HUMANITIES PROFESSORS’ CONCEPTIONS OF ASSESSMENT IN GENERAL EDUCATION

A doctoral thesis presented
By

Chad M. Rohrbacher

to
The School of Education

In partial fulfillment of the requirements for the degree of
Doctor of Education

in the field of
Education

College of Professional Studies
Northeastern University
Boston, Massachusetts
May 2015
Abstract

Assessment of general education is a complex process involving faculty and administrators from across campus and often is met with indifference or resistance. The purpose of this qualitative case study was to explore how humanities professors conceive of and use assessment in their general education courses. After collecting data from public documents and conducting interviews with seven faculty at two state universities, descriptive and structural coding was used to identify emergent themes which was followed by pattern matching to isolate its deep structure. Brown’s (2008) Conceptions of Assessment was used as the theoretical framework and data analysis resulted in themes and subthemes in the areas of assessment for learning, assessment as learning, assessment of learning, and assessment as irrelevant to learning. Participants described their experiences in general education assessment noting that their rationale behind assessment and its uses was complex and varied; however, all participants desired to use the data to improve student learning. Typically, the practical use of the assessment data beyond reporting was absent. The lack of shared governance in the process negatively affects faculty engagement with general education assessment. Faculty and administrators in higher education can find relevance in this study’s findings to inform general education structures and uses of general education assessment processes, to address some of the ambiguity around humanities learning outcomes, to assist teachers in teaching and learning, and administrators in garnering faculty buy-in and support for assessment initiatives. Additional research in humanities professors’ conceptions and uses of general education assessment is discussed.

Keywords: General education, general education assessment, humanities, conception of assessment, higher education assessment, teaching and learning.
Acknowledgements

Charles Mingus once said, “Making the simple complicated is commonplace; making the complicated simple, awesomely simple—that’s creativity.” My wife, Melanie, is the most creative person I know. She took on the complicated day-to-day experience of running a household and she made it seem utterly simple. The endless hours I spent researching and writing, she shouldered the responsibility of making our household speed along. She not only made this experience possible, but she was instrumental in its success. I also want to thank my daughters, Addison, Marin, and Katryn. They remained patient and understanding over the past several years and supported my efforts even when it might have meant being quiet on a Saturday morning or foregoing that sleepover so that daddy could work. When I was most stressed, they reminded me to stay balanced and to laugh. My grandmother, Evelyn Rohrbacher, also deserves a special mention. There was never a phone call that didn’t start with a question about my studies or end with a question asking me if I needed anything so I could finish. My mother, Joan Alverado, also encouraged me every step of the way. My family’s support kept me motivated and focused, and for that I am eternally grateful.

My thesis advisor, Dr. Corliss Brown-Thompson, was downright fantastic. As a teacher in higher education who is interested in teaching and learning, I have an idea of the amount of time and effort it takes to be successful. I found myself wanting to emulate Dr. Thompson after every encounter with her. I could not have asked for a more thorough, thoughtful, helpful person to assist me on this intellectual journey. Dr. Conn as my second reader asked great questions and prodded me to examine different perspectives. As my third reader, Dr. Robert Drake brought his wealth of practical assessment experience to the table and I am grateful he agreed to be a part of my educational experience.
I would like to thank Beth Kaufka, Dr. Joseph Graves, and Dr. Jim Crawford for their support, advice, encouragement, and friendship. Jen Bourassa was my partner in crime as we moved through the process together and I appreciate her insights and friendship. Rick Fisher deserves special thanks for his editing and proofreading assistance.
Table of Contents

Chapter One: The Research Problem ........................................................................7
Statement of the Problem ......................................................................................9
Justification for the Research Problem 11
Deficiencies in the Evidence 13
Significance of Research Problem 14
Research Questions ...............................................................................................
Positionality Statement 17
Theoretical Framework 19
Summary .............................................................................................................23

Chapter Two: Literature Review ...........................................................................24
Definition of Terms & Scope of Review 25
General Education ..................................................................................................27
Assessment ...........................................................................................................30
  Data Collection ...................................................................................................32
  Assessment of General Education ......................................................................33
  Findings from Assessment in General Education .............................................34
  Challenges in General Education Assessment ..................................................35
  Assessment Challenges in Humanities Courses 36
Conceptions of Assessment ....................................................................................38
Summary of Literature Review .............................................................................52

Chapter Three: Research Design ...........................................................................54
Methodology ..........................................................................................................55
Site and Participants ............................................................................................59
Data Collection .....................................................................................................67
Data Storage ..........................................................................................................68
Data Analysis .........................................................................................................69
Trustworthiness .....................................................................................................72
Validity ..................................................................................................................73
Reliability 73
Protection of Human Subjects ...............................................................................75
Summary ..............................................................................................................76

