

The University of Southern Mississippi  
**The Aquila Digital Community**

---

Dissertations

---

Fall 12-2013

**Multi-Group Invariance of the Conceptions of Assessment Scale  
Among University Faculty and Students**

Melanie Anne DiLoreto  
*University of Southern Mississippi*

Follow this and additional works at: <https://aquila.usm.edu/dissertations>



Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Higher Education Commons](#)

---

**Recommended Citation**

DiLoreto, Melanie Anne, "Multi-Group Invariance of the Conceptions of Assessment Scale Among University Faculty and Students" (2013). *Dissertations*. 214.  
<https://aquila.usm.edu/dissertations/214>

This Dissertation is brought to you for free and open access by The Aquila Digital Community. It has been accepted for inclusion in Dissertations by an authorized administrator of The Aquila Digital Community. For more information, please contact [Joshua.Cromwell@usm.edu](mailto:Joshua.Cromwell@usm.edu).

The University of Southern Mississippi

MULTI-GROUP INVARIANCE OF THE CONCEPTIONS OF ASSESSMENT SCALE  
AMONG UNIVERSITY FACULTY AND STUDENTS

by

Melanie Anne DiLoreto

Abstract of a Dissertation  
Submitted to the Graduate School  
of The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

December 2013

## ABSTRACT

### MULTI-GROUP INVARIANCE OF THE CONCEPTIONS OF ASSESSMENT SCALE AMONG UNIVERSITY FACULTY AND STUDENTS

by Melanie Anne DiLoreto

December 2013

Conceptions are contextual. In the realm of education, conceptions of various constituent groups are often shaped over a period of a number of years during which time these groups have participated in educational endeavors. Specifically, conceptions of assessment are influenced by beliefs, actions, attitudes, understandings, and past experiences. These conceptions can impact both teaching and learning, and ultimately student achievement. Based on the past work of Gavin T. L. Brown (2004, 2006, 2008, 2010, & 2011) and Fletcher, Meyer, Anderson, Johnston, and Rees (2011) concerning conceptions of assessment held by educators and students in environments with a low-stakes assessment culture, this study re-examines the factor structure of the Conceptions of Assessment III (CoA-III) for faculty and students of higher education within a high-stakes assessment culture. Five models were initially considered based on past research by the above-mentioned researchers. Upon examining model fit of these five models, results indicated an acceptable model fit to the data collected from faculty and students within the United States. Furthermore, invariance testing elicited differences in how faculty ( $N = 159$ ) and students ( $N = 404$ ) of higher education conceptualize the purpose of assessment. Specifically, faculty members report that a primary purpose of assessment is for improvement of both teaching and learning. This group also suggested that assessment is useful for ensuring student accountability. However, results indicate that

students associate assessment with accountability – both at the institutional and student levels. Furthermore, the data also suggest that a strong relationship between accountability and improvement exists. These results can be interpreted to mean that as accountability measures increase, there is a concomitant rise in the use of assessment for improvement purposes. Additional results of this study and implications of these findings for educational settings with high-stakes assessment cultures are discussed.

COPYRIGHT BY  
MELANIE ANNE DILORETO  
2013

The University of Southern Mississippi

MULTI-GROUP INVARIANCE OF THE CONCEPTIONS OF ASSESSMENT SCALE  
AMONG UNIVERSITY FACULTY AND STUDENTS

by

Melanie Anne DiLoreto

A Dissertation  
Submitted to the Graduate School  
of The University of Southern Mississippi  
in Partial Fulfillment of the Requirements  
for the Degree of Doctor of Philosophy

Approved:

Kyna Shelley \_\_\_\_\_  
Director

Rich Mohn \_\_\_\_\_

Forrest Lane \_\_\_\_\_

Lillian Hill \_\_\_\_\_

Susan A. Siltanen \_\_\_\_\_  
Dean of the Graduate School

December 2013

## ACKNOWLEDGMENTS

It is with sincere gratitude that I would like to thank my advisor, committee chair, personal and professional mentor, Dr. Kyna Shelley, for her support and encouragement throughout my doctoral studies and dissertation. Her expertise and wisdom across multiple situations has provided the grounding necessary for me to persist throughout the duration of my studies and this dissertation.

Special thanks go to my remaining committee members, Dr. Rich Mohn who always assisted without hesitation and intimidation; Dr. Forrest Lane, who challenged me in ways that improved my thinking and ultimately the end product; and Dr. Lillian Hill, who was available any time, and who provided prompt feedback and multiple ideas throughout our discussions.

Finally, I would like to express special thanks to my family and friends for their belief in me. Without them I would not have had the strength or the stamina to continue this journey!

## TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGMENTS.....	iv
LIST OF TABLES.....	vi
CHAPTER	
I. INTRODUCTION.....	1
Statement of the Problem	
Purpose of the Study	
II. REVIEW OF RELATED LITERATURE.....	13
Theoretical Framework	
Assessment in Higher Education in the United States	
Accountability in Education	
Teaching, Learning, and Assessment	
Conceptions of Assessment	
III. METHODOLOGY.....	55
IV. RESULTS.....	60
Data Analysis	
Confirmatory Factor Analysis	
Results: Research Questions 1 and 2	
Invariance of the CoA-III	
Results: Research Questions 3 and 4	
V. DISCUSSION.....	85
Model Fit	
Summary of Qualitative Findings	
Implications of the Findings	
Limitations and Suggestions for Future Research	
APPENDIXES.....	98
REFERENCES.....	109



## LIST OF TABLES

### Table

1.	Data Imputations using Multiple Regression.....	62
2.	Faculty and Student Mean Scores.....	63
3.	Faculty and Student Reliability Coefficients .....	63
4.	Omnibus Test of Goodness-of-Fit .....	67
5.	CoA-III Factor Loadings for Five Models Tested .....	68
6.	Omnibus Tests: Correlations between Latent Variables.....	72
7.	Correlations between Latent Variables for Faculty and Students.....	74
8.	CoA-III Factor Loadings for Faculty and Students .....	75
9.	First-Order Factor Chi-Square Differences .....	78
10.	Chi-Square Difference Tests for Items of First-Order Factors .....	79
11.	Chi-Square Difference Tests of Second-Order Factors .....	80
12.	Statistically Significant Items Different Across Groups.....	80
13.	Chi-Square Difference Tests for Correlations .....	81
14.	Types of Assessment Activities .....	83

## CHAPTER I

### INTRODUCTION

Assessment serves multiple purposes for students, faculty, and administrators of institutions of higher education. Further, assessment practices have evolved as a result of the demands of external stakeholders. One dilemma faced by stakeholders is the fact that the term assessment is often used within different contexts and with different meanings (Garfield, 1994). Harlen (2007) indicates that the term assessment describes a process by which evidence is collected for some purpose. Specifically, Harlen (2007) describes that term assessment refers to the evidence about what students know and can do and the judgments about their achievements. Wang and Hurley (2012) indicate that an assessment movement in higher education began in the 1980s with an emphasis on student learning. Since that time, accrediting agencies have required institutions of higher education to implement program-level and institution-level assessment procedures in addition to documenting student learning. Wang and Hurley (2012) found that the way assessment is perceived by faculty may impact student achievement. Anderson, Moore, Anaya, and Bird (2005) express their belief that the emphasis of assessment should be to focus on outcomes in a global sense.

Past research indicates that beliefs about assessment impact the way instructors teach and the way students learn (Brown, 2004; Struyven, Dochy, & Janssens, 2005). Furthermore, faculty members' attitudes toward and expertise in assessment impact the way they implement their own assessments in their classrooms. Additionally, students' attitudes toward and their experiences with assessment affect their personal approach to learning and their beliefs toward future successes as a learner (Fletcher et al., 2011).

Both sets of attitudes are affected by the way these individuals think, believe, and act. Of course, actions are the outgrowth of the desire to initiate a behavior. Consequently, the way one thinks about assessment greatly influences one's conceptualization and ultimately the behaviors associated with assessing. Indeed, planned behavior theory implies "what people believe, the amount of control they have or perceive they have, societal norms, and people's intentions interact to shape the behaviors and practices people carry out" (Chen, Brown, Hattie, & Millward, 2012, p. 938).

The term *thought* is a broad description of everything that comes to mind, that *goes through our heads*. The process of thought, how we think, and the process for training thoughts induce beliefs (Dewey, 1933). Furthermore, individuals' beliefs are impacted by their past experiences, reflective thoughts, and their evidence for their beliefs (Bandura, 1986; Dewey, 1933). Dewey (1933) stated, "The data at hand cannot supply the solution; they can only suggest it. . . Clearly past experience and prior knowledge" (p. 12) are the sources of the suggestion. Thoughts, as the centerpiece of beliefs which induce actions based on individual beliefs, affect teaching strategies, assessment practices and curricula used in classrooms and in courses.

Beliefs are meanings that are based on lived experiences and cultural norms from which sense is made about these experiences (Ekeblad & Bond, 1994). When used with educators, these meanings are oftentimes confusing due to the myriad of terms researchers use in their attempts to gain an understanding of how educators' actions are dictated by preconceived ideas versus knowledge. Also, the unknown impact of belief systems on the way educators teach and the way these beliefs impact student learning is problematic. Individual beliefs often do not require any type of general consensus that

might be required for validation in other areas like skills or knowledge. In fact, individual beliefs do not even need to be consistent within the educator's self-held belief system. In other words, educators do not put their beliefs up for debate or evaluation. However, an educator's knowledge and skill set might be open to critique. Furthermore, there is evidence to suggest that teachers' beliefs affect how they act outside of the classroom while their teaching behaviors are the result of their belief system being filtered through their experiences, which in turn have been affected by the accumulation of knowledge and skills (Pajares, 1992). Hence, beliefs, even though difficult to define and measure, play an important role in the understanding of conceptions of assessment.

Conceptions of various ideas, their importance and usefulness, are then in turn affected by the thoughts and belief system held by the individual. Generally speaking, conceptions are broad mental structures that encompass beliefs, representing how things work and are experienced (Brown, 2004; Pratt, 1992). Furthermore, conceptions can be thought of as mental constructs or representations of an individual's reality (Brown & Lake, 2006; Fodor, 1998; Kelly, 1991; Lakoff & Johnson, 2003; Thompson, 1992). Thus, according to White (1994), these conceptions are then "communicated in language or metaphors containing beliefs, meanings, preferences and attitudes that explains complex and difficult categories of experience such as assessment" (p. 2). Consequently, educators' conceptions are impacted by their belief systems and thus affect their pedagogical acts including teaching, learning, and assessment. Likewise, students' conceptions are affected by and filtered through their belief systems, thus affecting their views of assessment and its usefulness. Indeed, because conceptions are filtered through an individual's belief system, the conceptions of assessment held by students are often

different from those held by their teachers (Brown, 2004). In order to change an individual's conceptions, an awareness of the held conceptions must be present and then an argument for a different theoretical or explanatory framework must replace the earlier one (Vosniadou, 1992, 1994).

While it seems that the disparity of belief systems and their effects on the conceptions of assessment among the various stakeholders in education is real, it is also clear that the increased accountability pressure (often politicized and marketed as value added) to have institutions, schools, and teachers show improvement in student learning outcomes advocated by politicians, public policy, and parent populations, has impacted learning (Brown, 2011) and its measurement in various ways. The multifaceted purpose of assessment includes obtaining information about student learning, student progress, quality of teaching, as well as program and institutional accountability (Brown, 2010). Each facet of this purpose is affected by the beliefs of those who are implementing the assessments as well as those who are being assessed. The term assessment often signifies an ongoing process used to describe what students know and can do (Burger, n.d.). Also, there is an abundance of evidence that educators collect assessment data for each student (e.g., tests, assignments, etc.) and that these individual student results are useful when demonstrating that teachers and schools have fulfilled external expectations of increased student learning. Even so, according to Brown (2011), the pressures of external accountability can lead to conformity at the expense of learning the intended outcomes. Furthermore, due to these external pressures it is possible that both teachers and school administrators may inflate the results of high-stakes tests to demonstrate larger gains in student learning – without real learning taking place. Thus, both the disparity of belief

systems and the variety of external pressures impact the conceptions of assessment of students.

Educational policymakers in the United States over the past several decades have implemented many federal and state mandates requiring the use of assessments to meet external accountability demands. Assessments are often used to make high-stakes decisions in the United States. Past research indicates that the assessment practices implemented by faculty members are based on their conceptions of assessment and that student approaches to learning are affected by faculty's assessment practices (Cassidy, 2006; Struyven et al., 2005). Fletcher et al. (2011) states, "Attitudes towards and expertise in assessment by university faculty have an impact on the assessments they use, how assessments are incorporated into the teaching and learning process, and whether their assessment practices provide students with the opportunity to improve their performance" (p. 120). Fletcher et al. (2011) continues, "Assessment attitudes and experiences by students will affect their approach to learning, whether they utilise assessment feedback in their future study, and the extent to which they develop the skills and understandings to become self-assessing lifelong learners" (p. 121).

Brown (2004) indicates that simply mandating the use of assessments will not necessarily make teachers actually implement these assessments unless these policymakers take into consideration the alignment of teachers' conceptions of assessment and the new policy. Furthermore, Brown, Lake, and Matters (2011) report that differences in policy, cultures, and the nature of the stakes (high or low) attached to assessment results leads to differences in how assessment is conceptualized by various stakeholders. Specifically, Brown et al. (2011) hypothesize that when there are high-

stakes for students associated with the use of assessments, teachers and students will report a student-accountability purpose of assessment.

In summary, assessing what students know and can do requires the use of tools such as tests, projects, etc. As noted above, the pedagogical acts of, and assessments used by teachers are affected by their beliefs and conceptions (i.e., beliefs about teaching, learning, assessment, curriculum, and teacher efficacy). Furthermore, past research indicates that students' approach to learning and what they achieve may be affected by the assessments implemented by educators. Various stakeholders then view the results of these assessments through the lens developed and shaped based on their own personal belief system.

Brown (2004) argues that there are four conceptions of assessment based on the beliefs K-12 teachers hold regarding teaching and learning. These four main conceptualizations of assessment are: assessment makes schools accountable, assessment makes students accountable, assessment improves education, and assessment is irrelevant (Brown, 2004). Again, these four conceptualizations held by teachers are shaped by their belief systems. Students' belief systems may be different from teachers, however, so it makes sense that students may have a different view of the purpose of assessment. In fact, "Students conceive of assessment in at least four major ways (i.e., assessment makes students accountable; assessment is irrelevant because it is bad or unfair; assessment improves the quality of learning; and assessment is enjoyable)" (Brown & Hirschfeld, 2008, p. 3). Although there is some overlap between students' and teachers' conceptions of assessment, there are significant differences.

Research studies completed in New Zealand, a low-stakes assessment environment, confirm that faculty members' and students' conceptions of assessment differ. According to Fletcher et al. (2011), higher education faculty view assessment as an aid to the teaching and learning process whereas post-secondary education students view assessment as needed simply for accountability purposes or even irrelevant to the teaching and learning process. Furthermore, past research indicates that students' conceptions of assessment impact their approach to learning and studying (Struyven et al., 2005).

#### Statement of the Problem

Assessment serves multiple purposes for students and faculty of institutions of higher education, yet there is little known about how these various groups conceptualize the purpose of assessment. Although there is an abundance of information about best practices for assessment in higher education, there is little empirically-based research regarding the different purposes of assessment (Fletcher et al., 2011).

High-stakes decisions based on assessment results impact both students and faculty. For example, university or program admission decisions, program progression decisions, and graduation decisions are just a few that typically impact students and rely on results from assessments. With regards to faculty members, they are also evaluated based on assessment data. Pertinent examples include tenure and promotion decisions. These decisions are often influenced by the assessment data obtained from student feedback and annual evaluations. Furthermore, these tenure and promotion decisions are often further impacted by an assessment of the number and quality of publications produced by the faculty member. Furthermore, these types of decisions are associated



with high-stakes that impact student and faculty approaches to the various assessments being used. In particular, these assessments are viewed by students and faculty based on how they conceptualize the purpose of assessment.

Brown (2004) indicates that conceptions are contextual. Brown and Hirschfeld (2008) surmise that additional investigation is needed to determine if the context of a high-stakes testing culture impacts educators' conceptions of assessment. Research conducted in New Zealand (Brown, 2004, 2010; Fletcher et al., 2011), a low-stakes assessment environment on the conceptions of assessment held by faculty and/or teachers and students in both K-12 and higher education settings produced a four-latent-factor model purpose of assessment with the possibility of a two-factor model purpose of assessment. The problem is that conceptions are contextual; therefore, it is believed that data collected in a low-stakes assessment environment in New Zealand will elicit different conceptions than data collected in higher education in the United States where high-stake decisions are often made based on assessments. Thus, legitimate questions arise as to the validity of these models when determining the conceptions of assessment held by faculty and students at U.S. institutions of higher education.

#### Purpose of the Study

There are two primary objectives of this research. First, the researcher aimed to collect evidence of validity and reliability on a modified version of a previously published instrument. Second, the researcher used this questionnaire to explore the differences, if any, of how faculty members and undergraduate students of level V institutions of higher education conceptualize the purpose of assessment. Each of these level V institutions is located within the accreditation region of the Southern Association

of Colleges and schools (SACS) of the United States. Level V institutions are defined by SACS as institutions that offer three or fewer doctoral degrees as highest degrees.

### *Research Questions*

1. How do the existing models of conceptions of assessment derived from New Zealand data fit the data collected in the U.S.?
2. What are the differences (if any) of how university faculty and university students in the U.S. conceptualize assessment?
3. What themes, if any, do the participants report in their meaning of assessment?
4. When faculty and students think of the term assessment, what types of assessment activities come to mind?

### *Definitions*

*Belief* – A conscious or unconscious thought that is accepted as true by the individual which guides future actions and behaviors.

*Assessment* – “Any process that provides information about the thinking, achievement or progress of students” (Crooks, 2001, Defining Assessment section, para.

1). Boyce (2000) specifies that assessment is the “collection and use of data for the purpose of improvement” (p. 412).

*Accountability* – “The answerability for performance” (Romzek, 2000, p. 22) or the “obligation to report to others, to explain, to justify, to answer questions about how resources have been used, and to what effect” (Trow, 1996, p. 310).

*Educational Outcomes Assessment* – “Essentially the same as assessment, but may put more emphasis on assessing the outcome of a program rather than on how that outcome is developed” (Boyce, 2000, p. 412).

*Conceptions of Assessment* – “One’s beliefs, meanings, and understandings of assessment” (Fletcher et al., 2011, p. 120).

*Factors, constructs, latent variables* – for the purposes of this study these are: institutional accountability; student accountability; improvement; and irrelevant

#### *Delimitations of the Study*

The study was delimited to all four-year public level V doctoral degree-granting institutions that offer a minimum of one baccalaureate degree, located within the Southern Association of Colleges and Schools (SACS) accrediting region. Furthermore, this study was delimited to faculty members that are employed at and undergraduate students who attend these institutions of higher education.

#### *Assumptions*

This study relies on the assumption that faculty members and students who complete the questionnaire will provide an accurate depiction of their beliefs about the purpose of assessment. Furthermore, this study assumes that level V doctoral degree-granting public institutions in the SACS region have similar characteristics.

#### *Justification*

In an era of increased accountability, policymakers often use assessment data to determine student learning and to make high-stakes decisions reflected in educational policy. Accountability exists for faculty members and for students of higher education. Past accreditation requirements allowed administrators of higher education to determine

the expertise of the faculty member and assumed the faculty member was an expert able to make judgments about student work. These expert judgments are no longer enough to meet the ever-changing demands policymakers place on educators today. Through the accreditation process, institutions of higher education now face the potential to have their programs closed or funding removed if they are unable to provide assessment evidence of increased student learning. Furthermore, students must demonstrate their increased learning via course-based and program assessments. There are often high-stakes for students associated with these assessments, including their timely progression through and, ultimately, successful completion of a program resulting in graduation. Due to the high-stakes nature of such assessments, the pressures for accountability are even more evident and place it at the forefront of the assessment movement.

However, according to the report of the National Commission on Accountability in Higher Education (2005), a clear vision and purpose for assessment is lacking in higher education. As a result of this lack of vision and purpose, limited transparency exists (U.S. Department of Education, 2006). Faculty and students are aware of this lack of transparency, producing a fear or mystification of assessment. By determining the differences of how university faculty members and students conceptualize the purpose of assessment, educators may gain an understanding of the differences, if any, that may help guide future professional development and scholarship opportunities for various educational stakeholders. In addition, students may be able to change their approaches to learning based on the information – possibly resulting in higher achievement. Also, by confirming an appropriate model of conceptions of assessment, future research may

explore the differences across various groups in order to guide policymakers' decisions that influence accountability pressures.

