelines for writing and reviewing Likert style items as outlined by Wolfe and Smith (2007a). The original items (Janosik, 1987) remained the same with the exception of changing all negatively worded items as well as the addition of newly generated items.

Item Development

Items included on the pilot version of the SCS-R were constructed using a four-point Likert Scale (1932). Individual items consisted of statements that reflected developing competence within one of the three dimensions (intellectual, physical and manual, interpersonal). Respondents were asked to state the degree to which they disagreed or agreed with each statement. Items consisted of four ordinal categories for respondents to choose: (a) strongly disagree, (b) disagree, (c) agree, and (d) strongly agree. The Rasch Rating Scale Model (RSM) was used for the development of the additional items and validation of all items for the SCS-R. RSM was chosen to transform the ordinal data received from the Likert scale into an interval scale for the data to be useful (Wright & Mok, 2004). The number of categories included in a response scale is irrelevant, the response decision is always between adjacent categories. The decision is the point that the probability of selecting the next category is equal to that for the previous one which is the threshold.

Item Content

The SCS-R was also reviewed to meet the guidelines for Likert scale items as defined by Wolfe and Smith (2007a). In addition, newly generated items were created by this author in

Table 2

Intellectual Dimension Test Blueprint

	Intellectual Dimension	
		Total
Definition	Involves mastering content, gaining intellectual and aesthetic sophistication, and building a range of skills to comprehend, analyze and synthesize	
Personal Assessment	Student is secure in their intellectual decisions and actions. Student is realistic about one's own ability.	35 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

Table 3

Physical and Manual Dimension Test Blueprint

	Physical and Manual Dimension	
		Total
Definition	Involves athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline	
Personal Assessment	Student is secure in their physical and manual decisions and actions. Student is realistic about one's own ability.	26 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

Table 4

Interpersonal Dimension Test Blueprint

	Interpersonal Dimension	
		Total
Definition	Involves more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of a group, and to choose from a variety of strategies to help a relationship flourish or a group function	
Personal Assessment	Student is secure in their interpersonal decisions and actions. Student is realistic about one's own ability.	31 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

accordance with the guidelines for writing rating scale items (Crocker & Algina, 1986, Wolfe & Smith, 2007a). Content for the items was supported by Chickering and Reisser's (1993) theory for developing sense of competence.

Brainstorming Sessions

Groups of students were selected for brainstorming sessions to assist with the development of new items for the SCS-R. Two brainstorming sessions were held, the first session with three students and the second with nine students. Students expressed their thoughts about intellectual, physical and manual, and interpersonal abilities of college students. Notes were gathered from both meetings and assisted in the development of newly generated items for the SCS-R. Participants were asked to fill out blank concept maps by selecting tasks that described each dimension.

The items previously developed for the SCS-R were reviewed and field-tested accordingly. The items were also reviewed for content relevancy to the instrument blueprint. As noted earlier in Chapter One, current instruments measuring college students' competence either failed to exhibit consistency regarding the construct or did not include an all encompassing view of the sense of competence framework as defined by Chickering and Reisser (1993). Moreover, in Chapter Two a review of previously developed instruments (SDTLI, SDTLA, IDCI, & SCS) revealed areas within the framework that needed more attention (e.g. physical and manual skills).

Data Collection Procedure

Data collection required several steps. First, it was necessary to obtain approval from the Institutional Review Board for Research Involving Human Subjects (IRB) at the researcher's institution (Appendices K, L, & M). Once IRB approval was obtained, data collection was a multi-step process.

Groups were identified from five different sets of respondents who participated in either item editing, item content, item tryouts, pilot testing or the field- testing of the instrument. The groups assisted in assessing the items within the SCS-R for relevancy and for collecting accurate and significant data.

After the development of the initial pool of items was created according to instrument specifications (Wolfe & Smith, 2007a), a number of content reviews and studies took place to

address the relevancy of items. In addition, the reviews and studies assessed the capacity of each item for psychometrically sound data.

Item Editing

A group of five individuals were invited to review a draft set of the SCS-R inventory and provided comments regarding the items clarity and conciseness. In addition, the group reviewed the instrument for any areas illustrating bias. The group consisted of individuals who had some expertise in item writing and editing (e.g. faculty and doctoral students who had knowledge of measurement methodologies). The participants were given a draft set of items and were asked to make edits. Edits were then reviewed in conjunction with the participants. Participants were asked to provide feedback in response to the suggestions they made for the instrument. Input received from all participants was then considered and changes to the SCS-R were made accordingly.

Item Content Review

For content review a group of experts were selected that provided meaningful and useful information for this study. The group of experts included Chickering, Reisser, and additional faculty members with content proficiency. They were asked to review all items and provide feedback. Individual conference calls were conducted with each participant to provide explanations for their decisions and to give any suggestions regarding items that may not have been included. Items were added or deleted and recommendations from participants were taken into consideration as far as adding or eliminating items to the instrument. Items measuring other vectors of Chickering's (1969) theory were deleted. For example, I openly show my emotions better represents Vector 2 (managing emotions).

Item Tryout

The draft set of items was administered to a group of 10 students from the university where the instrument was being developed. During administration of the items, participants were asked to provide comments about how they felt while they completed the instrument and answered all items. Information gained during this phase was considered and items were edited as needed. One respondent stated, "Allowing more space between the questions would make the survey easier to read. Sometimes the lines ran together if I read too quickly." Another student stated, "It got a bit repetitive." The repetitive feel of the instrument was probably due to the

number of items on the preliminary version of the SCS-R. A review of the pilot study results illustrates additional support for content and structural validity.

Pilot Test

A preliminary version of the SCS-R, which contained the expanded set of draft items were administered to students who either lived on campus, participated in recreational sports, or were a member of the honors program at the same university. Open-ended questions were also included to receive feedback on the ease of item interpretation, the utility of the rating scale structure, and identification of potential bias.

Once the development phase was completed, the draft instrument was scripted into final form. The final form of the SCS-R was then reviewed again by the expert review committee to re-assess item quality, item content, construct validity and efficiency of the chosen rating scale structure. From the final expert review the SCS-R was drafted for distribution and validation. Discussion of the pilot study results are discussed in Chapter Four.

Instrument Validation (Field Test)

The sample was identified by respondents who were traditional age college students at a large Research I institution. Persons in the sample were asked to complete a demographic questionnaire that asked about their gender, ethnicity, and major. In addition, participants were asked brief questions about their involvement in student activities, academic engagement and physical and manual activities. Sampling strategies were based on projected response rates and the availability of respondents to accommodate a minimum requirement of 300 completed questionnaires.

For the purpose of this study, the entire sophomore class at a Research I institution was selected. The SCS-R includes multiple aspects (intellectual, physical and manual, interpersonal) of the dimension developing sense of competence, it was important to find students that exemplified the three aspects to determine whether dependence existed between responses beyond those reflected by the singular measure (developing sense of competence).

The primary instrumentation of the study was the SCS-R. The final instrument included 63 questions using a four-point response scale. The four categorizations were classified as 1 = (Strongly Disagree), 2 = (Disagree), 3 = (Agree), and 4 = (Strongly Agree). The items in the instrument were categorized in three sections: Intellectual, Physical and Manual, and

Interpersonal. Twenty-three of the 63 items were assigned to interpersonal, 18 items were assigned to physical and manual, and 22 items were assigned to intellectual. The items were categorized into three domains as shown in Table 5.

In addition, a demographic section was included to illustrate general traits of the respondents. The demographic section included items such as sex, ethnicity, level of education, and major to support a number of sub-group analyses for future research. Involvement was the second part of this section that included a total of seven items (GPA, student activities, recreational sports, athletics, artistic activities, wellness, hands-on learning). The demographic part of the questionnaire included dichotomous and polytomous items. Lastly, subjects were given seven qualitative questions to provide a personal response to better describe their experience. Those who were invited to provide responses included members of the populations for whom the instrument was developed, traditional age college students currently enrolled at a Research I institution.

The SCS-R was given in a web-based format to be completed when it was convenient for the student (e.g. after class, work, organizational meeting). In the web-based format, the revised instrument was sent to participants through a survey system set up by the university. The instrument included a cover letter, which asked participants to complete the instrument and submit responses by a specified date. Five follow-up emails were sent as a reminder to help provide for a larger response rate. To increase participation in the study students were given the opportunity to download their name in a separate survey for a drawing. The winner of the drawing received an opportunity to win one of six $50 gift certificates to the university bookstore or an Apple iPod. It is important to note that the names submitted for the drawing were not connected to the individuals' survey response.

Directions were provided within each section of the instrument, informing participants about the means by which they were to provide responses. A coding scheme was developed to facilitate the recording of data into a numerical format that was necessary for subsequent statistical analyses.

After facilitating the field test all completed questionnaires were placed in an excel document. Each item on the instrument was coded by a 1, 2, 3, or 4 for the Likert scale

Table 5

Standardized Residual Variance

Dimension	Total Items within Dimension	Items
Intellectual	22	1, 3, 9, 10, 14, 17, 18, 22, 23, 29, 30, 31, 32, 35, 36, 43, 50, 51, 52, 56, 57, 58
Physical and Manual	18	7, 11, 13, 15, 16, 20, 21, 26, 27, 28, 39, 40, 41, 42, 48, 49, 60, 61, 62
Interpersonal	23	2, 4, 5, 6, 8, 12, 19, 24, 25, 33, 34, 37, 38, 44, 45, 46, 47

Note. Total number of items here will be determined after the preliminary reviews and field test are completed.

responses. The excel document included all answers from the participants. The excel document was used to import into *WINSTEPS* (Linacre, 2006) software.

Data Analysis Procedure

Data analysis for this study consisted of several procedures. Since this study included a development and validation phase, select analyses were run to ensure reliability and validity. Responses from the pilot test and field test of data collection for the SCS-R were scaled using the Rasch measurement procedures (Rasch, 1980) contained in *WINSTEPS* (Linacre, 2006). The Rasch Rating Scale Model (RSM) was chosen as the measurement model to evaluate the SCS-R four-point Likert scale items.

Item Technical Quality (Content)

Developing items that measure each dimension (intellectual, physical and manual, and interpersonal) is important and yields high technical quality. The technical quality of items that appear on the SCS-R were explored using two sets of indices: item point-measure correlations (r_{pm}), and mean-square fit statistics (Wolfe & Smith, 2007b).

Field test responses were used to define the technical quality of each domain (intellectual, physical and manual, and interpersonal). A guideline for using RSM is that all items are scaled to be oriented in the same direction (Wolfe & Smith, 2007b). This guideline was explored using $r_{pm,}$ which is the Pearson Product Moment Correlation (r) between the vector of item scores and person measures. Items that have negative r_{pm} and/or values < .30 were flagged and investigated to determine the source of their failure to contribute effectively to the RSM.

The technical quality was examined through a review of the difference between the RSM's predicted responses and observed responses:

$$residual_{ni} = E_{ni} - X_{ni}$$

where:

E_{ni} = the expected response for participant n on item i, and

X_{ni} = empirically observed response for participant n on item i)

Analyses were examined using *WINSTEPS* (Linacre, 2006) software to determine the outfit and infit mean scores. Estimates of item weighted (infit) and unweighted (outfit) mean squares (*MS*) and the standardized version of the unweighted mean squares ($Z_{unweighted}$) were used to identify possible item mis-fit. Values for both infit ($MS_{weighted}$) and outfit ($MS_{unweighted}$) were

examined due to the fact that the $MS_{weighted}$ index is sensitive to non-extreme unexpected responses and the $MS_{unweighted}$ index is sensitive to extreme unexpected responses (Karabatsos, 2000). The expected value for both infit and outfit is +1.0, with possible values ranging from 0 to positive infinity (∞).

Values closer to zero and values < 1.0 indicate over-fit, where all observed responses are close to equal to the RSM's expected responses. Values > 1.0 indicate misfit, where observed responses differ from model expectations. Guidelines for the interpretation of these values were referenced in the review of estimates of item misfit, where: *MS* > 2.0 distorts the measurement system; *MS* = 1.5 - 2.0 are unproductive for the construction of measurement, but not degrading; *MS* = 0.5 - 1.5 are productive for measurement; and, *MS* < .05 are less productive for measurement, but not degrading, and may produce misleading reliabilities (Linacre, 2002).

Rating Scale Analysis (Substantive)

To determine substantive validity there was an examination of the rating scale structure of the SCS-R. The rating scale analysis for the SCS-R instrument used eight guidelines as defined by Linacre (2004).

Person Fit (Substantive)

The RSM defines the requirements for measurement and the expected values for each respondent-by-item combination. For example, person misfit with an item may occur due to a respondent's carelessness in answering items or guessing on items. Person fit indices are based on 0.00 and range from -∞ to ∞ where negative values indicate small residuals and positive values indicate large residuals. Wolfe and Smith (2007b), "suggest a value of 2.0 as a cut score for flagging respondents for misfit" (p. 212). Using this criterion, respondents associated with values of 2.0 and greater were flagged.

Dimensionality (Structural)

The Kaiser's eigenvalue was calculated to investigate unidimensionality. The eigenvalues were scaled on the unexplained variance that gave the Kaiser's eigenvalues. First, the researcher evaluated the degree to which the claim of unidimensionality was reasonable. Results for the dimensionality analyses were based on Kaiser's criterion. The criterion proposes that only eigenvalues that are at least equal to 1.00 should be retained (Cattell, 1966).

The analyses assisted in determining if the model is multidimensional rather than unidimensional. Intellectual, physical and manual, and interpersonal items were analyzed as separate constructs to determine if there was multidimensionality present.

Reliability (Generalizability)

The generalizability aspect of measures obtained from the SCS-R explored analyses of internal consistency. *WINSTEPS* (Linacre, 2006) was used for the analyses. Results of the field test are presented in Chapter Five (Manuscript Two).

Summary of Activities

In conclusion, the purpose of this study was to develop and validate the SCS-R. The methodology described in this chapter was deemed sufficient to address the research questions posed in the study.

As described within this chapter, several activities were performed pursuant to developing the additional items for the SCS-R and acquiring evidence of validity to support an appropriate interpretation and use of the instrument. Table 6 shows the reader the various activities that were performed in regards to the specific aspect of validity that were addressed.

Proposed Articles for Final Defense

As described in Chapter One of this document, two articles were created as part of this study, pursuant to fulfilling the requirements of the alternative option for the dissertation. As a reminder, manuscript one describes the development of the SCS-R and manuscript two focuses on the validity of the SCS-R.

Manuscript One

The manuscript describes the steps that were taken to develop the SCS-R in accordance with the validity framework outlined by Messick (1995). Within the manuscript, several of the aforementioned aspects to validity were cited, as well as the steps that were taken to provide evidence relating to those aspects. The manuscript begins with an introductory section that highlights student development theory and the importance of psychometrically sound instruments in the field of higher education. In addition, an explanation of two of Messick's six (1995) aspects of validity are described to construct a strong validity argument for the development of the SCS-R and results of the pilot study are shared.

Table 6

Instrument Development and Validation Activities Linked to Messick's (1995) Validity Aspects

	Validity Aspect			
	Content	Structural	Substantive	Generalizability
Type of Evidence	Instrument Purpose	Instrument Dimensionality	Rating Scale Function	Reliability
	Instrument Specifications		Person Fit	
	Item Development			
	Expert Reviews			
	Item Technical Quality			

Note. Adapted from "Instrument Development Tools and Activities for Measure Validation Using Rasch Models," by E. W. Wolfe & E. V. Smith Jr., 2007b, Journal of Applied Measurement, p. 244. Copyright 2007 by Maple Grove, MN.

Manuscript Two

The article describes the field test results of the analyses performed in *WINSTEPS* (Linacre, 2006). The analyses are presented and discussed through explanations, tables and figures. In addition, the implications for those results are explained. In addition, the manuscript includes discussion regarding limitations to the study, proposed applications for the SCS-R and suggested areas for continued research on both the SCS-R and the six remaining vectors of Chickering's (1969) student development theory.

Chapter 4

Manuscript 1

Development of the Sense of Competence Scale-Revised (SCS-R)

Abstract

The Sense of Competence Scale (SCS) (Janosik, 1987) was created to better understand college students' intellectual and interpersonal competency. In this study, a revised instrument titled Sense of Competence Scale-Revised (SCS-R), was developed to provide a measure of all three dimensions of Chickering's (1969) first vector, an important psychosocial construct. These dimensions include intellectual competence, physical and manual competence, and interpersonal competence. This article reviews the dimensions of the sense of competence construct and details the development of items for the preliminary version of the SCS-R through a description of item development and validity evidence. Content and structural validity evidence are described by an explanation of six phases that were conducted in sequential order to develop the SCS-R.

Keywords: student development theory, instrument development

Chapter 4

Manuscript 1

Development of the Sense of Competence Scale-Revised (SCS-R)

Student development is defined as the means by which students grow their developmental capacities as a result of enrollment in a college or university (Rodgers, 1990). Theory provides the foundation for the practice of student development in higher education (Evans, et al., 1998). The literature (Erikson, 1968, Chickering, 1969, Chickering & Reisser, 1993, & Pascarella & Terenzini, 2005) surrounding student development theory continues to grow and is researched by a number of individuals in higher education as well as other disciplines.

The literature often refers to a number of developmental theories with one being psychosocial theory. Psychosocial theory examines the personal and interpersonal lives of individuals (Evans, 1996). Following psychosocial development research conducted by Erik Erikson (1950, 1968), many scholars began developing theories (Chickering, 1969, Chickering & Reisser, 1993) and using identity models (Marcia, 1966, Josselson, 1987, Cross, 1991, Phinney, 1990, & D'Augelli, 1994) to assess college student development. Chickering's (1969) vector theory is one of these.

Chickering's (1969) seven vectors for establishing identity is one of the most popular student development theories used in student affairs practice but there are a few instruments that measure the constructs in Chickering's vectors. Four that do are the Student Development Task and Lifestyle Inventory (SDTLI, 1987), the Student Development Task and Lifestyle Assessment (SDTLA, 1999), the Iowa Student Development Inventories (ISDI, 1986) and the Sense of Competence Scale (SCS, 1987). However, these measures are based on Chickering's 1969 concepts of the seven vectors and do not include multiple content validity methods during the item development phase. The SDTLI (Winston, Miller, & Prince, 1987) for example, is based on Chickering's constructs in a general manner but does not conform specifically to the vectors proposed in 1969 (Moran, 2003; Winston, 1990). In addition, researchers should use caution when using the SDTLI to assess Chickering's theoretical constructs since a series of revisions have taken place based on the data collected from respondents (Evans, et al., 1998). Although, the SDTLA (Winston, Miller, & Cooper, 1999), the revised version of the SDTLI, reflects some of the vectors it does not measure sense of competence.

Faculty, staff, and students at the University of Iowa developed the ISDI (Hood, 1986), which is a compellation of six instruments that measure specific vectors of development. The instruments are based on Chickering's original definitions of the vectors (Evans, et al., 1998). IDCI (Iowa Developing Competency Inventory) (Hood & Jackson, 1983a) is one of the six instruments that measures competency, but is limited to measuring specific competencies such as self-confidence and competencies in writing and math. Potential issues for using IDCI are the instrument uses negatively scored items (see Wolfe & Smith, 2007a for further explanation) and the assessment of physical and manual competency is absent. The Sense of Competence Scale (SCS) (Janosik, 1987) was developed to measure students' sense of competence, however, the instrument centers on only two (intellectual and interpersonal) of the dimensions in Chickering's (1969) developing sense of competence vector.

In the development of instruments that attempt to measure Chickering's (1969; Chickering & Reisser, 1993) theory Arthur Chickering and Linda Reisser report that they have never been a part of the development of items for instruments reflecting the seven vectors of development or given the opportunity to provide feedback on the items created for instruments (A. Chickering and L. Reisser, personal communication, January 28, 2010). Methods using expert reviews and consultations with the target population enhance the validity of an instrument (Vogt, King, & King, 2004).

Better methods to assess development must be constructed. "Too many studies use 'home grown' assessment tools that may have face validity but have not been thoroughly examined for reliability or validity in a systematic way" (Evans, Forney, Guido, Patton, & Renn, 2010, p. 368). Validity is an evaluative decision of the degree to which empirical evidence support the interpretations of test scores derived from an instrument (Messick, 1995). Student affair practitioners should place more effort on the development of assessment instruments (Evans, Forney, and Guido-DiBrito, 1998) and follow-through with validation research (Moran, 2003). The development of psychometrically sound instruments is crucial to more meaningful research (Smart, 2005). Clearly, more work needs to be done, if researchers are interested in measuring Chickering's sense of competence.