Chapter Four: Findings and Analysis ...................................................................53
53
.................................................................................................................................54
.................................................................................................................................56
.................................................................................................................................58
.................................................................................................................................63
65
.................................................................................................................................66
.................................................................................................................................70
.................................................................................................................................74
Chapter Five: Discussion and Implications for Practice

Recommendations for Practice

Recommendations for Future Research

Summary

Conclusion

References

Appendix A – IRB Approval

Appendix B – IRB Modifications

Appendix C – Recruitment Email Survey

Appendix D – Recruitment Email Interview

Appendix E – Participant Unsigned Consent Survey

Appendix F – Participant Informed Consent

Appendix G – Interview Protocol
CHAPTER 1: INTRODUCTION

In higher education, multiple stakeholders such as students, parents, administrators, and legislators, are pressing for evidence of student learning (see Lederman 2008; Spellings Commission, 2006; Malandra, 2008, Miller & Malandra, 2006), especially in subjective areas where it may be difficult to quantify particular skills such as critical thinking or written communication. A member of the 2006 Spellings Commission, argues that it is imperative to translate those complex subjective skills into quantifiable numbers that stakeholders can easily understand (Lederman, 2006) so that they may be useful to all. Many of these subjective skills are nurtured in the humanities disciplines such as history, art, English, foreign language, religion, performing arts, music, and philosophy, among others.

Assessment of these subjective skills in general education is particularly difficult yet important to all university stakeholders. According to the Association of American Colleges and Universities, the goal of general education is to develop a student’s critical thinking, communication, social responsibility, and global citizenship skills, among other things (AAC&U, 2005; 2007). Typically, general education courses are required courses in a variety of disciplines that students share in their first and second year and many are in the humanities. In fact, general education courses command approximately 46 credit hours (Hart, 2009) or 30% of a student’s collegiate experience (Brint, Proctor, Murphey, Turk-Bicakci, & Hannerman, 2009). These courses are “the core of the undergraduate curriculum for all students regardless of major” (Allen, 2006, p. 1).

Assessing student learning is generally broken into three distinct theoretical areas dependent upon how the collected student data is used. These areas include assessment of, as, and for learning. Assessments of learning are part of institutional quality assurance and
accountability processes to validate the awarding of credentials (via grades, certificate, test, etc.) as well as the quality of disciplinary course offerings, professional training, and the student experience (Kuh, Kinzie, Shuh, & Whitt, 2005; Maki 2004; Ramsden 2003). Assessment as learning puts the onus on students by asking them to reflect on how well they are doing and how they can improve their learning (Carless et al. 2007; Gibbs 2006; Gibbs and Simpson 2004). Assessment for learning suggests assessment ultimately informs teaching by highlighting what is effective and ineffective in students achieving academic success and is most helpful if used for meaningful change or reform (Penn, 2011; Banta, Jones, & Black, 2009; Huba & Freed, 2000).

From my own experience as a professor, higher education institutions have engaged the assessment movement on many fronts. Internally, departments and programs are developing new tools and implementing new ways to assess student learning while accrediting bodies revise and clarify what institutions need. Externally, organizations like National Institute of Learning Outcomes Assessment, AAUP, and AACU share resources, run workshops and conferences, and publish scholarly work on assessment to promote its use and assist institutions in the process.

While the call for assessment is strong and the trends on campuses include more assessment being completed especially in general education, resistance remains. In the humanities, faculty resistance to assessment seems considerable (Clark and Filinson, 2009; Scobey, 2009; Brottman, 2009). There are some voices like Gerald Graf (2008), president of the Modern Language Association, who fully supports outcomes assessments, specifically shifting the way academics think from a teaching perspective to a learning one; however, there are many others feel the assessment movement is anathema to their work. Cary Nelson (2011), president of the American Association of University Professors wrote, “I am opposed to this (assessment) movement and to everything for which it stands,” (para. 1). He continued, “There are those who
urge us to seek compromise with the assessment and quantifiable-outcomes movement. Let’s own it and do it right, they urge. Not for what I do. Not for what I teach” (para. 8).

In part, this resistance seems to stem from the philosophical and pedagogical underpinnings of many humanists’ considerations of knowledge and education. As Blake and Ottoson (2009) note, it is important that evaluators make assumptions about knowledge use explicit, including assumptions by stakeholders and the evaluator (Blake & Ottoson, 2009). Further, they claim knowledge utilization serves as the natural starting point for assessment or evaluation. However, currently there is little systematically collected data in the assessment literature that describes how humanities professors conceive of and use general education assessment data.