## CHAPTER II

### REVIEW OF RELATED LITERATURE

#### Theoretical Framework

In order to understand the conceptions of assessment held by faculty members and students of higher education, it is necessary to first examine the relationship between learning and assessment. It is quite clear that practices and behaviors of individuals are influenced by their perceptions, past experiences, and beliefs (Bandura, 1986; Dewey, 1933); therefore, the way faculty members and students view, use, and engage in assessment activities associated with teaching and learning may differ. Consequently, the way these views become actualized and subsequently translated into various behaviors may be affected by beliefs held.

Reasoned action is a central factor in the theory of planned behavior. That is, the intention to perform a given behavior includes beliefs toward that particular behavior. In general, the stronger the intention to engage in a particular behavior, the more likely the individual is to actually engage in the behavior. Furthermore, both personal intentions and ability to perform a behavior influence the likelihood the individual will engage in such a behavior (Ajzen, 1991).

The literature, professional organizations, as well as many educators espouse the importance of learner-centered approaches to teaching, learning, and assessment. According to planned behavior theory, however, it does not necessarily matter what individuals are told to do in order to become effective in their teaching and/or learning because people's beliefs guide their thinking and action (Ajzen, 1991). For example, an instructor who does not believe in the learner-centered approach to teaching will likely

act in a way that is incongruent with learner-centered approaches – even if policymakers place external demands calling for such behaviors. On the other hand, instructors who believe that students learn by the act of doing are more likely to engage in a learner-centered approach to teaching as well as assessing student knowledge and skills.

Furthermore, planned behavior theory calls upon two aspects of behavioral control, actual and perceptual. Each of these aspects, actual as well as the perception of behavioral control, plays an important part in the translation of behaviors into action. Actual behavioral control is evident as it incorporates what an individual is actually capable of doing – which is self-evident. Perceived behavior control, however, refers to an individual's perception of the ease or difficulty of performing the behavior as it varies across situations and actions (Ajzen, 1991). While planned behavior theory explains both aspects, it is also necessary to consider motivation. Clearly, without motivation there is no impetus to act or behave in a certain way.

Thus, the theory of achievement motivation also has implications related to assessment in higher education. How one views the use and purpose of assessment is internally driven by one's hope for success and fear of failure. Learned-drive theory of achievement motivation is explained as the need for approval, belongingness, and achievement (Covington, 1984). The conflict between attempting success and avoiding failure and how individuals resolve this internal conflict is expressed in a need for achievement. Weiner (1972) posited that people's perceptions are the cause of their successes and failures that in turn influence their beliefs about their future achievement. The difference in how people view their successes and failures is precisely the essence of achievement motivation. For example, people motivated to *approach success* generally

attribute success as internally driven and they take personal responsibility for the success and their failures. Conversely, individuals who tend to be *failure-avoiding* typically attribute success to external factors and failure to their own personal inabilities (Covington, 1984).

It is evident that individuals are often judged based on their successes and accomplishments. The self-worth theory of achievement motivation purports that people associate behavior with self-worth based on how they conceptualize personal responsibility of the causation of successes and failures. Specifically, self-worth theory of achievement motivation incorporates one's drive for success and desire to avoid failure by associating failure with a sense of worthlessness and social disapproval. The direct and causal relationship that is perceived between one's accomplishment and self-worth is the foundation of self-worth theory of achievement motivation. Specifically, individuals believe that unless they are successful at some activity, a major source of their self-esteem will be missing (Covington, 1984).

According to Bandura (1994), in order for a person to perform tasks that ultimately influence the outcomes of specific events that are occurring or will occur, the person's self-efficacy must support such a belief of personal success. Thus, academic self-efficacy of an educator can significantly influence beliefs about personal ability to meet the demands of teaching in such a way as to positively impact the learning and achievement of students. Highly efficacious educators have a positive outlook with regard to overcoming obstacles that may seem to be impediments to teaching. Thus, an educator's academic self-efficacy as it is related to the teaching process and state-mandated assessments can significantly influence teaching and thus, student



performance. Further, educators' beliefs about assessment are impacted and these beliefs then impact their conceptions of assessment. Finally, assessment behaviors are then implemented based upon conceived type of control – actual or perceptual.

Consequently, in order to meet the demands of teaching and assessing effectively, positive self-efficacy needs to be developed. According to Bandura (1994), the best way to produce highly efficacious students is to engage in a variety of designed experiences that foster success through well-developed activities. By providing such experiences, the development of positive self-efficacy will be accomplished; however, experiences that are not well developed may cause failure and thus undermine progress toward positive self-efficacy (Bandura, 1994). As a result, academic self-efficacy increases by successfully engaging in and moving through a variety of well-constructed experiences. These mastery experiences might be carefully constructed activities, courses, or programs that build on each other. Thus, in general, providing extensive opportunities for success impacts one's ability to master one's experiences and become more confident in one's abilities. Specifically, these opportunities foster positive academic self-efficacy that can directly impact teaching effectiveness, with respect to both the delivery of content and assessment, as well as student success.

Indeed, a necessary component in the learning process is ongoing assessment. Holt and Willard-Holt (2000) indicate the importance of dynamic assessment – a way to assess the true potential of learners that differs from conventional tests. The interactive nature of the dynamic assessment process requires that the assessor, or instructor, engage in a meaningful dialogue with the learner, or student in order to (1) find out the learner's current level of performance or understanding on any given task,

and (2) discuss strategies for improving the learner's performance or understanding of future tasks. When viewed this way, it is clear that assessment and learning are two processes that should be considered as a whole. That is, it is difficult to separate assessment from the learning process. When assessment and learning are viewed as two equally necessary components of a dynamic process, the development and implementation of quality instructional practices will naturally and continually be fostered.

It is prudent to consider a constructivist view of assessment. Specifically, a social constructivist's view of assessment includes the notion that learning occurs through doing. As such, social constructivism encourages the learner (student) to arrive at a personalized version of the truth – which is influenced by personal background experiences and embedded worldviews. Furthermore, the student is at the center. The student has the responsibility of learning (Glaserfeld, 1989) and the motivation to learn is strongly dependent on confidence and an internal perspective about potential for learning. Vygotsky's (1978) zone of proximal development further supports that if students are successfully challenged within close proximity to, yet slightly above, their current level of development, they will gain the confidence and motivation needed to embark on more challenging endeavors. However, it is important to note that in order for students to gain confidence and to become or stay motivated, they must be continually challenged via a stretching of their zones of proximal development (Brownstein, 2001). These challenges should come in the form of tasks that require students to hone skills and acquire knowledge that have not yet been mastered. Furthermore, according to Derry (1999), the ideal situation is that tasks be selected in such a way as to be representative of

the learning environment in which students gain personal understanding through mastering skills and knowledge. These types of tasks will engage and challenge students in such a way that make the goal attainable while allowing the students to experience ownership of all aspects of the learning process. Effective assessment strategies such as the use of dynamic assessments as proposed by Holt and Willard-Holt (2000) can continually expand the learner's zone of proximal development, thus providing more confidence and motivation to continue learning.

Within a social constructivism environment, the approach to learning requires instructors act as facilitators. Students construct meaning via engaging in experiences that provide context within the learning environment. Within this context, the facilitator/instructor provides learning scenarios wherein the student becomes actively engaged in the learning process. These scenarios create an environment and opportunities for students to make sense of the content (Rhodes & Bellamy, 1999) instead of simply memorizing factual content. In order for the instructor to develop a sense of what the learner has gained, it is important that dialogue be at the center of assessment process. Consequently, acting as a facilitator, the instructor engages the students in activities that promote learning new content. Furthermore, within these activities, assessments are performed that actively engage the learner, that use dialogue, and that use performance-based components. Additionally, these assessments are dependent upon the conceptions of assessment held by the assessor. Thus, at the university level it is then the implementation of assessments that is driven by the faculty members' conceptions of assessment.

## Assessment in Higher Education in the United States

Institutions of higher education in the United States have an obligation to provide instruction, research, serve their communities and regions, observe ethical standards, provide a safe environment for students and employees, and comply with all federal and state health, safety, and employment regulations (National Commission on Accountability in Higher Education, 2005). The report indicates several concerns regarding the state of higher education within the United States. Specifically, the United States is no longer the leader in the world with respect to college completion rates. Also, the United States lags behind other countries in its ability to educate scientists and engineers in order to compete in the global economy. Furthermore, the number of minority students enrolling in college is rising in the United States; however, many of these same students do not graduate. Finally, the costs of higher education have consistently grown faster than the consumer price index; but, financial support (through grants) is lagging behind enrollment demand and inflation (National Commission on Accountability in Higher Education, 2005). It is noted, however, that there is not a problem with either the amount, or absence, of accountability. Clearly, universities are accountable to many stakeholders including but not limited to its student body, trustees, private financial supports, accreditors, and the government. However, it is the case that what matters most in improving student performance is the devotion of resources and the significance and clarity of assessment goals because these factors most directly impact faculty members and students (National Commission on Accountability in Higher Education, 2005).

The report of the Greater Expectations National Panel (Association of American Colleges and Universities [AAC&U], 2008) indicates that institutions of higher education must hold students to high standards when it comes to its outcomes. The report contains recommendations concerning the knowledge and skills that should be acquired by higher education students. A partial listing of the skills and knowledge recommended in the report includes that institutions at a minimum require students to develop strong analytical, communication, quantitative, and information skills; an understanding of, experience in, and inquiry into discipline-based knowledge about science, culture, and society; intercultural knowledge and collaborative problem-solving skills; responsibility for individual, civic, and social choices; and integrative ways of thinking and applying knowledge and skills in new settings (National Commission on Accountability in Higher Education, 2005, p. 25). The commission further indicates that providing the environment is not sufficient, assessing student performance on these outcomes is also important to ensuring increased learning.

There is a relatively large degree of autonomy given to institutions of higher education and to faculty who are responsible for establishing curriculum and program requirements. The varying institutional characteristics and the nature of the students who attend colleges and universities are all central components of the higher education system in the U.S. today. Consequently, the complexity associated with such diversity, provides a difficult challenge for stakeholders involved in developing or affecting educational reform agendas. The difficulty lies in the fact that in a complex, decentralized higher education system, there is no comprehensive strategy to provide effective public information including better data about real performance and learning (U.S. Department

of Education, 2006). “There is inadequate transparency and accountability for measuring institutional performance, which is more and more necessary to maintaining public trust in higher education” (U.S. Department of Education, 2006, p. 13).

### *Educational Standards Movement*

Educational reform is an ongoing topic among policymakers. Over the years, the emphasis of national and regional accrediting standards has shifted from inputs to processes to outcomes. Indeed, a noticeable shift from inputs to outcomes has occurred over the past two decades as can be seen in the requirements of both national and regional accrediting agencies. Due to this shift, institutions began to demonstrate that their faculty were not only qualified, but that they also used research-based best practices as well as demonstrated that their students attained the course learning outcomes. Thus, national standards reform efforts focus on evaluation and accountability of institutions of higher education with a current emphasis on outputs or outcomes of education rather than the inputs. This shift is evident in education reform from the late 1990s into the early 2000s wherein the focus turned to student learning outcomes. Specifically, Goals 2000, a key education initiative of the Clinton administration, encouraged states to develop content and performance standards that were demanding, shifting the focus to outcomes of education.

As mentioned above, historically speaking, institutions of higher education simply had to provide evidence that their faculty members were qualified and that courses were taught using research-based methodologies and strategies informed by best practices. Although faculty qualifications, teaching strategies, and methodology remain a central component of accrediting agencies’ requirements, these have each simply become a point

of compliance. It is clear, however, that the trend in national and regional standards is to assume that faculty members are qualified, methodologies are research-based, and clinical practices, where appropriate, are completed; therefore, these items have become a simple issue of compliance and no longer the primary barometer of an institution's worthiness of first-time or continued accreditation. Instead, the national and regional accrediting standards now emphasize the process by which students are educated and the outcomes they achieve. To sum up the current state of accreditation, it is the case that the emphasis of national and regional accrediting agencies standards is on students' experiences (processes) and their demonstration of competencies (outcomes) as they transition through programs instead of what is taught by instructors (inputs) (National Council of Accreditation of Teacher Education [NCATE], 2007; Southern Association of Colleges and Schools [SACS], 2012).

Educational policymakers consistently focus discussions and legislative mandates on institutional effectiveness defined as "the systematic, explicit, and documented process of measuring performance against mission in all aspects of an institution" (SACS, 2005). Typically, these policies are meant to be used as a way to encourage institutional accountability. Indeed, one measure often chosen by policymakers to determine the effectiveness of an institution is how well its students perform on various assessments. Thus, as the educational landscape continues to shift from inputs to outcomes, the need for improved performance on assessments becomes more evident. Consequently, high-stakes assessment results seem to have become the key measure of the outcomes in today's educational climate.

In summary, it is clear from the literature that standards-based educational reform has had a tremendous impact on university programs across the nation. Consequently, with the growing emphasis on accountability, it is as important as ever to ensure that graduates not only know the content but also are able to effectively perform in jobs post-graduation. Thus, assessment of students performed and designed by faculty and informed by administrators is an important component in the shifting accountability landscape. As stated above, practices and behaviors of individuals are influenced by their perceptions, past experiences, and beliefs (Bandura, 1986; Dewey, 1933), thus given that conceptions of assessment may vary depending on the belief system held by individuals, it is necessary to be able to quantify these various conceptions in order to effectively meet accountability mandates.

#### *K-12 and Higher Education Standards-Based Curriculum Reform*

Curriculum standards have become a formidable force affecting the reformation of education at all levels since the publication of the document *A Nation at Risk* (National Commission on Excellence in Education, [NCEE], 1983) and even more recently with the adoption of the Common Core Standards Initiative. These curriculum standards were written with an emphasis on experiences (processes) and outputs (outcomes). Using these curriculum standards as a major measure, educators began to adopt and use research-based teaching methods instigating a trend toward a hands-on constructivist approach to student learning, wherein the learning process is emphasized. This emphasis has had a major impact on current university teacher education students because they are being educated to become facilitators of the learning process rather than transmitters of knowledge.



Although many educators agree that standards-based education is the central driving force in educational reform today, there remains much debate on the meaning of standards-based education. For the most part, educators agree that content standards are subject-matter descriptions of what students should know or be able to do within specific grade levels; however, these are often confused with performance standards which are typically interpreted as expected performance on a test. Although policymakers emphasize systematic reform in both K-12 and higher education, it is unclear how standards-based reforms are expected to work (Anderson et al., 2005). To further complicate matters, policymakers and other stakeholders are calling for higher achievement and effectiveness at all levels, both in teaching and learning; however, it is unclear how to best assess these. Consequently, the reality of the implementation of standards-based education has resulted in a familiar policy of test-based accountability (Hauser & Koenig, 2011).

State and federal policymakers implement educational reform hoping to improve students' academic achievement (Schiller & Muller, 2003). Thus, institutions of higher education have a difficult challenge facing them - in an era of accountability, institutions of higher education have the added responsibility of ensuring their graduates are prepared to become effective professionals who will make a positive impact in their field. With the transition from inputs to outcomes, emphasis on effectiveness and added demands of accountability, faculty members of institutions of higher education today must not only be prepared to teach and lead their students, but also be prepared to be held accountable for the student learning outcomes resulting from their teaching and leadership methods.

## Accountability in Education

Accountability is not a new concept within the United States. The root of accountability movement can be traced back to the launch of Sputnik in 1957. In 1958, almost a year after the Soviets launched this satellite; the National Defense Education Act (NDEA) was passed under President Eisenhower's administration. This act provided increased federal funding for education, especially in the areas of mathematics and science. Barely seven years later, Congress passed another piece of legislation that included increased funding as well as requirements of accountability. This new legislation, titled the Elementary and Secondary Education Act (ESEA), was passed by Congress in 1965 under President Johnson's administration and was deemed to be the first shot in what has been termed the *war on poverty* as outlined by President Johnson in his January, 1964 State of The Union Address.

The Elementary and Secondary Education Act [ESEA] of 1965, the largest federal funding allocation to date, was primarily influenced by the disparity in educational opportunities and in student performance. After the passing of such a federal mandate, legislators demanded accountability for the funds distributed to various programs. As a result there was an enormous increase in the number of norm-referenced tests developed and published to determine student learning as a means to make the educational system accountable (Linn, 2005). In addition to the increased use of standardized tests, Linn (2000) reports that the Normal Curve Equivalent (NCE) was developed in order to standardize scores to coincide with National Percentile ranks at three points (1, 50, and 99). Furthermore, during the 1970's several state legislative bodies enacted minimum-competency testing requirements that were intended to ensure state-mandated standards

were being met. Indeed, the genesis of these minimum-competency tests was shortly after the initial implementations of norm-referenced standardized testing in the early 1970s.

In 1983, standardized testing in K-12 settings became the emphasis on the national forefront due to the publication of *A Nation at Risk*. Indeed, by the mid-1980s, every state in the nation had imposed some form of legislative mandate that required accountability measures in K-12 education. In the late 1980s and the early 1990s, various stakeholders, including policymakers and educators, used these standardized test results for accountability purposes. While taking on a variety of forms, they all shared in the aspect that there were increased stakes for educational administrators, instructors, and students in all K-16 educational settings. Given the increased pressures associated with using high-stakes testing for accountability purposes, educators consented to these pressures and began to explore and use a variety of means to improve the measures used to judge student performance (Meyer, 1996).

Such pressures make it even more necessary to work toward an accountability system that can be used in the complex and diverse system of higher education found in the United States. Indeed, according to the National Commission on Accountability in Higher Education (2005), accountability measures and approaches must be developed and used in order to improve student performance.

According to Louis, Febey, and Schroeder (2005), a gap exists between policy and practice in many institutions of higher education across the United States. This gap can be described as one of the major stumbling blocks to accountability reform. In fact,

the report of the National Commission on Accountability in Higher Education (2005) reports the following:

Too often accountability is the battleground between educators and policymakers. Many educators believe externally imposed accountability is a tool to place blame or avoid responsibility for inadequate financial support. Many policymakers, frustrated because existing investments are not producing better results, believe stronger external accountability is the only way to get improvement. In an atmosphere of resentment and mistrust, accountability initiatives produce more resistance than progress. (p. 11)

Clearly, such reports make a strong case that a better system of accountability is needed in the United States. In 2005, the National Commission on Accountability in Higher Education reported their findings in *Accountability for Better Results – A National Imperative for Higher Education*. One of the findings of this report indicated that a new approach to accountability is necessary. Specifically, if accountability were transparent and led by collaboration rather than intimidation and fear, performance might improve (National Commission on Accountability in Higher Education, 2005). Furthermore, according to the report, this new approach to accountability should take into consideration the diversity and complexity of the decentralized system of higher education that is found in the United States. Additionally, due to the autonomous status of U.S. institutions of higher education, this new system of accountability has the potential to establish individualized conditions for each institution related to its specific mission and should publicize an institution's costs, availability of coursework, graduation rates, and the assessment results of student learning outcomes (National Commission on

Accountability in Higher Education, 2005). Finally, according to the report, all stakeholders, including governors, state and federal legislators, state boards, trustees, accrediting agencies, higher education administrators, faculty, and students have a responsibility as well as an essential part in this transition to a new form of accountability.

Information published by the National Center for Higher Education Management Systems [NCHEMS] (2008), indicates that for every 100 ninth-graders, it is estimated that 70 of them will graduate from high school on time, 44 of them will immediately enroll in college, 30 will still be enrolled during their sophomore year, and 21 of them will graduate from college on time. Estimates such as these, point to an increased need of a system of accountability that meets the criteria suggested above by the National Commission on Accountability in Higher Education (2005). Indeed, the implementation of such a system of accountability measures might increase student performance and concurrently increase the on-time graduation rate.

Differences exist in how and what accountability measures should be established. For example, Louis et al. (2005) conclude that administrators at all levels in one school setting believe that legislators have no right to implement policies affecting accountability, but that educators do have such a right because they have the understanding of teaching and learning. Louis et al. (2005) report that teachers are angry and afraid due to the underhandedness of policy-makers instituting policies affecting education without the presence of the educators. Louis et al. (2005) indicates that teachers believe there is a disregard of their professional expertise. These same teachers,

however, believe that they can use their knowledge and expertise in order to retain their own interpretations of what is best for the organization.

In order to comply with a national push for transparent accountability and assessment outcomes, consensus concerning the goals of assessment and accountability of an institution need clarity (National Commission on Accountability in Higher Education, 2005). Clearly, one of the primary goals of assessment in higher education is to positively impact student learning. But, in order to positively impact student learning, stakeholders must have a shared belief in the attainment of this goal (Brown, 2004-2005; O'Donovan, Price, & Rust, 2008; Tynjala, 1997). Consequently, in order for administrators, faculty members, and students to meet accountability goals such as positively impacting student learning, individual stakeholder's beliefs about assessment must be acknowledged (Brown, 2011). Furthermore, interpretation of the purpose of assessment becomes problematic when there is a disconnect between the initiatives of policy makers and the conceptions held by faculty members and students of higher education. As indicated by Brown (2011), simply implementing assessments does not necessarily improve student learning. Thus, coming to an understanding of the similarities and differences of conceptions held by these three groups may make it possible to establish an accountability system based on the criteria suggested above by the National Commission on Accountability in Higher Education (2005).