The purpose of this article is to review the dimensions of the sense of competence construct and details the development of items for the preliminary version of the SCS-R through

a description of item development and validity procedures. The development and validation procedures of the SCS-R were conducted in accordance with the validity framework outlined by Messick (1995) and the analyses as outlined by Wolfe and Smith (2007). The article concludes with discussion and implications for future research.

Developing Sense of Competence

"Competence is a three-tined pitchfork. Intellectual competence, physical and manual skills, and interpersonal competence are the tines. But the handle is most important. Without it, no work can be done, no matter how sharp and sturdy the tines. A sense of competence stems from the confidence that one can cope with what comes and achieve goals successfully," (Chickering & Reisser, 1993, p. 53). The SCS-R attempts to measure college students' personal assessment of their sense of competence.

However, it is important to remember that a student's overall sense of competence is subjective; sense of competence stems from how individual students feel about personal achievement and can trust their own abilities. Some students may take their level of competence for granted by having strong interpersonal skills, while other students may think no matter what they achieve it is never enough leaving them unsure of their abilities (Chickering & Reisser, 1993). University administrators who strive on their college campuses to provide opportunities for students such as electives and co-curricular activities will create a foundation for students to build upon during their time in college. Through these "increments of growing mastery and assuredness…the development of competence occurs" (Chickering & Reisser, p. 82, 1993). The challenge is for university administrators to identify methods for measuring sense of competence to discern a student's development and self-confidence in their intellectual, physical and manual, and interpersonal skills.

Original Model: Sense of Competence Scale (SCS)

The foundation for the SCS-R rests on Janosik's SCS that developed more than 20 years ago after finding that there was no instrument at the time that dealt satisfactorily with Chickering's (1969) notion of sense of competence. As part of the instrument development process, a pilot test, factor analysis was conducted, 13items were assigned to the interpersonal dimension and remaining 10 items to the intellectual dimension.

Using Cronbach's (1951) Alpha (α) the reliability coefficient for the 23-item SCS was .78. For each subscale, intellectual .79 and interpersonal .76, the inter-item correlation was .14 and the correlation between the two subscales was .11. The results confirmed that the SCS was acceptably reliable and each item tended to represent a component of the respondent's sense of competence (Janosik, 1987).

The SCS-R was designed to assess all three dimensions of sense of competence as defined by Chickering (1969; Chickering & Reisser, 1993). For the SCS-R to accurately measure sense of competence a thorough review and implementation of content and structural validation methods was essential. I used Messick's (1995) interpretation of the content and structural aspect of construct validity along with the guidelines outlined by Wolfe and Smith (2007) to gain support for both aspects of construct validity.

Methods

Messick (1995) states, "Validity is an overall judgement of the degree to which empirical evidence and theoretical rationales support the adequacy and appropriateness of interpretations and actions on the basis of test scores or other modes of assessment" (p. 741). Messick (1995) believes that validity should not to be interpreted as a property of any instrument in question. Rather, validity should be viewed as a functional characteristic of the scores obtained from an instrument, the persons responding and the context of the assessment itself (Messick, 1995). Messick identifies six forms of evidence used to illustrate the validity of measures obtained from an instrument. These are content, substantive, structural, generalizability, external, and consequential aspects of construct validity. The six forms of evidence function as general standards for all educational measurement (Messick, 1995). The development procedures and validation analyses of the SCS-R pilot study concentrate on Messick's content and structural aspects as the primary focus. Messick's framework for validity was addressed through a series of activities designed specifically for instrument development and validation (Wolfe & Smith, 2007b). Table 1 provides an introduction to the activities that were performed in this regard, as well as a conceptual link to the specific aspect of validity each activity intended to address.

Content Validation

The content aspect of validity is based on the relevance and representativeness for which the items are developed and the technical quality of those items. Content validity is determining

Table 1

*Instrument Development and Validation Activities Linked to Validity Aspects**

	Validity Aspect	
	Content	Structural
Type of Evidence	Instrument Purpose	Dimensionality
	Instrument Specifications	
	Item Development	
	Expert Reviews	
	Item Technical Quality	

Note. Adapted from "Instrument Development Tools and Activities for Measure Validation Using Rasch Models," by E. W. Wolfe & E. V. Smith Jr., 2007b, Journal of Applied Measurement, p. 244. Copyright 2007 by Maple Grove, MN.

the knowledge, attitudes, skills and other attributes the researcher wishes to reveal from the instrument. Content relevance and representativeness are traditionally evaluated by, "expert professional judgment, documentation which serves to address the content aspect of construct validity" (Messick, 1995, p. 745). Evidence for this aspect takes form in the documentation of the purpose of the instrument along with the processes in place through which the construct was defined. Expert reviews as well as the technical quality of items should be documented to support and provide additional evidence relating to the development of the instrument (Wolfe & Smith, 2007b). The following sections describe the evidence in more detail by describing the instrument purpose and specification, item development and item technical quality. However, it is important to note that the development process was conducted in sequential order by following a number of phases. For purposes of content validation evidence I conducted a sequence of five phases prior to the pilot study.

Instrument purpose and specifications (Phase 1). The first task in the instrument development phase is to identify the purpose of the instrument. During this phase formal decisions are made based on the measures, constraints and frame of reference are identified, and a population is chosen against which measures will be compared (Wolfe & Smith, 2007a). It is important to identify the purpose as the first step in the instrument development phase to remain consistent throughout the development process. Creating instruments and thinking about the kinds of items or scores that are desired in the beginning of the development process is a common mistake among some test developers (Wolfe & Smith, 2007b). To avoid making this mistake one must identify the domain of inference and the type of inference. The domain may relate to instruction, cognitive theory, knowledge, skills or abilities required for achievement. In the case of this study the domain of inference is sense of competence and is defined by three dimensions (intellectual, physical and manual, and interpersonal). The type of inference relates to the description of participants or describes a group. An accurate description of a student's level of sense of competence during their time as a traditional age college student is the type of inference for this particular study. Outlining the domain and type assists the instrument developer with setting clear goals for the duration of the development process (Wolfe & Smith, 2007a).

Item development (Phase 2). Creating a detailed description of the construct is an important task in the development process. For the construction of the construct model I followed recommendations of Wolfe and Smith (2007a). I reviewed student development theory literature with a primary focus on Chickering's 1969 theory and the revised text (Chickering & Reisser, 1993). I reviewed instruments that measured the construct in question and those related to student development theory. In addition, I conducted expert reviews along with brainstorming sessions of the student population to gain insight of viewpoints related to sense of competence.

Items included on the preliminary version of the SCS-R were constructed using a four-point Likert (1932) scale. Individual items consisted of statements that reflected developing competence within one of the three dimensions (intellectual, physical and manual, interpersonal). Respondents were asked to state the degree to which they disagree or agree with each statement. Items consisted of four ordinal categories for respondents to choose: (a) strongly disagree, (b) disagree, (c) agree, and (d) strongly agree. The Rasch (1960, 1980) Rating Scale Model (RSM) was used for the development of the additional items and validation of all items for the SCS-R. The Rasch (1960, 1980) Rating Scale Model was chosen to analyze the data.

The RSM is an item response theory based approach to measurement that was developed by Rasch (1960, 1966, 1980). RSM proponents believe that, "a mental measure is the ratiocination of a discovery process," (Osterlind, 2006, p. 286). A test identifies a construct for consideration and prepares items that require just the relevant aspects of cognition. RSM explains a respondent's trait level by identifying a correct response to difficult items which means more trait (competency) is shown or incorrect responses with less trait (competency) shown (Osterlind, 2006). In general, the model specifies the probability that an examinee will respond to an item of a particular difficulty is a logistic function of the examinee's ability (θ) and the item's difficulty (δ). The model for constructing measures from responses using the RSM can be defined by the following equation:

$$\log (\pi_{nix} / \pi_{ni(x-1)} = \beta_n - (\delta_i + \tau_x)$$

where :

π_{nix} = the probability that participant *n* will respond in category *x* on item *i*

$\pi_{ni(x-1)}$ = the probability that participant *n* will respond in category *x-1* on item *i*

β_n = the ability of participant *n*

δ_i = the difficulty of item *i*

τ_x = the threshold that corresponds an equal probability for selecting category *x* and *x-1*

Response categories setup as Likert style items represent a respondent's increasing partiality towards the concept being questioned (Wright & Mok, 2004). The equation is described as:

$$\pi_{nix} = \frac{\exp \sum_{j=0}^{x} D\left(\theta_n - \delta_i - \tau_k\right)}{\sum_{k=0}^{m} \exp \sum_{j=0}^{k} D\left(\theta_n - \delta_i - \tau_j\right)}$$

For polytomous items, the RSM includes a threshold difficulty parameter that depicts the difficulty of moving from one scoring category to another. Using the RSM model it is assumed that the thresholds are to be constant across items. The Category Probability Curve (CPC) depicts the model-based probabilities of various outcomes, given the difference between measures and calibrations as illustrated in Figures 1 – 3.

Item Content. The SCS-R was also reviewed to meet the guidelines for Likert scale items as defined earlier by Wolfe and Smith (2007a). In addition, newly generated items were created in accordance with the guidelines for writing rating scale items (Crocker & Algina, 1986, Wolfe & Smith, 2007a). Content for the items was supported by reviewing Chickering's (1969; Chickering & Reisser, 1993) theory for developing sense of competence. Expert reviews and brainstorming sessions were conducted to accurately portray all three dimensions during the development phase.

Expert Reviews. Experts included those who had familiarity with the target population, psychometricians, those familiar with the construct in question. During the item development phase I conducted conference calls along with numerous email communications between Chickering and Reisser to gain insight and understanding of their descriptions of vector one. Throughout these conversations item edits, deletions, and additions occurred. Communication was facilitated via conference calls, electronic mailing and direct mailing. They were asked to review all items and provide feedback.

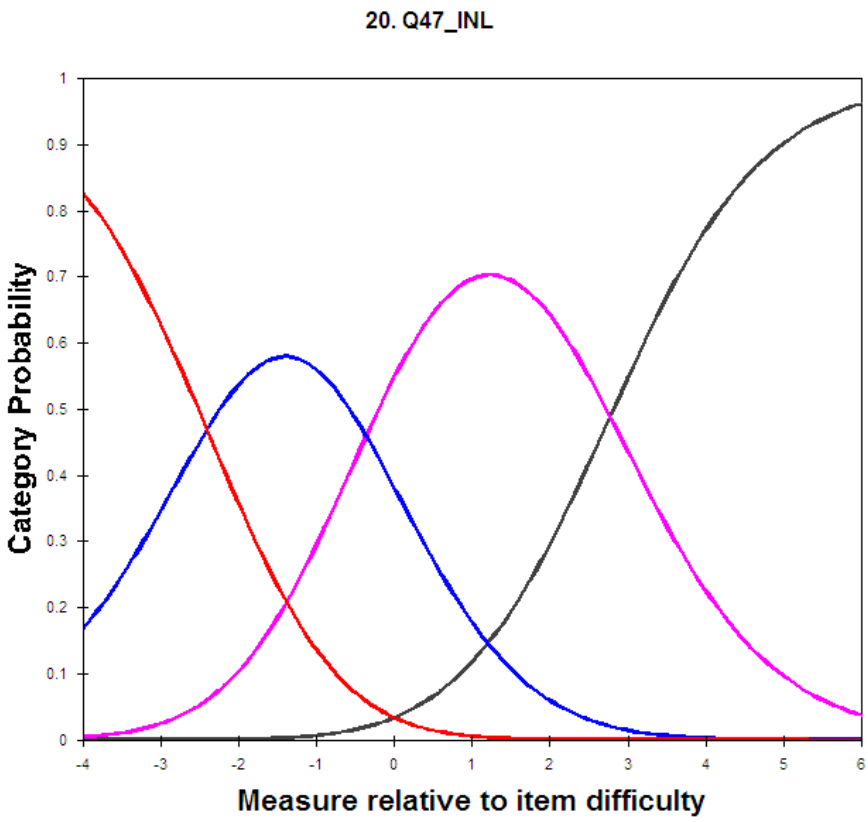

Figure 1. Intellectual Dimension Category Probability Curve

Graph depicts poor category probability curve for a four-point Likert scale.

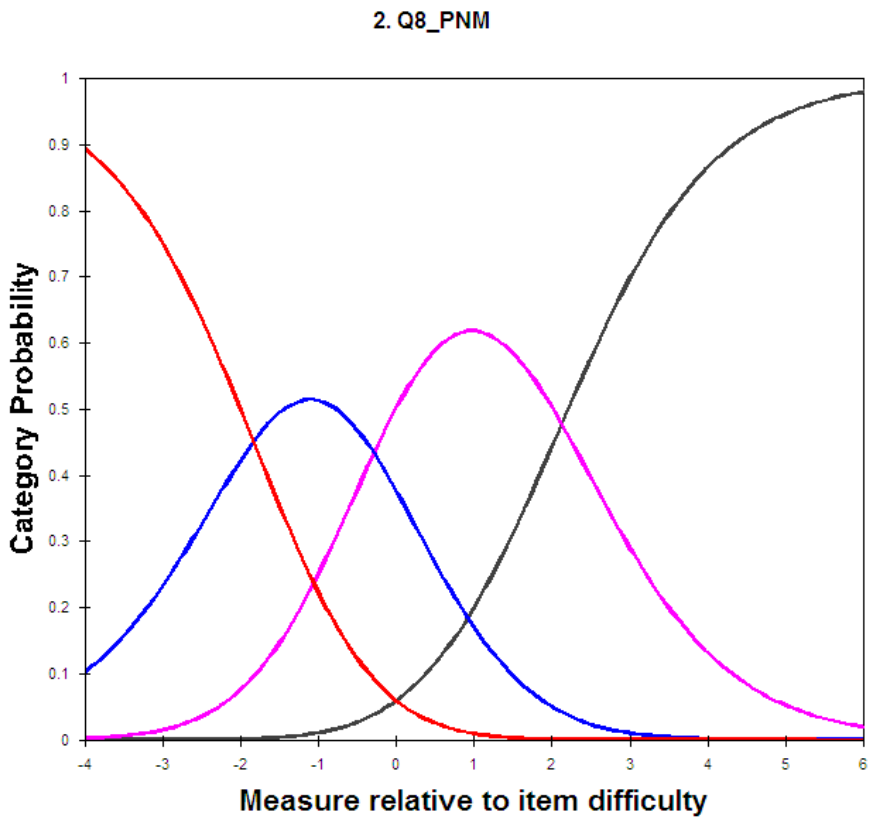

Figure 2. Physical and Manual Dimension Category Probability Curve

---

Graph depicts poor category probability curve for a four-point Likert scale.

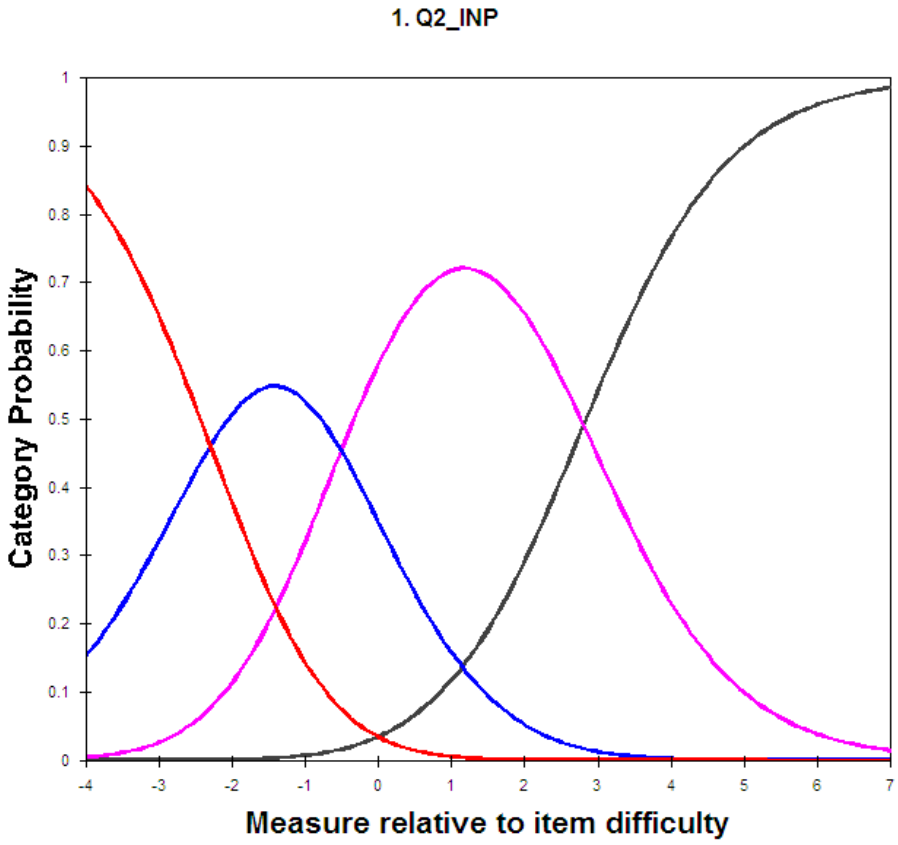

Figure 3. Interpersonal Dimension Category Probability Curve

Graph depicts poor category probability curve for a four-point Likert scale.

Conference calls or in-person meetings were conducted with reviewers to provide explanations for their decisions and to give any suggestions regarding items that may not have been included. Over an eight month period three phone conversations occurred with Chickering, two with Reisser and one conference call was conducted with all three of us to discuss changes to the preliminary version of the SCS-R. In addition to the conversations facilitated with Chickering and Reisser I met with faculty who had backgrounds in psychometrics and college student development theory. These discussions occurred in the beginning stages of item development. Following the item technical quality analyses a second expert review was conducted with Chickering and Reisser that is discussed later in the Results section.

***Brainstorming Sessions*.** Many authors have also suggested facilitating discussions with the target population in identifying specific constructs however researchers rarely follow through with the process (Vogt, et al., 2004, Haynes, et al., 1995; Smith & McCarthy, 1995). Two groups of traditional age college students were selected for brainstorming sessions to assist with the development of new items for the SCS-R. Two brainstorming sessions were held with a group of 3 students for the first session and 9 for the second session. Students expressed their thoughts about intellectual, physical and manual, and interpersonal abilities of college students by participating in a number of brainstorming exercises. Notes were gathered from both meetings and assisted in the development of items for the SCS-R.

After the development of the initial pool of items was created according to instrument specifications (Wolfe & Smith, 2007a), several content reviews and studies were conducted to address the relevancy of items. In addition, the reviews and studies assessed the capacity of each item for collecting accurate and significant data.

Item Editing (Phase 3). A group of five individuals were invited to review a draft set of the SCS-R and provided comments regarding the items clarity and conciseness. In addition, the group reviewed the instrument for any areas illustrating bias. The group consisted of individuals who had expertise in item writing and editing (e.g. faculty and doctoral students who had knowledge of measurement methodologies) or student development theory. The reviewers were given a draft set of items and were asked to make edits. Edits were then reviewed in conjunction with the reviewers. Reviewers were asked to provide feedback in response to the suggestions they made for the instrument. Input received from all reviewers was then considered and

changes to the SCS-R were made accordingly. For example one reviewer stated that the following item was considered to illustrate gender bias, I know how to use a set of socket wrenches.

Item Tryout (Phase 4). The draft set of items was administered to a group of 10 traditional age college students from the university where the instrument was being developed. Students that participated in the brainstorming sessions were asked to respond to the items and provide feedback. During administration of the items, students were asked to provide comments about how they felt while they completed the instrument and answered all items. Open-ended questions were also included to receive feedback on the ease of item interpretation and identification of potential bias. One student stated, "Allowing more space between the questions would make the survey easier to read. Sometimes the lines ran together if I read too quickly." Another student stated, "It got a bit repetitive." The repetitive feel of the instrument was probably due to the number of items on the preliminary version of the SCS-R.

Item technical quality (Phase 5). During the development phase of the instrument I followed the 12 guidelines for developing Likert scale items (Wolfe & Smith, 2007a). I reviewed items to avoid references to the past, factual statements and usage of absolutes such as 'all', 'none', 'always', and 'never.' Items were written in simple, short and clear statements and I restricted the use of words such as 'only', 'just', and 'merely.' Any one item that seemed to be asking two questions were broken down into separate items. Throughout the analysis I also avoided any suggestions to particular answers (Do you agree that….) as well as non-distinguishing items and negatively written items. All items were written at a reading level lower than the target population and any irrelevant items were discarded. To provide an accurate portrayl of the dimensions I wrote at least twice as many items that I would like to appear on the final instrument. Also, it is important to mention that Cronbach (1951) and Wolfe and Smith (2007a), recommend using an even number of categories leaving less room for error variance.