Humanities professors’ views of assessment affect how they approach classroom and learning, but also affects how they react to calls for assessment. In fact, there may be a disconnect or a divide between how they use assessment and how they perceive the administration’s need for data to “prove” learning took place. The research problem under review is a thorough exploration of the views of assessment by humanists in general education to assist in bridging the divide. This study is designed to explore those issues, guided by the research questions: How do humanities professors perceive general education assessment? In what ways do they use the assessment data? What is the rationale behind their use of assessment data?

**Research Problem Statement**

While general education in higher education has taken on many incarnations (see Stevens, 2001; Bowen, 2004; Allen, 2006), in every manifestation all students are required to take general education courses regardless of major; therefore, it has been a chief focus for change
in assessment practices. Indeed, the Spellings Commission (2006) argued for, among other things, institutional accountability and proven student achievement specifically through “value added” measures. The Spelling Commission had a direct effect on institutional direction and focus in terms of assessment. For example, in one AAUP survey 89 percent of the 433 Chief Executive Officers stated that their institutions were in the process of evaluating or modifying their general education programs with most focusing on common learning outcomes (Hart Research and Associates, 2009).

How institutions use that collected data is also a significant issue. In a National Institute for Learning Outcomes Assessment report, provosts and academic officers from over 1500 institutions reported that assessment at their institutions was most commonly used for accreditation (Kuh & Ickenberry, 2009). However, as noted in the literature, data collection needs to be tied to reform for student learning if it is to be helpful (Banta, Jones, & Black, 2009; Huba & Freed, 2000). Data in such a way can be described as “closing the feedback loop” (Wehlburg, 2007; Watson, 2003). Vincent Tinto claimed there was plenty of information on student success, but what was missing was the ability to transform it into practical knowledge (as cited in Malandra 2008). Blaich and Wise (2011) substantiated Tinto’s claim, stating that part of the problem is that many institutions may have a hard time analyzing what the data means and how to use it effectively. These findings are not new (Olson, 2003; Reeves, 2007). Essentially, the apparent lack of using concrete data to assist in teaching and learning keeps curricular reform in stasis (Popham, 2006a, 2006b; Gaff, 1980). If universities do not use data to inform teaching and learning practices, there may be little student improvement in general education learning outcomes as Gaff (1980) noted over 30 years ago. Ultimately, then, how humanities general
education assessment data is perceived and used may have significant effects on both student and institutional success.

**Justification for the Research Problem**

Current data collection on student learning is widespread and thorough. Typically, accrediting bodies like The Schools and Middle States Commission on Higher Education require data beyond reported grades, such as evidence of students meeting specific learning outcomes (http://www.msche.org/). In their national report on the current state of data collection and use, Kuh and Ickenberry (2009) noted that most colleges and universities collected multiple types of data. Blaich and Wise (2011) found similar results and showed a trend of student learning evidence being collected. After the data was collected, Ewell, Paulson, and Kinzie (2011) found the primary use of assessment results at the program level is for program review, followed by instructional improvement, and institutional accreditation.

While many higher education faculty recognize the importance and utility of assessment, there are still deeply held convictions within academe. For example, Sadler (2005) related the conception of assessment as evaluating a student’s overall performance and generating assumptions about her learning including the quality or achievement in tests, projects, reports, etc. On the other hand, Dunn and Mulvenon (2009) showed how formative assessments facilitated improvement in instruction, methodology, curriculum, and assisted in student performance.

More broadly, though, faculty have many reservations when considering assessment. Many chairs and administrators view the idea of using assessment for learning as worthwhile, but too time consuming to warrant pursuing (Blaich & Wise, 2011; Wang & Hurley, 2012; Kuh & Ickenberry, 2009). The lack of time, resources, and support for using assessment for learning
is widespread and disconcerting. Other studies noted that, in general, faculty believe pursuing assessment strategies would not benefit them in professional goals such as tenure, etc. (Jones, 2010; Gaff & Wasecha, 2001; Gillmore, 2004).

While administrators recognize the importance of faculty involvement in developing assessment protocols (Hutchins, 2010; Walvoord, 2004), they find that faculty lack engagement in the process (Kuh & Ickenberry, 2009). Some of this may be explained by Hutchins’ (2010) delineation of faculty perspectives about assessment including that it:

- Is purely an administrative activity;
- Is something required by outside stakeholders (government, accrediting bodies, etc.);
- Is completely separate from what actually takes place in the classroom;
- Is part of a management culture influencing academe;
- Is something that is not supported in higher education as evidenced by lack of training;
- And has unclear benefits to faculty (for example, it is not valued or rewarded in faculty performance reviews).

In other words, Hutchins (2010) suggests, faculty across all disciplines may resist assessment because they view it as an unnecessary task that neither benefits faculty nor is the responsibility of faculty.