Given that belief systems influence behavior and motivation, it may be necessary to consider how social expectations of peers and cultural transmission may be effective ways to transform beliefs. Specifically, Pajares (1992) indicates that beliefs are only changed when they are no longer satisfactory to an individual and they become

unsatisfactory when they are challenged to the point where they cannot be assimilated. Changes in beliefs are possible when supported by the culture of an organization. Thus, administrators inherit a responsibility, because of their distinct roles and their institutional knowledge, to respond to policy initiatives in such a way as to take into account the current beliefs held by both faculty and students (Marks & Printy, 2003; Spillane, Diamond, Burch, Hallett, Jita, & Zoltners, 2002). Allowing time for faculty members to discuss implications of new policy initiatives increases the likelihood of a change in beliefs of these faculty members. Furthermore, a policy's message can be thought of as an external representation that demonstrates that problems have been targeted, delineated, and well defined. It thus can be argued that the clarity in the policy's message has an impact on the conceptualizations of those whom are affected by the policy (Waite, Boone, & McGhee, 2001). Indeed, inconsistencies and changes in state policies create a sense of anxiety about accountability policies (Louis et al., 2005). Thus, organizational learning fostered by transparent discussions may be a critical component to changing current beliefs about assessment (Marks & Printy, 2003; Spillane et al., 2002).

The importance of meeting the criteria suggested above becomes evident when in an era of increased accountability, policymakers often use assessment data to determine overall institutional effectiveness as well as specific student learning outcomes. It is the case that accountability measures exist for faculty members and students of higher education and that there are often high-stakes decisions made as a result of these accountability measures. For example, faculty members are expected to demonstrate strong teaching skills and the capacity to publish. Decisions about students' progression through a program and ultimately graduation are made based on individual student

performance. Sometimes, however, these high-stakes decisions are contradictory to the intended outcomes of initial accountability mandates and accompanying assessments. Linn (2005) concludes that the use of high-stakes accountability often confounds the intended positive effects of accountability.

### *High-Stakes Assessment in K-12 and Higher Education*

Over the past 50 years, the accountability movement has spawned an ever-growing focus on assessment. James B. Conant, past President of Harvard University who was known for his egalitarian vision of education, provided a rationale for differentiated instructional programs based on a common core in the 1950s. Linn (2000) reports, “Tests were seen as important tools to support the implementation of Conant’s conceptualization of the educational system, both for purposes of selecting students for higher education and for identifying students for gifted programs within comprehensive high schools” (p. 5). By 1986, all 50 states were under some form of a legislative mandate requiring assessment of K-12 students (Wolf, 1990). Although regional accrediting agencies of institutions of higher education were formed in the 1880s with a focus on educational standards and admissions procedures, this focus began to shift in the 1980s. Beginning in the 1980s and continuing into the present, regional accrediting agencies began emphasizing the outcomes of education (El-Khawas, 2001).

Although assessment is a term often used to refer to the data collected by specific tools and measures that are used to meet external accountability requirements established by various accrediting agencies for both K-12 and higher education, Wolf (1990) contends that assessment consists of a composition of both quantitative and qualitative data as well as the interpretation of those data. Dwyer, Millet, and Payne (2006) also



recommend that assessment be comprehensive with an iterative cycle of measuring progress at multiple points in time. The National Council of Teachers of Mathematics [NCTM] (1995) added to the dialogue of recommended uses of assessment by stating that the primary purpose of any student assessment should be to improve student learning. Furthermore, NCTM (1995) contends that there are several secondary purposes for assessment. Specifically, these secondary purposes include: to provide individual feedback to students about their learning; to provide information to the instructor about how well the class understands a particular topic, what additional activities might need to be introduced, or whether it is time to move on to another topic; to provide diagnostic information to instructors about individual students' understanding or difficulties with understanding new material; to provide information to teachers about students' perceptions and reactions to the class, the material, the subject matter, or particular activities; to provide an overall indicator of students' success in achieving course goals; and to help students determine their overall strengths and weaknesses in learning the course material. Fletcher et al. (2011) indicates that assessment in higher education is part of institutional quality and accountability processes in addition to measuring student learning.

In order to meet the demands of accountability, policymakers have focused on the use of high-stakes testing and assessment for the past 25 years. Linn (2000) purports that tests and assessments are relatively inexpensive compared to other reforms, can be externally mandated, can be rapidly implemented, and the results are fairly immediate and visible. As such, state-wide standardized assessments are often used in both K-12 settings as well as higher education as a means to determine grade and program

progression as well as admittance and graduation. As a result, tests and assessments foci remain at the forefront of the educational landscape. Mandates such as No Child Left Behind [NCLB] (2002), ESEA (1965), and the Higher Education Act [HEA] (1965) often require K -12 school systems and institutions of higher education to provide evidence of meeting the policy requirements by providing various data and assessment results.

Using data to understand problems and implement effective curricular changes is prevalent in higher education due in part to external accreditation standards (Association to Advance Collegiate Schools of Business [AACSB], 2012; NCATE, 2007; SACS, 2012), policy mandates of the state board rules, and the federal government's ESEA (1965) and HEA (1965). However, federal and state policymakers, personnel at institutions of higher education, and district leaders in K-12 school districts may have their own guidelines for what is meant by the term and what is to be included in assessment (Wolf, 1990).

One dilemma faced by stakeholders is the fact that the term assessment is often used within different contexts and with different meanings (Garfield, 1994). Anderson et al. (2005) emphasize the need for assessment to focus on outcomes in a global sense. Dwyer et al. (2006) however, recommend that assessment should focus primarily on student learning and the academically-related activities that students engage in during their tenure at an academic institution, specifically analyzing a graduates' workforce and general education skills, domain-specific knowledge, soft skills such as teamwork and creativity, and student engagement. However, faculty members' attitudes toward and expertise in assessment impacts the way they implement their own assessments in higher education. Additionally, students' attitudes toward and their experiences with assessment

affect their personal approach to learning and their beliefs toward future successes as a learner (Fletcher et al., 2011). Hinchey (2010) states that the development and implementation of an assessment system relies heavily on a clear purpose of assessment and that all stakeholders be included in the design of any assessment system.

Nonetheless, assessment serves many purposes for faculty members and students of institutions of higher education. It should be noted again that assessment practices have evolved as a result of the demands of external stakeholders related to the accountability movement of the past two decades. As mentioned earlier, measuring performance can be a difficult task in a decentralized system of higher education. There are enormous arrays of instruments that focus on the outcomes of higher education; however, due to the complexity of many of these measures, it is difficult to decipher how and for what purposes to use the data. The National Commission on Accountability in Higher Education (2005) indicates that public data systems cannot reliably answer some basic questions such as: How many students who enter higher education ultimately complete one or more degrees or certificates? What is the pattern of student persistence in higher education? On average, how long does it take students to reach different levels of attainment? What happens when students transfer? Do these students tend to encounter delays or additional costs in getting a degree? Can the transfer process be improved? Does it take students longer to accomplish their educational goals if they do not receive sufficient financial aid? Are student aid resources adequate to support low and moderate-income students? How much student aid comes from different funding sources? What is the actual net price of attending a college or university after grants and loans are taken into account? How fast is the net price increasing? Are students learning

what they need to know to be successful in life and work? Institutions as well as policymakers have a responsibility to establish clear goals in order to answer these questions and assessment data must be collected and reported in a way that allows enough evidence for valid decision making regarding these goals.

More is not always better in terms of assessment data. Actually, the contrary is often the case. Institutions of higher education are asked to collect a plethora of assessment data at many levels ranging from individual student, course, program, department, college, institution, performance, achievement, completion rates, enrollment, financial, and satisfaction. Collecting too much assessment data for accountability purposes actually limits the usefulness of these data (National Commission on Accountability in Higher Education, 2005) because it makes it difficult to decipher actual results about faculty and student performance.

Student learning is frequently the central core of the assessment movement in higher education. Faculty members assess learning in the context of course-specific student learning outcomes. However, as noted above, these assessments are heavily influenced by the conceptions of assessment held by these faculty members which affects students' approach to learning. To further exacerbate the problem of lack of clarity, the goals and outcomes related to the general education curriculum are vague. "More explicit instructional goals and disciplined, transparent learning assessment will likely enhance student learning, institutional practice, and public confidence" (National Commission on Accountability in Higher Education, 2005, p. 24). Additionally, Dwyer et al. (2006) distinguish four classes of assessment that have utility for one or more purposes: competencies of students at admissions, performance of students as they

progress through their degree programs, competencies of students at the completion of their program, and the value-added to their education based on their experiences in college. The result of this lack of clarity is an accountability system that lacks specificity of measurable goals that can be assessed.

There is a push to use more performance-based measures of assessment than the traditional standardized tests used to determine admissions decisions. Using performance-based measures of assessment allows both educators and policymakers to follow national trends on learning as opposed to performance on one standardized test. This strategy incorporates various methods of measuring performance which provides all students, including diverse students, to perform at high levels, with matched objectives to their specified learning needs (Linn, 2000).

According to the National Commission on Accountability in Higher Education (2005), five states are participating in a project that uses portfolios (incorporating a variety of external assessment measures) to determine levels of knowledge and skills acquired by college students. Institutions should be cautioned, however, that these particular external assessments should not be used to determine institutional effectiveness as these may not be valid measures of the preparation provided by an institution as students often attend a variety of institutions throughout their collegiate careers and there are widely varying standards for assessing among various institutions (Baker, O'Neil, & Linn, 1993). Thus, there would be no way to determine specifically where and when the learning occurred; but rather, only what students know and are able to do at a particular point in time. Consequently, the effectiveness of the impact of institutions of higher

education on student learning should be determined using internal measures of student performance.

### *Faculty Accountability and Assessment*

There is an ongoing tension between faculty and administrators in higher education related to accountability and assessment (Welsh & Metcalf, 2003). Often, faculty members consider accountability and assessment an administrative task rather than a professional responsibility. Administrators often feel that it is the responsibility of each faculty member to be accountable for their own and their students learning. This notion is further emphasized when the term accountability is confused with the term assessment. Faculty members resist the use of assessment measures and assessment data as a result of a fear of accountability (Welsh & Metcalf, 2003). Indeed, there is a fairly common notion held by faculty members that if they collect and report assessment data indicating their students' performance, then there will be consequences if these results do not appear satisfactory (Schilling & Schilling, 1998). On the other hand, if the assessment results seem inflated, then there will somehow be the inclination to believe the faculty member is too easy on students.

In order to determine if real learning occurs within each classroom, it is important to clearly define accountability and assessment-related common language and terms at each institution. A clear, concise definition for accountability can help validate assessment results. As such, stronger evidence of validity can lead to appropriate decisions being made based on the assessment results. This transparency within the process in turn can reduce the amount of resistance exhibited by faculty members (Schilling & Schilling, 1998). "People achieve excellence because they want to, not

because they have to. . . accountability for better results is different from accountability for minimum standards” (National Commission on Accountability in Higher Education, 2005, p. 11). The National Commission on Accountability in Higher Education (2005) also indicates that in order to improve results, a shared vision for success, how to define success, and how to measure success must be evident.

The success of an institution is often judged by a reputation that is based on alumni and current student recommendations and faculty prestige, instead of learning, student achievement, quality scholarship and service (National Commission on Accountability in Higher Education, 2005). According to the National Commission on Accountability in Higher Education (2005), an intensive focused dialogue on the importance of accountability priorities at the national, state, and institutional levels, can move accountability from a political perspective focusing on compliance issues to a meaningful, effective movement emphasizing strategies for improvement using reliable measures to make valid decisions.

Several research studies have provided results that indicate a need for changes in perspective concerning the value and types of assessments. For example, a research study conducted in China (Chen et al., 2012) indicates that educators tend to use traditional forms of teaching and also traditional assessments including school examinations. However, in order to guide real transformation of educational reforms toward a more learner-centered approach to teaching, learning, and assessment, educators must realize the need for a new approach and challenge their traditional views of teaching and learning (Chen et al., 2012). Furthermore, Liu (1995) claims that changing the conceptions of teaching and learning is a key factor to improving the quality of students’

learning and in order to make this transition, administrators of higher education have the responsibility of changing teachers' conceptions of teaching and learning. Additionally, Prosser, Trigwell, and Taylor (1994) report that university faculty members conceptualize learning as: (a) accumulating more information to satisfy external demands, (b) acquiring concepts to satisfy external demands, (c) acquiring concepts to satisfy internal demands, (d) conceptual development to satisfy internal demands, (e) conceptual change to satisfy internal demands. Thus, it should not be surprising that Fletcher et al. (2011) contend if students' understandings of learning as well as their performance on assessments are in fact influenced by faculty members' assessment practices, then it is critical to determine the conceptions held by faculty members since these conceptions can have a significant impact on the intended educational outcomes of their students.

#### *Student Accountability and Assessment*

Students of higher education are at the center of the accountability and assessment movement. How students understand assessment and their attitudes toward assessment might contribute significantly to learning behavior and academic achievement (Brown & Hirschfeld, 2008). Improved student preparation in K-12 settings is imperative in order to improve learning. Fostering a smooth transition from K-12 to higher education requires collaboration between the two environments (Smith & Zhang, 2009).

There is much research about how students view conceptions of learning (Boulton-Lewis, 1994; Brown, n.d.; Shepard, 2000). Marton, Dall'Alba, and Beaty (1993) report that learning is conceptualized as: (a) increasing ones knowledge, (b) memorizing and reproducing, (c) applying, (d) understanding, (e) seeing something in a different way and (f) changing as a person. In the United States, K-12 students have been



experiencing various mandated assessments for years. Having had these experiences has certainly shaped their views and beliefs of learning as well as about how their learning is assessed.

There is a relatively small amount of research emphasizing students' beliefs about assessment (Brown & Harris, 2012). It is known, however, that Australian students became increasingly negative toward literacy assessment in their first-year of high school based on a study by Moni, van Kraayenoord, and Baker (2002). Moni et al. (2002) imply that this shift in students' attitudes was the result of their increasing awareness in the "volume and difficulty of assessment, alongside perceptions that teacher assessment decisions were subjective" (Brown & Harris, 2012, p. 47). Furthermore, Brown and Harris (2012) indicate that students' negativity may be the result of students becoming more aware of the high-stakes decisions made by the results of assessment. Specifically, there are personal consequences of assessments for students.

Students' beliefs concerning assessment may vary depending on their current level of schooling (Brown & Harris, 2012). Specifically, Brown and Harris (2012) report that students' attitude toward assessment become more negative as they progress through school, become more aware of the personal consequences, and as they become more experienced with high-stakes assessments. Brown and Harris (2012) report that a plausible explanation for these changes in attitudes toward assessment, although more negative, might be linked to an increase in an appreciation of the importance of assessment due to the fact they see the need for it (Brookhart & Bronowicz, 2003).

*Making the Connection: Policy and Practice*

There is a lack of compatibility between K-12 and higher education policies, practices and priorities. Dwyer et al. (2006) indicate that a national “systemic, data-driven, comprehensive approach to understanding the quality of two-year and four-year postsecondary education, with direct, valid and reliable measures of student learning” is needed based on the assessment driven educational landscape of today (p. 1). The recommendations include the need for clarity, simplicity, and common language to be used within higher education as well as by its stakeholders. In order for policymakers, legislators, or educators to develop a clear, simple system of assessment that includes a common language, a common purpose of assessment must be developed. Brown (2004) indicates that in order to develop this common purpose of assessment, a better understanding of the conceptions of assessment held by various stakeholders affected by such mandates and policies, must be obtained and analyzed. Brown (2004) provides evidence of four conceptions of assessment that include assessment as a means to improve teacher instruction and student learning by providing quality information for decision-making, assessment as a way to make students accountable for their learning, that teachers or schools are made accountable through the use of assessment, and that assessment is irrelevant altogether to the work of teachers and to the life of students.

Educational policymakers in the United States over the past several decades have implemented many federal and state mandates requiring more accountability.

Assessments are often used to make high-stakes decisions regarding institutional accountability (AACSB, 2012; HEA, 1965; NCATE, 2007; SACS, 2012). Although the accountability and assessment movement has been at the forefront of the educational

reform landscape, there has been little movement from educators toward embracing the change (Alexander, 2000). Brown (2004) contends that in order to meet the external accountability demands, policymakers must attend to teachers' conceptions as much as they deal with declarative or procedural knowledge requirements. Specifically, in order to have success, the complexity of these conceptions must be taken into account when implementing any new assessment policy (Brown, 2004). Furthermore, Brown (2004) argues that simply introducing or mandating a new assessment policy will not necessarily achieve the intended policy objectives unless these policymakers take into consideration teachers' conceptions of assessment and the alignment to the new assessment policy. Learning is a complex set of interactions between students and skilled educators (Brown, 2011) and effective administrators foster an environment that promotes hands-on activities, engagement of students, and communication with students. If faculty and students conceptions of assessment do not align to the implemented mandate, then administrators have the responsibility of facilitating the process of working toward a consensus, if they want the new policy to be effective (Brown, 2011). If consensus is not obtained, there is the risk of unintended outcomes which can ultimately lead to decreased student learning.

Brown et al. (2011) propose that differences in policy and cultures lead to differences in how assessment is conceived by various stakeholders. They hypothesize that cultures that emphasize low-stakes decisions associated with assessments, such as New Zealand, compared to cultures that emphasize high-stakes decisions associated with assessments, such as the United States, will have a difference in the way instructors and students conceptualize assessment. Specifically, Brown et al. (2011) imply that it is

conceivable for teachers and students within cultures that have low-stakes associated with assessment would conceptualize the purpose of assessment for improvement reasons and that teachers and students within cultures of high-stakes associated with assessment would conceptualize the purpose of assessment for student accountability reasons.

### Teaching, Learning, and Assessment

Assessment is one of the most important things that educators can do to help students learn (Brown 2004-2005). Furthermore, assessment drives student learning (Ramsden, 2003); however, students are often confused about what is asked of them regarding assessment. Rust, O'Donovan, and Price (2005) argue for a social constructivist approach to assessment explaining that when students are actively engaged with every step of the teaching, learning, and assessment process, they develop a deeper understanding of their course content and may subsequently produce improved work. Indeed, if students are engaged in the entire learning process, learning will be positively impacted (Brown, n.d.). The various aspects of the learning process should take into account pertinent dialogue concerning the development and use of assessments. Through this engagement, students will be allowed to construct their own meanings and thereby providing a more effective understanding of the subject matter at hand.

It is clear from past research (Brown & Hirschfeld, 2008; Shepard, 2000) that a more effective understanding of a concept or process induces deeper learning and personal ownership of learning. As was stated earlier, when student learning and assessment are viewed as inseparable components of a dynamic process, the ongoing development and continual improvement of instructional practices will naturally occur (Holt & Willard-Holt, 2000). Brown (2011) however, argues that teachers often make

instructional decisions apart from assessment data which diminishes the intended uses of assessment as a diagnostic or formative tool to improve teaching and learning – not to meet the demands of external pressures of accountability. Furthermore, Peck, Gallucci, and Sloan (2010) report that external policy mandates often accompanied by negative rhetoric, may undermine the motivation qualities required to successfully implement any external policy. Deepening the rhetoric is the notion that educators are simply passive deliverers of curriculum and instructional policies that are externally created without acknowledging the importance of contextual aspects that affect the implementation of any new policy (Peck et al., 2010).

Assessment literature lacks a specific focus which is congruent with the notion that there are many purposes for assessment in education. Ajzen's (1991) theory of planned behavior suggests that peoples' beliefs are one of the predictors of behaviors, undergirds much of the research related to assessment practices concerned with teaching and learning. Brown (2004) promotes the idea that personal beliefs about assessment affect conceptions of assessment and that these conceptions impact how teachers teach and how students approach learning. Furthermore, Fletcher et al. (2011), using Brown's (2004, 2006) past research and Conceptions of Assessment inventory, explored how these conceptions of assessment differed among university faculty and students. Chen et al. (2012) further supports the research by indicating that in order to change the way teachers teach, educators must first explore the underlying beliefs of those teachers.