Using a four-point Likert scale is supported by Wolfe and Smith (2007a) in that, "fewer rating scale categories (between three and five) tend to provide better fit to the Rasch Rating Scale Model" (p. 13). Researchers (Crocker & Algina, 1986, Wolfe & Smith, 2007) recommend avoiding negatively worded items. "Analyses of negatively worded items may result in less reliable data, may demonstrate more rater bias and response sets, and may result in more

homogeneous item response distributions due to respondents being less willing to use extreme rating categories, particularly negative ones" (Wolfe & Smith, 2007a, p. 20).

Pilot test responses were also used to define the technical quality of each dimension (intellectual, physical and manual, and interpersonal) within an instrument. During the evaluation process the item point-measure correlation and the mean-squared fit indices were considered. The item point-measure correlation indicates the extent to which scores on an item are consistent with the average score across all other items (Wolfe & Smith, 2007b). The purpose of the fit analysis is to predict potential problematic response patterns as well as flag any misfit items. The mean square statistics can be interpreted as either over-fit, fit, or misfit. Overfit determines all observed responses are equal to the expected responses while misfit is the opposite all observed responses differ greatly from expected responses. The fit interpretation observed responses will differ from the expected responses based on the amount predicted (Wolfe & Smith, 2007b). The item point-measure correlation and fit analyses appear in the Results section.

Structural Validation

The structural aspect of validity evaluates "the fidelity of the scoring structure to the structure of the construct domain," (Wolfe & Smith, 2007b, p. 213). Evidence is illustrated in the form of correlation analyses of response consistencies and indicators. Messick (1995) states for the structural aspect of validity, "the theory of the construct domain should guide not only the selection or construction of relevant assessment tasks but also the rational development of construct-based scoring criteria and rubrics," (p. 746). The structural validation evidence was conducted by an analysis of dimensionality in the last phase, phase six.

Dimensionality analysis (Phase 6). Dimensionality is defined as the number of aspects or dimensions that represent a construct (Wolfe & Smith, 2007a). Dimensionality may be addressed in several ways. The Rasch Model requires unidimensionality and local independence to assess structural validity. Local independence occurs when all items are independent of one another for examinees of the same ability on a latent tendency (Osterlind, 2006). When analyzing unidimensionality all items of the instrument are scaled equally. A principle component analysis is conducted on the standardized residuals from the previous equal scaling. Eigenvalues from the equalized scaling and principal component analysis of the residuals are

placed onto a common scale. The analyses assist in determining if the model is multidimensional rather than unidimensional (Wolfe & Smith, 2007b). A review of the pilot study results illustrates additional support for content and structural validity.

Pilot Study

The preliminary version of the SCS-R (Appendices B, C, & D) contained 92 items which were administered to sophomore students who participated in recreational sports, were a member of the honors program, or lived on campus. A convenience sample was used to collect the data. Administrators from the honors program, recreational sports and residence life departments on-campus were asked to send the SCS-R to students who were of sophomore status. A total of 211 students responded to the questionnaire. A demographic instrument was included to illustrate general traits of the respondents as seen in Table 2.

The demographic section included items such as sex, ethnicity, housing and involvement. The demographic instrument included dichotomous and polytomous items. Students answered qualitative questions about how their experiences at the university have contributed to their intellectual, physical and manual, and interpersonal development. Lastly, subjects were asked to write any open-ended comments about the questionnaire. Those who were invited to provide responses included members of the populations for whom the instrument was developed, traditional age college students currently enrolled at a Research I institution. The results were relatively representative of the university's demographics for the sophomore class with 53.4% female (45% university total representation) and 46.6% male respondents (55% university total representation). As far as ethnicity, the results were within one to two percentage points for a few categories for representation of the sophomore class, for example Black, Non-Hispanic was 2% (4% university total representation), Hispanic was 2.4% (3.2% university total representation), and Non-Resident Alien was 0 (1.7% university total representation).

The SCS-R was administered as a web-based questionnaire to be completed when it was convenient for the student (e.g. after class, work, organizational meeting). In the web-based format, the draft instrument was sent to participants through a survey system set up by the university. The instrument included a cover letter, which asked participants to complete the instrument and submit responses by a specified date. A second, follow-up email was sent as a reminder to help provide for a larger response rate. To increase participation in the study

Table 2

Demographic Characteristics of Pilot Test Respondents

Variable	Category	*N*	% of those responding
Gender	Female	110	53.4
	Male	96	46.6
	Missing	6	--
Group in which you Identify with most	American Indian/Native American	3	1.5
	Asian or Pacific Islander	17	8.3
	Black, Non-Hispanic	4	2.0
	White, Non-Hispanic	168	82.0
	Hispanic	5	2.4
	Non-Resident Alien	0	0
	Other	8	3.9
	Missing	6	--
Live on-campus	Yes	119	58.0
	No	86	42.0
	Missing	2	--
Member of honors program	Yes	75	36.6
	No	130	63.4
	Missing	7	--
Participate in extra-curricular activities	Yes	192	93.2
	No	14	6.8
	Missing	6	--
Participate in intramurals	Yes	150	72.8
	No	56	27.2
	Missing	6	--

students were given the opportunity to download their name in a separate survey for a drawing. The winner of the drawing received a $50 gift certificate to the university bookstore. Names submitted for the drawing were not connected to the individuals' survey response.

Directions were provided within each section of the instrument, informing participants about the means by which they were to provide responses. A coding scheme was developed to facilitate the recording of data into a numerical format that was amenable to subsequent statistical analyses.

Results

After facilitating the pilot study all completed questionnaires were placed in an excel document. Each item on the instrument was coded by a 1, 2, 3, or 4 for the Likert scale responses. The excel document included all answers from the participants. The excel document was used to import data into *WINSTEPS* (Linacre, 2006) software to evaluate item technical quality and dimensionality of the SCS-R.

Item Technical Quality. Developing items that measure each dimension (intellectual, physical and manual, and interpersonal) is important and yields high technical quality. The technical quality of items that appear on the SCS-R were explored using two sets of indices: item point-measure correlations (r_{pm}), and mean-square (MS) fit statistics (Wolfe & Smith, 2007b).

Pilot test responses were used to define the technical quality of each domain (intellectual, physical and manual, and interpersonal). A guideline for using the Rasch (1960) Model is that all items are scaled to be oriented in the same direction (Wolfe & Smith, 2007b). This guideline was explored using r_{pm}, which is the Pearson Product Moment Correlation (r) between the bearing of item scores and person measures. Items that have negative r_{pm} and/or values < .30 were flagged and investigated to determine the source of their failure to contribute effectively to the Rasch (1960) Model.

The technical quality was examined through a review of the difference between the Rasch (1960) Model's predicted responses and observed responses:

$$residual_{ni} = E_{ni} - X_{ni}$$

where:

E_{ni} = the expected response for participant n on item i, and

X_{ni} = empirically observed response for participant n on item i

Analyses were examined using *WINSTEPS* (Linacre, 2006) software to determine the outfit and infit mean scores. Estimates of item weighted (infit) and unweighted (outfit) mean squares (*MS*) and the standardized version of the unweighted mean squares ($Z_{unweighted}$) were used to identify possible item misfit. Values for both infit ($MS_{weighted}$) and outfit ($MS_{unweighted}$) were examined due to the fact that the $MS_{weighted}$ index is sensitive to non-extreme unexpected responses and the $MS_{unweighted}$ index is sensitive to extreme unexpected responses (Karabatsos, 2000). The expected value for both infit and outfit is +1.0, with possible values ranging from 0 to positive infinity (∞).

Values closer to zero and values < 1.0 indicate overfit, where all observed responses are close to equal to the Rasch (1960) Model's expected responses. Values > 1.0 indicate misfit, where observed responses differ from model expectations. Guidelines for the interpretation of these values were referenced in the review of estimates of item misfit, where: *MS* > 2.0 distorts the measurement system; *MS* = 1.5 - 2.0 are unproductive for the construction of measurement, but not degrading; *MS* = 0.5 - 1.5 are productive for measurement; and, *MS* < .05 are less productive for measurement, but not degrading, and may produce misleading reliabilities (Linacre, 2002).

Table 3 shows item technical quality and lists a total of 17 items exhibiting misfit where *MS* > 2.0 distorts the measurement system. Of the 17 items, the decision was made to delete nine items representing intellectual, while the physical and manual and interpersonal had four items each exhibiting misfit. Figure 4 and 5 illustrate an example of one item exhibiting poor fit and another exhibiting good fit.

Following the pilot test a total of 17 items (5 intellectual, 6 physical and manual, and 6 interpersonal) out of 92 items were deleted based on discussions facilitated with Chickering and Reisser (personal communication, January 28, 2010). Again it is important to note that three of the 17 items were deleted based on the item technical quality analyses and discussions with Chickering and Reisser. Table 4 shows a total of 31 items were deleted from the 92 item preliminary version of the SCS-R. Deletions were made if items related more to one of the other six vectors (constructs). In addition, items reflecting social desirability were deleted. It is important to note that some of the items suggested for deletion by Chickering and Reisser were also items illustrating misfit during the item technical quality analyses. Chickering and Reisser

Table 3

Item Misfit from Pilot Test

Item Text	Dimension	MS_{infit}	$Z_{weighted}$	MS_{outfit}	$Z_{unweighted}$	r_{pm}	*Decision*
9. I will need tutoring in a specific course.	Intellectual	1.24	2.4	1.24	2.4	.40	Delete
15. I have spent time participating in artistic activities.	Physical & Manual	1.50	4.6	1.50	4.6	.44	Delete
17. I will graduate with honors.	Intellectual	1.34	3.3	1.40	3.7	.43	Delete
22. I will fail at least one college course.	Intellectual	1.54	4.9	1.50	4.6	.39	Delete
28. I dedicate time to regular exercise workouts.	Physical & Manual	1.43	3.9	1.42	3.9	.42	Delete
31. I am afraid of making mistakes.*	Intellectual	1.24	2.6	1.28	3.0	.45	Delete
37. I share my personal weaknesses with close friends.	Interpersonal	1.61	4.9	1.60	4.8	.45	Delete
39. I am frustrated when I have to work in groups.	Interpersonal	1.40	3.7	1.45	4.1	.47	Delete
46. I have difficulty organizing information to write a paper.	Intellectual	1.85	7.7	1.93	8.2	.45	Delete
51. I know where to go on campus if I need assistance with my writing.	Intellectual	1.62	5.1	1.65	5.4	.42	Delete
57. I talk with family members about my college experience.	Interpersonal	1.42	3.7	1.42	3.8	.43	Delete
62. I appreciate the arts.	Physical & Manual	1.42	3.8	1.42	3.8	.39	Delete
64. I stay up to date on current affairs.	Intellectual	1.34	3.3	1.37	3.6	.44	Delete
66. I speak up in class to voice my opinion.	Intellectual	1.20	2.1	1.22	2.3	.44	Delete
75. I am intimidated by my peers' academic abilities.	Intellectual	1.37	3.4	1.36	3.3	.43	Delete
86. I feel it is alright to fail.*	Interpersonal	1.69	6.7	1.78	7.2	.50	Delete
89. I dedicate time to volunteering with diverse people in the community.*	Physical & Manual	1.31	3.0	1.34	3.3	.44	Delete

*Items were also requested to be deleted by Chickering and Reisser (personal communication, January 28, 2010).

Table 4

Items Deleted from the Preliminary Version of the SCS-R

	Item	Deleted
1.	I read fast enough to handle assignments at college.	
2.	I am comfortable meeting new people.	
3.	I have little difficulty with college work.	
4.	I feel empathy for others.	
5.	I respond to the needs of others.	
6.	I make friends easily.	
7.	I am competent using computers.	
8.	I enjoy participating in physical activity.	
9.	**I will need tutoring in a specific course.**	**X**
10.	College work is easy.	
11.	I trust the decisions I make.	
12.	**I can cope with daily life challenges.**	**X**
13.	I get along with most of the people I meet.	
14.	I believe it is important to live a healthy lifestyle.	
15.	**I have spent time participating in artistic activities.**	**X**
16.	I feel overwhelmed with academic work.	
17.	**I will graduate with honors.**	**X**
18.	I am clumsy when it comes to participating in physical activity.	
19.	**I openly show my emotions.**	**X**
20.	**I am comfortable with my body.**	**X**
21.	I get frustrated when using new computer software.	
22.	**I will fail at least one college course.**	**X**
23.	I will achieve my academic goals.	
24.	Professors underrate my ability.	
25.	**I am afraid to show my emotions.**	**X**
26.	I converse casually with my friends.	
27.	I am competent using social networking sites.	
28.	**I dedicate time for regular exercise workouts.**	**X**
29.	I get frustrated when reading instructions to put something together.	
30.	**I understand my physical limitations for my body type.**	**X**
31.	**I am afraid of making mistakes.**	**X**
32.	I welcome criticism as an opportunity for growth.	
33.	**I feel comfortable with how I prioritize my time.**	**X**
34.	I have difficulty analyzing problems to identify solutions.	
35.	**I get distracted easily during class lectures.**	**X**
36.	I have serious conversations with friends.	
37.	**I share my personal weaknesses with close friends.**	**X**
38.	I talk with professors easily about my concerns.	
39.	**I am frustrated when I have to work in groups.**	**X**
40.	I feel better when I participate in physical activity.	
41.	I learn from hands on experiences in the classroom.	
42.	I am confident in my ability to use technology when designing a project for class.	
43.	I find answers with well-formed questions.	
44.	I can identify a problem and explain it in a way others understand.	
45.	I am accepting of new ideas.	
46.	**I have difficulty organizing information to write a paper.**	**X**
47.	I have efficient not taking skills.	
48.	I am comfortable with sharing my own personal beliefs about issues.	
49.	I am accepting of others who are different from me.	
50.	**I can resist peer pressure from other college students.**	**X**
51.	**I know where to go to on campus if I need assistance with my writing.**	**X**
52.	I am self-motivated when starting class projects.	

53.	**I can identify areas where I need improvement.**	**X**
54.	I know how to find scholarly resources on a specific topic area.	
55.	**I have a willingness to learn.**	**X**
56.	I am afraid to ask for a professor's help.	
57.	**I talk with family members about my college experience.**	**X**
58.	**I stand up for what I believe to be important.**	**X**
59.	I spend time weekly participating in a relaxing activity.	
60.	I make healthy food choices.	
61.	I adapt well to new technology.	
62.	**I appreciate the arts.**	**X**
63.	I get an adequate amount of sleep.	
64.	**I stay up to date on current affairs.**	**X**
65.	I create my own perspective from information I learned in class.	
66.	**I speak up in class to voice my opinion.**	**X**
67.	I know how to conduct myself in the classroom environment.	
68.	I am afraid to ask questions in class.	
69.	I introduce myself to new people I meet.	
70.	I share my problems with others.	
71.	I know when I need to take a break.	
72.	I try new physically challenging activities.	
73.	I apply information I have learned in class.	
74.	I participate often in class.	
75.	**I am intimidated by my peers' academic abilities.**	**X**
76.	I understand assigned readings for my classes.	
77.	I respect others' decisions.	
78.	I am comfortable saying no to my peers.	
79.	I can balance my relationships with school priorities.	
80.	**I can deal with rejection.**	**X**
81.	**I know how to navigate campus.**	**X**
82.	**I balance school work with free time.**	**X**
83.	I can relate ideas learned across subject areas.	
84.	I write efficiently for my area of study.	
85.	I use information I learn in class.	
86.	**I feel it is all right to fail.**	**X**
87.	I can work on class projects with difficult people.	
88.	I have established new friendships.	
89.	**I dedicate time to volunteering with diverse people in the community.**	**X**
90.	I find time for breaks during the day.	
91.	I am confident in my ability to assemble a book shelf.	
92.	I have adjusted well to living with a roommate.	

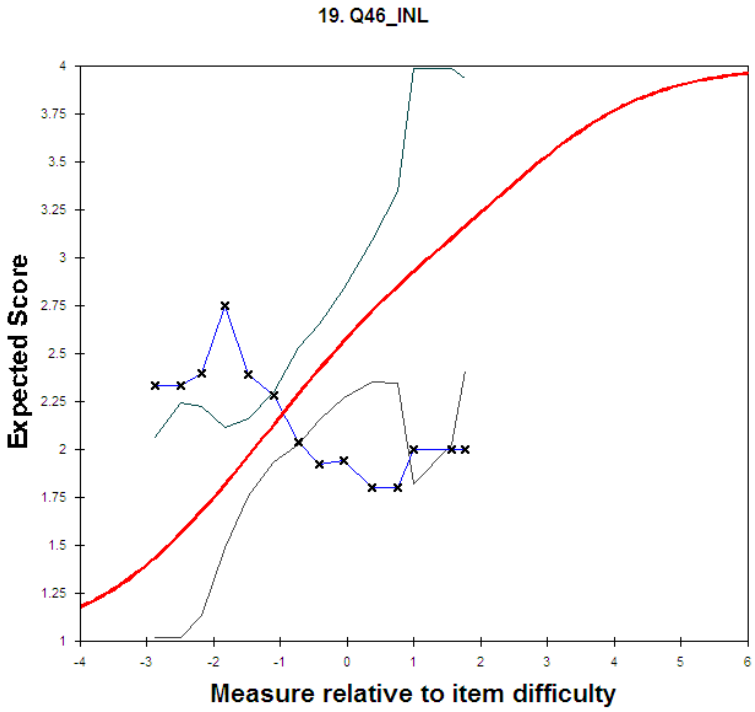

Figure 4. Item Characteristic Curve

Graph depicts poor fit for item 46, I have difficulty organizing information to write a paper.

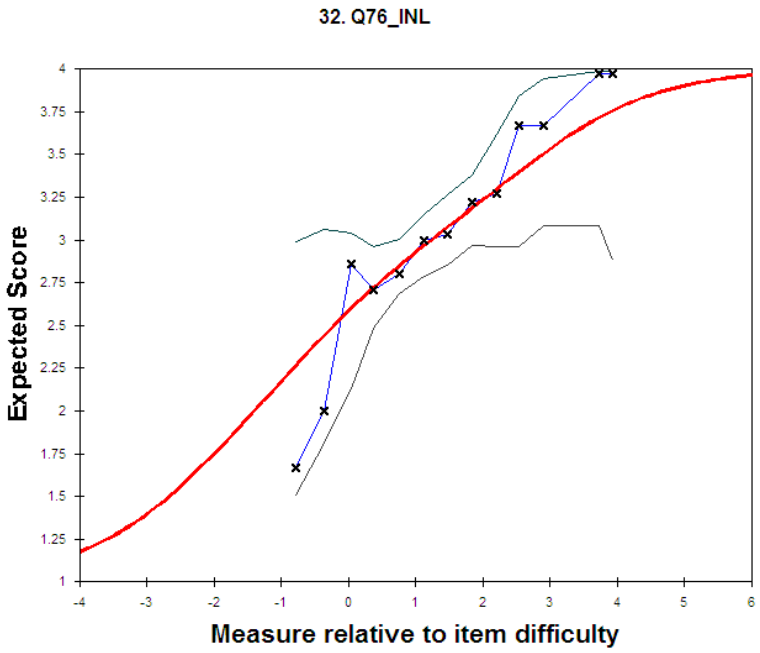

Figure 5. Item Characteristic Curve

Graph depicts good fit for item 76, I understand assigned readings for my classes.

suggested the creation of three to five additional items for the interpersonal dimension to illustrate group functioning and listening skills as well as rewriting a few items to better express vector one (A. Chickering & L. Reisser, personal communication, January, 28, 2010).

As requested by Chickering and Reisser (personal communication, January 28, 2010) two additional items were added to represent aspects of group work within the interpersonal dimension. The following two items were added; I listen to group member's ideas when working on class projects and I respect other's decisions when working on class projects.

Dimensionality. Evidence relating to the structural aspect of validity was gathered through an exploration of dimensionality analyses of the preliminary version of the SCS-R. Intellectual, physical and manual, interpersonal dimensions were run as separate constructs to define sense of competence. Data analyses for each dimension were conducted separately in *WINSTEPS* (Linacre, 2006) to identify unidimensionality of each aspect of the overarching construct sense of competence.

WINSTEPS (Linacre, 2006) contains a feature in which a principal components analysis (PCA) of the standardized residuals is performed after extracting the Rasch dimension. *WINSTEPS* (Linacre, 2006) constructs a series of "contrasts," where each reflects opposing response patterns across items by persons. An analysis of the strength of those contrast components, expressed in eigenvalues units, is often useful in determining the degree to which opposing factors are providing structure in standardized residuals after data have been scaled to a model. With respect to the interpretation of the magnitude of a contrast, an eigenvalue of 4, for example, may be interpreted as a residual component holding as much strength as 4 items. In terms of the relative importance of contrasts versus the modeled dimension, Linacre (2006) suggested using the following Rules of Thumb: (a) percentage of variance explained by measures > 60% is good, (b) unexplained variance explained by the 1st contrast – eigenvalue < 3.0 is good, and (c) unexplained variance explained by 1st contrast < 5% is good. However, it is important to realize that these are only guidelines.