When considering humanities faculty specifically, information on the views of assessment is limited. For the American Academy of Arts and Sciences, Townsend (2010) describes results from a study from 1044 department chairs representing six humanities disciplines and found that about two-thirds of the humanities departments are using some method to assess overall undergraduate student learning. While not focused on assessment, the survey
results indicated that assessment practices including portfolio building and standardized tests had penetrated the practices of humanities disciplines with between 48 and 75 percent of the departments in each discipline reported employing them. Further, literature by professional associations (Graf, 2008), departments (Borish, Kechum, Lyon-Jennes, 2009), and individual faculty (Savagian, 2009; Cole & Demaio, 2009; Murphy, 2010) describe the process of collecting, analyzing, and using data for a variety of purposes, but do not explore how they conceive of assessment or use the data after it is collected.

**Deficiencies in the Evidence**

While institutions are paying more attention to general education, many of which are humanities courses, and assessment in those courses, an AACU survey indicated that the assessment of general education over multiple courses remains a concern with just over half (52%) of the respondents collecting and using that data (Hart & Associates, 2009). It is currently unclear if the collected general education assessment data is actually used to improve student learning and development.

Although the intent of assessment is to focus on and improve learning for students, little is known regarding humanities professors’ conceptions of assessment, particularly in their general education courses. Because faculty are responsible for assessment in their courses, it is important to know what they believe about assessment especially given that almost half of general education courses are not being assessed (Hart & Associates, 2009). Specific faculty and student perceptions and uses of assessment data were addressed in a few international studies (Iqbal & Azam, 2009; Fletcher, Meyer, Anderson, Johnson, & Rees, 2011; Carless, 2006), but there is nothing in the literature concerning United States professors. It is noteworthy that all three studies came to the same conclusion: faculty and students were not in agreement
concerning assessment data and how it is used to facilitate learning. Though the studies provide faculty perception, given the difference in context between countries, an understanding of U.S. professors’ beliefs on assessment is still missing. Additionally, the samples were about professors across all disciplines, and they were not focused on general education or humanities professors.

There is a lack of research on humanities professors’ beliefs about assessment and there is also a gap concerning how humanities professors actually use collected student data in general education courses and their rationale for using data in those ways. A few studies focused on chairs or other administrators’ perceptions of what assessments were being used and for what purpose(s) (Blaich & Wise, 2011; Wang & Hurley, 2012; Kuh & Ickenberg, 2009; Townsend, 2010). However, since faculty were not included, it is unclear if they share their chairs’ perceptions. There is a need for an understanding of how humanities, faculty members in the U.S. view and use assessment.

**Significance of Research Problem**

All higher education stakeholders are affected by a lack of understanding about how humanities professors who teach general education view and use assessment. For example, faculty may miss multiple opportunities to assist students, themselves pedagogically, and administratively. The single most positive influence on a student’s learning is the teacher (Wenglinsky 2002), and academic performance early on during college years is the single strongest predictor for degree attainment (Pascarella & Terenzini 1991, 2005) which points to the importance of strong instructors who are able to ensure that students learn necessary content. Ainsworth and Viegut claim that without systematic analysis to inform instruction, valuable information is untapped (as cited in Guskey, 2007). The untapped information may assist
teachers in helping each student meet his/her academic goals. Indeed, assessment for learning is a complex educational and social process that, when properly used, can help students improve their learning (Torrance & Pryor 2001). Ultimately, then, how data is used (or not used) can have significant effects on student and institutional success both inside and outside the classroom.

By merely reporting data, administration and faculty absolve themselves of responsibility and abandon students in the area of academic success. Reeves (2007) showed that when faculty and administration claimed they had “no control over student achievement”, 44 percent of students scored as proficient or higher; however, when faculty and administration believed they had “more influence” than those external factors such as the environment or parent involvement, student achievement was 65 percent (p. 246). Those teachers who took responsibility for their students’ learning instead of blaming something else recognized noteworthy gains in their students’ academic success. In other words, faculty and school administration who collect and use data to inform instruction, are able to improve student performance. Faculty and administrators both are responsible for assessment in higher education. While faculty are responsible for conducting courses, informed by assessment data, administrators are responsible for things such as program, departmental, institutional effectiveness, and accreditation.