Brown (2004) began investigating various conceptions of assessment held by teachers and students in K-12 settings. He indicates that all pedagogical acts of teachers are affected by their beliefs and conceptions about teaching, learning, assessment,

curriculum, and teacher efficacy. As such, Brown (2004) concludes that although past research relied heavily on the notion that there are three primary conceptions of assessment (Torrance & Pryor, 1998; Warren & Nisbet, 1999; Webb, 1992), a fourth conception should be added to the model. Specifically, Brown (2004) indicates that assessment may be viewed as completely irrelevant to the life and work of teachers and students.

Fletcher et al. (2011) indicate that there are many purposes of assessment in higher education and that faculty and students' attitudes toward assessment differ in that faculty view assessment as an aid to the teaching and learning process whereas students view assessment as needed for accountability but irrelevant to the teaching and learning process. They conclude that the modified version of the abridged *Conceptions of Assessment III* (Brown, 2006) needs further investigation in higher education in order to determine that this instrument is a valid means for collecting information regarding the purpose of assessment in institutions of higher education. Furthermore, Brown and Hirschfeld (2008) purport that conceptions are contextual and that conceptions of assessment may differ in high-stakes environments versus low-stakes environments.

Deneen, Brown, Lam, and Tsui (2012) examined student knowledge of course-specific assessments and the importance of including students in the process of determining assessment processes relevant to practices. Specifically, Deneen et al. (2012) conclude that students feel that there should be greater time spent on practice and implementation of local curriculum and assessment theories as opposed to international comparisons. Furthermore, they conclude that students are often excluded from the assessment process when faculty members and administrators continually react to

legislative mandates. This omission can keep faculty members and administrators from obtaining curriculum-related insights held by students that can enhance the assessment of what they are actually learning.

In order to make changes in the way teachers teach, Chen et al. (2012) concludes that it is first necessary to ascertain how teaching practices are viewed by teachers. Underlying Chen's et al. (2012) intention for this study is the theory of planned behavior. Specifically, "what people believe, the amount of control they have or perceive they have, societal norms, and people's intentions interact to shape the behaviors and practices people carry out" (p. 3); therefore, in order to make change, understanding the beliefs held by individuals' provide a foundation for the starting place to implement the change. Fullen (1993) adds that it is only when individuals engage in actions to alter their own situations that results in a chance for change implying that each person can and should be acting as a change agent in order for real change to occur.

Collective sense making is an important perspective to individual responses to externally mandated accountability and assessment models. Louis et al. (2005) report evidence to support the need to understand a policy's effects on the implementation of the policy. Teachers from three schools with diverse perspectives were studied and it was determined that high school teachers adapted their instruction as a result of many factors including shared professional experiences with other teachers, cultural assumptions about teaching and learning, collective norms and values from their individual schools, and mediating behaviors of administrators in their schools and from their districts (Louis et al., 2005).

Henkel (1997) suggests that a struggle related to public policy that started in the 1980s continues through the 1990s between the government and the universities in England, Sweden, and Norway. Faculty members and administrators indicate that a shift of power from the institution to centralized managers has occurred and there is ongoing momentum to centralize regulatory authority of academic work. Representatives of higher education in Europe question whether or not institutional autonomy can be maintained by centralizing authority. Rust (2002) reports an espoused rhetoric by members of higher education focuses not on teaching; but rather, on learning at institutions of higher education in the United States. The United Kingdom universities, however, have placed a greater emphasis on the development of life skills and lifelong-learning.

Conversely, Rust (2002) reports that in the United States, the assessment of student learning is the same as it was in the 1940s, 1960s, and 1980s with little change in measuring what students know and can do. In order to shift toward an approach to assessment that focuses on outcomes-based student learning, there is a need for explicit and transparent processes shared with students via a constructivist approach which involves the active engagement of students in the learning and assessment process (Rust, 2002). In recent years, the emphasis on quality is the center of performance measurement in institutions of higher education in Hong Kong as a result of external accountability mandates (Pounder, 2000). A study of seven institutions of higher education related to quality revealed that comparative inter-institutional performance evaluation is difficult at best and the challenge for higher education is to determine



concepts that complement quality and add precision to institutional performance assessment in higher education (Pounder, 2000).

Brown and Hirschfeld (2008) detail students' conceptions of assessment and their relationship to academic achievement and provide evidence of a four-conceptions of assessment model held by students: assessment makes schools accountable, assessment makes students accountable, assessment is ignored, and assessment is fun. The school accountability conception had an inverse relationship with student achievement, indicating that if assessment policies are presented to students as an external mechanism for accountability, then it is likely that achievement will go down. Instead, if assessment policies are presented as a measure of individual student learning (and students believe this), then achievement scores are more likely to go up. These results, however, were collected and reported from New Zealand, which has a low-stakes environment for assessment results, meaning that scores on tests have no impact on either students or schools. Brown and Hirschfeld (2008) indicate that additional investigation is needed to determine if the context of a high-stakes testing culture impacts the results of this type of study. Furthermore, these results indicate that students who conceptualize assessment in terms of personal accountability (not external accountability) achieve more. Some of the students who completed this study may eventually become educators (teachers or university instructors) and past experiences with assessment may impact these future educators' conceptions of assessment. Brown and Hirschfeld (2008) contend that these results align with self-regulation theory and that if taking responsibility of an individual's role in the learning and assessment process increases student learning, then educators

need to develop better mechanisms to help students transition into their roles as teachers/instructors, which can over time, improve achievement of all students.

A national study of external and internal influences on institutional approaches to student assessment by Peterson and Augustine (2000) explores state-level characteristics, institutional type, and accrediting association of 885 public institutions' purposes for conducting assessment as well as their approaches to student assessment. Results indicate that research universities are the least likely to conduct assessments, doctoral and master's institutions are most likely to stress that they conduct assessment to meet accrediting requirements, and associate of arts institutions are most likely to report that they conduct assessment both for internal purposes and to meet state requirements. Furthermore, the results indicate that external demands from state or accrediting agencies is related only minimally to collecting student assessment data. Institutions that report data collection for internal purposes of improving the teaching and learning process as opposed to data collection for external accrediting requirements are more likely to collect data about its students. Peterson and Augustine (2000) find that the approach that an institution takes to meet the demands of external accrediting agencies makes a substantial difference in how institutions approach student assessment. For example, institutions located in the North Central region are most likely to collect data on cognitive competencies whereas intuitions in the Southern region are most likely to gather data on student affective competencies. Institutions in the Middle States regions are most likely to collect data on former students whereas institutions in the Western region are least likely to collect data extensively. Finally, institutions in the New England region where student assessment requirements came later than to the other reported regions, scored

relatively low on approaches to collecting data on student cognitive competencies, student affective competencies, and former students.

Peterson and Augustine (2000) report that intuitions within the Southern region (accredited by SACS) are the most likely to report internal continuous improvement reasons for collecting assessment data whereas institutions in the Western, Northwest, and New England regions are least likely to report internal reasons for conducting assessments. There are no differences, however, by accrediting region in institutions reporting that meeting accreditation standards is an important purpose for collecting assessment data. Finally, when external mandates for assessment practices are implemented at institutions of higher education, stakeholders of various institutions are more responsive to these mandates if they are provided authority to develop their own indicators and outcomes which continues to foster institutional autonomy.

Meyer et al. (2010) indicate that institutions of higher education in New Zealand rely heavily on assessment data to make admissions decisions, student progression through programs and for degree completion, and to document that students master learning outcomes. Furthermore, faculty members rely on assessment data to provide information about teaching effectiveness and student achievement. On the other hand, students rely on feedback about their learning in order to determine what they must do in order to meet faculty expectations. Accrediting agency staff and government officials require that members of institutions of higher education provide documentation related to assessments to ensure quality of its programs and graduates. Meyer et al. (2010) provide the results of a mixed-methods project investigating whether and how attitudes toward and experiences with assessment held by academic staff and their students are

represented in official institutional assessment policy and policy guidelines. Results concluded that a dichotomy exists between assessment *of* learning and assessment *for* learning and that institutional staff generally reported more positive attitudes toward assessment than students. Meyer et al. (2010) report, “Students seemed suspicious of assessment to the extent that, surprisingly, most preferred external moderation for consistency even if grading on a curve resulted in overruling teaching staff’s judgments about mastery of learning outcomes” (p. 346). The study findings are worrisome because it is known that student achievement and student approaches to learning are affected by overall attitudes about assessment (Brown & Harris, 2012; Meyer et al., 2010).

Past experiences as well as cultural norms impact conceptions of assessment which affect future behaviors (Brown, n.d.; Brown, 2004; Brown et al., 2011; Pajares, 1992; Tynjala, 1997). Contextual factors may also impact the conceptions held by individuals. Brown (2008) emphasizes the notion that contextual factors with assessment greatly impact individual’s conceptions of assessment. Thus, educators and students having previous experiences in a high-stakes assessment environment might view the purpose of assessment differently than those having past experiences in a low-stakes assessment environment. Brown and Hirschfeld (2008) stress the importance of knowing the conceptions held by educators and students in order to improve student learning since these conceptions can impact student achievement. Furthermore, Fletcher et al. (2011) indicate that in higher education, the conceptions held by faculty members and students may differ from those of teachers and students in K-12 environments; therefore, there is a need to further investigate the differences in faculty members’ and students’ conceptions based on environmental and contextual factors.

## Conceptions of Assessment

Brown (2006) used the CoA-III to measure conceptions of students in New Zealand's and Queensland's K-12 educational setting. Based on his results, Brown (2006) suggests that four conceptions of assessment exist, three of which are described as purposes and one of which is described as a non-purpose:

1. Assessment is used for improving teaching and learning (Improvement);
2. Assessment is for making schools and teachers accountable (School Accountability);
3. Assessment is for making students accountable for their learning (Student Accountability); and
4. Assessment is irrelevant to the lives and work of teachers and students (Irrelevant).

Fletcher et al. (2011) used an abridged CoA-III to determine if Brown's (2006) model of the conceptions of assessment could generalize to higher education. Fletcher et al. (2011) tested Brown's (2006) model of conceptions of assessment to determine if the model fit the data collected from higher education faculty and students in New Zealand. Based on the data collected by Fletcher et al. (2011) in New Zealand, the researchers determined that Brown's (2006) model of conceptions of assessment did not fit the data well for higher education. Fletcher et al. (2011) proposed the possibility that within a low-stakes assessment culture of higher education in New Zealand, only two conceptions of assessment exist, positive and negative conceptions. The current study aims to determine if either of the models of the conceptions of assessment developed by Brown (2006) and Fletcher et al. (2011) fit the data collected from higher education faculty and

students within high-stakes assessment environment located within the southeastern region of the United States.

Formative assessments are often used within higher education settings in order to improve teaching and learning. Past research suggests that this use of assessment has a positive impact on learning and achievement (Popham, 2001; Popham, Cruse, Rankin, Sandifer, & Williams, 1985; Struyven et al., 2005).

The institutional accountability purpose of assessment is viewed as a means for institutions of higher education to prove that they are meeting the requirements set-forth by external stakeholders such as accreditors, legislators, policymakers, or the educational community as a whole. Often there are consequences imposed such as reduced or diminished funding or accreditation status revoked or suspended if these set mandates are not met by institutions of higher education.

The student accountability conception of assessment focuses on student accountability for their learning. Student accountability measures are often comprised of course and assignment grades as well as results of standardized tests which demonstrate to parents, potential employers, and others that the student met the requirements set forth by educational constituents. Often, there are high-stakes associated with consequences associated if students don't meet these requirements. Examples of these high-stakes are failing coursework, not progressing through a degree program, and possible the inability to graduate from a program of study.

The irrelevant conception of assessment is grounded in the view that external demands placed on schools, instructors, and students are inadequate, inaccurate, and even irrelevant (Brown, 2011). Research suggests that high-stakes accountability mandates

have negative consequences for both teachers and students (Darling-Hammond, 2003).

Linn (2000) suggests that high-stakes testing does not improve the quality of education.

## CHAPTER III

### METHODOLOGY

#### *Participants*

All undergraduate students and all full and part-time faculty members who teach at level V institutions of higher education with a minimum of one bachelor's degree located within the accreditation region of the Southern Association of Colleges and Schools (SACS) were asked to participate in this study. Level V doctoral-degree granting institutions are defined by SACS as institutions that offer three or fewer doctoral degrees as highest degrees. For the purposes of this study, faculty members were identified as university employees whose primary duty is classroom teaching, research, department chairpersons, academic deans, and program coordinators. Additionally, students were identified as undergraduate students attending one of the institutions within this region.

#### *Instrument*

A modified version of Brown's (2006) abridged *Conceptions of Assessment III* (CoA-III) instrument was originally used to determine faculty members' and undergraduate students' self-reported conceptions of assessment. The CoA instrument, developed by Brown (2004), was used to measure self-reported conceptions of assessment held by primary school teachers in the compulsory school sector in New Zealand and Queensland (Brown, 2006). Brown (2004) used a six-point positively packed rating scale (strongly disagree, mostly disagree, slightly agree, moderately agree, mostly agree, strongly agree) in order to elicit variability in responses. An abridged version of the instrument was developed after collecting additional evidence of validity



confirming reasonable psychometric properties (Brown, 2006). The current abridged version of the CoA, CoA-III, includes 27 items using a four factor measurement model: student accountability, school accountability, improvement, and irrelevant. Fletcher et al. (2011) modified the CoA-III using a six-point Likert scale ranging from strongly disagree to strongly agree for use in tertiary environments in New Zealand. Specifically, Fletcher et al. (2011) used the modified CoA-III in a survey of faculty who taught undergraduate students and undergraduate students attending tertiary institutions in New Zealand. The current study used a modified version of the CoA-III (Appendix A). Fletcher et al. (2011) indicate that the nine conceptions measured by the CoA-III are: assessment makes institutions accountable, assessment makes students accountable, assessment describes improvements in student abilities, assessment improves student learning, assessment improves teaching, assessment is valid, assessment is irrelevant and bad, assessment is irrelevant and ignored, and assessment is irrelevant and inaccurate.

In the present study, participants responded to 27 items using a six-point Likert scale, ranging from strongly disagree to strongly agree. In order to identify trends and to further explore faculty members' and undergraduate students' beliefs about the definition of assessment, an open-ended question developed by the researcher was added to the modified abridged version of the CoA-III developed by Fletcher et al. (2011). Specifically, participants were asked what the term assessment means to them. Furthermore, participants were asked to select from a list of possible responses, what types of activities come to mind when they think of the term assessment. These additional questions were used to gain further insight into faculty members' and

undergraduate students' conceptions of assessment within level V institutions of higher education in the SACS accreditation region.

### *Design*

A cross-sectional design was employed using survey methodology. A cross-sectional design utilizes different groups of individuals who may differ along the variable of interest but share some common characteristics (Williams, 2007). In this study, the dependent variable of interest is the conceptions of assessment and the common characteristics shared are those associated with being in higher education as a faculty member or as a student. Thus, the independent variable for this study was group membership. This independent variable has two levels: faculty members and undergraduate students.

The cross-sectional design allows for some inferences to be drawn that relate to the levels of the independent variable of interest in this study. In order to accomplish this, Brown's (2006) abridged CoA-III instrument, as modified by Fletcher et al. (2011), was used to determine how faculty members and undergraduate students conceptualize assessment. Permission to use this instrument was granted by both the original developer, Dr. Gavin Brown, Associate Professor, University of Auckland and by Dr. Richard Fletcher of Massey University. The self-reported belief scores on the abridged CoA-III about the purpose of assessment acted as the components that comprise the dependent variable, conceptions of assessment. A phenomenological approach was used to analyze the open-ended question in order to determine any trends in the responses provided by members of the three groups.

*Procedures*

The researcher received permission to proceed with the study from the Institutional Review Board (Appendix B) at the University of Southern Mississippi. The researcher then obtained pertinent information for each institution's primary contact (e.g. Director of Institutional Research, Director of Enrollment Services, Provost, or President). The researcher then asked these primary contacts to forward an invitation letter, including a link to the electronic questionnaire and a formal request to participate, to all members of the two groups (faculty members and undergraduate students). The researcher complied with the requirements set forth by each Institutional Review Board at participating institutions.

All level V institutions of higher education that offer bachelor's degrees located within the accreditation region of the Southern Association of Colleges and Schools (SACS) were contacted by the researcher. A representative from each institution was asked to email the link to the questionnaire, the informed consent, and an explanation of the purpose of the study to all faculty members teaching at and all undergraduate students attending these institutions. Participants were offered an invitation to be included in a raffle for one of the newest versions of the Apple iPad as an attempt to increase participant response rate. The researcher sent multiple reminders to each institution's primary contact in order for this individual to follow-up with participants.

Prior to sending the link to the questionnaire housed in Qualtrics to the institutions' primary contact, permission was requested from the University of Southern Mississippi's Human Subjects Protection Review Board. All information obtained directly from participants was kept confidential. Participant anonymity was protected

unless the participant chose to provide contact information for the Apple iPad raffle. In the case that the participant provided contact information, all identifiable information was stripped by a volunteer research associate prior to sending the data to the researcher.

Participation was completely voluntary and individuals could withdraw from the study at any point with no consequences or penalty.

### *Analysis*

Because an existing model (Figure 1) for the CoA-III exists for data collected by Brown (2006) and Fletcher et al. (2011) in New Zealand and Queensland, responses for the United States' faculty members and students were analyzed using confirmatory factor analysis and multi-group invariance testing. The goal was to determine whether the proposed model from New Zealand fit the United States data well and to what extent the model is equivalent for both groups (faculty and students). The model was tested using four fit indexes including Chi-square statistic, root mean square error of approximation (RMSEA), Tucker-Lewis Index (TLI) and comparative fit index (CFI). Reliability estimates were determined for both students and faculty responses.

The researcher also analyzed the open-ended question in order to determine any related themes among the definitions provided by participants. Using a phenomenological approach to data analysis, the researcher attempted to uncover deeper meanings of the conceptions of assessment as well as to capture the essence of the experiences associated with assessment for members of each of the two groups.

## CHAPTER IV

### RESULTS

In this chapter, results of the analyses of the data are provided. These results are presented as answers to each of the research questions posed. The research questions that drove this study were as follows.

1. How do the existing models of conceptions of assessment derived from New Zealand data fit the data collected in the U.S.?
2. What are the differences (if any) of how university faculty and university students in the U.S. conceptualize assessment?
3. What themes, if any, do the participants report in their meaning of assessment?
4. When faculty and students think of the term assessment, what types of assessment activities come to mind?

In order to answer the research questions, the researcher postulated a statistical model based on the related theory and empirical evidence of the conceptions of assessment from past research (Brown, 2004, 2006, 2008; Fletcher et al., 2011). Data were analyzed using a structural equation modeling framework in order to determine structural validity of the model before analyzing where the differences, if any, exist between faculty members' and students' conceptions of assessment. After determining the best fitting model, invariance analysis began. To this end, the best fitting model was tested for invariance across groups, faculty and students. And finally, the open-ended question "What does the term assessment mean to you" was analyzed to determine if

there were any trends in the responses in addition to the types of activities that come to mind when faculty and students think of the term assessment.

### Data Analysis

Of the initial 111 institutions, a total of 10 institutional primary contacts agreed that their institutions would participate in the study. Because the data collection process was anonymous, the researcher was unable to positively determine whether or not the primary contacts did in fact forward the invitation to participate. Further, the total number of potential participants is unknown.

In order to answer each research question, representatives of these 10 participating institutions of higher education across the southeastern United States agreed to invite their faculty members and undergraduate students to complete the abridged *Conceptions of Assessment III* (CoA-III) instrument. .

A total of 870 individuals opened the electronic questionnaire housed in Qualtrics. Of these, 563 valid questionnaires were used in the analyses after deleting those with more than 15% missing responses (Mertler & Vannatta, 2001) and those who reported administrator, staff, or other status as their primary role at the institution. Data imputations, using multiple regression procedures, were used to replace any missing items on the remaining questionnaires. The number of imputations completed per item is listed in Table 1.

Table 1

*Data Imputations using Multiple Regression*

Item #	# of Imputations	Item #	# of Imputations	Item #	# of Imputations
Item #1	0	Item #10	0	Item #19	0
Item #2	3	Item #11	1	Item #20	2
Item #3	1	Item #12	2	Item #21	0
Item #4	2	Item #13	2	Item #22	1
Item #5	1	Item #14	2	Item #23	1
Item #6	3	Item #15	1	Item #24	4
Item #7	0	Item #16	0	Item #25	1
Item #8	1	Item #17	1	Item #26	2
Item #9	0	Item #18	1	Item #27	1

The data were analyzed to determine if any required statistical test assumptions were violated. It was found that the data violated both skewness and kurtosis; however, these violations were deemed minor due to the large dataset and thus the analyses proceeded. The mean scores and standard deviations for each of the four correlated factors are included in Table 2. As evident from Table 2, there are numerical mean differences between faculty and students on three of the four latent constructs; however, additional testing was required in order to determine the statistical significance of these differences (see Table 2).