Table 5 illustrates the performance of items based on each dimension as well as the performance of the 61-item set (total of items after Chickering and Reisser's review and item technical quality analyses). Based on the preliminary analyses into the dimensionality of the pilot test items suggest that there are three dimensions.

Table 5

Performance Indices for Dimensions of Items

	Intellectual	Physical & Manual	Interpersonal
Number of Items	22	17	22
Rel_{θ}	.82	.74	.83
Cronbach's α	.98	.98	.99
$r_{pm} \geq .50$	9 items	3 items	7 items
$r_{pm} \geq .40$	10 items	9 items	12 items
$r_{pm} \geq .30$	2 items	4 items	3 items
$MS_{infit} \geq 1.5$	0 items	0 items	0 items
$Z_{unweighted} \geq 2.0$	0 items	0 items	0 items
Total Var_{θ} explained by Rasch Dimension	39%	43.4%	42.8%
Eigenvalue for 1st Contrast	2.3	3.4	2.6
Var_{θ} unexplained for 1st Contrast	6.3%	11.4%	6.5%

Note. Item 32 (intellectual) is not noted in the rpm because it is represented by a low r_{pm} of 1.8.
Item 41 (physical and manual) is not noted in the rpm because it is represented by a low r_{pm} of 2.3.

Discussion and Implications

The objective of this research was to develop and validate the preliminary version of the SCS-R that could be used to measure a student's sense of competence. The preliminary version of the 92 item instrument was piloted to determine if the SCS-R represented all three dimensions of sense of competence. The pilot study data illustrated that there were three dimensions. Validity was assessed by administering the preliminary version of the instrument to sophomores at a Research I institution and then using the scores to better understand the sense of competence construct.

Content validity was evaluated by discussions with experts that revealed some items reflected other vectors (i.e. managing emotions, establishing identity) of Chickering's (1969; Chickering & Reisser, 1993) theory and supported the decision for certain items to be deleted. The scores of the pilot were used to determine item technical quality of the items and determine which items illustrated evidence for demonstrating misfit. From the scores item technical quality analyses revealed 17 items exhibiting misfit. During the content validation phase a total of 30 items were deleted prior to the dimensionality analyses.

Structural validity was evaluated by conducting dimensionality analyses. The analyses revealed low variance scores explained by the Rasch model. This suggests that a field test of the final version of the SCS-R should be conducted with a larger sample and response rate. Eigenvalues were below 3.0 on two of the dimensions (intellectual and interpersonal), however the physical and manual dimension expressed an eigenvalue of 3.5. Future investigations can improve the SCS-R in several ways. First, additional dimensionality analyses, such as calculating the Kaiser's eigenvalue to investigate unidimensionality, should be conducted by exploratory factor analysis. The criterion proposes that only eigenvalues of 1.00 and factors above the elbow in a scree plot should be retained (Cattell, 1966).

A second way the SCS-R can be strengthened is by conducting a larger pilot test, to test other aspects of Messick's validity framework. For example, a review of substantive, generalizability, external and consequential aspects of construct validity could be analyzed. Substantive validity provides empirical evidence of response consistencies and generalizability validity addresses the degree to which measures maintain their meaning across measurement contexts. External validity illustrates the degree to which measures are related to external

measures of the same construct, similar constructs, and other constructs. Lastly, consequential validity addresses actual and potential consequences of how test scores will be used, in regard to sources of invalidity such as fairness or bias. Future research is planned to test the remaining aspects validity analyses outlined by Messick (1995) to form the final version of the SCS-R.

Conclusion

Conducting thorough content and structural validity methods are crucial when developing an instrument to measure a specific construct. Although time-consuming and potentially costly, taking time to meet with experts and the target population to define the construct being measured is imperative during the item development phase (Vogt, et al., 2004).

Determining content and structural validity are only two steps in the validation process. Instruments that are developed with a strong content validity stage will have a well defined purpose and possess stronger psychometric properties than those instruments developed in a less rigorous manner.

The preliminary version of the SCS-R provides student affairs practitioners and university administrators with a tool for understanding college students' sense of competence. Furthermore, examination of the three dimensions may reveal to administrators that students at their university are not confident in their skills informing staff that programs need to be created or adjusted to enhance the learning and social environment for their students.

More attention to the development of instruments is needed in the field of higher education to better assess college and university environments and enhance the student experience. Using strong content validity tools and methods such as expert reviews, brainstorming sessions with the target population, and item technical quality are imperative in the development stage.

Clearly, it is important for all those involved in the instrument development process to remember that content and structural aspects of validity are only part of the process for creating psychometrically sound instruments in construct validation. Instruments that are content valid for their intended purpose and reflect other psychometric properties have the potential to make a substantial contribution to the higher education literature. More researchers along with student affairs practitioners should take the time to develop an instrument that properly measures an important construct such as sense of competence.

References

Catell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chickering, A. (1969). *Education and identity*. San Francisco: Jossey-Bass.

Chickering, A., & Reisser, L. (1993). *Education and Identity* (2nd ed.). San Francisco: Jossey-Bass.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Thomson Learning.

Cronbach, L. J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika, 16*, 297-334.

Cross, W. E., Jr. (1991). *Shades of black: Diversity in African American identity*. Philadelphia: Temple University Press.

D'Augelli, A. R. (1994). Identity development and sexual orientation: Toward a model of lesbian, gay, and bisexual development. In E. J. Trickett, R. J. Watts, & D. Birman (Eds.), *Human diversity: Perspectives on people in context* (pp. 312-333). San Francisco: Jossey-Bass.

Erikson, E. H. (1950). *Childhood and society*. New York: Norton.

Erikson, E. H. (1968). *Identity: Youth and crisis*. New York: Norton.

Evans, N. J. (1996). Theories of student development. In S. R. Komives, D. B. Woodard Jr., & Associates, *Student services: A Handbook for the profession* (3rd ed., pp. 164-187). San Francisco: Jossey-Bass.

Evans, N. J., Forney, S. F., & Guido-DiBrito, F. (1998). *Student development in college: Theory, research, and practice*. San Francisco: Jossey-Bass.

Evans, N. J., Forney, S. F., Guido, F., Patton, L. D., & Renn, K. A. (2010). *Student development in college: Theory, research, and practice*. (2nd Ed.) San Francisco: Jossey-Bass.

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

Hood, A. B. (Ed.). 1986. The Iowa Student Development Inventories. Iowa City, Iowa:

Hitech Press.

Hood, A. B., & Jackson, L. M. (1983). *Assessing the development of competence.* Technical report. Iowa City: University of Iowa, College of Education.

Janosik, S. M. (1987). Relationship of residence hall environments and student sense of competence and academic achievement (Doctoral dissertation, Virginia Tech, 1987).

Josselson, R. (1987). Identity diffusion: A long-term follow-up. *Adolescent Psychology, 14*, 230-258.

Karabatsos, G. (2000). A critique of Rasch residual fit statistics. *Journal of Applied Measurement, 1*(2), 152-176.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 44-53.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 266-283.

Linacre, J. M. (2006) WINSTEPS. A User's guide to Winsteps: Rasch Model computer programs. Chicago, IL: Retrieved February 21, 2009 from www.winsteps.com/a/winsteps.pdf.

Marcia, J. E. (1966). Development and validation of ego-identity status. *Journal of Personality and Social Psychology, 3*, 551-558.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.

Miller, T. K., & Winston, R. B., Jr. (1990). Assessing development from a psychosocial perspective. In D. G. Creamer & Associates, *College student development: Theory and practice for the 1990's* (pp. 99-126). Alexandria, VA: American College Personnel Association.

Moran, C. (2003). Conceptualizing identity development: Unmasking the assumptions within inventories measuring identity development. NASPA Journal, 46, 3, p. 469 – 481.

Osterlind, S. J. (2006). *Modern measurement: Theory, principles, and applications of mental appraisal.* Upper Saddle River, NJ: Pearson Education.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of*

research (Vol. 2). San Francisco: Jossey-Bass

Phinney, J. S. (1990). Ethnic identity in adolescents and adults: Review of research. *Psychological Bulletin, 108*, 499-514.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment test*. Copenhagen: Denmark Pedogogiske Institute. (Reprinted 1980 by University of Chicago.)

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (revised and expanded ed.). Chicago: The University of Chicago Press. (Originial work published 1960.)

Rodgers, R. F. (1990). Recent theories and research underlying student development. In D. Creamer & Associates (Eds). *College student development: Theory and practice for the 1990s* (pp. 27-790). Alexandria, VA: American College Personnel Association.

Shostrom, E. L. (1963). *Personal orientation inventory*. Educational and Industrial Testing Service. San Diego, CA

Smart, J. C. (2005). Attributes of exemplary research manuscripts employing quantitative analyses. *Research in Higher Education, 46*, 461-477.

Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment, 7*, 300-308.

Vogt, D. S., King, D. W., & King, L. A. (2004). Focus groups in psychological assessment: Enhancing content validity by consulting members of the target population. *Psychological Assessment, 16*, 3, 231-243.

Winston, R. B., Jr. (1990). The Student Developmental Task and Lifestyle Inventory: An approach to measuring students' psychosocial development. *Journal of College Student Development, 31*, 108-120.

Winston, R. B., Jr., Miller, T. K., & Prince, J. S. (1987). *Student Developmental Task and Lifestyle Inventory.* Athens, GA: Student Development Associates.

Winston, R. B., Jr., Miller, T. K., & Prince, J. S. (1999). *Student Developmental Task and Lifestyle Assessment.* Athens, GA: Student Development Associates.

Wright, B. D., & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. In E.V. Smith, Jr. & R. M. Smith (Eds.), *Introduction to Rasch Measurement* (pp. 1-24). Maple Grove, MN: JAM Press.

Wright, B. D., & Stone, M. H. (2003). *Managing observations, inventing constructs, crafting yardsticks, examining fit.* Chicago: Phaneron Press.

Wolfe, E. W., & Smith, E. V., Jr. (2007a). Instrument development tools and activities for measure validation using Rasch Models: Part I-Instrument development tools. In E. V. Smith & R. M. Smith (Eds.), *Rasch Measurement: Advanced specialized applications* (pp. 2-27). Maple Grove, MN: Journal of Applied Measurement Press.

Wolfe, E. W., & Smith, E. V., Jr. (2007b). Instrument development tools and activities for measure validation using Rasch Models: Part II-Instrument development tools. In E. V. Smith & R. M. Smith (Eds.), *Rasch Measurement: Advanced specialized applications* (pp. 204-234). Maple Grove, MN: Journal of Applied Measurement Press.

Chapter 5

Manuscript 2

Development and Validation of the Sense of Competence Scale-Revised (SCS-R)

Abstract

The development and validation of the instrument, titled Sense of Competence Scale-Revised (SCS-R), was conducted to provide a measure of Chickering's (1969) first vector, an important psychosocial construct. A valid and reliable measure of sense of competence would allow administrators to modify an institution's academic and social environment and enhance the development of the intellectual, physical, and interpersonal competencies of college students. The author discusses three dimensions of the sense of competence construct and details the development and validation of measures using Rasch Models. Various aspects of validity as outlined by Messick (1995) are addressed and supported by activities outlined by Wolfe and Smith (2007b). This article describes the activities and documents content, substantive, structural and generalizability forms of validity evidence. Results support intellectual, physical and manual and interpersonal constructs as three unidimensional measures that measure college students'personal assessment of their sense of competence.

Keywords: student development, instrument validation, Rasch measurement

Chapter 5

Manuscript 2

Development and Validation of the Sense of Competence Scale-Revised (SCS-R)

Better methods to assess student development must be constructed, "too many studies use 'home grown' assessment tools that may have face validity but have not been thoroughly examined for reliability or validity in a systematic way" (Evans, Forney, Guido, Patton, & Renn, 2010, p. 368). Researchers in higher education need to dedicate time and effort in the development of psychometrically sound instruments (Smart, 2005). Smart (2005) contends that lack of attention to detail creates a deficiency in higher education research. Because of this deficiency, there are only a few reliable or valid instruments available to measure student development and those that do exist are hard to administer (Evans, Forney, DiBrito, 1998).

Student affairs practitioners use psychosocial theories frequently in their work. In this area of inquiry, Arthur Chickering's Vector Theory of Student Development (Chickering, 1969) is one of the most popular. The questionnaires that attempt to measure this development suffer from the shortcomings previously mentioned here. Inventories such as the Student Development Task and Lifestyle Inventory (SDTLI), 1987; the Student Development Task and Lifestyle Assessment (SDTLA) updated version of the SDTLI, 1999, the Iowa Student Development Inventories (ISDI), 1986; and the Sense of Competence Scale (SCS), 1987 suffer from a lack of focus or lack of sufficient psychometric study. The absence of reliable and valid measures of student development in the higher education literature provides support for the development of the Sense of Competence Scale-Revised (SCS-R), which would measure Chickering's first vector named Sense of Competence, an important psychosocial construct.

The purpose of this study is to develop the SCS-R as a measure of a student's sense of competence, as defined by Chickering and Reisser (1993). Administrators could use data from such an instrument to modify an institution's academic and social environment to enhance the development of the intellectual, physical, and interpersonal competencies of college students.

I begin with an overview of the construct (sense of competence) and then discuss the three dimensions (intellectual, physical and manual, interpersonal skills) that define the construct. After a brief description of the data set used in the analyses, a sample of respondents from a large Research I institution, I describe the methods and findings.

The following were the research questions of inquiry:

1. What is the item technical quality of the items? (Content)
2. Is there evidence that the theoretical model explains the observed stability among item responses (person item fit)? (Substantive)
3. What dimensional structure best depicts traditional age college students' sense of competence? (Structural)
4. Which measurement model and rating scale configuration best depicts the rating scale structure of these data? (Substantive)
5. Do the three dimensions (intellectual, physical and manual, and interpersonal) of the revised SCS-R establish a reliable measurement to measure a student's personal assessment of their intellectual, interpersonal and physical and manual competence? (Generalizability)

Developing Sense of Competence

"Competence is a three-tined pitchfork. Intellectual competence, physical and manual skills, and interpersonal competence are the tines. But the handle is most important. Without it, no work can be done, no matter how sharp and sturdy the tines. A sense of competence stems from the confidence that one can cope with what comes and achieve goals successfully," (Chickering & Reisser, 1993, p. 53). The SCS-R plans to measure college students' personal assessment of their intellectual, physical and manual and interpersonal skills.

Dimensions

After review of literature, speaking with experts in the field and conducting brainstorming sessions with the target population I developed items to represent the three dimensions of sense of competence (intellectual, physical and manual, interpersonal).

Intellectual. The intellectual ability involves, "mastering content, gaining intellectual and aesthetic sophistication, and, most important, building a repertoire of skills to comprehend, analyze, and synthesize," (Chickering & Reisser, p. 45). Most importantly, intellectual skills comprise the ability to reason, solve problems and participate in active learning opportunities (Chickering & Reisser, 1993). Chickering and Reisser (1993) urge faculty members and college administrators not to strictly define intellectual competence as "skills at passing tests or mastery of some 'essential' knowledge," (p. 63) but also as the ability to listen, question, reflect, and

communicate (Chickering & Reisser, 1993). In addition, students should be an active participant in searching for knowledge rather than using a more passive approach (Chickering & Reisser, 1993).

Physical and Manual. Physical and manual abilities contribute to a student's sense of competence. These skills come from a variety of activities being either athletic or artistic in nature. These skills are derived from participation in athletics and recreational sports, attention to wellness, and involvement in performing arts, tangible creations, and hands-on learning (Evans, et al., 1998, Chickering & Reisser, 1993). For a few students, participation in such activities become a vocation while for others the skills become an avocation (Chickering & Reisser, 1993). Vocation is defined as a career pursuit or routine, while avocational activities are described as a hobby or leisure pursuit. Physical skill development seems obvious when one learns to kick a soccer ball, take photographs, dance or sculpt, however little research exists illustrating the development of these skills while in college (Chickering & Reisser, 1993).

Interpersonal. Along with physical ability as a component of developing competence comes the facility to interact with others. A student's interaction with others contributes to their level of interpersonal competence. Interpersonal skills include things like listening, self-disclosing, and participating in dialogue that brings insight and satisfaction (Chickering & Reisser, 1993). These skills are, "a prerequisite for building successful friendships and intimate relationships," (Chickering & Reisser, 1993, p. 77) and "for playing one's role as a citizen" (Chickering & Reisser, 1993, p. 77). Intellectual, physical and manual, and interpersonal abilities are all components of developing competence.

However, it is important to remember that a student's overall sense of competence is subjective; sense of competence stems from how an individual student feels about their achievements and can trust their own abilities. Some students may take their level of competence for granted by having strong interpersonal skills, while other students may think no matter what they achieve it is never enough leaving them unsure of their abilities (Chickering & Reisser, 1993). University administrators who strive to provide opportunities for students such as electives and extracurricular activities will create a foundation for students to build upon during their time in college. Through these "increments of growing mastery and assuredness…the development of competence occurs" (Chickering & Reisser, 1993, p. 82). The challenge is for

university administrators to identify methods for measuring sense of competence to discern a student's development and self-confidence in their abilities.

Original Measure

Developing sense of competence, the first vector of Chickering's (1969) theory was the focus of this study. Competence was defined as intellectual, physical and manual, and interpersonal skills (Chickering, 1969; Chickering & Reisser, 1993). Instruments such as the Student Development and Lifestyle Inventory (SDTLI) and Iowa Developing Competency Inventory (IDCI) measure areas of Chickering's (1969) model; however, there is no instrument that specifically measures sense of competence as defined by Chickering and Reisser (1993).

The SCS developed by Janosik (1987) comes the closest but it does not include the third skill, physical and manual as described in Chickering and Reisser's (1993) vector of achieving sense of competence. The development of the SCS took place more than 20 years ago after finding that there was no instrument at the time that dealt satisfactorily with Chickering's (1969) notion of sense of competence. With the absence of an existing instrument, the SCS was developed and consisted of 23 items to measure intellectual and interpersonal competence of college students (Janosik, 1987). Items were developed after a review of instruments that measured self-esteem and other personal characteristics of college students.

I developed the SCS-R to assess all three dimensions of developing sense of competence as defined by Chickering (1969) and Chickering & Reisser (1993). For the SCS-R to accurately measure sense of competence, extensive development and validation procedures were followed as outlined by Wolfe and Smith (2007b). In the proceeding section I highlight the developmental procedures and follow up with results from the validation analyses.

Data Source and Methodology

The research questions stated will be answered using field test data from the SCS-R. This web-based survey was administered to all enrolled undergraduate students holding sophomore class standing at a Research I institution. The sophomore class was represented by approximately 5460 students. Approximately 17% (916 respondents) of students with sophomore class standing responded to the questionnaire. The survey asked students about their intellectual, physical and manual and interpersonal competence as well as demographic information. Overall, students were asked to respond to 63 Likert style questions.

I used Messick's (1995) validity framework along with activities outlined by Smith and Wolfe (2007b) to develop and validate the SCS-R as outlined in Table 1. This methodology includes a variety of activities that occur in the development and validation of items for an instrument. Items were added or deleted from the SCS (Janosik, 1987) with permission to develop the SCS-R.

Dimensions and Concept Maps

Developing sense of competence consists of three dimensions: intellectual, physical and manual, and interpersonal. Figure 1 diagrams this conceptual model. Each dimension is seen as being a contributor to the construct (sense of competence), no one particular dimension is more important than the other. Answers to the items on the SCS-R reflect students' assessment of their capabilities among all three dimensions. Figures 2, 3 and 4 illustrate the individual conceptual map for each dimension (intellectual, physical and manual, interpersonal).

During the development phase, each dimension (intellectual, physical and manual, interpersonal) of sense of competence was operationalized by a series of items to be a representative sample of all possible indicators. However, through literature and expert review as well as brainstorming sessions with the target population an appropriate number of items were developed. A set of 26 original items were created for the physical and manual dimension. Along with the original items from the SCS (Janosik, 1987) additional items were developed for the intellectual and interpersonal dimensions for the SCS-R based on expert reviews and brainstorming sessions held with the target population. This lead to a final total of 35 items for the intellectual scale and 31 items for the interpersonal scale.