Higher education administration can also suffer from a lack of knowledge concerning humanities professors and their views of assessment. As the Spellings Commission noted in 2006, student achievement is inextricably connected to institutional success. Retention from first to second year is a serious concern with many institutions implementing high impact practices to improve student success and retention. If administrators had a better understanding of how humanities professors viewed assessment, they may be able to better navigate its complexities to improve its policy development and implementation.
Even though faculty and administrators are key to conducting and using the data from assessment, the practice of assessment is most beneficial for higher education students. The National Association of Colleges and Employers study (Job Outlook, 2011) noted that employers look at coursework, but value student’s ability to work in teams, to communicate well, and to discuss and employ transferable critical thinking skills. Employers want students to identify and discuss the transferable skills gained in the classroom (Martin, Johnson, Hansen, Dangor, & Pietryla, n.d.). General education is the first, and sometimes only, opportunity students get to develop these transferable skills. For example, in first year rhetoric and composition, students get their first real exposure to close reading, analyzing, and synthesizing information and practice in communicating ideas clearly. When faculty and their institutions only use assessment as a means to rank the school and students, they miss out on the benefits such as curriculum reform, that focuses on learning outcomes (Guskey, 2006); they also miss out on the data regarding student needs, and support practices that assists them in improving over time (Stiggins, 2007).

The lack of quality assessment or use of data concerning student success or progress is also detrimental to their employment opportunities after a degree is earned. Consider data collected by a biology department that reveals the following: majors can restate memorized information extremely well; however, when students were tasked with synthesizing data and communicating analysis, assessment reveal gaps in their knowledge and skills. By not using the data to inform instruction, the faculty do a disservice to the students. If those students then graduate from the program believing they have the complex skills needed to succeed in the job market, but find they cannot complete their job requirements, it reflects poorly on the institution and may lead to the student losing gainful employment.
While all general education programs have had similar goals to create a well-rounded, educated individual, there is a pervasive divide between collecting data and understanding the various uses and purposes of it. An understanding of student learning, which can be obtained through assessment (Allen, 2006) is particularly important in general education courses. Given that many of the general education courses are in the humanities, but little is known about humanities professors’ perceptions of knowledge, it is important to explore humanities professors’ beliefs and uses of assessment. A more complete articulation of what comprises assessment and its uses may assist humanities faculty and administrators in understanding how assessment can be used to improve the learning experiences of students.

**Positionality Statement**

Currently I am an Associate Professor of English trained in literature and creative writing. As such, I have been trained by humanists in certain ways of reading, thinking, and interpreting texts, and my experiences influence the way I approach assessment. Helm (2000) shares that humanists are in the business of clarifying, making sense of, and articulating the meaning of texts. It is understood that the messiness of reality and complexities of human life cannot be reduced to answers on a test or quantified in a meaningful way. I understand the aversion to assessment, as some knowledge takes the form of knowing what questions to ask, and some growth is accomplished by taking steps backwards. However, I am also a teacher and a student. As a teacher and a student I also understand the invaluable role assessment plays in the construction of knowledge.

I believe assessment has a place in the humanities. Ideally, the methods for assessing are as diverse as the students being assessed. I am in a firm believer of formative assessment. I believe learning happens on a discipline’s highway as the student speeds through his/her courses,
and in the back alleys of books where the student inches along with ideas, following side roads that lead to conversations in classrooms and coffee shops, and gravel driveways that never take the student home. Assessment helps the student become a cartographer of his/her education and mark where s/he has been.

Generally, the researcher should acknowledge his or her values, experiences, biases, etc. and clearly indicate them within the study (Ponterotto, 2005; Creswell, 2013; Yin, 1994). As a researcher studying my own institution, I deem it important to situate myself within the organization. This type of research is considered “inside research” or sometimes referred to as “endogenous research” (Trowler, 2011). While being an insider might be considered advantageous because one is known, may obtain easy access to key figures and/or documents, has ready participants, and has previous knowledge of the organization (Asselin, 2003) and as Geertz argues, one is “…empowered to offer a thick description of lived realities” (as qtd. in Trowler, 2011), however, there are disadvantages as well. Trowler (2011) reports that it may be more difficult to get a neutral, ‘etic’ accounts, it may be harder to see some of the social realities because they are “normal” and your colleagues may change their responses due to preformed expectations of you as a professional. Further, he cites multiple researchers who detail the ethical and methodological issues interviewing those who are more powerful, or less powerful, than you and who may be your peers (2011).

As an associate professor in the humanities who has been tasked with working on assessment, it is my desire to learn how humanities professors view assessment, how they know or recognize that students have achieved proficiency or competency in selected criteria, and what they do with that data after it is collected. As such, the purpose of this study is to understand and explore the perceptions and uses of general education assessment by humanities professors. The
following research questions will guide my study: How do humanities professors perceive
general education assessment? In what ways do they use the assessment data? What is the
rationale behind their use of assessment data? These questions will be answered through the lens
of constructivist theory and Brown’s conception of assessment.