Table 2

*Faculty and Student Mean Scores*

Factor	Faculty		Student		MD
	M	SD	M	SD	
Institutional Accountability	3.52	1.21	3.80	1.10	.28
Student Accountability	3.80	.96	4.24	.93	.44
Improvement	4.01	.98	3.97	.87	.04
Irrelevant	3.44	.83	3.44	.76	0

The overall internal consistency of the CoA-III is  $\alpha = .815$  when faculty and students are analyzed together. When analyzed separately, the reliability estimate for faculty was  $\alpha = .780$  and for students was  $\alpha = .824$ . Additionally, the factor reliability estimates for faculty ranged from .463 to .906 and from .538 to .809 for students (see Table 3).

Table 3

*Faculty and Student Reliability Coefficients*

Factors	Faculty $\alpha$	Students $\alpha$
Overall Questionnaire	.780	.824
Institutional Accountability	.906	.779
Student Accountability	.520	.524
Improvement/Describe	.775	.753
Improvement/Valid	.867	.787



Table 3 (continued).

Factors	Faculty $\alpha$	Students $\alpha$
Improvement/Teaching	.737	.612
Improvement/Student Learning	.879	.809
Irrelevant/Inaccurate	.463	.538
Irrelevant/Ignore	.665	.592
Irrelevant/Bad	.790	.725

#### Confirmatory Factor Analysis

In order to answer research questions 1 and 2, it was first necessary to perform a confirmatory factor analysis.. This analysis allowed this researcher to analyze the underlying constructs of the CoA-III. Thus, in particular, before analyzing the data to determine if there were differences in the conceptions of assessment held by faculty and students in the United States, it was necessary to analyze the structural elements of the underlying theoretical construct measured by the CoA-III. Consequently, confirmatory factor analyses procedures were used in a structural equation modeling framework to determine the best fitting model to the data collected in the United States. The researcher of this study tested five competing models based on the research of Brown (2006) and Fletcher et al. (2011). The five models include:

1. A correlated four factor model specified by Brown (2006) with two first-order factors (institutional accountability and student accountability) and two second-order factors (improvement and irrelevant). The improvement factor consisted of improvement/describe, improvement/valid,

improvement/teaching, improvement/student learning. The irrelevant factor consisted of irrelevant/inaccurate, irrelevant/ignore, and irrelevant/bad

2. A second correlated four factor model specified by Brown (2006) with a path added between improvement/student learning and irrelevant/inaccurate
3. Fletcher et al. (2011) tested four competing models, one specified by Brown (2006) and three of their own. The three new models tested by Fletcher et al. (2011) are described below:
  - a. A three factor second-order model with nine first-order factors specifying conceptions of assessment for accountability (institutional and student), positive conceptions of assessment (improvement/description, improvement/validity, improvement/teaching, improvement/students) and negative conceptions of assessment (irrelevant/inaccurate, irrelevant/ignore, and irrelevant/bad)
  - b. A two factor second-order model with the nine first-order factors specifying positive and negative conceptions of assessment. The positive conceptions were comprised of institutional accountability, student accountability, improvement/description, improvement/students, improvement/teaching, and improvement/validity. The negative conceptions of assessment were comprised of irrelevant/bad, irrelevant/ignored, and irrelevant/inaccurate.
  - c. A one factor second-order model with nine first-order factors including institutional accountability, student accountability, improvement/description, improvement/students, improvement/teaching,

improvement/validity, irrelevant/bad, irrelevant/ignored, and  
irrelevant/inaccurate

*Results of the Confirmatory Factor Analysis*

An analysis of model fit for each of the models described above using three fit indexes, the comparative fit index (CFI), the Tucker-Lewis Index (TLI), and the root-mean-square error of approximation (RMSEA) was completed. The researcher began the model fit testing using Brown's (2006) hypothesized model consisting of two second-order factors and nine first-order factors (Appendix C). The model was inadmissible due to negative variance for e\_bad (-.025). To correct the negative variance, the researcher constrained e\_inacc, the error variable within the same second-order latent construct with the closest absolute value of the error, and e\_bad to be equal (Byrne, 2009). This imposition allowed the model to be statistically identified with a model fit of  $\chi^2_{(312)} = 1350.34, p < .01$ ; CFI = .848; TLI = .829; RMSEA = .077.

Second, the researcher continued the model fit testing using Brown's (2006) published four factor model that included a path between improvement/student learning and irrelevant/inaccurate without equalizing the error terms e\_inacc and e\_bad (Appendix D). The model showed moderate fit to the data collected in the United States,  $\chi^2_{(310)} = 1345.97, p < .01$ ; CFI = .849; TLI = .829; RMSEA = .077.

Third, the researcher continued the omnibus test of model fit by testing all models that were initially proposed by Fletcher et al. (2011) starting with a three factor second-order model and nine first-order factors (Appendix E). The model was inadmissible due to negative variance for e\_IrrBad (-.028). As in the case above, the researcher imposed a constraint setting e\_IrrBad equal to e\_IrreInacc which had the closest absolute value to

e\_IrrBad (Byrne, 2009). Imposing this constraint allowed the model to converge with an overall model fit of  $\chi^2_{(314)} = 1364.49$ ,  $p < .01$ ; CFI = .847; TLI = .828; RMSEA = .077.

The fourth model tested by the researcher was proposed by Fletcher et al. (2011). This fourth model consisted of two second-order factors with nine first-order factors specifying positive and negative conceptions of assessment (Appendix F). The model was inadmissible due to negative variance for e\_Bad (-.021). The researcher imposed a constraint equalizing e\_IB and e\_IInacc (Byrne, 2009). This imposition allowed the model to converge resulting in an overall model fit of  $\chi^2_{(315)} = 1370.60$ ,  $p < .01$ ; CFI = .846; TLI = .828; RMSEA = .077.

The fifth and final model the researcher tested was another proposed by Fletcher et al. (2011). This fifth model was comprised of one second-order factor and nine first-order factors (Appendix G). The model fit indexes resulted in a poor fit to the data  $\chi^2_{(315)} = 1575.49$ ,  $p < .01$ ; CFI = .816; TLI = .795; RMSEA = .084. Table 4 illustrates the model fit indexes for all five models tested in these analyses. Factor loadings for all models are listed in Table 5. Furthermore, correlations between latent variables for all models are provided in Table 6.

Table 4

*Omnibus Test of Goodness-of-Fit*

Model Description	<i>df</i>	$\chi^2$	CFI	TLI	RMSEA
1. Two second-order factors; nine first-order factors	312	1350.34*	.848	.829	.077

Table 4 (continued).

Model Description	<i>df</i>	$\chi^2$	CFI	TLI	RMSEA
2. Two second-order factors; nine first-order factors with path from improvement/student learning to irrelevant/inaccurate	310	1345.97*	.849	.829	.077
3a. Three second-order factors; nine first-order factors	314	1364.49*	.847	.828	.077
3b. Two second-order factors; nine first-order factors	315	1370.60*	.846	.828	.077
3c. One second-order factor; nine first-order factors	315	1575.49*	.816	.795	.084

Note. \*significant at .05

Table 5

*CoA-III Factor Loadings for Five Models Tested*

Factors and CoA-III items	Brown (2006)		Fletcher et al. (2011)		
	Model 1 (constrained error)	Model 2	Model 3a (constrained error)	Model 3b (constrained error)	Model 3c
<i>Institutional Accountability</i>					
Assessment is an accurate indicator of an institution's quality	.731	.731	.786	.730	.729
Assessment is a good way to evaluate an institution	.870	.870	.876	.869	.870

Table 5 (continued).

Factors and CoA-III items	Brown (2006)		Fletcher et al. (2011)		
	Model 1 (constrained error)	Model 2	Model 3a (constrained error)	Model 3b (constrained error)	Model 3c
Assessment provides information on how well institutions are doing	.750	.751	.768	.752	.752
<i>Student Accountability</i>					
Assessment is assigning a grade or level to student work	.332	.332	.337	.323	.321
Assessment determines if students meet qualification standards	.850	.849	.876	.878	.884
Assessment places students into ranks	.289	.289	.290	.271	.268
<i>Improvement/Describe</i>					
Assessment establishes what students have learned	.759	.759	.775	.759	.759
Assessment measures students' higher order thinking skills	.690	.689	.706	.691	.691
Assessment is a way to determine how much students have learned from teaching	.721	.723	.737	.720	.721

Table 5 (continued).

Factors and CoA-III items	Brown (2006)		Fletcher et al. (2011)		
	Model 1 (constrained error)	Model 2	Model 3a (constrained error)	Model 3b (constrained error)	Model 3c
<i>Improvement/Student Learning</i>					
Assessment feeds back to students their learning needs	.805	.805	.816	.806	.805
Assessment helps students improve their learning	.821	.820	.833	.821	.822
Assessment provides feedback to students about their performance	.735	.837	.749	.735	.734
<i>Improvement/Teaching</i>					
Assessment is integrated with teaching practices	.630	.630	.645	.628	.628
Assessment information modifies ongoing teaching of students	.646	.646	.662	.647	.647
Assessment allows different students to get different instruction	.607	.606	.623	.608	.608
<i>Improvement/Validity</i>					
Assessment results are trustworthy	.795	.795	.807	.795	.795

Table 5 (continued).

Factors and CoA-III items	Brown (2006)		Fletcher et al. (2011)		
	Model 1 (constrained error)	Model 2	Model 3a (constrained error)	Model 3b (constrained error)	Model 3b
Assessment results are consistent	.679	.679	.693	.680	.679
Assessment results can be depended on	.818	.818	.828	.818	.818
<i>Irrelevant/Bad</i>					
Assessment forces instructors to teach in a way that is against their beliefs	.653	.650	.657	.651	.649
Assessment is unfair to students	.767	.764	.772	.770	.738
Assessment interferes with teaching	.699	.691	.701	.651	.720
<i>Irrelevant/Ignored</i>					
Assessment results are filed and ignored	.811	.815	.813	.812	.881
Assessment has little impact on teaching	.391	.390	.394	.391	.361
Instructors conduct assessments but make little use of the results	.634	.632	.635	.633	.595
<i>Irrelevant/Inaccurate</i>					
Assessment results should be treated cautiously because of measurement error	.450	.508	.454	.458	.713



Table 5 (continued).

Factors and CoA-III items	Brown (2006)		Fletcher et al. (2011)		
	Model 1 (constrained error)	Model 2	Model 3a (constrained error)	Model 3b (constrained error)	Model 3b
Assessment is an imprecise process	.657	.628	.657	.652	.395
Instructors should take into account the error and imprecision in all assessment	.312	.350	.314	.311	.520

Table 6

*Omnibus Tests: Correlations between Latent Variables*

Model Description and Latent Variables	<i>r</i>
Model 1: Two second-order and nine first-order factors with constrained error terms	
Institutional accountability with student accountability	.765
Student accountability with improvement	.857
Improvement with irrelevant	-.545
Student accountability with irrelevant	-.330
Institutional accountability with irrelevant	-.400
Institutional accountability with improvement	.822
Model 2: Two second-order factors and nine first-order factors with path from improvement/student learning to irrelevant/inaccurate	
Institutional accountability with student accountability	.766
Student accountability with improvement	.858
Improvement with irrelevant	-.543
Student accountability with irrelevant	-.333
Institutional accountability with irrelevant	-.397
Institutional accountability with improvement	.822

Table 6 (continued).

Model Description and Latent Variables	<i>r</i>
Model 3a: Three second-order factors with nine first-order factors with constrained error terms	
Positive conceptions with accountability	.956
Positive conceptions with negative conceptions	-.556
Accountability with negative conceptions	-.437
Model 3b: Two second-order factors with nine first-order factors with constrained error terms	
Positive Conceptions with Negative Conceptions	-.521

#### Results: Research Questions 1 and 2

1. How do the existing models of conceptions of assessment derived from New Zealand data fit the data collected in the U.S.?
2. What are the differences (if any) of how university faculty and university students in the U.S. conceptualize assessment?

To answer research questions 1 and 2, a multi-group study using confirmatory factor analysis, conducted within structural equation modeling with 563 participants (159 faculty members; 404 students) was performed on five competing models using AMOS Version 21. The analyses indicate slight differences in the fit indexes of the five measurement models tested (see Table 4). Although the differences in fit indexes were slight; Brown's (2006) correlated four factor model with two first-order factors (school accountability and student accountability) and two second-order factors (improvement and irrelevant) with a path between student learning and inaccurate was retained on the basis of statistical best fit without the necessity of imposing additional constraints due to

negative variance associated with irrelevant/bad ( $\chi^2_{(310)} = 1345.967, p < .01$ ; CFI = .849; TLI = .829; RMSEA = .077). Although the comparative fit index (CFI) and the Tucker-Lewis Index (TLI) are slightly lower than recommended (Meyers, Gamst, & Guarino, 2006), the root mean square error of approximation (RMSEA) indicates moderate fit to the data at .077.

It is interesting to note that Fletcher et al. (2011) found that the two-factor higher order model was the best fitting model. Fletcher et al. (2011) specifically chose not to retain Brown's (2006) model in part due to high correlations for students between student accountability and improvement ( $r = 0.76$ ). As evidenced in Table 7, the data collected in this current study also elicited high correlations between student accountability and improvement for both faculty and students ( $r = 0.79$  and  $r = 0.86$ ). Furthermore, there were high correlations between institutional accountability and improvement ( $r = 0.72$  and  $r = 0.90$ ) for both faculty and students (see Table 7).

Table 7

*Correlations between Latent Variables for Faculty and Students*

Latent Variable Correlations	Overall <i>r</i>	Faculty <i>r</i>	Students <i>r</i>
Institutional Accountability with Student Accountability	.766	.675	.799
Student Accountability with Improvement	.858	.790	.862
Improvement with Irrelevant	-.543	-.693	-.431
Student Accountability with Irrelevant	-.333	-.373	-.305
Institutional Accountability with Irrelevant	-.397	-.432	-.374
Institutional Accountability with Improvement	.822	.720	.901

In this study, however, the researcher found the two-factor higher order model retained by Fletcher et al. (2011) to be inadmissible due to negative variance for the irrelevant/bad error term. Thus, even though the highly correlated factors were found, Brown's (2006) four factor model was retained due to its history of empirical evidence found in multiple studies (Brown, 2004; Brown, 2006; Brown & Lake, 2006).

#### Invariance of the CoA-III

The results for the CoA-III initial model fit indicated a moderate fit to the data collected in the United States ( $\chi^2_{(310)} = 1345.97, p < .01$ ; CFI = .849; TLI = .829; RMSEA = .077). Factor loadings for both faculty and students are listed in Table 8. For this model, the factor loadings ranged from .296 to .897 for faculty and from .284 to .860 for students (see Table 8). This model was retained and an omnibus test of equality of covariance structures was performed in order to determine if the model is equivalent across faculty and students.

Table 8

#### *CoA-III Factor Loadings for Faculty and Students*

Factors and CoA-III items	Faculty	$\alpha$	Students	$\alpha$
<i>Institutional Accountability</i>		.906		.779
Assessment is an accurate indicator of an institution's quality	.895		.650	
Assessment is a good way to evaluate an institution	.909		.825	
Assessment provides information on how well institutions are doing	.818		.747	
<i>Student Accountability</i>		.520		.524
Assessment is assigning a grade or level to student work	.296		.352	

Table 8 (continued).

Factors and CoA-III items	Faculty	$\alpha$	Students	$\alpha$
Assessment determines if students meet qualification standards	.896		.860	
Assessment places students into ranks	.327		.284	
<i>Improvement/Describe</i>		.775		.753
Assessment establishes what students have learned	.764		.753	
Assessment measures students' higher order thinking skills	.730		.672	
Assessment is a way to determine how much students have learned from teaching	.736		.721	
<i>Improvement/Student Learning</i>		.879		.809
Assessment feeds back to students their learning needs	.870		.780	
Assessment helps students improve their learning	.840		.808	
Assessment provides feedback to students about their performance	.813		.712	
<i>Improvement/Teaching</i>		.737		.612
Assessment is integrated with teaching practices	.818		.537	
Assessment information modifies ongoing teaching of students	.675		.651	
Assessment allows different students to get different instruction	.615		.602	
<i>Improvement/Validity</i>		.867		.787
Assessment results are trustworthy	.847		.780	
Assessment results are consistent	.744		.654	
Assessment results can be depended on	.897		.785	

Table 8 (continued).

Factors and CoA-III items	Faculty	$\alpha$	Students	$\alpha$
<i>Irrelevant/Bad</i>		.790		.725
Assessment forces instructors to teach in a way that is against their beliefs	.797		.574	
Assessment is unfair to students	.717		.815	
Assessment interferes with teaching	.728		.648	
<i>Irrelevant/Ignored</i>		.665		.592
Assessment results are filed and ignored	.792		.842	
Assessment has little impact on teaching	.478		.364	
Instructors conduct assessments but make little use of the results	.692		.591	
<i>Irrelevant/Inaccurate</i>		.463		.538
Assessment results should be treated cautiously because of measurement error	.356		.756	
Assessment is an imprecise process	.700		.409	
Instructors should take into account the error and imprecision in all assessment	.325		.517	

### *Metric Invariance*

Having established Brown's (2006) correlated four factor model with two first-order factors (school accountability and student accountability) and two second-order factors (improvement and irrelevant) with a path between student learning and inaccurate as the baseline model in this study, the two groups were then tested to determine if participants responded the same to the items and their respective underlying constructs. This was accomplished by the researcher imposing a series of tests using a sequence of increasingly stringent constraints across groups (Byrne, 2009). First, all first-order latent

factors together were constrained in order to determine if there were differences between groups on these factors. These constraints elicited a statistically worse fitting model  $\Delta\chi^2_{(18)} = 55.337, p = < .01$  than the retained model, indicating that there are differences across groups. Thus, metric invariance was not achieved between faculty and students.

The next step in the process was to determine which of the first-order factors were statistically different across groups. Chi-square difference tests were completed in order to determine where these differences actually existed. As a result of these tests, statistically significant differences across faculty and student groups were determined to exist in four of the nine first-order factors (see Table 9).

Table 9

*First-Order Factor Chi-Square Differences*

First-Order Factors	<i>df</i>	$\Delta\chi^2$	<i>p</i>
Institutional Accountability	2	6.089	.048
Improvement/Teaching	2	16.463	.001
Irrelevant/Inaccurate	2	12.369	.002
Irrelevant/Bad	2	11.729	.003

Upon completion of the first step in metric invariance testing that determines which of the first-order factors are significantly different between groups, the researcher continued the hierarchical approach of constraints in order to determine which items within each first-order factor are statistically significant. In order to determine which items are different across groups within the significantly different first-order factors, chi-square difference tests were completed on a total of eight items by constraining each item

to be equal to one another across groups. The analyses resulted in statistical differences in five items across groups (see Table 10).

Table 10

*Chi-Square Difference Tests for Items of First-Order Factors*

Item	<i>df</i>	$\Delta\chi^2$	<i>p</i>
Assessment provides information on how well institutions are doing	1	4.054	.044
Assessment is a good way to evaluate an institution	1	5.322	.021
Assessment information modifies ongoing teaching of students	1	14.304	.001
Assessment allows different students to get different instruction	1	8.980	.003
Assessment is an imprecise process	1	12.369	.001

Once invariance testing was completed for all first-order factors and corresponding items, chi-square difference testing continued by the researcher. The researcher constrained all second-order factors to determine if model fit the data worse. In order to make this determination, a chi-square difference test yielded a significantly worse fitting model,  $\Delta\chi^2_{(5)} = 37.635, p < .001$ . Next, in order to determine specifically which of the second-order factors are significantly different across groups, chi-square difference tests were run for each of the second-order factors alone. This analysis revealed that both second-order factors were significantly different across groups (see Table 11).



Table 11

*Chi-Square Difference Tests of Second-Order Factors*

Second-Order Factors	<i>df</i>	$\Delta\chi^2$	<i>p</i>
Improvement	3	30.989	.001
Irrelevant	2	6.705	.035

After determining that neither second-order factors are invariant across groups, each item that measures each second-order factor was tested individually in order to determine which items are not invariant. These analyses yielded two statistically significant items (see Table 12).

Table 12

*Statistically Significant Items Different Across Groups*

Item	<i>df</i>	$\Delta\chi^2$	<i>p</i>
Improvement/teaching	1	30.668	.001
Irrelevant/ignore	1	5.290	.021

Finally, the researcher ended invariance testing by constraining all covariances in order to determine if the model resulted in worse fit to the data. A chi-square difference test was performed on a model with all covariances constrained. The analysis yielded a model that was not statistically significantly worse  $\Delta\chi^2_{(6)} = 6.123$ ; therefore, the revised model with the covariances constrained to be equal was retained.