Table 2 provides a reference for how items were created and/or adapted in accordance with the instrument blueprint. The number of items for physical and manual skills listed in Table 2 represents a proposed outcome for the SCS-R, and not necessarily the number of items that were on the final instrument. At least twice as many items were created for physical and manual skills. Items for intellectual and interpersonal include original items from the SCS as well as newly generated items. All items of the SCS-R were refined and field tested leading to the final instrument. Tables 3-5 illustrate the test blueprint for each dimension. Each dimension description includes a definition of the skill and an item breakdown for measuring a student's assessment of their capabilities.

Table 1

Instrument Development and Validation Activities Linked to Validity Aspects

	Validity Aspect			
	Content	Structural	Substantive	Generalizability
Type of Evidence	Instrument Purpose Instrument Specifications Item Development Expert Reviews Item Technical Quality	Instrument Dimensionality	Rating Scale Function Person Fit	Reliability

Note. Adapted from "Instrument Development Tools and Activities for Measure Validation Using Rasch Models," by E. W. Wolfe & E. V. Smith Jr., 2007b, Journal of Applied Measurement, p. 244. Copyright 2007 by Maple Grove, MN.

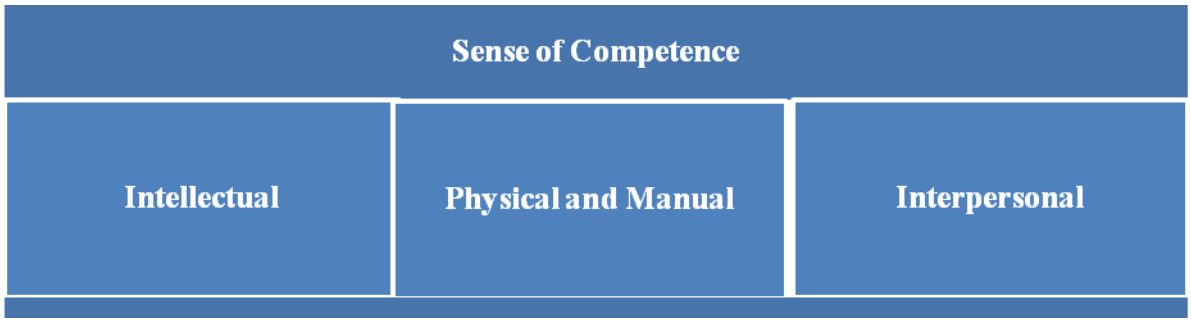

Figure 1. Conceptual Model of Sense of Competence-Revised (SCS-R)

Each dimension is a contributor to the construct sense of competence no one particular dimension is more important than the other.

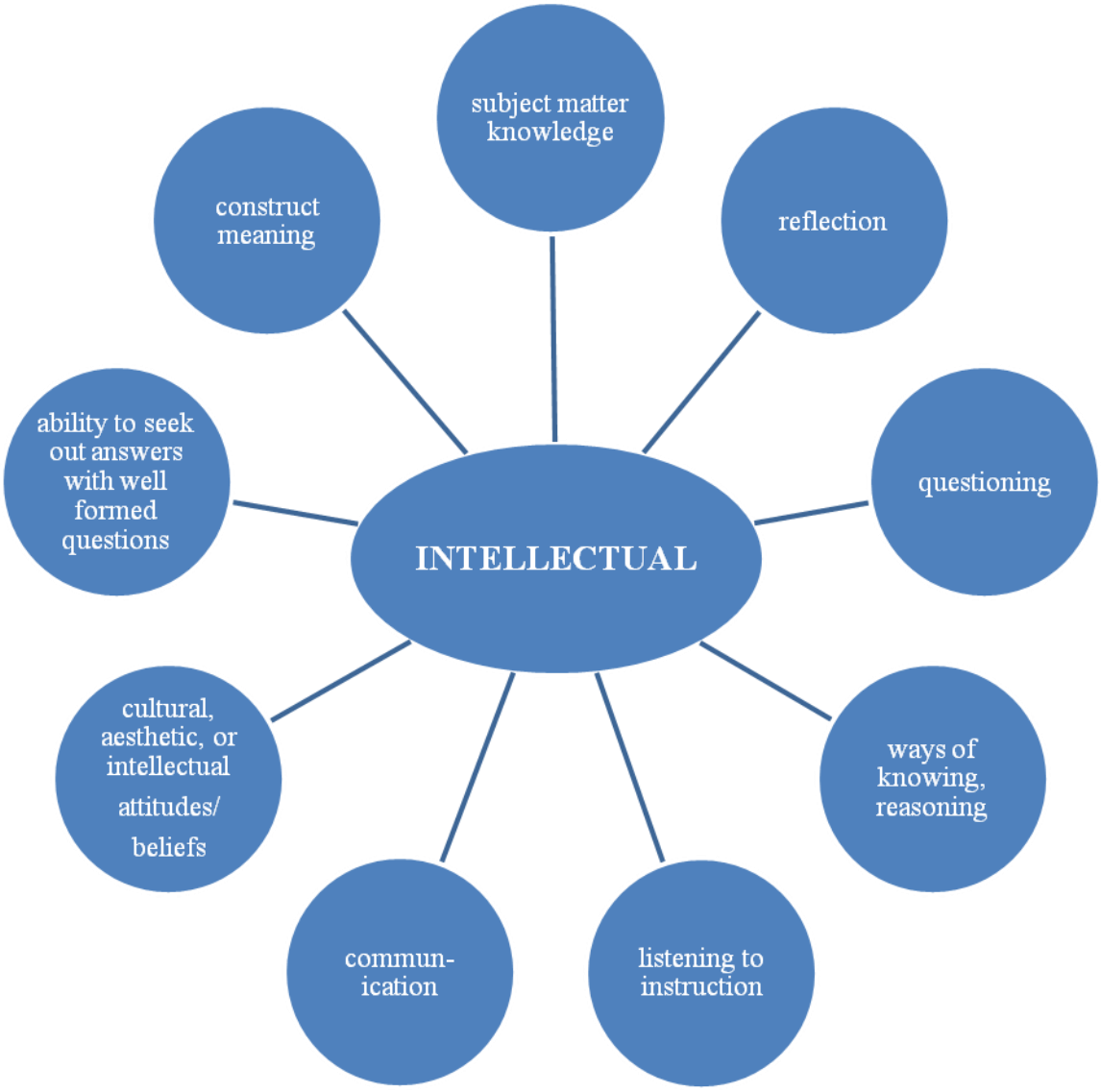

Figure 2. Intellectual Concept Map

Circles depict attributes of the dimension based on Chickering and Reisser (1993).

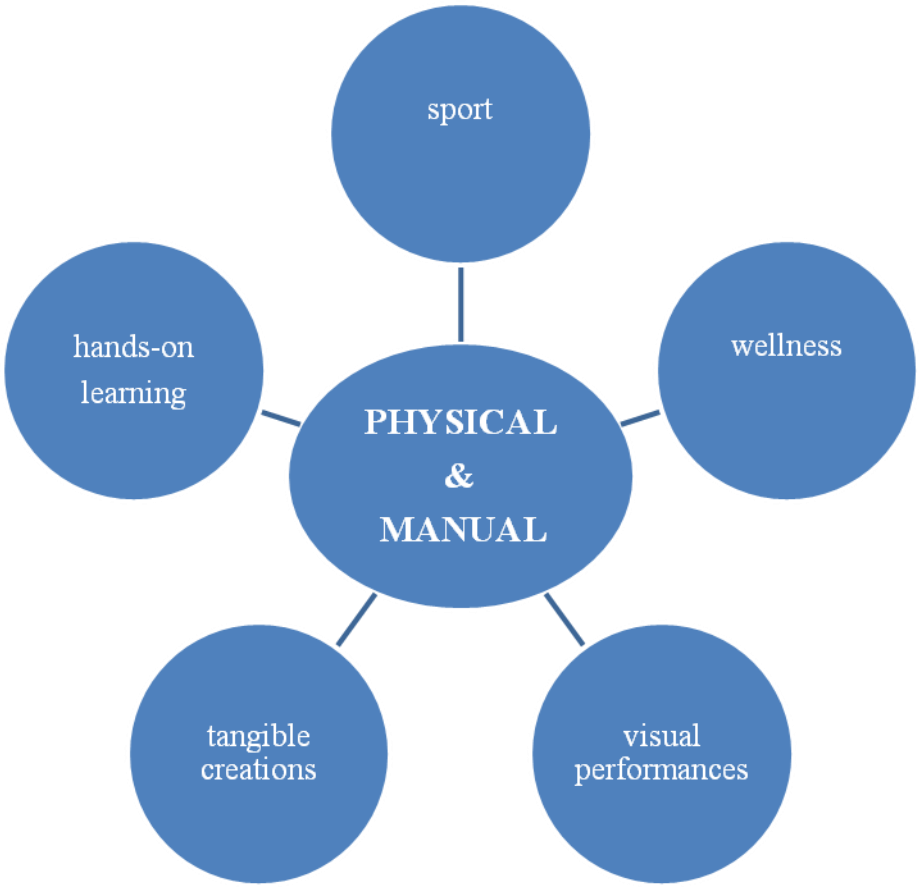

Figure 3. Physical and Manual Concept Map

Circles depict attributes of the dimension based on Chickering and Reisser (1993).

Figure 4. Interpersonal Concept Map

Circles depict attributes of the dimension based on Chickering and Reisser (1993).

Table 2

Sense of Competence-Revised (SCS-R) Instrument Blueprint

Sense of Competence-Revised Dimensions		
		Total
Dimension 1 Intellectual	Involves mastering content, gaining intellectual and aesthetic sophistication, and building a range of skills to comprehend, analyze and synthesize	35 items
Dimension 2 Physical and Manual	Involves athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline	26 items
Dimension 3 Interpersonal	Involves more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of a group, and to choose from a variety of strategies to help a relationship flourish or a group function	31 items
Total		92 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

Table 3

Intellectual Dimension Test Blueprint

	Intellectual Dimension	
		Total
Definition	Involves mastering content, gaining intellectual and aesthetic sophistication, and building a range of skills to comprehend, analyze and synthesize	
Personal Assessment	Student is secure in their intellectual decisions and actions. Student is realistic about one's own ability.	35 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

Table 4

Physical and Manual Dimension Test Blueprint

	Physical and Manual Dimension	
		Total
Definition	Involves athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline	
Personal Assessment	Student is secure in their physical and manual decisions and actions. Student is realistic about one's own ability.	26 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

Table 5

Interpersonal Dimension Test Blueprint

	Interpersonal Dimension	
		Total
Definition	Involves more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of a group, and to choose from a variety of strategies to help a relationship flourish or a group function	
Personal Assessment	Student is secure in their interpersonal decisions and actions. Student is realistic about one's own ability.	31 items

Note. Definitions adapted from "Education and Identity (2nd ed.)," by A. Chickering and L. Reisser, 1993, San Francisco: Jossey Bass.

After items were drafted, Chickering and Reisser authors of the revised version (Chickering & Reisser, 1993) of Chickering's (1969) student development theory were asked to review the instrument blueprint to assess its relevance to the construct (developing sense of competence). In addition to Chickering and Reisser, a psychometrician and the developer of the SCS (Janosik, 1987) reviewed the items. I reviewed and revised all items using the general guidelines for writing and reviewing Likert (1932) style items as outlined by Wolfe and Smith (2007a). The original items (Janosik, 1987) remained the same with the exception of changing all negatively worded items as well as the addition of newly generated items.

Pilot Study

The piloted version of the SCS-R included 92 items and was conducted through a web-based format for the purpose of assessing potential item performance. Five open-ended questions allowed respondents to provide additional comments regarding their impressions of their intellectual, physical and manual and interpersonal competence. In addition, two questions asked about their involvement in curricular and co-curricular activities. The entire sophomore class at a large Research I institution were sent an e-mail invitation to participate in the study along with a URL link to the survey. In total, 211 students responded to the pilot test. The results were relatively representative of the university's demographics for the sophomore class with 53.4% female (45% university total representation) and 46.6% male respondents (55% university total representation). As far as ethnicity, the results were within one to two percentage points for a few categories for representation of the sophomore class, for example Black, Non-Hispanic was 2% (4% university total representation), Hispanic was 2.4% (3.2% university total representation), and Non-Resident Alien was 0 (1.7% university total representation). After the data were cleaned analyses of item fit and dimensionality were conducted using parameters from the Rasch Rating Scale Model (Rasch, 1980; Wright & Masters, 1982).

Preliminary analyses of the pilot test data entailed principal components analyses of standardized residuals to detect systematically beyond a respective modeled dimension. The analyses suggested a distinction between intellectual, physical and manual, and interpersonal dimension. At the completion of the pilot test analyses (item technical quality and dimensionality analyses), a total of 61 items were selected for future development activities. I drafted two

additional items based on feedback during the expert review to represent the interpersonal dimension, yielding a final set of 63 Likert (1932) style items.

Measurement Model

I tested the dimensions using a unidimensional model. This section presents the Rasch Polytomous Models employed in the analyses, the software used in their estimation and the statistical analyses completed.

Response data for each phase of this study were scaled according to variations of the Rasch Polytomous Model (Luo, 2007; Rasch, 1980). The decision to use Rasch analyses during the development of the SCS-R was based on the presumed importance of providing construct-related evidence of validity for interpreting measures obtained from instrumentation (Messick, 1995). Rasch models are used as instruments of construct validity (Fisher, 1994) by working cooperatively with instrument developers to establish the scope of gathered data and the accuracy of measuring a hypothesized trait. If accuracy is established based on the analyses, strong inferences are made that state the dimensions are in fact expressions of that trait (Bond & Fox, 2007). Having empirical data adhere to a Rasch model's specifications instead of creating structures to model response data, stronger confidence can be placed in the supposition that an instrument's measures are indeed expressions of some trait(s) of interest.

Because each item on the field test form of the SCS-R employed an identical response format, responses were scaled to the Rasch Rating Scale Model (Andrich, 1978) given by:

$$\ln(P_{nix} / P_{ni\,x-1}) = B_n - (D_i + F_x)$$

where:

P_{nix} = the probability that participant *n* will respond in category *x* on item *i*

$P_{ni\,x-1}$ = the probability that participant *n* will respond in category *x-1* on item *i*

B_n = the ability for participant *n*

D_i = the difficulty for item *i*

F_x = the step that corresponds to an equal probability of selecting category *x* and *x-1*

I chose the RSM as the measurement model to evaluate the SCS-R four-point Likert (1932) scale items. Response categories setup as Likert (1932) style items represent a respondent's increasing partiality towards the concept being questioned (Wright & Mok, 2004). For polytomous items, the RSM includes a threshold difficulty parameter that depicts the difficulty of moving from one scoring category to another.

Results

The revised form of the SCS-R (Appendices E, F, & G) contained 63 items designed to measure three distinct dimensions that measure sense of competence. Analyses of the field test data suggested the SCS-R were more accurately gauging three dimensions, which were subsequently modeled as individual subscales labeled: intellectual, physical and manual, and interpersonal.

Field Test

Participants for the field test and measure validation phase of this study were college students with sophomore status at a Research I institution. The entire class of approximately 5,460 students was sent an e-mail invitation that contained a secure web-based survey link. A total of 916 (16.8%) students responded to the questionnaire as noted in Table 6. The results were relatively representative of the university's demographics for the sophomore class with 54.7% female (45% university total representation) and 45% male respondents (55% university total representation). As far as ethnicity, the results were within one to two percentage points for a few categories for representation of the sophomore class, for example Black, Non-Hispanic was 2.4% (4% university total representation), Hispanic was 2% (3.2% university total representation), and Non-Resident Alien was 0 (1.7% university total representation). Continued software updates and advances in technology the use of web-based instruments potentially can affect sample size by limiting the ability of people who do not have the most updated software or hardware to respond (Dillman, 2000). This may have affected the response rate but there is no way to calculate the extent to which this may have occurred. A total of five follow-up emails were sent as reminders to help increase the response rate. Incentives were also provided.

To increase participation in the study students were given the opportunity to download their name in a separate survey for a drawing. The winner of the drawing received an opportunity to win one of six $50 gift certificates to the university bookstore or an Apple

Table 6

Demographic Characteristics of Field Test Respondents

Variable	Category	*N*	% of those responding
Gender	Female	492	54.7
	Male	405	45
	Transgender	1	0.0
	Other	1	0.0
	Missing	17	--
Group in which you Identify with most	American Indian/Native American	7	1.0
	Asian or Pacific Islander	71	8.0
	Black, Non-Hispanic	22	2.4
	White, Non-Hispanic	756	84
	Hispanic	18	2.0
	Non-Resident Alien	2	0.0
	Other	26	3.0
	Missing	14	--
Live on-campus	Yes	437	49
	No	455	51
	Missing	24	--
Member of honors program	Yes	114	12.7
	No	783	87.3
	Missing	19	--
Participate in extra-curricular activities	Yes	772	85.7
	No	129	14.3
	Missing	15	--
Participate in intramurals	Yes	320	35.7
	No	577	64.3
	Missing	19	--

iPod. Names submitted for the drawing were not connected to the individuals' survey response so the data remained anonymous.

Directions were provided within each section of the instrument, informing participants about the means by which they were to provide responses. A coding scheme was developed to facilitate the recording of data into a numerical format that was amenable to subsequent statistical analyses.

After facilitating the field test, data from all completed questionnaires were placed in a *Microsoft Office Excel* (2007) document. Each item on the instrument was coded by a 1, 2, 3, or 4 for the Likert (1932) scale responses. The *Excel* document included all answers from the participants. The document was used to import into *WINSTEPS* (Linacre, 2006) software.
In this section of the article I present the results and analyses aimed at authenticating intellectual, physical and manual, and interpersonal skills as the dimensions that measure college students' sense of competence.

Item Technical Quality

I examined the item technical quality using *WINSTEPS* (Linacre, 2006) to gain thorough evidence of the information about average difficulty estimate for items, the unweighted standardized mean square fit statistics ($Z_{unweighted}$) and the correlation (r_{pm}) with measures produced by the dimensions provided in Tables 7 and 8. In general, most items fit within the criteria designated for determining adequate fit and useful contribution. Across the three dimensions, all items exhibited $r_{pm} > .30$ with the exception of eight items and mean square fit statistics that fell within the recommended .5 and 1.5 values. In total, 17 items were flagged based on standardized outfit values (i.e., $MS_{outfit} > 2.0$). In most cases, however, these items exhibited adequate fit and contribution according to the other indices employed. Where potential item misfit occurred, the rationale for why an item was retained is described in Table 7.

Dimensionality

I tested structural validity evidence of the instrument using *WINSTEPS* (Linacre, 2006) for data analyses. Dimensionality is defined as the number of aspects that constitute a construct. Dimensionality may be addressed in a couple of ways. The RSM requires unidimensionality and

Table 7

Summary of Item Quality Indicators

Dimension	Statistic	Logit	$Z_{unweighted}$	r_{pm}
1. Intellectual	Mean	.00	-.3	--
	SD	1.00	5.2	--
	Minimum	-1.34	-9.3	.09
	Maximum	2.75	7.2	.60
	N Flagged	--	8[a]	--
2. Physical & Manual	Mean	.00	.1	--
	SD	.72	3.6	--
	Minimum	-1.4	-6.1	.32
	Maximum	1.45	9.5	.59
	N Flagged	--	2[b]	--
3. Interpersonal	Mean	.00	-.1	--
	SD	.94	4.9	--
	Minimum	-1.3	-9.3	.36
	Maximum	1.98	9.9	.60
	N Flagged	--	7[c]	--

Note. [a]Items 1, 3, 14, 18, 23, 32, 35, and 51 all exhibited $Z_{unweighted}$ > 2.5, however each item illustrated strong correlation with the modeled dimension (r_{pm} > .35), with the exception of item 18 (r_{pm} > .09). [b]Items 15 and 42 all exhibited $Z_{unweighted}$ > 2.5, however each item illustrated strong correlation with the modeled dimension (r_{pm} > .35), with the exception of item 42 (r_{pm} > .32). [c]Items 4, 8, 25, 38, 45, 47, 63 all exhibited $Z_{unweighted}$ > 2.5, however each item illustrated strong correlation with the modeled dimension (r_{pm} > .35)

Table 8

Performance Indices for Dimensions of Items

	Intellectual	Physical & Manual	Interpersonal
Number of Items	22	18	23
Rel_θ	.82	.76	.81
Cronbach's α	.88	.82	.88
$r_{pm} \geq .50$	13 items	4 items	10 items
$r_{pm} \geq .40$	5 items	10 items	10 items
$r_{pm} \geq .30$	1 items	4 items	3 item
$MS_{infit} \geq 1.5$	0 items	0 items	0 items
$Z_{unweighted} \geq 2.0$	8 items	2 items	7 items
Total Var_θ explained by Rasch Dimension	41.00%	39.00%	41.30%
Eigenvalue for 1st Contrast	2.5	3.6	2.5
Var_θ unexplained for 1st Contrast	6.80%	12.20%	6.30%

Note. Item 18 (intellectual) is not noted in the rpm because it represented a low r_{pm} of .09. Item 22 (intellectual) is not noted in the rpm because it represented a low r_{pm} of .29.

local independence to assess structural validity. Local independence is when all items are independent of another for examinees of the same ability on a latent tendency (Osterlind, 2006). I evaluated the degree to which the claim of unidimensionality was reasonable.