**Theoretical Framework**

Constructivist theory and Brown’s (2004, 2008) Conception of Assessment will provide
the theoretical framework for this study. In the broadest sense, constructivism explains how
learning takes place; learning is the process of constructing a conceptual framework (Cobern,
1993). Constructivism involves a meaningful negotiation and interpretation of one’s reality that
is influenced by one’s prior knowledge (Cobern, 1993). Brown’s Conception of Assessment is a
model that is based on the accepted uses for assessment: *for* learning (improvement of teaching
and learning), *as* learning (making students accountable), *of* learning (accountability of schools
and teachers), and *irrelevant to* learning (it can be safely ignored even if it must be used, and/or
it is inaccurate). The overall framework can be seen in figure 1.

![Theoretical framework for this study.](image)

**Figure 1.** Theoretical framework for this study.
Constructivism

Historically, knowledge was viewed as objective and stable. Essentially there was some correct answer that could be observed or arrived at. When considering general education, one might see more disciplines like those in science and math employing this view of knowledge. Oftentimes, instructors with this view of knowledge may rely on lecture to impart information and assessments such as multiple-choice exams to test students’ knowledge acquisition. Murphy, Sharp, and Whitelegg claim that when knowledge is viewed in this way and students are seen as passive recipients, any lack of learning is attributed to the innate qualities of the learner (as cited in Willis, 2007). The problem of viewing knowledge in this manner is that when the teacher places all the responsibility on the learner, there is no reason to use collected data to assist in revising one’s pedagogy or curriculum. However, Hood (2002) described a shift in the late 1970s when knowledge began to be considered more dynamic and constructed.

The constructivist philosophy focuses on learning and student knowledge creation versus the teaching. According to constructivist theory, knowledge is developmental, internally constructed, and socially and culturally mediated (Hood, 2002). Shepard (2000) explains that in traditional environments, learning is viewed as a mechanistic process of breaking knowledge into small units for students to absorb and memorize. Students who participate in a constructivist learning environment, however, assemble their own meanings of knowledge that depend on the social and cultural context of a learning situation. While young children may not be able to fully assemble meanings from experiences, older students like those in higher education have a wealth of educational and life experiences to draw upon to synthesize with new information to create knowledge. Based on my personal experience as a humanist, humanities professors view knowledge as dynamic and constructed. Describing humanities professors’ perceptions of how
knowledge is constructed will assist in analyzing how it informs their assessment choices concerning their disciplines.

While there are those in the humanities such as those who argue in support of new assessment initiatives (Graf, 2010), it seems that there is still considerable resistance to the assessment movement. In large part, this aversion many humanists have toward assessment is due to the very nature of knowledge and how it is constructed. Scobey (2009) explains, “the most important effects of a humanities education resist measurement” (para. 2). Scoby continues:

Humanists tend also to look askance at the abbreviated time-frame of many assessment tools, whether these tools test student performance at a single moment or mark change in a “formative-summative” sequence. To the contrary (our experience tells us), the most powerful learning in the humanities takes place in ways that are meandering, iterative, self-reflexive, and unpredictable. (para. 3)

Thus, for many humanists, assessment is philosophically conflicting at its very core. The conception of learning and knowledge in the humanities is as Scoby (2009) suggests, “resist[ant] to measurement” making it important to have a model that is broad and nuanced enough to describe and assess learning in the humanities. Using constructivism as a theory of learning will help guide the analysis of professors’ description of how they expect their students to build knowledge.

Conception of Assessment

Brown (2008) theorizes that conceptions “express human purposes and intentions, by categorizing in an organized manner the world around us…how groups of people with shared experiences understand, represent, and intend to act within a domain” (p. 7-8). In other words,
conceptions are malleable, so a professor’s perspective on assessment may change depending on his/her experience and current reality.

The study of instructors’ conceptions of assessment is important because evidence exists that instructors’ conceptions of teaching, learning, and curricula influence strongly how they teach and what students learn or achieve (Clark & Peterson, 1986; Calderhead, 1996). As previously stated, researchers have suggested that three major purposes for assessment exist: for learning (improvement of teaching and learning), as learning (making students accountable), and of learning (accountability of schools and teachers). In addition to the three conceptions, Brown (2004) included a fourth conception: assessment is fundamentally irrelevant to the life and work of teachers and students. This conception included the beliefs that assessment is bad for teachers and students, it can be safely ignored even if it must be used, and/or it is inaccurate. It was with this more complex categorization, that Brown’s Conception of Assessment (CoA) was identified as a strong model to classify humanities professors’ conceptions. Originally used for secondary education, Brown (2006) modified the model for higher education application. It was tested by Fletcher et al. (2011) with respectable success.