As a final check, the researcher created a new baseline model that contained all constraints that had resulted in non-significant differences between groups. All

significant factor constraints were removed from the model before completing one final chi-square difference test. This new baseline model, when compared to the original default model, is not statistically significantly worse  $\Delta\chi^2_{(10)} = 9.170, p = .516$ ; therefore, this became the final model.

Due to two instances of highly correlated factors for both faculty and student groups (student accountability and improvement  $r = 0.79$  and  $r = 0.86$ ); (institutional accountability and improvement  $r = 0.72$   $r = 0.90$ ), the researcher performed an additional chi-square difference test after creating direct paths from improvement to student accountability and institutional accountability. The researcher then added institutional accountability as a first-order factor (see Appendix H). The chi-square difference test revealed that the model became significantly worse. Because the results of the chi-square difference test indicated a statistically significantly worse model, the previous baseline model was retained (see Table 13).

Table 13

*Chi-square Difference Tests for Correlations*

Correlation	<i>df</i>	$\Delta\chi^2$	<i>p</i>
Student accountability with improvement	15	69.591	.001
Institutional accountability with improvement	2	28.501	.001

### Results: Research Questions 3 and 4

3. What themes, if any, do the participants report in their meaning of assessment?
4. When faculty and students think of the term assessment, what types of assessment activities come to mind?

In order to better understand what the term assessment means to the participants in this study and to answer research question three, participants were asked to respond to an open-ended item on the questionnaire, “What does the term assessment mean to you?” The word test, testing, quiz, and/or exam appeared 13 times in the 146 faculty responses (9%). These same words appeared in 142 of the 394 student responses (36%).

Faculty mentioned the terms evaluation or evaluating in either program evaluation contexts or student learning contexts 40 times in the 146 responses (27%). Students, however, mentioned evaluation only 77 times out of the 394 responses (20%). The vast majority of the evaluation related responses for both faculty and students referred to evaluating students’ knowledge and skills. Faculty indicated more often than students that assessment is the evaluation of a course or program. Students mentioned that assessment is the evaluation of a situation or student learning more often than faculty. In nearly all of the cases that the term evaluation was mentioned, the context referred to an external accountability requirement for demonstrating knowledge, skills, teaching abilities, and quality. Neither faculty nor students mentioned formative assessment, feedback purposes, or improvement purposes often. Specifically, eleven responses from students and fourteen responses from faculty indicate that the purpose of assessment is to provide feedback or to improve teaching and learning. In the overwhelming majority of

statements, the term assessment was defined as meeting external demands imposed by someone within or outside of education.

In order to answer research question 4, a crosstab analysis was completed on the following question, “When you think of the term assessment, what types of activities come to mind?” Participants were asked to check all that apply from a list of 15 items (standardized test, self-reflection, program evaluation, oral questions/answers, portfolios, homework, course grades, written reports/research, conferencing, teacher made tests, tenure and/or promotion dossier, performance evaluation, accreditation, student evaluation, other). Table 14 contains the frequency of responses to each item by faculty and students.

Table 14

*Types of Assessment Activities*

Item	Faculty		Students	
	# of Responses	Percentage of Faculty ( <i>n</i> = 158)	# of Responses	Percentage of Students ( <i>n</i> = 404)
Standardized tests	122	<b>77</b>	357	<b>88</b>
Program evaluation	118	<b>75</b>	262	65
Performance evaluation	107	68	301	<b>75</b>
Student evaluation	110	<b>70</b>	262	65
Course grades	103	65	277	<b>69</b>
Teacher made tests	114	72	253	63

Table 14 (continued).

Item	Faculty		Students	
	# of Responses	Percentage of Faculty ( <i>n</i> = 158)	# of Responses	Percentage of Students ( <i>n</i> = 404)
Written reports/research	109	<b>69</b>	191	<b>47</b>
Homework	100	63	174	43
Oral questions/answers	98	62	185	46
Portfolios	98	62	139	34
Self-reflection	90	57	141	35
Accreditation	90	57	121	30
Tenure and/or promotion dossier	53	34	42	10
Conferencing	52	33	70	17
Other	15	9	15	4

Note. Top three percentages for both students and faculty are in boldface.

## CHAPTER V

### DISCUSSION

There were two primary purposes of this research study: 1) Confirm a model of the conceptions of assessment based on the past research about the purpose of assessment, and 2) Understand if and how students and faculty differ in their conceptions of assessment. The cross-sectional analysis using survey methodology provided a one-time snapshot of information from both faculty and students of higher education within the southeastern United States. Furthermore, the results of the data analyses offer insight into and possible answers for each of the four research questions.

1. How do the existing models of conceptions of assessment derived from New Zealand data fit the data collected in the U.S.?
2. What are the differences (if any) of how university faculty and university students in the U.S. conceptualize assessment?
3. What themes, if any, do the participants report in their meaning of assessment?
4. When faculty and students think of the term assessment, what types of assessment activities come to mind?

This study confirms some of the findings from other studies, in particular Brown (2004, 2006, 2008) and Fletcher et al. (2011). Additionally, it provides useful information that adds to the understanding of the complexity and importance of context in measuring conceptions of assessment. Furthermore, this information supports that there are many individualized purposes of assessment that are related to the attitudes, understandings, beliefs about, and experiences with assessment.

### Model Fit

The first research question aimed to confirm one of the two existing theoretical models of conceptions of assessment previously developed by Brown (2006) and Fletcher et al. (2011). The measure used in this study was an abridged version of Brown's (2006) CoA-III consisting of 27 items. Reliability estimates provided in Table 3 indicate some low levels of reliability, similar to the findings of Fletcher et al. (2011). Additionally, these results corroborate those of Fletcher et al. (2011) in that rewording some of the items and adding items to the constructs might be necessary in order to effectively measure these constructs more reliably in higher education settings. Similar to Fletcher et al. (2011), Brown's (2006) four correlated factor model did not have great fit to the data. This model, however, did have moderate fit to the data; therefore, Brown's (2006) four correlated factor model with two first-order factors and two second-order factors with a path between student learning and inaccurate was retained.

A high correlation among factors was noted in past studies (Brown, 2004; Brown, 2006; Fletcher et al., 2011). This finding was corroborated by the results of this study. Specifically, Brown (2004) found a strong negative relationship between the factors of improvement and irrelevance. This finding was also noted in the current study although the relationship was not quite as strong. This might be explained by the notion that within high-stakes assessment cultures, accountability mandates are at the forefront. As such, as faculty and students report that assessment is for improvement purposes, the data also suggest that assessment is relevant. Furthermore, like Brown (2004; 2006), the results of this study suggest a positive relationship between improvement and institutional accountability. Again, within the United States, there is a demand for

accountability. Both faculty and students associate these two constructs very highly with one another. Finally, as Brown (2004, 2006) reported a strong positive relationship between institutional accountability and student accountability as do the results of this study suggest. The relationship between these two constructs may be due to the number of years that external demands have been placed on institutions, faculty, and students. These stakeholders view institutional accountability and student accountability very closely related. Often, proponents of high-stakes accountability testing have argued that this would happen; therefore, high-stakes accountability mandates are viewed positively (Brown, 2006).

The data collected in the United States also represent some interesting differences from past research. Specifically, the U.S. data suggest a strong positive relationship between improvement and student accountability ( $r = .86$ ). Brown (2004) reports a weaker relationship between these two constructs. This difference is likely caused by the differences in the level of stakes involved with assessment and accountability in the U.S. versus New Zealand. Specifically, faculty and students report that improvement and student accountability are almost synonymous. This might be explained by New Zealand's student-centered philosophies whereas the U.S. places much more emphasis on faculty-centered approaches to teaching and learning that emphasizes accountability and standardized testing.

The results of this study also indicate that there are fairly strong negative relationships between accountability and irrelevance (both institutional and student). Brown (2004; 2006) does not find this same relationship in the data collected in New Zealand and Queensland. Specifically, Brown (2004; 2006) reports positive relationships



in the data collected from Queensland. He concludes that assessment is thought to be irrelevant when it is linked to student accountability and it is more likely to be acceptable if it is related to improvement of teaching and learning. The U.S. data indicate that when the purpose of assessment is linked to either institutional or student accountability then it is not irrelevant. The pattern found in the data in New Zealand and Queensland suggests that teachers report that assessment is either for improvement or for student accountability – although more complicated than this. The U.S. data indicate that these are not dichotomous constructs, but rather, improvement and student accountability are strongly related. Again, these relationships are likely explained through the differences in cultures in New Zealand and Queensland where there are very low-stakes associated with assessment. Here in the U.S., there are very high stakes-associated with assessment.

Although Brown's (2006) two second-order factor and nine first-order factor model was retained with a moderate fit to the data collected in this study, it appears that this theoretical model does not clearly generalize to either higher education populations or high-stakes environments. Consequently, additional research in higher education settings as well as high-stakes environments needs to be conducted in order to determine the best fitting theoretical model for these settings.

The data collected in this study from both faculty members and students, when analyzed together using Brown's (2006) model of conceptions of assessment, demonstrated moderate fit. The two-factor conceptions of assessment model of Fletcher et al. (2011), however, was inadmissible due to negative variance for the error term relating to the purpose of assessment as irrelevant/bad. Thus, Brown's (2006) model was retained and analyzed for specific differences across groups. The moderate fit to

Brown's (2006) model might be due to the fact that these data in the present study do come from a high-stakes assessment culture. Given that Fletcher's et al. (2011) model was not admissible certainly causes one to think that, indeed, these data are different due to the assessment culture here in the southeastern United States.

The difference in the  $\chi^2$  for the unconstrained model was statistically significant for faculty and students indicating that the parameters differed by more than chance. Similar patterns (directions and relative strengths of the relationships) were seen among the conceptions of faculty and students. It is possible that the model differences represented real differences in faculty and students' conceptions rather than inadequacies in the model. Conceptions are contextual; therefore, it is possible that the nature of the high-stakes culture within the United States is a plausible explanation for the differences between students and faculty conceptions (Brown, 2004, 2006, 2008).

Nonetheless, Brown's (2006) model of the conceptions of assessment has a moderate fit to the data collected here in the United States. The overall fit indexes are acceptable; however, there are both differences and similarities when the results from the present study are compared to those of Brown (2006) and Fletcher et al. (2011). One difference emerged when the present correlation between irrelevant conception and student accountability was compared with the same from Brown (2006). In particular, Brown (2006) reports that the student accountability conception positively correlated with irrelevant ( $r = .36$ ); however, the data collected in the United States indicate a negative correlation ( $r = -.33$ ). In fact, these correlations are almost the same except for their directions. Thus, a reasonable conclusion that might be drawn from this comparison is that in the high-stakes assessment culture of the United States, the irrelevant

conception and student accountability have a negative relationship. A similarity to Brown's (2006) findings was a high correlation between the latent variables of institutional accountability and improvement. Brown (2006) reported a correlation of  $r = .75$  and these same latent variables in the present study were also highly positively correlated ( $r = .82$ ). This suggests that both students and faculty have reasons to think that institutional accountability is strongly related to improving teaching and learning. Fletcher et al. (2011) reported a strong positive correlation between the latent variables student accountability and improvement ( $r = .76$ ). Similarly, in this current study, high correlations between student accountability and improvement for both faculty and students ( $r = 0.79$  and  $r = 0.86$ ) were found. These results suggest that although negativity is often associated with external accountability mandates (Darling-Hammond, 2003; Linn, 2000), there is evidence that both faculty and students associate accountability mandates with improved teaching and learning.

#### *Differences between Faculty and Students*

The second research question was, "What are the differences (if any) of how university faculty and university students in the U.S. conceptualize assessment?" This question sought to establish if there are differences across two groups, students and faculty, measured by the CoA-III. A sequence of increasingly stringent constraints imposed across faculty and student groups were used in this study (Byrne, 249). Results of invariance testing provided evidence that students conceptualize assessment differently than faculty. Specifically, there are significant differences between faculty and student responses to four of the nine first-order factors: institutional accountability, improvement/teaching, irrelevant/inaccurate, and irrelevant/bad. Students indicated that

assessment is a good way to evaluate an institution whereas faculty members did not necessarily concur. Also, students reported that assessment determines if students meet qualification standards but faculty members indicated that they do not concur. Faculty members reported that assessment information modifies ongoing teaching of students whereas students do not appear to share this notion. Furthermore, contrary to what students reported, faculty members suggested that assessment allows different students to receive different instruction. Finally, faculty members deemed important that instructors should take into account the error and imprecision in all assessment whereas students did not rate this item as highly. These findings mesh well with the issues discussed by Meyer (1996). Specifically, Meyer (1996) suggested that educators explored ways to assess their students' learning and attempted to find a variety of means to improve the measures used to judge student performance. Additionally, these results corroborate those of Fletcher et al. (2011). Their findings indicated that assessment in higher education is part of institutional quality and accountability processes in addition to measuring student learning.

These findings corroborate results from previous research studies. Specifically, Brown (2004) found that the conceptions of assessment held by students are often different from those held by their teachers. Furthermore, the findings of the present study compare favorably to those of Fletcher et al. (2011). Specifically, Fletcher et al. (2011) indicated that there are many purposes of assessment in higher education and that faculty and students' attitudes toward assessment differ in that faculty view assessment as an aid to the teaching and learning process whereas students view assessment as needed for accountability but irrelevant to the teaching and learning process.

### Summary of Qualitative Findings

The third-research question sought to determine if there are underlying themes in responses of faculty and students regarding the term assessment. When asked to write, “What does the term assessment mean to you”, faculty and students responded very differently from each other. Students used the word(s) test, testing, quizzes, and exams nearly 30% more often than faculty. However, this information alone should not be surprising. Indeed, undergraduate students, have most likely recently experienced the K-12 environment where standardized tests are used to make high-stakes decisions. On the other hand, faculty members most likely have not attended a K-12 school in several years; therefore, they are unlikely to have experienced the pressures of the high-stakes decisions associated with assessment in K-12 environments. Clearly, there is a massive difference in the accountability demands placed on K-12 teachers and students today compared to just 10 or 15 years ago. As such, many traditional undergraduate students are extremely familiar with the expectations for them to perform well on standardized tests. Thus, it is reasonable to conclude that the findings of the present study have been impacted by the students’ recent emersion in a high-stakes assessment culture.

The fourth research question asked faculty and students to select the types of activities that come to mind when they think of the term assessment. It is interesting to note that although only 7% of the faculty responses related to tests, quizzes, and exams in the open-ended question “What does the term assessment mean to you?”, 77% of the faculty marked standardized testing as an activity that comes to mind when they think of the term assessment. Standardized tests are indicated as an activity that comes to mind for 88% of the students. Consequently, for both groups, standardized testing is the most

frequently selected assessment activity that comes to mind when they think of the word assessment. Conversely, faculty members did not use terms associated with testing in their responses to the open-ended question. This is possibly an indication that although faculty use various ways to describe assessment(s), the high-stakes testing culture apparent in the United States today nonetheless influences the deep rooted meaning of the activities associated with the term assessment.

### Implications of the Findings

The term assessment has many contexts and meanings to multiple stakeholders of higher education. Understanding attitudes toward and about the purpose of assessment can inform policy makers, faculty, students, administrators, and education constituents as a whole about the disparities between practice and public policies. As evidenced by the significant differences in responses of students and faculty on the CoA-III, these differences are most likely due to real differences in beliefs. Indeed, students view assessment as a means of accountability, whereas faculty members are more likely to associate assessment with improving teaching and learning. Nonetheless, both faculty members and students are faced with the consequences of high-stakes associated with results on various assessments. Past studies indicate that faculty and students' beliefs about and attitudes toward assessment have a profound impact on learning and achievement (Brown & Hirschfeld, 2008; Struyven et al., 2005).

The current culture within the United States today is to increasingly hold institutions, faculty, and students accountable. As such, policymakers and legislators demand increased learning and preparation for employment of graduates in a 21<sup>st</sup> century globalized economy. No longer is it satisfactory to rely on the expertise of faculty and

the graduation rates of college students. Instead, faculty members are held accountable and must meet much higher standards by having to demonstrate effective and appropriate outcomes. Knowing that attitudes, beliefs, and past experiences (Ajzen, 1991; Bandura, 1986) with assessment can affect future learning and outcomes of students, and that the assessment practices of instructors can improve student outcomes (Brown & Hirschfeld, 2008; Struyven et al., 2005), it is important for policymakers to take into consideration the conceptions of both faculty and students if in fact they want these implemented policies to have a positive impact on learning and achievement. Furthermore, it is clear that differences in students' and faculty members' responses concerning assessment and accountability in the present study can be reasonably attributed to beliefs affected by the high-stakes assessment culture in the United States. It is important to recognize that even though students and faculty seem to view assessment and accountability differently, the ongoing use of high-stakes assessments must be recognized as an important factor that will continually shape the conceptions of students as well as faculty.

As evidenced in past research, this study supports the notion that faculty often support certain values about assessment that are often contradicted by actual practice (Fletcher et al., 2011). The finding of the open-ended question related to the meaning of the term assessment represents this contradiction. Faculty indicated that assessment is a form of testing and/or evaluation of either students or programs. Unlike the open-ended question, faculty reported improvement purposes of assessment, however, to the closed-ended items on the questionnaire. Fletcher et al. (2011) indicates that "divergent views about assessment among faculty and students may be the unfortunate consequence of the

absence of clear principals for policy and practice in this area as well as an immature evidence-base” (p. 131).

As indicated throughout this study, conceptions are contextual and they incorporate past experiences; therefore, the more exposure to how assessment works and can help when used appropriately, the better aligned practice will be to policy. Furthermore, external accountability demands won't be the driving force; but rather, they will simply be a check-off for faculty, students, and administrators of higher education. If students develop an understanding of assessment in a way that is not regarded as unfair and/or inconsistent, then these students may become self-actualized learners which can improve the educational outcomes for all stakeholders involved.

#### Limitations and Suggestions for Future Research

Although this study produced a model of conceptions of assessment for faculty and students of higher education, the results should be interpreted with caution. A limitation of this study is that all participants were recruited from only the southeastern geographic region of the United States. Furthermore, there is a large disparity between the number of faculty members ( $n = 159$ ) and undergraduate students ( $n = 404$ ) who participated in the study. Future research should aim to collect a larger number of faculty members to provide additional support for the model of conceptions of assessment.

Results of the present study suggest several other avenues of research. First, analyses revealed that Brown's (2006) model of the conceptions of assessment has a moderate but not good fit to the data collected in the southeastern United States. It is apparent that additional evidence of factorial and structural validity needs to be collected in order to determine whether or not the model is appropriate for higher education and/or



high-stakes environments. It is recommended that data be collected outside the southeastern United States to include other geographic locations within the United States. Furthermore, it is recommended that data be collected from both private and public institutions of higher education to determine if there are differences in the conceptions of assessment based on the type of institution the faculty member teaches at and students attend. Such data could strengthen model fit and help determine if differences are indeed due to institution type, faculty type, and student type. Additionally, in order to determine if the model fits high-stakes environments such as the United States, it is suggested that data be collected from K-12 environments within the U.S. This should help determine whether or not Brown's (2006) model used in this study would better fit K-12 environments than higher education. It could also help determine if the moderate model fit of the data from the present study is related to the stakes associated with assessment or to the level of schooling (higher education versus K-12). Finally, the qualitative aspect of this study suggests a need to further explore the thinking of these two groups as it relates to assessment types.

The results of this study indicate two highly correlated constructs – institutional accountability with improvement and student accountability with improvement. These correlations are not found in the conceptions of assessment evidence collected in K-12 environments (Brown, 2004, 2006); however, they are similar to those found in a past study completed within a low-stakes higher education environment (Fletcher et al., 2011). As such, a further area of study includes the highly correlated constructs. Determining how and why these constructs are highly correlated could provide important information that would help to revise the model. Lastly, replication of this study using an exploratory

rather than a confirmatory approach may elicit a better theoretical model for use in higher education where there is evidence of a high-stakes culture. Within the highly correlated constructs, there is also a notably low factor loading. Specifically, the item, assessment places students into ranks, has a factor loading of .289. This study used a confirmatory approach. It is possible, however, that an exploratory approach might have been justified due to the highly correlated constructs as well as the low factor loadings. Future research using a combination of an exploratory and confirmatory approach might lend itself to better model fit as well as to eliminating the highly correlated constructs and low factor loadings.

## APPENDIX A

## ABRIDGED CONCEPTIONS OF ASSESSMENT QUESTIONNAIRE (CoA-III)

This section asks about your beliefs and understandings about ASSESSMENT, whatever that terms means to you. Please answer the questions using your own understanding of assessment. Please give your rating of each statement based on YOUR opinion about assessment by filling in the most appropriate response option.