The Kaiser's eigenvalue was calculated to investigate unidimensionality. I used *Microsoft Office Excel* (2007) to set up the formula for Kaiser's eigenvalue. The eigenvalues were retrieved from the standardized residual variance table in *WINSTEPS* (Linacre, 2006) and placed into the excel document to calculate the Kaiser's eigenvalue. As a reminder, intellectual, physical and manual, and interpersonal items were analyzed as separate constructs that measure sense of competence. My goal was for the analyses to identify additional factors within each dimension. Each dimension's eigenvalues were scaled on the unexplained variance that gave the Kaiser's eignevalues shown in Tables 9-11. Based on the transformed eigenvalues I focused the analyses for intellectual as a unidimensional construct and not multidimensional, meaning there was only one factor present. This is congruent with interpersonal as well as the physical and manual dimension. The scree plots located in Figures 5 – 7 illustrate the unidimensionality of the data. Rasch is explained by the first data point on the graph. The second point through the fifth point expresses the value of each residual. The scree plots confirm that first is higher than the second, third, fourth and fifth. From the plots we can see that the second, third, fourth and fifth residuals fall closer to the Kaiser's rule line drawn in the graph. The scree plot significantly flattens beginning with second residual with the exception of the physical and manual dimension where it flattens around the third residual. The scree plots were created in *Microsoft Office Excel* (2007).

Rating Scale Function

An initial review of response category frequencies demonstrated that some items failed to receive the recommended number of ten responses (Linacre, 2004) in all end-point categorical options (i.e., *Strongly Disagree* and *Strongly Agree*). As such, the interpretation of rating scale efficacy, in addition to the accuracy and stability of less-centralized parameter estimates, would have to be met with a certain degree of caution. With that in mind, the adequacy of the item response scale for each of the instrument's three dimensions was examined in accordance with Linacre's (2004) criteria, the results of which are presented in Table 12.

Table 9

Eigenvalues for Intellectual Dimension

Intellectual Dimension	Eigenvalue Units (Residual)	Kaiser's Eigenvalues (Scaled)
Total Variance	37.5	
Unexplained Variance (total)	22.0	
Rasch Dimension	15.5	9.09
Residual 1	2.5	1.47
Residual 2	1.7	1.00
Residual 3	1.6	.94
Residual 4	1.5	.88
Residual 5	1.3	.76

Table 10

Eigenvalues for Physical and Manual Dimension

Physical and Manual Dimension	Eigenvalue Units (Residual)	Kaiser's Eigenvalues (Scaled)
Total Variance	29.6	
Unexplained Variance (total)	18.0	
Rasch Dimension	11.6	7.05
Residual 1	3.6	2.19
Residual 2	2.2	1.34
Residual 3	1.3	.79
Residual 4	1.3	.79
Residual 5	1.1	.67

Table 11

Eigenvalues for Interpersonal Dimension

Interpersonal Dimension	Eigenvalue Units (Residual)	Kaiser's Eigenvalues (Scaled)
Total Variance	39.2	
Unexplained Variance (total)	23.0	
Rasch Dimension	16.2	9.51
Residual 1	2.5	1.47
Residual 2	2.2	1.29
Residual 3	1.8	1.06
Residual 4	1.5	0.88
Residual 5	1.4	0.82

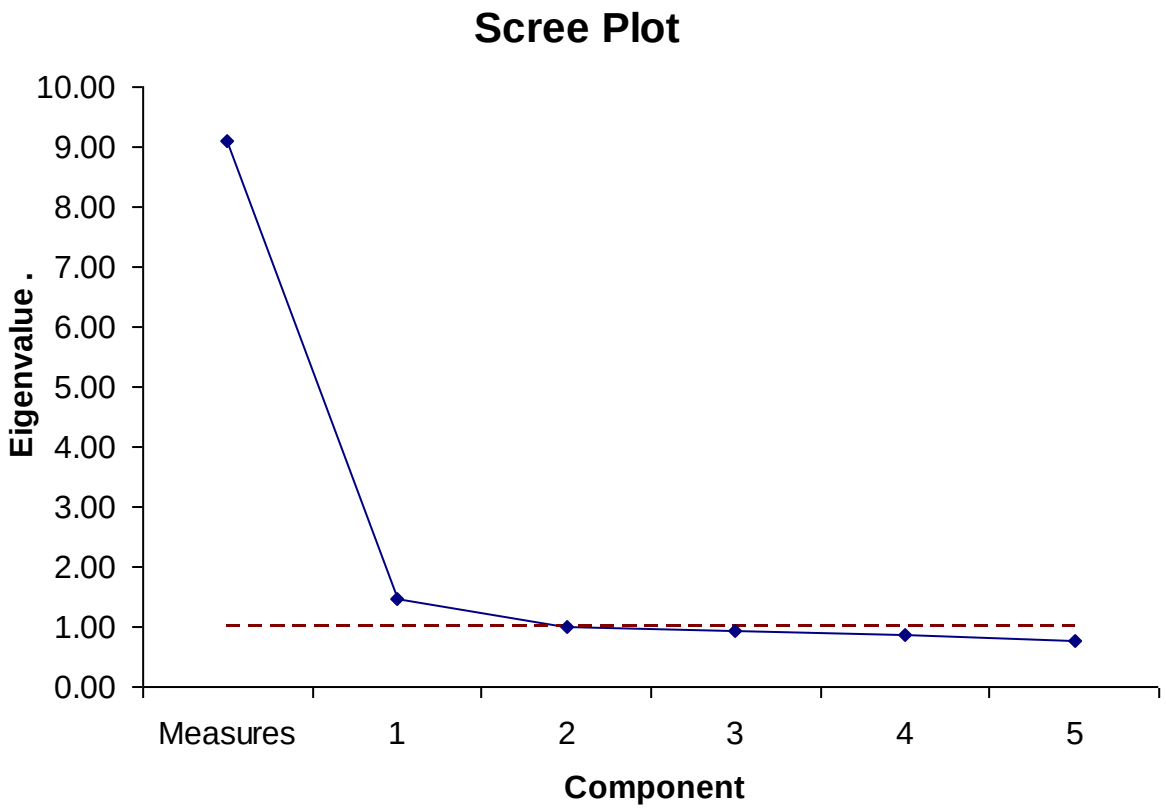

Figure 5. Intellectual Dimension Scree Plot

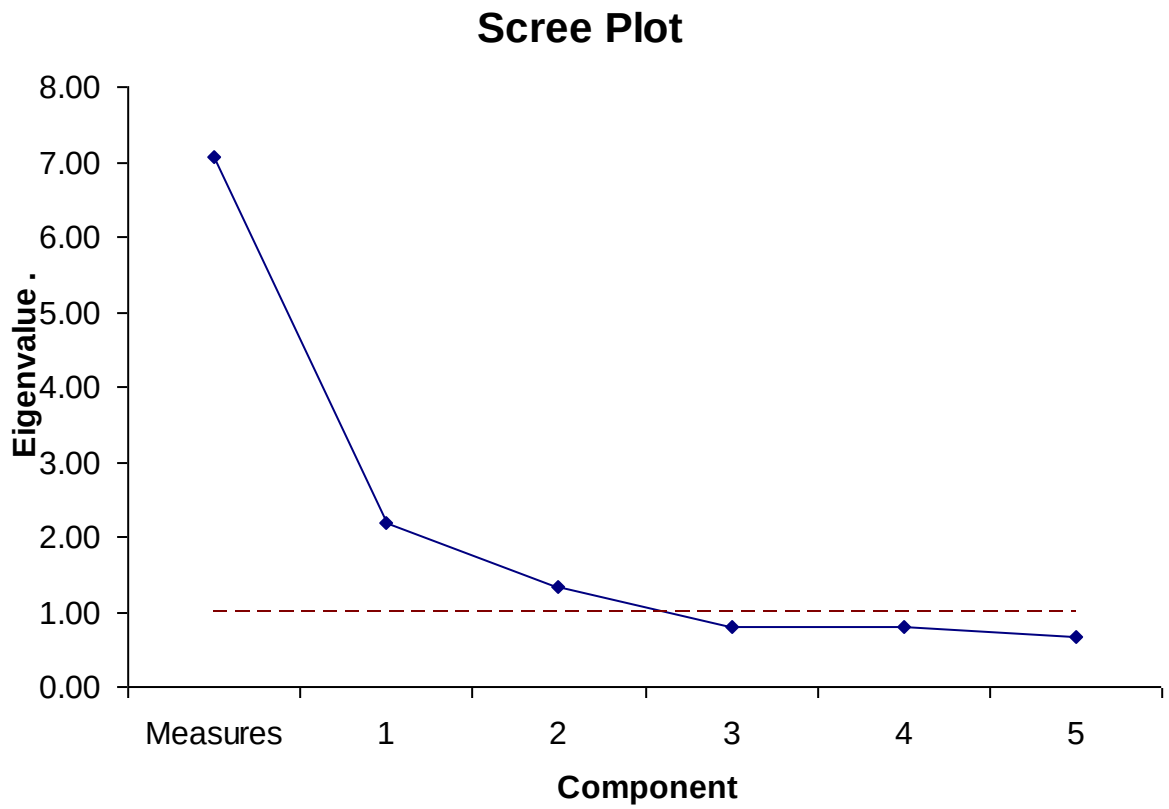

Figure 6. Physical and Manual Dimension Scree Plot

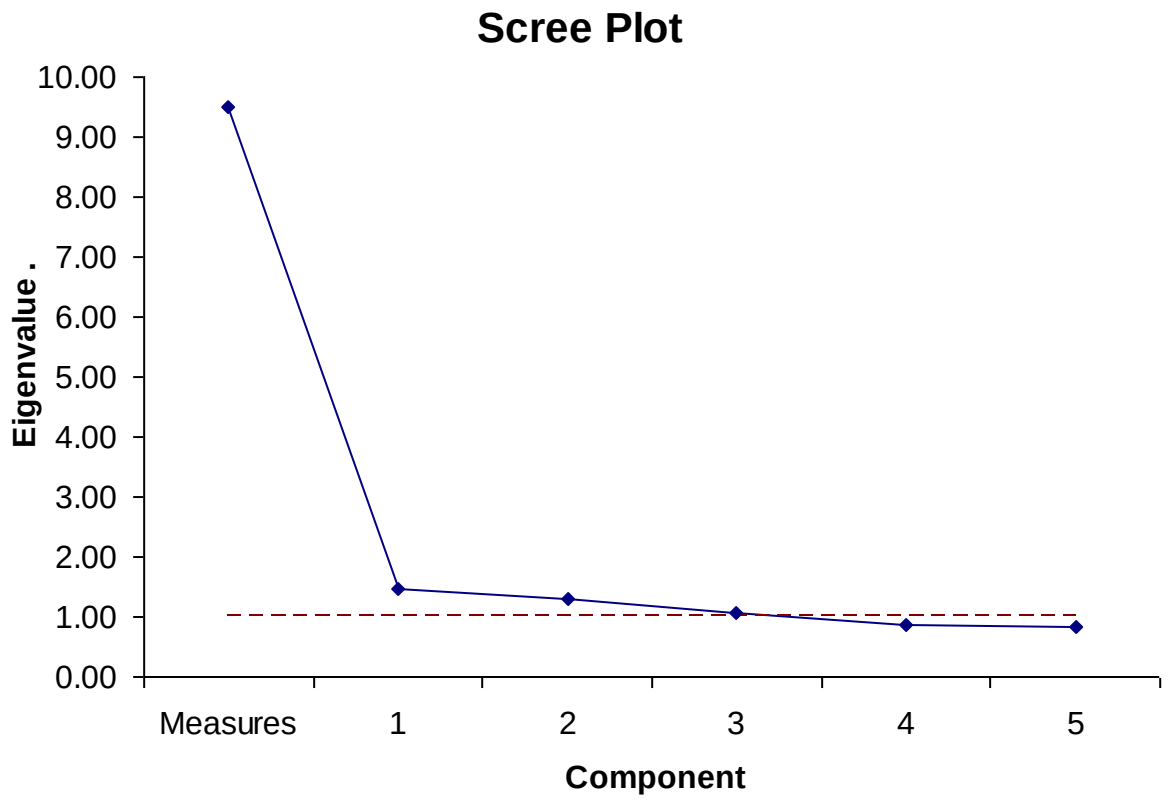

Figure 7. Interpersonal Dimension Scree Plot

Table 12

Rating Scale Analyses

	Intellectual	Physical & Manual	Interpersonal
MS_{outfit}			
Person	.82	.76	.81
Item	1.00	.99	1.00
Linacre's Criteria			
$N_g > 10$	☒	☒	☒
Distributions	☑	☑	☑
$M(\theta)$	☑	☑	☑
$MS_{unweighted}$	☑	☑	☑
τ increase > 1.4	☑	☑	☑
τ distance < 5.0	☑	☑	☑
$Coherence_{Measures}$	☑	☒	☒
$Coherence_{Categories}$	☒	☒	☒

The real person reliability which is equivalent to Coefficient Alpha is greater than .70 for all three dimensions suggesting that the instrument discriminates between the respondents well. The real item separation reliability of .99 and 1.00 indicates that the items create a well defined variable. The majority of the Linacre's (2004) suggested criteria were met. However, concerns with the rating scale structure became apparent in the coherence for measures and categories Linacre (2004) suggest coherence of 40% or greater as adequate. For coherence with measures implying categories the interpersonal (25%) and physical and manual (16%) dimensions' Strongly Disagree category did not meet the criteria. The intellectual dimension for categories implying the measure two of the rating scale categories were less than 40% (Strongly Disagree, 7% and Strongly Agree, 32%). Physical and manual also did not meet the criteria for categories implying the measure of two of the rating scale categories (Strongly Disagree, 1% and Disagree, 26%). Lastly, interpersonal dimension failed to meet the criteria for categories implying measures with 1% for Strongly Disagree and 32% for Disagree. Linacre's (2004) criteria are suggested criterion. Condensing these categories to a three-point Likert Scale could increase coherence. The majority of Lincacre's (2004) criteria were met establishing the SCS-R as a reliable measure of sense of competence.

Person Fit

Table 13 illustrates information about the degree to which cases of person misfit were observed across each of the instrument's dimensions. The results indicate that respondents demonstrating observed misfit was likely due to individuals with low ability measures endorsing one or two difficult items or those with high ability failed to endorse one or two items that were closely matched to them along the latent trait measured by the respective dimension.

To identify respondents exhibiting misfit ($Z_{unweighted} > 2.5$), I removed all persons exhibiting extreme misfit from the original data set and model parameters were re-calibrated for each of the three dimensions. As a reminder each dimension was run as a separate data set in WINSTEPS (2006). For example, Table 13 illustrates that a total of 59 of 916 respondents were removed from the intellectual dimension data set. The removal of these persons improved the dimension reliability.

Table 13

Person Fit

Dimension	Mean $Z_{unweighted}$	Minimum $Z_{unweighted}$	Maximum $Z_{unweighted}$	Number Flagged > 2.5	%
1. Intellectual	-.2	-4.5	6.4	59	6.5%
2. Physical & Manual	-.2	-3.9	5.1	51	5.6%
3. Interpersonal	-.2	-4.7	6.3	69	7.6%

Reliability

Regarding reliability, Table 14 demonstrates the internal consistency of Rasch measures. As well as raw scores, by reporting indices of Rasch Person separation reliability (R_P), R_P after deleting persons exhibiting extreme misfit and coefficient alpha (α). All dimensions produced reliability estimates that were greater than .70. Data analyses from the field test of the SCS-R suggest that the reliability separation for persons provided by the instrument dimensions is sufficient for the implementation of future research.

Scoring

I developed a scoring arrangement based on the person measures from the field test. Scoring was created for two audiences, researchers and practitioners. For researchers, I conducted a linear transformation of the logit measures. The linear transformation (scale score) of the logit measures was calculated by taking the raw score and subtracting the person measures mean for the intellectual dimension and dividing by the person measure standard deviation creating positive numbers for reporting scores. Linear transformations were also conducted for the physical and manual and interpersonal dimensions. For practitioners, I calculated the percentile rank for the raw and scale score. Table 15 illustrates the scoring arrangement.

Discussion

The responses to the SCS-R were better modeled as three unidimensional constructs including intellectual, physical and manual, and interpersonal competence. I conceptualized sense of competence as being comprised of three important abilities with an understanding that the SCS-R measures how a college student assesses their own sense of competence. For those who plan to use the instrument, it is important to remember the subjectivity of measuring sense of competence. For example, students who are academically successful may score low on their assessment of their intellectual abilities because they are not confident.

The results suggest that the SCS-R dimensions provide reasonable estimates of students' sense of competence. Empirical evidence for the dimensionality of the instrument supports the three dimensions of sense of competence as outlined by Chickering and Reisser (1993). I anticipate these dimensions to serve as useful tools for future research studies.

Table 14

Internal Consistency of Measures Produced by Dimensions (Reliability)

Dimension	R_P	R_P [a]	α	α [b]
1. Intellectual	.82	.82	.88	.89
2. Physical & Manual	.76	.78	.82	.84
3. Interpersonal	.81	.82	.88	.90

Note. [a]R_P after removal of persons exhibiting extreme mis-fit (i.e., $Z_{unweighted} > 2.5$). [b]α after removal of persons exhibiting extreme mis-fit (i.e., $Z_{unweighted} > 2.5$).

Table 15

SCS-R Scoring Arrangement

	Raw Score	Scale Score	Percentile Rank
Intellectual Dimension	75	70	95
	65	55	50
	50 or less	30 or less	1
Physical and Manual Dimension	65	65	95
	55	55	50
	45 or less	45 or less	1
Interpersonal Dimension	84	72	95
	71	58	50
	62 or less	50 or less	1

Note. Interpretation of raw score and scale scores' percentile ranks are provided in the table. A complete description of the raw scores, scale scores and percentile ranks are located in Appendices H, I and J.

Future Research

The initial results of validity for the SCS-R are good. The SCS-R should be used to assess college students' personal growth during their time at the university. For example practitioners could use the SCS-R to conduct a pre-test and post-test at the beginning and end of students' sophomore year. However, there is more work to be done to establish concurrent validity. Future research should find any related measure that measures the same construct and correlate both instruments. These efforts should be grounded in psychosocial theory, which examines the personal and interpersonal lives of individuals. Reviewing multiple instruments that measure the same construct would help to understand to what degree that the SCS-R measures what I intended.

Since the items all had a four-point Likert (1932) scale, I chose the RSM to elicit the same rating structure across items. Future research needs to explore the rating scale categories for the SCS-R. I found the categories to be weak for some of the items. Potentially collapsing the categories into two or three categorical choices may create a better response pattern.

In addition, sense of competence may vary by demographic group. For example, Chickering and Reisser (1993) suggest that students of all ages and backgrounds transition through the seven vectors during their time enrolled in college. All the participants in the study were students at the same university who had sophomore class status. It is possible that students designated as freshmen, junior or senior status would differ in some important ways from students who were classified as sophomores at the institution. It is also possible that students at the institution differed in some way from students at other colleges and universities. With the disparity of research conducted on under-represented groups and the increase of older adults returning to education more research is needed to better understand these multi-faceted subpopulations. Continued focus will be dedicated to refining the SCS-R through the use of larger cross-validation samples in which respondents represent diverse target populations for a range of student populations independent of dominant culture representation (Evans, Forney, Guido, Patton, and Renn, 2010).

Additionally, future research in the development and validation of instruments that measure the remaining six vectors of Chickering's (1969; Chickering & Reisser, 1993) theory as it relates to multiple sub-populations within the student body is imperative. Collectively, these

opportunities for future research will allow for exploration in the development of students during their time with the university. Administrator's can use data from the research to modify their institution's academic and social environment to enhance the development of students' intellectual, physical and manual, and interpersonal skills.

References

Bond, T. G. & Fox, C. M. (2007). Applying the Rasch model: Fundamental measurement in the human sciences (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Chickering, A. (1969). *Education and identity*. San Francisco: Jossey-Bass.

Chickering, A., & Reisser, L. (1993). *Education and Identity* (2nd ed.). San Francisco: Jossey-Bass.

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method*. New York: John Wiley & Sons.

Evans, N. J., Forney, S. F., & Guido-DiBrito, F. (1998). *Student development in college: Theory, research, and practice*. San Francisco: Jossey-Bass.