Some instruments allude to perceptions of assessments via how data is used (e.g. see The National Institute for Learning Outcomes Assessment survey, Institutional Climate for Student Assessment); however, Brown’s (2008) Conception of Assessment is the only model of its kind that specifically examines how professors perceive assessment and it offers a good starting point to identify the humanities professors’ perspectives. Implementing a model that has been previously used offers the study some reliability as a valid measure. Fletcher et al. (2011) claimed establishing factorial validity of the CoA was an important aspect of their study as it provided the basis for understanding how relationships function and how groups may differ in
those understandings. The difficulty, as McLellan (2004) and Iqbal and Azam (2009) found, is when the faculty espoused values concerning assessment differ from their actual use. This type of discrepancy is not easily identified and is not accounted for in the model. However, through use of inductive coding, I will be able to extract perceptions from within the framework and beyond the framework.

**Chapter Summary**

The way that an instructor approaches assessment reflects the teacher’s beliefs and assumptions about what it means to know or understand, and it therefore shapes the learner’s own beliefs about learning (Willis, 2007). These assumptions are often not held as fully informed practices, but can be “intuitive, rudimentary theories” (Black & Wiliam, 2006). Faculty construct their conceptions of assessments based on prior knowledge, not current reality. Faculty have been shown to have mixed opinions on the purposes of assessment based on their attitudes about teaching and learning (McLellan, 2004). Those who viewed teaching and learning as the transmission of knowledge were likely to view assessment as a method to test students’ ability to reproduce information. In contrast, those who saw teaching and learning as facilitating critical thinking viewed assessment as an integral part of the learning process (Samuelowicz & Bain, 2002). The concepts that comprise the framework for this study attempt to make humanities professors’ assumptions and beliefs about knowledge and assessment visible. After establishing the instructor’s construction and usage of assessment, Brown’s Conceptions of Assessment model (2004, 2008) should offer a strong theoretical framework for the proposed study.

Multiple stakeholders such as students, parents, administrators, and legislators, are pressing for evidence of student learning, especially in subjective areas where it may be difficult to quantify. With resources being tied to assessment results, the importance of a sound,
comprehensive assessment plan is necessary. However, for a meaningful assessment plan to emerge administrators need to garner faculty “buy in”. Anecdotal evidence suggests resistance to the assessment movement is particularly high in the humanities (Scobey, 2009). Faculty concerns about assessment need to be recognized. A qualitative study designed to better understand how humanities professors see assessment and use it in their respective courses may assist in both of those areas. Understanding how humanities general education assessment data is perceived and identifying whether data is used concerning student learning may assist teachers in teaching and learning, humanities departments in clarifying goals and rationales, and administrators in garnering faculty buy in and support for assessment initiatives.
Chapter 2: REVIEW OF SELECTED LITERATURE

Introduction

Both assessment and general education are complex and complicated in their construction and use. Both terms hold multiple meanings and purposes for various stakeholders. General education is a significant part of a student’s collegiate experience. While there are a number of editorials critiquing the current assessment movement as applied to humanities disciplines, there are not any studies focused specifically on how those in the humanities view and use assessment. The purpose of this study is to understand the common perceptions of general education assessment by humanities professors, including how they use that data and why they use it in those identified ways. This review is structured into the following sections:

- Introducing significant terms and the scope of the review;
- Contextualizing general education, since its purpose, description, and structure are conceived and implemented differently depending on the institution;
- Exploring general education assessment, as it has a unique position within the university; therefore there are some specialized issues when considering assessing general education courses;
- Analyzing in-depth the four distinct theoretical areas of faculty’s Conception of Assessment (Brown, 2008) that serves as a foundation for this study. These four areas include assessments of, as, and for learning, and assessment as irrelevant to learning;
- And examining the humanities faculty’s documented resistance to assessment to situate them in a broader, more robust assessment landscape.

The research questions that direct this study are: How do humanities professors perceive general education assessment? In what ways do they use the assessment data? What is the
rationale behind their use of assessment data? At the conclusion of the literature review, the claim for a qualitative case study that offers a more nuanced and complicated exploration of general education humanities professors’ conceptions and use of assessment will be made, and this study will be located in the current body of general education assessment literature.

**Definition of Terms & Scope of Review**

**Terms**

**Assessment:** Scriven (1967) said that assessment was a single process that should serve multiple purposes. Currently, collecting assessment data is a primary concern for professors, departments, and institutions and serves multiple purposes including: 1) program, service, and individual improvement 2) accountability to internal and external stakeholders and 3) progress toward institutional effectiveness (Brown, Bull, & Pendebury, 1997; Ohia, 2011).

**Conception:** Brown (2008) describes conception as “Concepts, and by extension, conceptions, express human purposes and intentions, by categorizing in an organized manner the world around us” (p. 7).

**Constructivism:** Knowledge is developmental, internally constructed, and socially and culturally mediated (Hood, 2002).

**Direct Assessment:** Direct assessments of student performance include such things as exams of all types, projects, papers, and exhibitions, and provide an opportunity for students to actively practice their skills synthesizing their knowledge (Aloi et al., 2003).