For each of the statements below, indicate your level of agreement or disagreement using the following options:

	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
Assessment is a way to determine how much students have learned from teaching						
Assessment places students into ranks						
Assessment provides information on how well institutions are doing						
Assessment provides feedback to students about their performance						
Assessment is integrated with teaching practices						
Assessment results are trustworthy						
Assessment forces instructors to teach in a way that is against their beliefs						
Instructors conduct assessments but make little use of the results						
Assessment results should be treated cautiously because of measurement error						
Assessment is an accurate indicator of an institution's quality						
Assessment is assigning a grade or level to student work						
Assessment establishes what students have learned						
Assessment feeds back to students their learning needs						
Assessment information modifies ongoing teaching of students						
Assessment results are consistent						

	Strongly Disagree	Mostly Disagree	Slightly Disagree	Slightly Agree	Mostly Agree	Strongly Agree
Assessment is unfair to students						
Assessment results are filed and ignored						
Instructors should take into account the error and imprecision in all assessment						
Assessment is a good way to evaluate an institution						
Assessment determined if students meet qualification standards						
Assessment measures students' higher order thinking skills						

Please rate the following statement based on your opinion about assessment. Indicate how much you agree or disagree with each statement. Choose the one response per item that is closest to describing your opinion.

	Strongly Disagree	Mostly Disagree	Slightly Disagree	Moderately Agree	Mostly Agree	Strongly Agree
Assessment helps students improve their learning						
Assessment allows different students to get different instruction						
Assessment results can be depended on						
Assessment interferes with teaching						
Assessment has little impact on teaching						
Assessment is an imprecise process						

What does the term assessment mean to you?

When you think of the term assessment, what types of activities come to mind? (check all that apply)

- Standardized Test
- Self-Reflection
- Program Evaluation
- Oral Questions/Answers
- Portfolios
- Homework
- Course Grades
- Written Reports/Research
- Conferencing
- Teacher Made Test
- Tenure and/or Promotion Dossier
- Performance Evaluation
- Accreditation
- Student Evaluation
- Other

What type of training and/or professional development in educational assessment have you completed? (check all that apply)

- None
- Attended conference session(s)
- Attended training offered by my institution
- Completed coursework in assessment
- Other

Please indicate your primary role at the institution.

- Student
- Faculty Member
- Administrator
- Staff
- Other

Please indicate your primary teaching responsibilities at the institution.

- Teach mostly undergraduate students
- Teach mostly graduate students
- Teach only undergraduate students
- Teach only graduate students
- Teach approximately the same number of undergraduate and graduate students
- I am not currently teaching
- Other

Please indicate your administrative role at the institution.

- Academic Dean
- Department Chair
- Program Coordinator
- Other

Please indicate your student classification.

- Freshman
- Sophomore
- Junior
- Senior
- Other

Please indicate your primary role.

- Student Services
- Accreditation and/or Assessment
- Other

Would you like to be included in the raffle for a free Apple iPad?

- Yes
- No

Please provide your email address for the Apple iPad raffle.

## APPENDIX B

## INSTITUTIONAL REVIEW BOARD NOTICE OF COMMITTEE ACTION



**INSTITUTIONAL REVIEW BOARD**  
 118 College Drive #5147 | Hattiesburg, MS 39406-0001  
 Phone: 601.266.6820 | Fax: 601.266.4377 | [www.usm.edu/irb](http://www.usm.edu/irb)

**NOTICE OF COMMITTEE ACTION**

The project has been reviewed by The University of Southern Mississippi Institutional Review Board in accordance with Federal Drug Administration regulations (21 CFR 26, 111), Department of Health and Human Services (45 CFR Part 46), and university guidelines to ensure adherence to the following criteria:

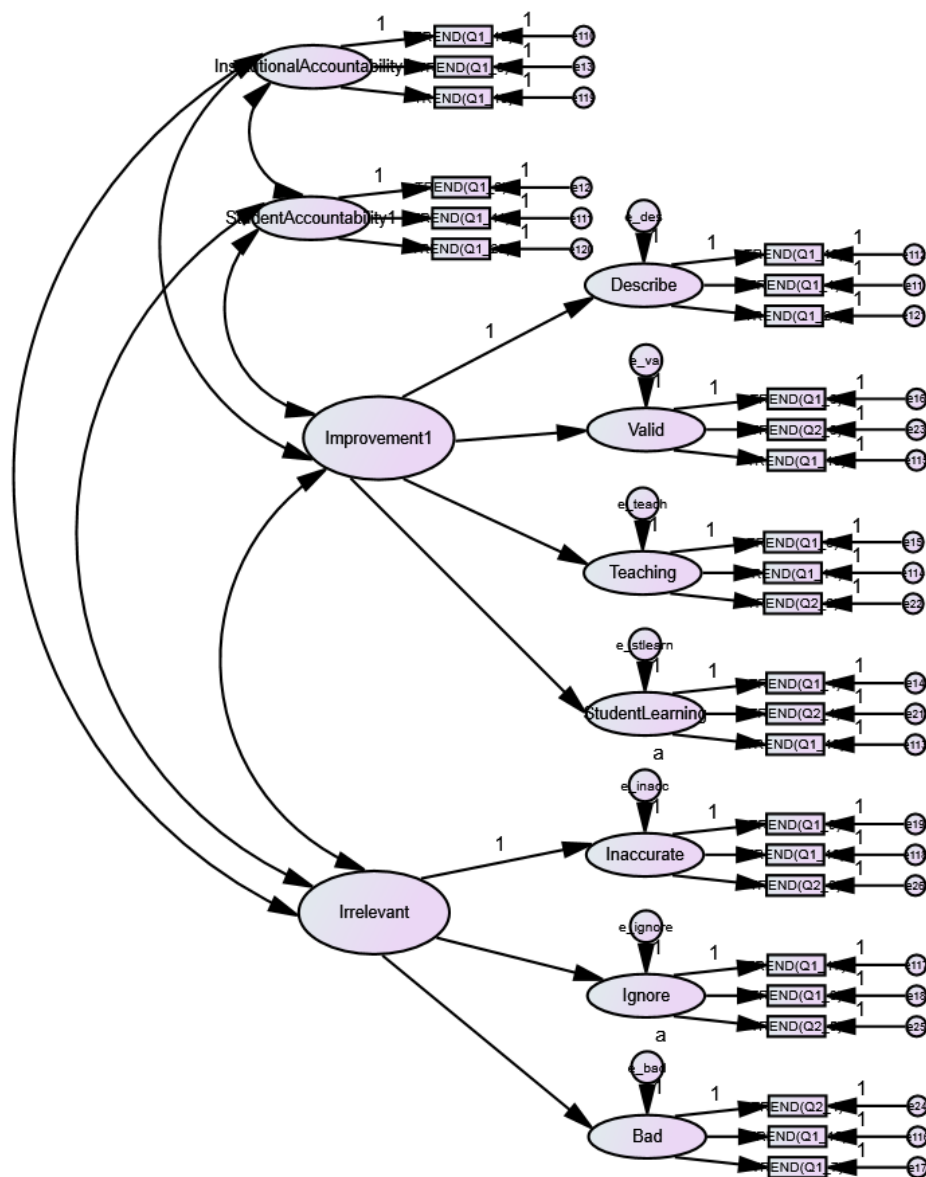
- The risks to subjects are minimized.
- The risks to subjects are reasonable in relation to the anticipated benefits.
- The selection of subjects is equitable.
- Informed consent is adequate and appropriately documented.
- Where appropriate, the research plan makes adequate provisions for monitoring the data collected to ensure the safety of the subjects.
- Where appropriate, there are adequate provisions to protect the privacy of subjects and to maintain the confidentiality of all data.
- Appropriate additional safeguards have been included to protect vulnerable subjects.
- Any unanticipated, serious, or continuing problems encountered regarding risks to subjects must be reported immediately, but not later than 10 days following the event. This should be reported to the IRB Office via the "Adverse Effect Report Form".
- If approved, the maximum period of approval is limited to twelve months.  
 Projects that exceed this period must submit an application for renewal or continuation.

PROTOCOL NUMBER: **13030102**  
 PROJECT TITLE: **University Faculty Members' and Undergraduate Students' Conceptions of Assessment**  
 PROJECT TYPE: **Dissertation**  
 RESEARCHER(S): **Melanie DiLoreto**  
 COLLEGE/DIVISION: **College of Education & Psychology**  
 DEPARTMENT: **Educational Studies & Research**  
 FUNDING AGENCY/SPONSOR: **N/A**  
 IRB COMMITTEE ACTION: **Expedited Review Approval**  
 PERIOD OF APPROVAL: **03/04/2013 to 03/03/2014**

**Lawrence A. Hosman, Ph.D.**  
**Institutional Review Board**

APPENDIX C

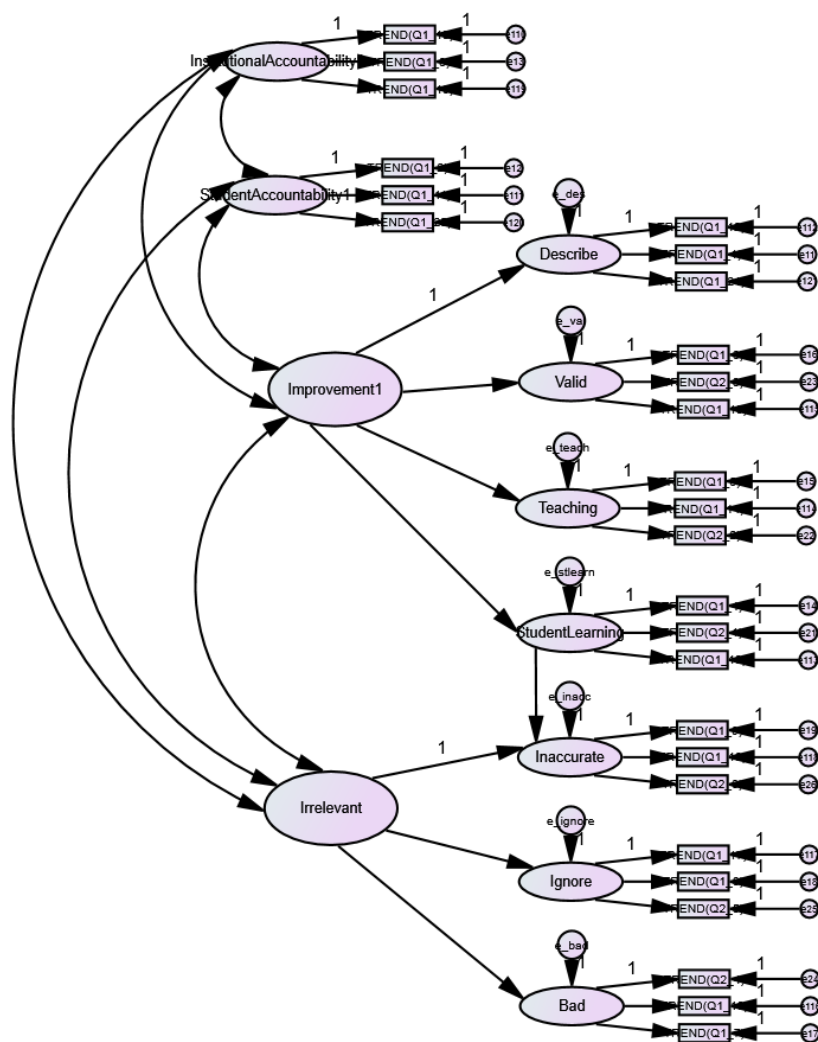
MODEL 1: TWO SECOND-ORDER FACTORS AND NINE FIRST-ORDER FACTORS WITH CONSTRAINED ERROR TERMS





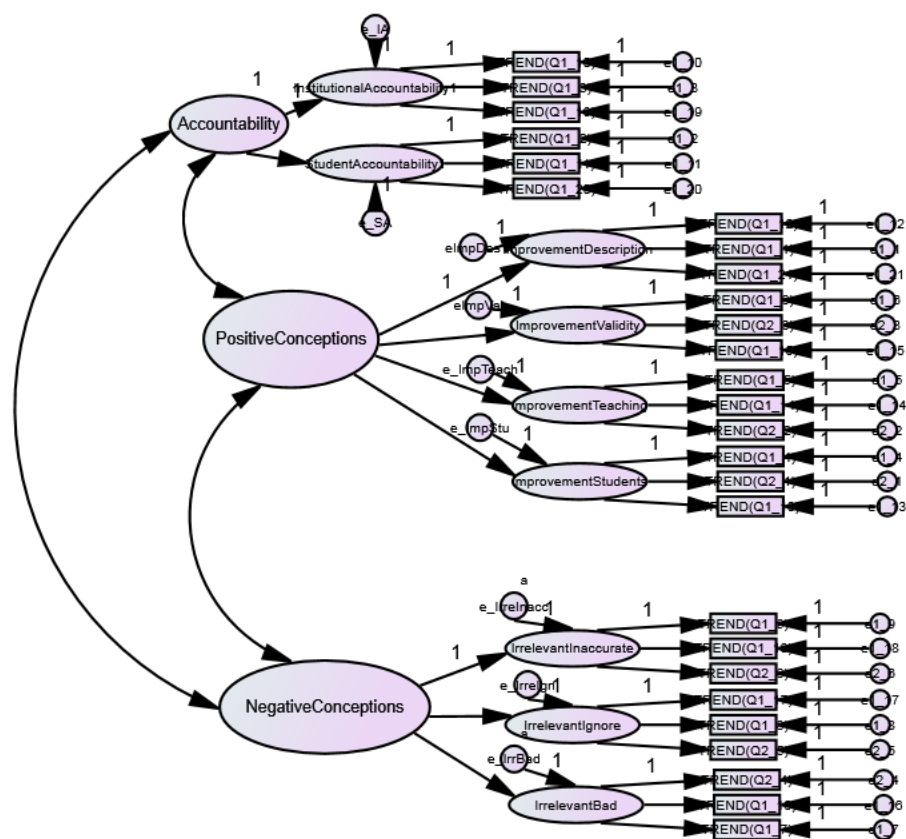
## APPENDIX D

MODEL 2: TWO SECOND-ORDER FACTORS AND NINE FIRST-ORDER FACTORS WITH PATH FROM IMPROVEMENT/STUDENT LEARNING TO IRRELEVANT/INACCURATE



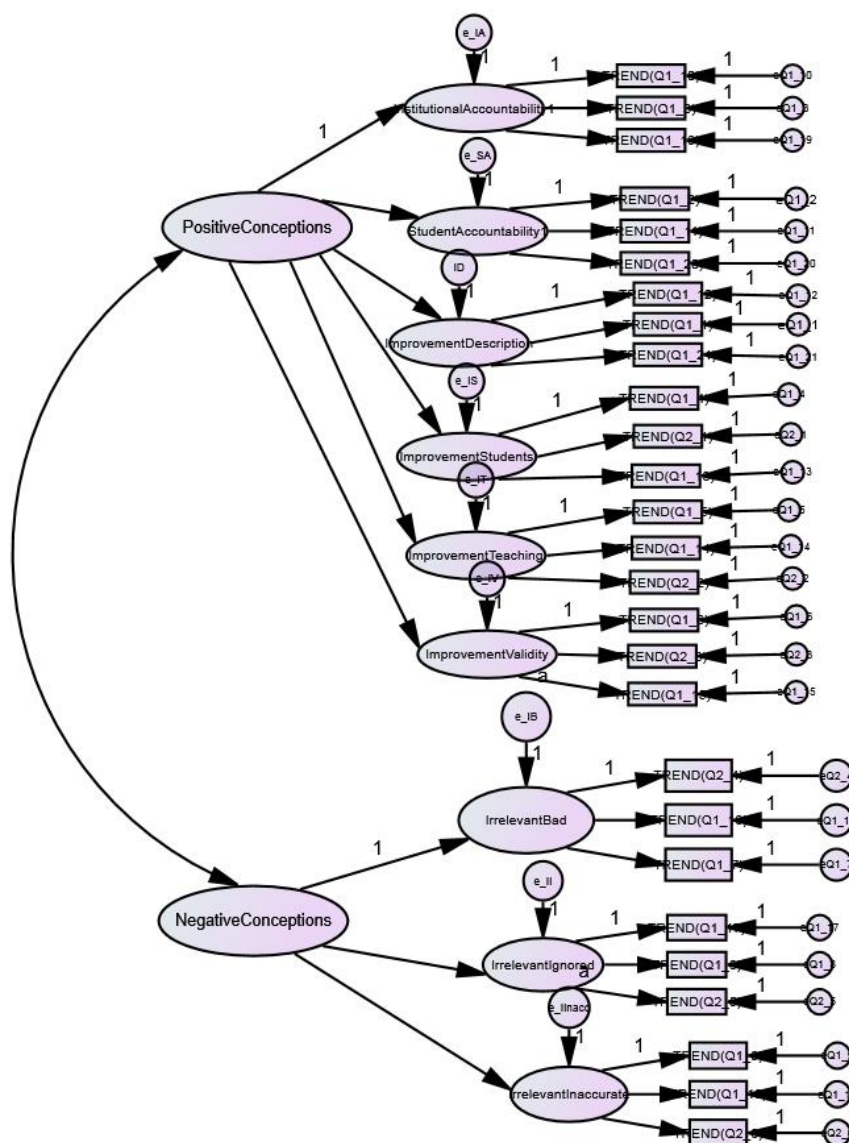
APPENDIX E

MODEL 3a: THREE SECOND-ORDER FACTORS WITH NINE FIRST-ORDER FACTORS WITH CONSTRAINED ERROR TERMS



APPENDIX F

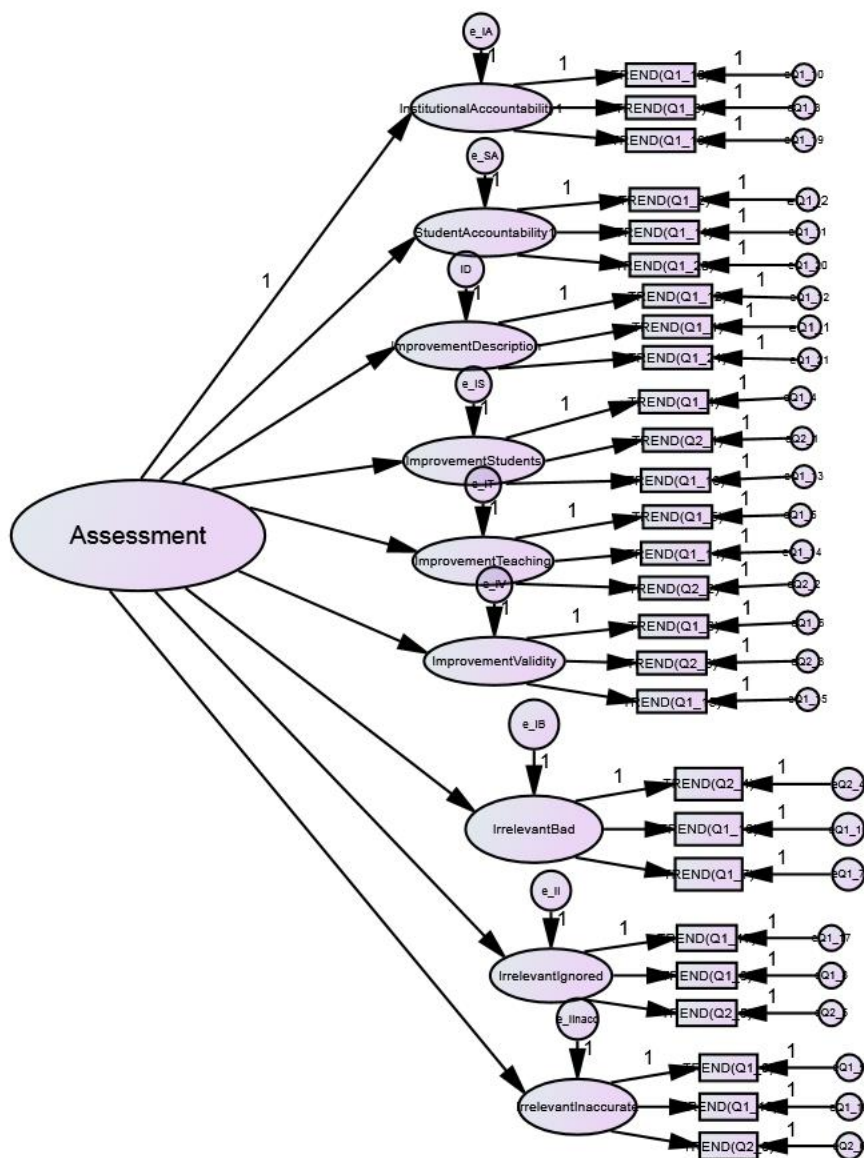
MODEL 3b: TWO SECOND-ORDER FACTORS WITH NINE FIRST-ORDER FACTORS WITH CONSTRAINED ERROR TERMS