Evans, N. J., Forney, S. F., Guido, F., Patton, L. D., & Renn, K. A. (2010). *Student development in college: Theory, research, and practice*. (2nd Ed.) San Francisco: Jossey-Bass.

Fisher, W. P., Jr. (1994). The Rasch debate: Validity and revolution in educational measurement. In M. Wilson (Ed.), Objective measurement: Theory and practice (Vol. 2, pp. 36-72). Norwood, NJ: Ablex.

Janosik, S. M. (1987). Relationship of residence hall environments and student sense of competence and academic achievement (Doctoral dissertation, Virginia Tech, 1987).

Linacre, J. M. (1998). Detecting multidimensionality: Which residual type works best? *Journal of Outcome Measurement, 2*(3), 266-283.

Linacre, J. M. (2006) WINSTEPS. A User's guide to Winsteps: Rasch Model computer programs. Chicago, IL: Retrieved October 4, 2007 from www.winsteps.com.

Luo, G. (2007). The relationship between the Rating Scale and Partial Credit Models and the implication of disorderd thresholds of the Rasch models for polytomous responses. In Smith, E.V. & Smith, R. M. (Eds.), *Rasch measurement: Advanced specialized applications* (pp. 181-201). Maple Grove, MN: Journal of Applied Measurement Press.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.

Microsoft Office Excel. (2007). Microsoft.

Moran, C. (2003). Conceptualizing identity development: Unmasking the assumptions within inventories measuring identity development. NASPA Journal, 46, 3, p. 469 – 481.

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests* (revised and expanded ed.). Chicago: The University of Chicago Press. (Originial work published 1960.)

Smart, J. C. (2005). Attributes of exemplary research manuscripts employing quantitative analyses. *Research in Higher Education, 46*, 461-477.

Wolfe, E. W., & Smith, E. V., Jr. (2007a). Instrument development tools and activities for measure validation using Rasch Models: Part I-Instrument development tools. In E. V. Smith & R. M. Smith (Eds.), *Rasch Measurement: Advanced specialized applications* (pp. 2-27). Maple Grove, MN: Journal of Applied Measurement Press.

Wolfe, E. W., & Smith, E. V., Jr. (2007b). Instrument development tools and activities for measure validation using Rasch Models: Part II-Instrument development tools. In E. V. Smith & R. M. Smith (Eds.), *Rasch Measurement: Advanced specialized applications* (pp. 204-234). Maple Grove, MN: Journal of Applied Measurement Press.

Wright, B. D. & Masters, G. N. (1982). *Rating scale analyses*. Chicago: MES Press.

Wright, B. D., & Mok, M. M. C. (2004). An overview of the family of Rasch measurement models. In E.V. Smith, Jr. & R. M. Smith (Eds.), Introduction to Rasch Measurement (pp. 1-24). Maple Grove, MN: JAM Press.

References

Bond, T. G. & Fox, C. M. (2007). Applying the Rasch model: Fundamental measurement in the human sciences (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Catell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research, 1*, 245-276.

Chafin, C. N. (2006). The impact of a living learning community and inquiry guided learning on first year students' emotional intelligence and academic achievement. *Electronic Theses and Dissertations*. (NCST No. 06162006-160811)

Chickering, A. (1969). *Education and identity*. San Francisco: Jossey-Bass.

Chickering, A., & Reisser, L. (1993). Education and Identity (2nd ed.). San Francisco: Jossey-Bass.

Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Thomson Learning.

Cronbach, L. J. (1951). Coefficient alpha and internal structure of tests. *Psychometrika, 16*, 297-334.

Cross, W. E., Jr. (1991). *Shades of black: Diversity in African American identity*. Philadelphia: Temple University Press.

D'Augelli, A. R. (1994). Identity development and sexual orientation: Toward a model of lesbian, gay, and bisexual development. In E. J. Trickett, R. J. Watts, & D. Birman (Eds.), *Human diversity: Perspectives on people in context* (pp. 312-333). San Francisco: Jossey-Bass.

Dillman, D. A. (2000). *Mail and internet surveys: The tailored design method*. New York: John Wiley & Sons.

Erikson, E. H. (1950). *Childhood and society*. New York: Norton.

Erikson, E. H. (1968). *Identity: Youth and crisis*. New York: Norton.

Evans, N. J. (1996). Theories of student development. In S. R. Komives, D. B. Woodard Jr., & Associates, *Student services: A Handbook for the profession* (3rd ed., pp. 164-187). San Francisco: Jossey-Bass.

Evans, N. J., Forney, S. F., & Guido-DiBrito, F. (1998). *Student development in college: Theory, research, and practice*. San Francisco: Jossey-Bass.

Evans, N. J., Forney, S. F., Guido, F., Patton, L. D., & Renn, K. A. (2010). *Student development in college: Theory, research, and practice*. (2nd Ed.) San Francisco: Jossey-Bass.

Fisher, W. P., Jr. (1994). The Rasch debate: Validity and revolution in educational measurement. In M. Wilson (Ed.), Objective measurement: Theory and practice (Vol. 2, pp. 36-72). Norwood, NJ: Ablex.

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

Hood, A. B. (Ed.). (1986). The Iowa Student Development Inventories. Iowa City, Iowa: Hitech Press.

Hood, A. B., & Jackson, L. M. (1983). *Assessing the development of competence.* Technical report. Iowa City: University of Iowa, College of Education.

Janosik, S. M. (1987). Relationship of residence hall environments and student sense of competence and academic achievement (Doctoral dissertation, Virginia Tech, 1987).

Josselson, R. (1987). Identity diffusion: A long-term follow-up. *Adolescent Psychology, 14*, 230-258.

Karabatsos, G. (2000). A critique of Rasch residual fit statistics. *Journal of Applied Measurement, 1*(2), 152-176.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 140*, 44-53.

Linacre, J. M. (2004). Optimal rating scale category effectiveness. In E. V. J. Smith & R. M. Smith (Eds.), *Introduction to Rasch Measurement* (pp. 258-278). Maple Grove, MN: JAM Press.

Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement, 3*(1), 266-283.

Linacre, J. M. (2006) WINSTEPS. A User's guide to Winsteps: Rasch Model computer programs. Chicago, IL: Retrieved February 21, 2009 from www.winsteps.com/a/winsteps.pdf.

Luo, G. (2007). The relationship between the Rating Scale and Partial Credit Models and the

implication of disorderd thresholds of the Rasch models for polytomous responses. In Smith, E.V. & Smith, R. M. (Eds.), *Rasch measurement: Advanced specialized applications* (pp. 181-201). Maple Grove, MN: Journal of Applied Measurement Press.

Marcia, J. E. (1966). Development and validation of ego-identity status. *Journal of Personality and Social Psychology, 3*, 551-558.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist, 50*, 741-749.

Microsoft Office Excel. (2007). Microsoft.

Miller, T. K., & Winston, R. B., Jr. (1990). Assessing development from a psychosocial perspective. In D. G. Creamer & Associates, *College student development: Theory and practice for the 1990's* (pp. 99-126). Alexandria, VA: American College Personnel Association.

Moran, C. (2003). Conceptualizing identity development: Unmasking the assumptions within inventories measuring identity development. NASPA Journal, 46, 3, p. 469 – 481.

Osterlind, S. J. (2006). *Modern measurement: Theory, principles, and applications of mental appraisal.* Upper Saddle River, NJ: Pearson Education.

Pace, C. R., & Kuh, G. D. (1979). *The college student experiences questionnaire assessment program*. Indiana University Center for the Study of Postsecondary Research, Bloomington.

Pascarella, E. T., & Terenzini, P. T. (2005). *How college affects students: A third decade of research* (Vol. 2). San Francisco: Jossey-Bass

Phinney, J. S. (1990). Ethnic identity in adolescents and adults: Review of research. *Psychological Bulletin, 108*, 499-514.

Rasch, G. (1966). An item analysis which takes individual differences into account. *British Journal of Mathematical and Statistical Psychology, 19*, 49-57.

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment test*. Copenhagen: Denmark Pedogogiske Institute. (Reprinted 1980 by University of Chicago.)

Rasch, G. (1980). *Probabilistic models for some intelligence and attainment tests*

(revised and expanded ed.). Chicago: The University of Chicago Press. (Originial work published 1960.)

Rodgers, R. F. (1990). Recent theories and research underlying student development. In D. Creamer & Associates (Eds). *College student development: Theory and practice for the 1990s* (pp. 27-790). Alexandria, VA: American College Personnel Association.

Schmidtlein, F. A., & Berdahl, R. O. (2005). Autonomy and accountability: Who controls academe? In P. G. Altbach, R. O. Berdahl, & P. J. Gumport (Eds.), *American higher education in the twenty-first century: Social, political and economic challenges* (pp. 71-90). Baltimore, MD: The Johns Hopkins University Press.

Shostrom, E. L. (1963). *Personal orientation inventory*. Educational and Industrial Testing Service. San Diego, CA

Smart, J. C. (2005). Attributes of exemplary research manuscripts employing quantitative analyses. *Research in Higher Education, 46*, 461-477.

Smith, G. T., & McCarthy, D. M. (1995). Methodological considerations in the refinement of clinical assessment instruments. *Psychological Assessment, 7*, 300-308.

Stone, M. H. (1998). Rating scale categories: Dichotomy, double dichotomy and the number two. *Popular Measurement, 1*, 61-65.

Vogt, D. S., King, D. W., & King, L. A. (2004). Focus groups in psychological assessment: Enhancing content validity by consulting members of the target population. *Psychological Assessment, 16*, 3, 231-243.

Winston, R. B., Jr. (1990). The Student Developmental Task and Lifestyle Inventory: An approach to measuring students' psychosocial development. *Journal of College Student Development, 31*, 108-120.

Winston, R. B., Jr., Miller, T. K., & Prince, J. S. (1987). *Student Developmental Task and Lifestyle Inventory.* Athens, GA: Student Development Associates.

Winston, R. B., Jr., Miller, T. K., & Prince, J. S. (1999). Student Developmental Task and Lifestyle Assessment. Athens, GA: Student Development Associates.

Wright, B. D., & Stone, M. H. (2003). *Managing observations, inventing constructs, crafting yardsticks, examining fit.* Chicago: Phaneron Press.

Wolfe, E. W., & Smith, E. V., Jr. (2007a). Instrument development tools and activities

for measure validation using Rasch Models: Part I-Instrument development tools. In E. V. Smith & R. M. Smith (Eds.), *Rasch Measurement: Advanced specialized applications* (pp. 2-27). Maple Grove, MN: Journal of Applied Measurement Press.

Wolfe, E. W., & Smith, E. V., Jr. (2007b). Instrument development tools and activities for measure validation using Rasch Models: Part II-Instrument development tools. In E. V. Smith & R. M. Smith (Eds.), *Rasch Measurement: Advanced specialized applications* (pp. 204-234). Maple Grove, MN: Journal of Applied Measurement Press.

Wright, B. D. & Masters, G. N. (1982). *Rating scale analyses*. Chicago: MES Press.

Zumeta, W. (2001). Public policy and accountability in higher education: Lessons for the past and present for the new millennium. In D. E. Heller (Ed.), *Public higher education policy: Affordability, access, and accountability* (pp. 155-197). Baltimore, MD: The Johns Hopkins University Press.

APPENDIX A. SENSE OF COMPETENCE SCALE (SCS)

(Janosik, 1987)

The following questions are designed to assess your sense of intellectual and interpersonal competence. Intellectual competence refers to your ability for knowing yourself and accomplishing tasks which require mental ability. Interpersonal competence refers to your capacity to relate or to interact with others.

Please select the appropriate response which best represents your feelings about each statement.

I meet people comfortably.
I read fast enough to handle assignments at college.
I have little difficulty with college work.
I feel empathy for others.
I will need tutoring in a specific course.
I respond to the needs of others.
I feel swamped with academic work.
I will graduate with honors.
I will fail at least one college course.
I make friends easily.
I will achieve my academic goals.
I get along with most of the people I meet.
Professors underrate my ability.
I am afraid of making mistakes.
I trust the decisions I make.
I can cope with life's ups and downs.
I welcome criticism as an opportunity for growth.
I am not afraid to show my emotions.
I converse casually and seriously with friends.
I don't mind if friends know my weaknesses.
I talk with professors easily about my concerns
Rate your intellectual capacity
Rate your interpersonal skill level

APPENDIX B. SENSE OF COMPETENCE SCALE – REVISED (SCS-R)

SCS-R Pilot Test

The purpose of this study is to develop an instrument to measure the sense of competence of traditional age college students across the dimensions that define the construct. Administrators can use data from the instrument to modify an institution's academic and social environment to enhance the development of the intellectual, physical, and interpersonal competencies of college students.

Important Definitions:

1. **Sense of Competence Scale –Revised (SCS-R): measures a student's intellectual, physical and manual, and interpersonal skills.**
2. **Intellectual competence: "involves mastering content, gaining intellectual and aesthetic sophistication, and, most important, building a repertoire of skills to comprehend, ananlyze and synthesize" (Chickering &Reisser, 1993, p. 45)**
3. **Physical and manual competence: "can involve athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline" (Chickering & Reisser, 1993, p. 46).**
4. **Interpersonal competence: "entails not only the skills of listening, cooperating, and communicating effectively, but also the more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of the group, and to choose from a variety of strategies to help a relationship flourish or a group function" (Chickering & Reisser, 1993, p. 46).**

Please select the appropriate response which best represents your feelings about each statement.

1. I read fast enough to handle assignments at college.
2. I am comfortable meeting new people.
3. I have little difficulty with college work.
4. I feel empathy for others.
5. I respond to the needs of others.
6. I make friends easily.
7. I am competent using computers.
8. I enjoy participating in physical activity.
9. I will need tutoring in a specific course.
10. College work is easy.
11. I trust the decisions I make.
12. I can cope with daily life challenges.
13. I get along with most of the people I meet.
14. I believe it is important to live a healthy lifestyle.
15. I have spent time participating in artistic activities.
16. I feel overwhelmed with academic work.
17. I will graduate with honors.
18. I am clumsy when it comes to participating in physical activity.
19. I openly show my emotions.
20. I am comfortable with my body.
21. I get frustrated when using new computer software.
22. I will fail at least one college course.
23. I will achieve my academic goals.
24. Professors underrate my ability.
25. I am afraid to show my emotions.
26. I converse casually with my friends.
27. I am competent using social networking sites.
28. I dedicate time for regular exercise workouts.

29. I get frustrated when reading instructions to put something together.
30. I understand my physical limitations for my body type.
31. I am afraid of making mistakes.
32. I welcome criticism as an opportunity for growth.
33. I feel comfortable with how I prioritize my time.
34. I have difficulty analyzing problems to identify solutions.
35. I get distracted easily during class lectures.
36. I have serious conversations with friends.
37. I share my personal weaknesses with close friends.
38. I talk with professors easily about my concerns.
39. I am frustrated when I have to work in groups.
40. I feel better when I participate in physical activity.
41. I learn from hands on experiences in the classroom.
42. I am confident in my ability to use technology when designing a project for class.
43. I find answers with well-formed questions.
44. I can identify a problem and explain it in a way others understand.
45. I am accepting of new ideas.
46. I have difficulty organizing information to write a paper.
47. I have efficient not taking skills.
48. I am comfortable with sharing my own personal beliefs about issues.
49. I am accepting of others who are different from me.
50. I can resist peer pressure from other college students.
51. I know where to go to on campus if I need assistance with my writing.
52. I am self-motivated when starting class projects.
53. I can identify areas where I need improvement.
54. I know how to find scholarly resources on a specific topic area.
55. I have a willingness to learn.
56. I am afraid to ask for a professor's help.
57. I talk with family members about my college experience.
58. I stand up for what I believe to be important.
59. I spend time weekly participating in a relaxing activity.
60. I make healthy food choices.
61. I adapt well to new technology.
62. I appreciate the arts.
63. I get an adequate amount of sleep.
64. I stay up to date on current affairs.
65. I create my own perspective from information I learned in class.
66. I speak up in class to voice my opinion.
67. I know how to conduct myself in the classroom environment.
68. I am afraid to ask questions in class.
69. I introduce myself to new people I meet.
70. I share my problems with others.
71. I know when I need to take a break.
72. I try new physically challenging activities.
73. I apply information I have learned in class.
74. I participate often in class.
75. I am intimidated by my peers' academic abilities.
76. I understand assigned readings for my classes.
77. I respect others' decisions.
78. I am comfortable saying no to my peers.
79. I can balance my relationships with school priorities.
80. I can deal with rejection.
81. I know how to navigate campus.
82. I balance school work with free time.

83. I can relate ideas learned across subject areas.
84. I write efficiently for my area of study.
85. I use information I learn in class.
86. I feel it is all right to fail.
87. I can work on class projects with difficult people.
88. I have established new friendships.
89. I dedicate time to volunteering with diverse people in the community.
90. I find time for breaks during the day.
91. I am confident in my ability to assemble a book shelf.
92. I have adjusted well to living with a roommate.

Overall Competence

93. Rate your intellectual competence
94. Rate your physical and manual competence
95. Rate your interpersonal competence

Demographics

1. Gender:

   Female
   Male
   Transgender
   Other

2. Group with which you identify the most:

   American Indian/Native American
   Asian or Pacific Islander
   Black, Non-hispanic
   White, Non-hispanic
   Hispanic
   Non-Resident Alien
   Other

3. Which academic college do you belong to:

   College of Agriculture and Life Sciences
   College of Architecture and Urban Studies
   College of Engineering
   College of Liberal Arts and Human Sciences
   College of Natural Resources
   Pamplin College of Business
   College of Science
   University Studies
   Other

4. Is English your native language?

   Yes
   No

5. Are you a first-generation college student?

   Yes
   No

6. Do you have time to participate in extra-curricular activities?

   Yes
   No

7. Do you participate in intramural sports?

   Yes
   No

8. Do you participate in sport clubs?

   Yes
   No

9. Are you an athlete for a particular sport at the university?

   Yes
   No

10. Are you a member of a club or organization at the university?

    Yes
    No

11. Do you participate in volunteer activities?

    Yes
    No

12. Do you participate in artistic activities? (such as photography, painting, sculpting, building models,ll theatre, music)?

    Yes
    No

13. Is your cumulative GPA higher than a 3.0?

    Yes
    No

14. Are you a member of the honors program?

    Yes
    No

15. Do you live on campus?

    Yes
    No

Qualitative Questions

1. What at the university has made you feel confident in your intellectual abilities?
2. What at the university has made you feel confident in your physical and manual abilities?
3. What at the university has made you feel confident in your interpersonal abilities?
4. How have you changed since freshmen year?
5. What types of academic activities are you involved in on campus?
6. What types of out of class activities are you involved in on campus?

APPENDIX C. SENSE OF COMPETENCE SCALE-REVISED (SCS-R)

Recruitment E-mail for Pilot Test

Dear Student,

I am writing to ask for your help in a study about college students and their personal perceptions of their own sense of competence. The purpose of this study is to understand intellectual, interpersonal and physical competence of undergraduate students who are of sophomore class standing.

Your participation in this study is voluntary and will not be restricted on the basis of race, gender, age or any other characteristic. By submitting this survey, you consent to participate. You may withdraw from this study at anytime and may refuse to answer any questions that make you feel uncomfortable. It is anticipated that the survey will take approximately 20-25 minutes to complete.

By participating in the study your freely and voluntarily, and without element of force or coercion, consent to be a participant in the research study. The main purpose of the study is to better understand how confident undergraduate students' feel about their intellectual, interpersonal and physical competence. All of your answers to the questions will be kept confidential and your name will not appear on any of the results. No individual responses will be reported, only group findings.

All responses will be kept entirely confidential and will be securely maintained with limited access. Your answers to the questionnaire will be used only in combination with those of other respondents in the sample to develop an understanding of undergraduates' sense of competence. By filling out the questionnaire you will see a link at the end for a second survey. The second survey is an incentive form, **you will automatically be entered into a prize drawing for a $50 gift card to the university bookstore**. The second survey is not linked to the questionnaire.

The research is being conducted in fulfillment of degree requirements at Virginia Tech to receive a Doctor of Philosophy in Educational Leadership and Policy Studies. Please complete the survey at the following link: by Friday, November 13, 2009. Thank you very much for helping with this study.

Sincerely,

Cara McFadden
Doctoral Candidate, Higher Education
Educational Leadership and Policy Studies
Virginia Tech
cmcfadd@vt.edu
Principal Investigator: Steven Janosik, Ph.D.
Secondary Investigators: Gary Skaggs, Ph.D., Nancy Bodenhorn, Ph.D., Laura Welfare, Ph.D.