APPENDIX G

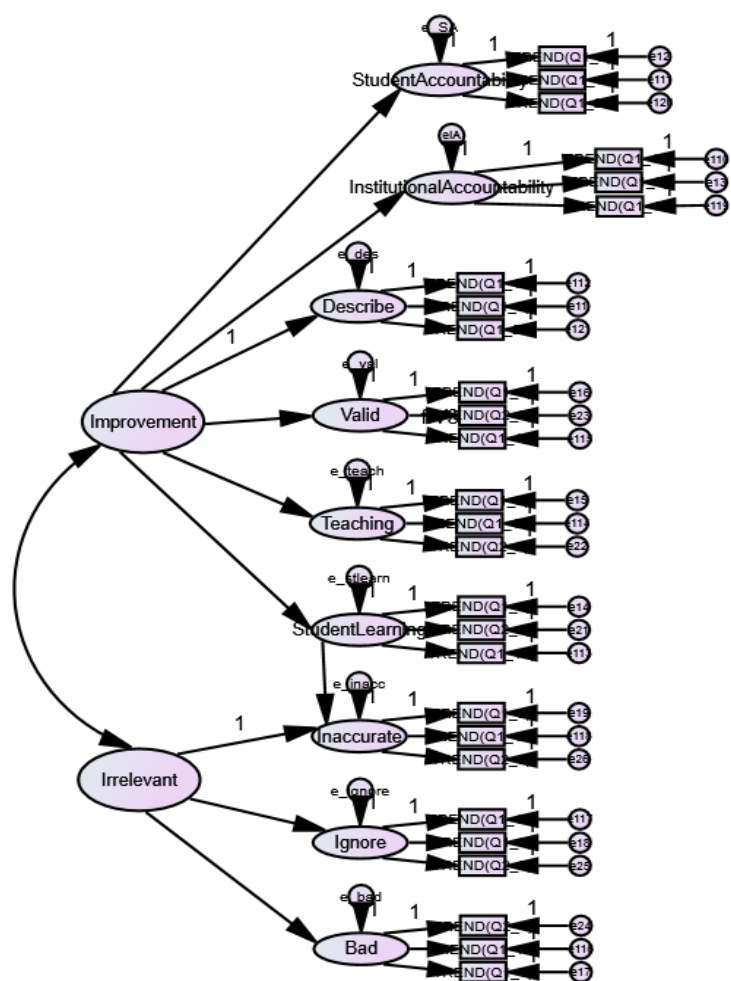
MODEL 3c: ONE SECOND-ORDER FACTOR WITH NINE FIRST-ORDER FACTORS

FACTORS



## APPENDIX H

## THREE SECOND-ORDER FACTORS WITH EIGHT FIRST-ORDER FACTORS



## REFERENCES

- Alexander, F. K. (2000). The changing face of accountability: Monitoring and assessing institutional performance in higher education. *The Journal of Higher Education, 71*(4), 411-431.
- Anderson, H. M., Moore, D. L., Anaya, G., & Bird, E. (2005). Student learning outcomes assessment: A component of program assessment. *American Journal of Pharmaceutical Education, 69*(2), 256-268.
- Association to Advance Collegiate Schools of Business. (2012). *AACSB standards 2012 update*. Retrieved from: <http://www.aacsb.edu/accreditation/standards/>
- Association of American Colleges and Universities. (2008). *Our students' best work: A framework for accountability worthy of our mission*. Retrieved from [www.aacu.org](http://www.aacu.org)
- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes, 50*, 179-211.
- Baker, E. L., O'Neil, H. F., & Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist, 48*(2), 1210-1218.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.
- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran, *Encyclopedia of human behavior, 4*, 71-81. New York, NY: Academic Press.
- Boulton-Lewis, G. (1994). Tertiary students' knowledge of their own learning and a SOLO taxonomy. *Higher Education, 28*, 387-402.

- Boyce, E. G. (2000). Development of an educational assessment plan for doctor of pharmacy programs. *Journal of Pharmacy Practice, 13*(5), 411-419. doi: 10.1106/3Q28-NVJ7-6QDK-67HL
- Brookhart, S. M., & Bronowicz, D. L. (2003). 'I don't like writing. It makes my fingers hurt': Students talk about their classroom assessments. *Assessment in Education, 10*(2), 221-242.
- Brown, G. T. L. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education, 11*(3), 301-318. doi: 10.1080/0969594042000304609
- Brown, G. T. L. (2006). Teachers' conceptions of assessment: Validation of an abridged instrument. *Psychological Reports, 99*, 166-170.
- Brown, G. T. L. (2008). *Conceptions of assessment: Understanding what assessment means to teachers and students*. New York, NY: Nova Science Publishers.
- Brown, G. T. L. (2010). Assessment: Principles and Practice. In R. Cantwell, & J. Scevak, *An Academic Life: A Handbook for New Academics*. Melbourne, Australia: ACER Press.
- Brown, G. T. L. (2011). *Leading school-based assessment for educational improvement: Rethinking accountability*. Keynote address presented to the Third International Conference 'Excellence in School Education', Indian Institute of Technology, New Delhi, India.
- Brown, G. T. L. (n.d.). *Self-regulation of assessment beliefs and attitudes: A review of the Students' Conceptions of Assessment inventory*. Unpublished manuscript, The Hong Kong Institute of Education, Hong Kong, China.

- Brown, G. T. L., & Harris, L. (2012). Student conceptions of assessment by level of schooling: Further evidence for ecological rationality in belief systems. *Australian Journal of Educational and Developmental Psychology*, 12, 46-59.
- Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy and Practice*, 15(1), 3-17.
- Brown, G. T. L., & Lake, R. (2006). *Queensland Teachers' Conceptions of Teaching, Learning, Curriculum and Assessment: Comparisons with New Zealand Teachers*. Paper presented at the Annual Conference of the Australian Association for Research in Education (AARE). Adelaide.
- Brown, G. T. L., Lake, R., & Matters, G. (2011). Queensland teachers' conceptions of assessment: The impact of policy priorities on teacher attitudes. *Teaching and Teacher Education*, 27, 210-220.
- Brown, S. (2004-2005). Assessment for learning. *Learning and Teaching in Higher Education*, 1, 81-89.
- Brownstein, B. (2001). Collaboration: The foundation of learning in the future. *Education*, 22(2), 240-247.
- Burger, D. (n.d.). *Assessment and accountability*. U.S. Department of Education. Honolulu, HI: Pacific Resources for Education and Learning.
- Byrne, B. (2009). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Cassidy, S. (2006). Learning style and student self-assessment skill. *Education and Training*, 48(2/3), 170-177.



- Chen, J., Brown, G. T. L., Hattie, J. A. C., & Millward, P. (2012). Teachers' conceptions of excellent teaching and its relationships to self-reported teaching practices. *Teaching and Teacher Education, 28*, 936-947.
- Covington, M. V. (1984). The self-worth theory of achievement motivation: Findings and implications. *The Elementary School Journal, 85*(1), 4-20.
- Crooks, T. (2001). *The validity of formative assessments*. Retrieved from Education-Line: <http://www.leeds.ac.uk/educol/documents/00001862.htm>
- Darling-Hammond, L. (2003). Standards and assessments: Where we are and what we need. *Teachers College Record*. Retrieved from: <http://www.tcrecord.org>
- Deenen, C. C., Brown, G. T. L., Lam, B. H., & Tsui, K. T. (2012). *Student evaluation of curriculum content for effective beginning teaching*. Paper presented at the 2012 AERA Annual Meeting of the American Educational Research Association conference, Vancouver, B.C., Canada.
- Derry, G. N. (1999). *What science is and how it works*. Princeton, NJ: Princeton University Press.
- Dewey, J. (1933). *How we think*. Boston, MA: D.C. Heath.
- Dwyer, C. A., Millett, C. M., & Payne, D. G. (2006). *A Culture of Evidence: Postsecondary Assessment and Learning Outcomes, Recommendations to Policymakers and the Higher Education Community* [Issue paper]. Princeton, NJ: ETS.
- Ekeblad, E., & Bond, C. (1994). The nature of a conception: questions of context. In R. & Ballantyne, *Phenomenography: philosophy and practice*, 343-353. Brisbane, Australia: Center for Applied Environmental and Social Education Research.

- Elementary and Secondary Education Act of 1965 (P.L. 89-10). (1965). Retrieved from U.S. Department of Education website:  
<http://www.scribd.com/doc/49149656/Elementary-and-Secondary-Education-Act-of-1965>
- El-Khawas, E. (2001). *Accreditation in the USA: Origins, developments and future prospects*. International Institute for Educational Planning. Retrieved from [www.unesco.org/iiep](http://www.unesco.org/iiep)
- Fletcher, R. B., Meyer, L. H., Anderson, H., Johnston, P., & Rees, M. (2011). Faculty and Students Conceptions of Assessment in Higher Education. *Springer Science+Business Media B.V.*, 64, 119-133. doi: 10.1007/s10734-011-9484-1
- Fodor, J. A. (1998). *Concepts: Where cognitive science went wrong*. Oxford, UK: Clarendon Press.
- Fullen, M. G. (1993). *Change forces: Probing the depth of educational reform*. London, UK: Falmer Press.
- Garfield, J. B. (1994). Beyond testing and grading: Using assessment to improve student learning. *Journal of Statistics Education*, 2(1). Retrieved from:  
<http://www.amstat.org/publications/jse/v2n1/garfield.html>
- Glaserfeld, E. V. (1989). Cognition, construction of knowledge, and teaching. *Synthese*, 80(1), 121-140.
- Harlen, W. (2007). *Assessment of Learning*. London, UK: Sage Publications Ltd.
- Hauser, R. M. & Koenig, J. A. (2011). Committee for Improved Measurement of High School Dropout and Completion Rates: Expert Guidance on Next Steps for Research and Policy Workshop, National Research Council, National Academy of

- Education. *High school dropout, graduation, and completion rates: Better data, better measures, better decisions*. Washington, DC: The National Academies Press. Retrieved from: [http://www.nap.edu/catalog.php?record\\_id=13035](http://www.nap.edu/catalog.php?record_id=13035)
- Henkel, M. (1997). Teaching quality assessments: Public accountability and academic autonomy in higher education. *Evaluation, 3*, 9-23.
- Higher Education Act of 1965 (P.L. 89-329). (1965). Retrieved from U.S. Department of Education website: [http://www.house.gov/legcoun/Comps/HEA65\\_CMD.pdf](http://www.house.gov/legcoun/Comps/HEA65_CMD.pdf)
- Hinchey, P. H. (2010). *Getting Teacher Assessment Right: What Policymakers Can Learn from Research [Brief]*. Boulder, CO: National Education Policy Center. Retrieved from <http://nepc.colorado.edu/publication/getting-teacher-assessment-right>
- Holt, D. G., & Willard-Holt, C. (2000). Let's get real: Students finding authentic corporate problems. *Educational Technology Research and Development, 45*(1), 65-94.
- Kelly, G. A. (1991). *The psychology of personal constructs: A theory of personality*. (Vol. 1). London, UK: Routledge.
- Lakoff, G., & Johnson, M. (2003). *Metaphors we live by*. Chicago, IL: Chicago University Press.
- Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(4), 4-16. doi: 10.3102/0013189X029002004
- Linn, R. L. (2005). Conflicting demands of no child left behind and state systems: Mixed messages about school performance. *Education Policy Analysis Archives, 13*(33). Retrieved from: <http://epaa.asu.edu/epaa/v13n33/>

- Louis, K. S., Febey, K., & Schroeder, R. (2005). State-mandated accountability in high schools: Teachers' interpretations of a new era. *Educational Evaluation and Policy Analysis*, 27(2), 177-204. doi: 10.3102/01623737027002177
- Liu, B. (1995). Some ideas about quality education. *People's Education Beijing*, 7, 8-12.
- Marks, H. M., & Printy, S. (2003). Principal leadership and school performance: An integration of transformational and instructional leadership. *Educational Administration Quarterly*, 39, 370-397.
- Marton, F., Dall'Alba, G., & Beaty, E. (1993). Conceptions of learning. *International Journal of Educational Research*, 19, 277-300.
- Mertler, C. A., & Vannatta, R. A. (2001). *Advanced and multivariate statistical methods: Practical application and interpretation*. Los Angeles, CA: Pyrczak
- Meyer, R. H. (1996). Comments on chapters two, three, and four. In H. F. Ladd (Eds.), *Holding schools accountable: Performance-based reform in education* (pp. 137-145). Washington, DC: The Brookings Institution.
- Meyer, L. H., Davidson, S., McKenzie, L., Rees, M., Anderson, H., Fletcher, R., & Johnston, P. M. (2010). An investigation of tertiary assessment policy and practice: alignment and contradictions. *Higher Education Quarterly*, 64(3), 331-350. doi: 10.1111/j.1468-2273.2010.00459
- Meyers, G. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research*. Thousand Oaks, CA: Sage.
- Moni, K. B., van Kraayenoord, C. E., & Baker, C. D. (2002). Students' perceptions of literacy assessment. *Assessment in Education: Principles, Policy and Practice*, 9(3), 319-342.

National Council for Accreditation of Teacher Education. (2007). *Professional standards for the accreditation of teacher preparation institution*. Retrieved from:

<http://www.ncate.org/Standards/NCATEUnitStandards/tabid/123/Default.aspx>

National Center for Higher Education Management Systems. (2008). Retrieved from:

<http://higheredinfo.org/dbrowser/index.php?submeasure=119&year=2008&level=nation&mode=data&state=0>

National Commission on Accountability in Higher Education. (2005). *Accountability for better results: A national imperative for higher education*, SHEEO. Retrieved from [www.sheeo.org](http://www.sheeo.org)

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Retrieved from:

[http://datacenter.spps.org/uploads/SOTW\\_A\\_Nation\\_at\\_Risk\\_1983.pdf](http://datacenter.spps.org/uploads/SOTW_A_Nation_at_Risk_1983.pdf)

National Council of Teachers of Mathematics. (NCTM). (1995). *Assessment standards for School Mathematics*. Reston, VA: NCTM.

No Child Left Behind Act of 2001 (P.L. 107-110). § 115, Stat 1425 (2002). Retrieved from U.S. Department of Education website:

<http://www2.ed.gov/policy/elsec/leg/esea02/107-110.pdf>

O'Donovan, B., Price, M., & Rust, C. (2008). Developing student understanding of assessment standards: a nested hierarchy of approaches. *Teaching in Higher Education, 13*(23), 205-217.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Clearing up a messy construct. *Review of Educational Research, 62*(3), 307-332.

- Peck, C. A., Gallucci, C., & Sloan, T. (2010). Negotiating implementation of high-stakes performance assessment policies in teacher education: From compliance to inquiry. *Journal of Teacher Education, 61*, 451-463.
- Peterson, M. W., & Augustine, C. H. (2000). External and internal influences on institutional approaches to student assessment: Accountability or improvement? *Research in Higher Education, 41*(4), 443-479.
- Popham, W. J. (2001). Teaching to the test? *Educational Leadership, 58*(6), 16-20.
- Popham, W. J., Cruse, K. L., Rankin, S. C., Sandifer, P. D., & Williams, P. L. (1985). Measurement-driven instruction: It's on the road. *The Phi Delta Kappan, 66*(9), 628-634.
- Pounder, J. S. (2000). Evaluating the relevance of quality to institutional performance assessment in higher education. *Evaluation, 6*(1), 66-78. doi: 10.1177/13563890022209127
- Pratt, D. D. (1992). Conceptions of teaching. *Adult Education Quarterly, 42*(4), 203-220.
- Prosser, M., Trigwell, K., & Taylor, P. (1994). A phenomenographic study of academics' conceptions of science learning and teaching. *Learning and Instruction, 4*, 217-231.
- Ramsden, P. (2003). *Learning to teach in higher education*. 2nd ed. London, UK: Routledge Falmer.
- Rhodes, L. K., & Bellamy, G. T. (1999). Choices and consequences in the renewal of teacher education. *Journal of Teacher Education, 50*(1), 17.
- Romzek, B. S. (2000). Dynamics of public accountability in an era of reform. *International Review of Administrative Sciences, 66*(1), 21-44.

- Rust, C. (2002). The impact of assessment on student learning: How can the research literature practically help to inform the development of departmental assessment strategies and learner-centred assessment practices? *Active Learning in Higher Education*, 3, 145-158. doi: 10.1177/1469787402003002004
- Rust, C., O'Donovan, B., & Price, M. (2005). A social constructivist assessment process model: How the research literature shows us this could be best practice. *Assessment and Evaluation in Higher Education*, 30, 231-240.
- Schiller, K. S., & Muller, C. (2003). Raising the bar and equity? Effects of state high school graduation requirements and accountability policies on students' mathematics course taking. *Educational Evaluation and Policy Analysis*, 25(3), 299-318.
- Schilling, K. M., & Schilling, K. L. (1998). Proclaiming and sustaining excellence: Assessment as a faculty role. *ASHE-ERIC Higher Education Report*, 26(3).
- Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher*, 29(7), 4-14.
- Smith, W. L., & Zhang, P. (2009). Students' perceptions and experiences with key factors during the transition from high school to college. *College Student Journal*, 43(2), 643-657.
- Southern Association of Colleges and Schools Commission on Colleges. (2005). *Resource Manual for the Principles of Accreditation*. Retrieved from: <http://www.sgc.edu/president/departments/IE/files/SACSPinciplesofInstitutionalEffectiveness.pdf>

- Southern Association of Colleges and Schools Commission on Colleges. (2012). *The principles of accreditation: Foundations for quality enhancement* (5th ed). Retrieved from: <http://www.sacscoc.org/pdf/2012PrinciplesOfAccreditation.pdf>
- Spillane, J. P., Diamond, J. B., Burch, P., Hallett, T., Jita, L., & Zoltners, J. (2002). Managing in the middle: School leaders and the enactment of accountability policy. *Educational Policy*, 16, 731-762. doi: 10.1177/089590402237311
- Struyven, K. D., Dochy, F., & Janssens, F. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment and Evaluation in Higher Education*, 30(4), 127-146.
- Thompson, A. G. (1992). Teachers' beliefs and conceptions: A synthesis of the research. In D. A. Grouws, *Handbook of research on mathematics teaching and learning*, 127-146. New York, NY: Macmillan.
- Torrance, H., & Pryor, J. (1998). *Investigating formative assessment: teaching, learning and assessment in the classroom*. Buckingham, UK: Open University Press.
- Trow, M. (1996). Trust, markets and accountability in higher education: A comparative perspective. *Higher Education Policy*, 9(4), 309-324.
- Tynjala, P. (1997). Developing education students' conceptions of the learning process in different learning environments. *Learning and Instruction*, 7(3), 277-292.
- U.S. Department of Education. (1998). 1998 Amendments to the Higher Education Act of 1965 (P.L. 105-244). Retrieved from: <http://www2.ed.gov/policy/highered/leg/hea98/HR6.pdf>
- U.S. Department of Education. (2006). *A Test of Leadership: Charting the Future of U.S. Higher Education*. Washington, D.C.



- Vosniadou, S. (1992). Knowledge acquisition and conceptual change. *Applied Psychology: An International Review*, 41, 347-357.
- Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction*, 4, 45-69.
- Vygotsky, L. (1978). *Interaction between learning and development*. Cambridge, MA: Harvard University Press.
- Waite, D., Boone, M., & McGhee, M. (2001). A critical sociocultural view of accountability. *Journal of School Leadership*, 11, 182-203.
- Wang, X., & Hurley, S. (2012). Assessment as a scholarly activity? Faculty perceptions of a willingness to engage in student learning assessment. *The Journal of General Education*, 61(1), 1-15.
- Warren, E., & Nisbet, S. (1999). The relationship between the purported use of assessment techniques and beliefs about the uses of assessment. *22nd Annual Conference of the Mathematics Education and Research Group of Australia*, 515-521. Adelaide.
- Webb, N. L. (1992). Assessment of students' knowledge of mathematics: Steps toward a theory. In D. A. Grouws (Eds.), *Handbook of research on mathematics teaching and learning* (pp. 661-683). New York, NY: Macmillan.
- Weiner, B. (1972). *Theories of motivation: From mechanism to cognition*. Chicago, IL: Markham.
- Welsh, J. F., & Metcalf, J. (2003). Faculty and administrative support for institutional effectiveness activities: A bridge across the chasm? *The Journal of Higher Education*, 74(4), 445-468.

- White, R. T. (1994). Commentary: Conceptual and conceptional change. *Learning and Instruction, 4*(1), 117-121.
- Williams, C. (2007). Research Methods, *Journal of Business and Economic Research, 5*(3), p. 65-72.
- Wolf, B. (1990). The assessment movement in higher education: Implications for special education teacher preparation programs. *Teacher Education and Special Education, 13*(3-4), 246-251.