APPENDIX D. SENSE OF COMPETENCE SCALE-REVISED (SCS-R)

E-mail Reminder for Pilot Test

Dear Student,

Last week a questionnaire was sent to you via email seeking your own personal perception regarding your sense of competence as a sophomore. The following is the original message. Please take a few minutes to fill out the survey below if you agree to participate in the study.

Original Message:
I am writing to ask for your help in a study about college students and their personal perceptions of their own sense of competence. The purpose of this study is to understand intellectual, interpersonal and physical competence of undergraduate students who are of sophomore class standing.

Your participation in this study is voluntary and will not be restricted on the basis of race, gender, age or any other characteristic. By submitting this survey, you consent to participate. You may withdraw from this study at anytime and may refuse to answer any questions that make you feel uncomfortable.
It is anticipated that the survey will take approximately 20-25 minutes to complete.

By participating in the study your freely and voluntarily, and without element of force or coercion, consent to be a participant in the research study. The main purpose of the study is to better understand how confident undergraduate students' feel about their intellectual, interpersonal and physical competence. All of your answers to the questions will be kept confidential and your name will not appear on any of the results. No individual responses will be reported, only group findings.

All responses will be kept entirely confidential and will be securely maintained with limited access. Your answers to the questionnaire will be used only in combination with those of other respondents in the sample to develop an understanding of undergraduates' sense of competence. By filling out the questionnaire you will see a link at the end for a second survey. The second survey is an incentive form, **you will automatically be entered into a prize drawing for a $50 gift card to the university bookstore**. The second survey is not linked to the questionnaire.

The research is being conducted in fulfillment of degree requirements at Virginia Tech to receive a Doctor of Philosophy in Educational Leadership and Policy Studies. Please complete the survey at the following link: by Friday, November 13, 2009. Thank you very much for helping with this study.

Sincerely,

Cara McFadden
Doctoral Candidate, Higher Education
Educational Leadership and Policy Studies
Virginia Tech
cmcfadd@vt.edu
Principal Investigator: Steven Janosik, Ph.D.
Secondary Investigators: Gary Skaggs, Ph.D., Nancy Bodenhorn, Ph.D., Laura Welfare, Ph.D.

APPENDIX E. SENSE OF COMPETENCE SCALE-REVISED (SCS-R)

SCS-R Field Test

The purpose of this study is to develop an instrument to measure the sense of competence of traditional age college students across the dimensions that define the construct. Administrators can use data from the instrument to modify an institution's academic and social environment to enhance the development of the intellectual, physical, and interpersonal competencies of college students.

Important Definitions:

1. **Sense of Competence Scale –Revised (SCS-R): measures a student's intellectual, physical and manual, and interpersonal skills.**
2. **Intellectual competence: "involves mastering content, gaining intellectual and aesthetic sophistication, and, most important, building a repertoire of skills to comprehend, ananlyze and synthesize" (Chickering &Reisser, 1993, p. 45)**
3. **Physical and manual competence: "can involve athletic and artistic achievement, designing and making tangible products, and gaining strength, fitness and self-discipline" (Chickering & Reisser, 1993, p. 46).**
4. **Interpersonal competence: "entails not only the skills of listening, cooperating, and communicating effectively, but also the more complex abilities to tune-in to another person and respond appropriately, to align personal agendas with the goals of the group, and to choose from a variety of strategies to help a relationship flourish or a group function" (Chickering & Reisser, 1993, p. 46).**

Please select the appropriate response which best represents your feelings about each statement.

1. I read fast enough to handle assignments at college.
2. I am comfortable meeting new people.
3. I have little difficulty with college work.
4. I feel empathy for others.
5. I respond to the needs of others.
6. I make friends easily.
7. I am competent using computers.
8. I communicate daily with peers using social networking sites.
9. College work is easy.
10. I trust the decisions I make.
11. I participate daily in physical activity.
12. I get along with most of the people I meet.
13. I believe it is important to live a healthy lifestyle.
14. I feel overwhelmed with academic work.
15. I am clumsy when it comes to participating in physical activity.
16. I get frustrated when using new computer software.
17. I will achieve my academic goals.
18. Professors underrate my ability.
19. I converse casually with my friends.
20. I am competent using social networking sites.
21. I get frustrated when reading instructions to put something together.
22. I welcome criticism as an opportunity for growth.
23. I have difficulty analyzing problems to identify solutions.
24. I have serious conversations with friends.
25. I talk with professors easily about my concerns.
26. I feel better when I participate in physical activity.
27. I learn from hands on experiences in the classroom.

28. I am confident in my ability to use technology when designing a project for class.
29. I find answers with well-formed questions.
30. I can identify a problem and explain it in a way others understand.
31. I am accepting of new ideas.
32. I have efficient not taking skills.
33. I am comfortable with sharing my own personal beliefs about issues.
34. I am accepting of others who are different from me.
35. I am self-motivated when starting class projects.
36. I know how to find scholarly resources on a specific topic area.
37. I listen to group member's ideas when working on class projects.
38. I am afraid to ask for a professor's help.
39. I spend time weekly participating in a relaxing activity.
40. I make healthy food choices.
41. I adapt well to new technology.
42. I get an adequate amount of sleep.
43. I create my own perspective from information I learned in class.
44. I know how to conduct myself in the classroom environment.
45. I am afraid to ask questions in class.
46. I introduce myself to new people.
47. I share my problems with others.
48. I know when I need to take a break.
49. I try new physically challenging activities.
50. I apply information I have learned in class.
51. I participate often in class.
52. I understand assigned readings for my classes.
53. I respect others' decisions when working on class projects.
54. I am comfortable saying no to my peers.
55. I can balance my relationships with school priorities.
56. I can relate ideas learned across subject areas.
57. I write efficiently for my area of study.
58. I use information I learn in class.
59. I can work on class projects with difficult people.
60. I have established new friendships.
61. I find time for breaks during the day.
62. I am confident in my ability to assemble a book shelf.
63. I have adjusted well to living with a roommate.

Overall Competence

64. Rate your intellectual competence
65. Rate your physical and manual competence
66. Rate your interpersonal competence

Demographics

1. Gender:

   Female
   Male
   Transgender
   Other

2. Group with which you identify the most:

   American Indian/Native American

Asian or Pacific Islander
Black, Non-hispanic
White, Non-hispanic
Hispanic
Non-Resident Alien
Other

3. Which academic college do you belong to:

College of Agriculture and Life Sciences
College of Architecture and Urban Studies
College of Engineering
College of Liberal Arts and Human Sciences
College of Natural Resources
Pamplin College of Business
College of Science
University Studies
Other

4. How long have you attended Virginia Tech?

1-2 semesters
3-4 semesters
5-6 semesters
7 or more semesters

5. Is English your native language?

Yes
No

6. Are you a first-generation college student?

Yes
No

7. Are you a transfer student?

Yes
No

8. Do you have time to participate in extra-curricular activities?

Yes
No

9. Do you participate in intramural sports?

Yes
No

10. Do you participate in sport clubs?

    Yes
    No

11. Are you an athlete for a particular sport at the university?

    Yes
    No

12. Are you a member of a club or organization at the university?

    Yes
    No

13. Do you participate in volunteer activities?

    Yes
    No

14. Do you participate in artistic activities? (such as photography, painting, sculpting, building models,ll theatre, music)?

    Yes
    No

15. Is your cumulative GPA higher than a 3.0?

    Yes
    No

16. Are you a member of the honors program?

    Yes
    No

17. Do you live on campus?

    Yes
    No

Qualitative Questions

1. What at the university has made you feel confident in your intellectual abilities?
2. What at the university has made you feel confident in your physical and manual abilities?
3. What at the university has made you feel confident in your interpersonal abilities?
4. How have you changed since freshmen year?
5. What types of academic activities are you involved in on campus?
6. What types of out of class activities are you involved in on campus?
7. What have your learned from your involvement at the university?

APPENDIX F. SENSE OF COMPETENCE SCALE-REVISED (SCS-R)

Recruitment E-mail for Field Test

Dear Student,

I am writing to ask for your help in a study about college students and their personal perceptions of their own sense of competence. The purpose of this study is to understand intellectual, interpersonal and physical competence of undergraduate students who are of sophomore class standing. The request for participation in the study was sent through the sophomore class listserv. My hopes are to have good representation of the Class of 2012.

Your participation in this study is voluntary and will not be restricted on the basis of race, gender, age or any other characteristic. By submitting this survey, you consent to participate. You may withdraw from this study at anytime and may refuse to answer any questions that make you feel uncomfortable. It is anticipated that the survey will take approximately 20-25 minutes to complete.

By participating in the study you freely and voluntarily, and without element of force or coercion, consent to be a participant in the research study. The main purpose of the study is to better understand how confident undergraduate students' are about their intellectual, interpersonal and physical competence. All of your answers to the questions will be kept confidential and your name will not appear on any of the results. No individual responses will be reported, only group findings.

All responses will be kept entirely confidential and will be securely maintained with limited access. Your answers to the questionnaire will be used only in combination with those of other respondents in the sample to develop an understanding of undergraduates' sense of competence. By filling out the questionnaire you will see a link at the end for a second survey. The second survey is an incentive form, you will automatically be entered into a prize drawing for an IPOD or a chance of winning a $50 gift card to the university bookstore (a total of 7 prizes will be awarded). The second survey is not linked to the questionnaire.

The research is being conducted in fulfillment of degree requirements at Virginia Tech to receive a Doctor of Philosophy in Educational Leadership and Policy Studies. **Please complete the survey at the following link: https://survey.vt.edu/survey/entry.jsp?id=1266346033974 by Friday, March 5, 2010.** Thank you very much for helping with this study.
Sincerely,

Cara McFadden
Doctoral Candidate, Higher Education
Educational Leadership and Policy Studies
Virginia Tech
cmcfadd@vt.edu
Principal Investigator: Steven Janosik, Ph.D.
Secondary Investigators: Gary Skaggs, Ph.D., Nancy Bodenhorn, Ph.D., Laura Welfare, Ph.D.

APPENDIX G. SENSE OF COMPETENCE SCALE-REVISED (SCS-R)

E-mail Reminder for Field Test

Dear Student,
Last week a questionnaire was sent to you via email seeking your own personal perception regarding your sense of competence as a sophomore. The following is the original message. Please take a few minutes to fill out the survey below if you agree to participate in the study.

Original Message:
I am writing to ask for your help in a study about college students and their personal perceptions of their own sense of competence. The purpose of this study is to understand intellectual, interpersonal and physical competence of undergraduate students who are of sophomore class standing. The request for participation in the study was sent through the sophomore class listserv. My hopes are to have good representation of the Class of 2012.
Your participation in this study is voluntary and will not be restricted on the basis of race, gender, age or any other characteristic. By submitting this survey, you consent to participate. You may withdraw from this study at anytime and may refuse to answer any questions that make you feel uncomfortable.

It is anticipated that the survey will take approximately 20-25 minutes to complete.
By participating in the study you freely and voluntarily, and without element of force or coercion, consent to be a participant in the research study. The main purpose of the study is to better understand how confident undergraduate students' are about their intellectual, interpersonal and physical competence. All of your answers to the questions will be kept confidential and your name will not appear on any of the results. No individual responses will be reported, only group findings.

All responses will be kept entirely confidential and will be securely maintained with limited access. Your answers to the questionnaire will be used only in combination with those of other respondents in the sample to develop an understanding of undergraduates' sense of competence. By filling out the questionnaire you will see a link at the end for a second survey. The second survey is an incentive form, you will automatically be entered into a prize drawing for an IPOD or a chance of winning a $50 gift card to the university bookstore (a total of 7 prizes will be awarded). The second survey is not linked to the questionnaire.

The research is being conducted in fulfillment of degree requirements at Virginia Tech to receive a Doctor of Philosophy in Educational Leadership and Policy Studies. **Please complete the survey at the following link: https://survey.vt.edu/survey/entry.jsp?id=1266346033974 by Friday, March 5, 2010.** Thank you very much for helping with this study.

Sincerely,

Cara McFadden
Doctoral Candidate, Higher Education
Educational Leadership and Policy Studies
Virginia Tech
cmcfadd@vt.edu
Principal Investigator: Steven Janosik, Ph.D.
Secondary Investigators: Gary Skaggs, Ph.D., Nancy Bodenhorn, Ph.D., Laura Welfare, Ph.D.

APPENDIX H. INTELLECTUAL DIMENSION SCORING FOR THE SCS-R

Raw Score	Logit Measure	Percentile Rank	Scale Score	Raw Score	Logit Measure	Percentile Rank	Scale Score
22	-7.74	0	0.00	59	0.37	16	49.63
23	-6.49	0	7.65	60	0.54	22	50.67
24	-5.74	0	12.24	61	0.71	28	51.72
25	-5.27	0	15.12	62	0.89	36	52.82
26	-4.92	0	17.26	63	1.07	44	53.92
27	-4.63	0	19.03	64	1.25	51	55.02
28	-4.39	0	20.50	65	1.43	59	56.12
29	-4.17	0	21.85	66	1.61	67	57.22
30	-3.96	0	23.13	67	1.8	73	58.39
31	-3.77	0	24.30	68	1.97	78	59.43
32	-3.6	0	25.34	69	2.15	81	60.53
33	-3.43	0	26.38	70	2.33	84	61.63
34	-3.26	0	27.42	71	2.5	86	62.67
35	-3.11	0	28.34	72	2.68	88	63.77
36	-2.95	0	29.32	73	2.85	90	64.81
37	-2.8	0	30.23	74	3.03	92	65.91
38	-2.65	0	31.15	75	3.21	93	67.02
39	-2.51	0	32.01	76	3.39	95	68.12
40	-2.37	1	32.87	77	3.58	96	69.28
41	-2.22	1	33.78	78	3.78	97	70.50
42	-2.08	1	34.64	79	3.98	97	71.73
43	-1.95	1	35.44	80	4.21	98	73.14
44	-1.81	1	36.29	81	4.45	99	74.60
45	-1.67	1	37.15	82	4.72	99	76.26
46	-1.53	1	38.01	83	5.02	99	78.09
47	-1.4	1	38.80	84	5.38	99	80.30
48	-1.26	1	39.66	85	5.81	99	82.93
49	-1.12	1	40.52	86	6.37	99	86.35
50	-0.99	1	41.31	87	7.25	100	91.74
51	-0.85	1	42.17	88	8.59	100	99.94
52	-0.71	2	43.02				
53	-0.56	3	43.94				
54	-0.42	4	44.80				
55	-0.27	5	45.72				
56	-0.12	6	46.64				
57	0.04	8	47.61				
58	0.2	11	48.59				

APPENDIX I. PHYSICAL AND MANUAL DIMENSION SCORING FOR THE SCS-R

Raw Score	Logit Measure	Percentile Rank	Scale Score	Raw Score	Logit Measure	Percentile Rank	Scale Score
18	-6.33	0	8.63	56	1.39	55	55.88
19	-5.09	0	16.22	57	1.55	61	56.86
20	-4.35	0	20.75	58	1.72	67	57.90
21	-3.9	0	23.50	59	1.89	72	58.94
22	-3.56	0	25.58	60	2.06	77	59.98
23	-3.29	0	27.24	61	2.25	82	61.14
24	-3.05	0	28.70	62	2.44	86	62.30
25	-2.85	0	29.93	63	2.63	89	63.47
26	-2.66	0	31.09	64	2.84	92	64.75
27	-2.49	0	32.13	65	3.07	94	66.16
28	-2.33	0	33.11	66	3.31	96	67.63
29	-2.18	0	34.03	67	3.58	97	69.28
30	-2.03	1	34.95	68	3.89	98	71.18
31	-1.9	1	35.74	69	4.26	99	73.44
32	-1.76	1	36.60	70	4.75	99	76.44
33	-1.63	1	37.39	71	5.53	99	81.21
34	-1.51	1	38.13	72	6.79	99	88.92
35	-1.38	1	38.92				
36	-1.26	1	39.66				
37	-1.14	1	40.39				
38	-1.02	1	41.13				
39	-0.9	1	41.86				
40	-0.78	1	42.60				
41	-0.66	1	43.33				
42	-0.54	1	44.07				
43	-0.42	1	44.80				
44	-0.29	2	45.60				
45	-0.17	3	46.33				
46	-0.04	4	47.13				
47	0.09	6	47.92				
48	0.22	8	48.72				
49	0.35	11	49.51				
50	0.49	16	50.37				
51	0.63	22	51.23				
52	0.78	28	52.14				
53	0.92	35	53.00				
54	1.07	42	53.92				
55	1.23	48	54.90				

APPENDIX J. INTERPERSONAL DIMENSION SCORING FOR THE SCS-R

Raw Score	Logit Measure	Percentile Rank	Scale Score	Raw Score	Logit Measure	Percentile Rank	Scale Score
23	-7.49	0	1.53	60	0.07	3	47.80
24	-6.25	0	9.12	61	0.21	5	48.66
25	-5.5	0	13.71	62	0.36	7	49.57
26	-5.04	0	16.53	63	0.51	9	50.49
27	-4.7	0	18.61	64	0.66	12	51.41
28	-4.42	0	20.32	65	0.81	16	52.33
29	-4.18	0	21.79	66	0.97	21	53.31
30	-3.96	0	23.13	67	1.13	27	54.29
31	-3.77	0	24.30	68	1.29	33	55.26
32	-3.59	0	25.40	69	1.45	39	56.24
33	-3.42	0	26.44	70	1.61	45	57.22
34	-3.26	0	27.42	71	1.77	51	58.20
35	-3.11	0	28.34	72	1.94	57	59.24
36	-2.97	0	29.19	73	2.1	63	60.22
37	-2.83	0	30.05	74	2.26	67	61.20
38	-2.69	0	30.91	75	2.43	72	62.24
39	-2.56	0	31.70	76	2.6	76	63.28
40	-2.43	0	32.50	77	2.77	79	64.32
41	-2.3	0	33.29	78	2.94	82	65.36
42	-2.18	0	34.03	79	3.12	85	66.46
43	-2.06	0	34.76	80	3.3	87	67.57
44	-1.94	0	35.50	81	3.49	89	68.73
45	-1.81	0	36.29	82	3.68	91	69.89
46	-1.69	0	37.03	83	3.89	94	71.18
47	-1.58	1	37.70	84	4.11	96	72.52
48	-1.46	1	38.43	85	4.35	97	73.99
49	-1.34	1	39.17	86	4.61	98	75.58
50	-1.22	1	39.90	87	4.9	98	77.36
51	-1.1	1	40.64	88	5.23	99	79.38
52	-0.97	1	41.43	89	5.62	99	81.76
53	-0.85	1	42.17	90	6.13	99	84.89
54	-0.73	1	42.90	91	6.94	99	89.84
55	-0.6	1	43.70	92	8.21	99	97.62
56	-0.47	1	44.49				
57	-0.34	1	45.29				
58	-0.21	1	46.08				
59	-0.07	2	46.94				

APPENDIX K. IRB EXEMPT APPROVAL

DATE: November 4, 2009

MEMORANDUM

TO: Steven M. Janosik
Gary E. Skaggs
Nancy E. Bodenhorn

FROM: Carmen Green

SUBJECT: **IRB Exempt Approval:** "Instrument Development and Validation of the Sense of Competence Scale - Revised (SCS-R)" , IRB # 09-939

I have reviewed your request to the IRB for exemption for the above referenced project. The research falls within the exempt status, CFR 46.101(b) category(ies) 2.

Approval is granted effective as of November 4, 2009.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in the research protocol. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.

2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

APPENDIX L. IRB AMENDMENT 1 APPROVAL

DATE: November 12, 2009

MEMORANDUM

TO: Steven M. Janosik
Cara McFadden
Nancy E. Bodenhorn

FROM: Carmen Green

SUBJECT: **IRB Amendment 1 Approval:** "Instrument Development and Validation of the Sense of Competence Scale - Revised (SCS-R)" , IRB # 09-939

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on November 4, 2009. You subsequently requested permission to amend your IRB application.

Approval has been granted for the requested protocol amendment, effective as of November 12, 2009.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in the research proposal. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.

2. Report promptly to the IRB any injuries or other unanticipated or adverse involving risks or harms to human research subjects or others.

APPENDIX M. IRB AMENDMENT 2 APPROVAL

DATE: February 23, 2010

MEMORANDUM

TO: Steven M. Janosik
Cara McFadden
Nancy E. Bodenhorn

FROM: David M. Moore

SUBJECT: **IRB Amendment 2 Approval:** "Instrument Development and Validation of the Sense of Competence Scale - Revised (SCS-R)" , IRB # 09-939

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on November 4, 2009. You subsequently requested permission to amend your IRB application.

Approval has been granted for the requested protocol amendment, effective as of February 23, 2010.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in the research protocol. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.

2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.