**Formative Assessment:** A practice where evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited (Black and Wiliam, 1998). It is generally
not included in a final summative grade (Taras, 2005).

**General Education:** General education is a central tenet of the American baccalaureate degree. General education assures that all students, regardless of specialization or intended career, become acquainted with history, culture, science, and mathematics. It is also a major vehicle for cultivating capacities such as communication, critical thinking, quantitative reasoning, and integration of knowledge (Ratcliff, Johnson, La Nasa, & Gaff, 2001).

**Humanities:** Humanities includes the following subject areas according to the National Endowment of the Humanities (1965): “Language, both modern and classical; linguistics; literature; history; jurisprudence; philosophy; archaeology; comparative religion; ethics; the history, criticism and theory of the arts…aspects of social sciences which have humanistic content and employ humanistic methods; and the study and application of the humanities to the human environment with particular attention to reflecting our diverse heritage, traditions, and history and to the relevance of the humanities to the current conditions of national life” (http://www.neh.gov/about).

**Indirect Assessment:** surveys, interviews, or focus groups that ask students to reflect on what they have learned and experienced rather than demonstrating knowledge and skills (Aloi et al., 2003).

**Summative Assessment:** generally the predominant version used in our educational system, suggests Better-Reed et al. (2008). Summative data is any data that depicts a judgment that encapsulates all the evidence up to a given point and is seen as final (Taras, 2005).

**Scope of the Review**

Much of the identified literature was garnered through publications and reports published by well-established organizations in the academic field like the Association of American
Colleges and Universities, the National Institute for Learning Outcomes Assessment, and Wabash College’s Center of Inquiry in the Liberal Arts. Searching EBSCO databases for terms such as “general education,” “assessment,” “conceptions of assessment,” “formative assessment,” “faculty and assessment,” “assessment implementation,” “assessment and arts/humanities,” “general education and assessment,” and “program assessment evaluation” identified further studies. Qualitative studies were gathered by using terms such as “assessment implementation” and “program assessment evaluation.” Many of these were case studies of individuals, programs, departments, or institutional experiences using assessment to improve student learning or evaluating institutional, disciplinary, or programmatic success. Examining various studies’ references identified further sources. Studies looking at institutional effectiveness and accreditation issues were excluded since this study is examining faculty perceptions of assessment in general education. The *Journal of General Education* and *Assessment & Evaluation in Higher Education* were particularly useful.

### General Education

To better understand the current state of assessing general education, one should have a sense of general education’s purpose within higher education and its significance within academe. General education comprises close to one third of an undergraduate’s university experience (Brint et al., 2009). While the debate between the core model (Smelser & Schnudson, 2004) and the distributive/elective model continues unabated (Gaff, 1983), most all agree that general education is about developing a well-rounded individual who has developed a “habit of mind” (Stearns, 2002, p. 44). This would also include developing critical thinking, communication, social responsibility, and global citizenship skills, among other things (AACU, 2005, 2006).
Historically, general education manifested itself in many forms and served numerous purposes. Mulcahey noted that the foundational elements for many general education programs can be traced back to their liberal education roots developed by Aristotle, Plato, and other Greek philosophers in the fourth century (as cited in Penn, 2011). The Yale Report of 1828 took an early position on the meaning of liberal arts, stating that students should be required to take a variety of topics so their “whole mind” would be exercised (Cohen, 1998). The Morrill Act of 1862 spurred higher education to take a more utilitarian focus and students began to elect courses they wanted without any disciplinary requirements. Miller, author of *The Meaning of General Education: The Emergence of a Curriculum*, stated at the turn of the century that many state colleges and universities focused on educating students to deal with the problems of society (as cited in Brint et al., 2009), thus serving a more practical purpose rather than the “whole mind” approach. In 1901, Yale took a concentration and distribution curricular structure that was adopted by other institutions up through the 1920s. Generally speaking, the literature shows that reforms at private institutions were largely tied to the “cultivation of the person and preparation of leadership” (Smelser & Schnudson, 2004, p. 9).

Not until the 1940s and 1950s was “general education” connected to “breadth” requirements as more universities assumed the “curricular organization of majors, distribution requirements, and electives,” says Rudolph (as cited in Brint et al., 2009, p. 607). During the 1960s and 1970s, general education saw another major shift as the cafeteria model bloomed due to student activism. In the interest of student choice, the required, structured major was disassembled (Smelser & Schudson, 2004). However, in 1977 a report labeled general education a “disaster area,” and similar reports that followed (Gaff, 1983) made it clear that general education was primed for reform. In response to the disorder of undergraduate education, Stark
and Lattuca contend, many universities (over 85% in 1990 according to one study) adopted the core curriculum (as cited in Boning, 2007). The core curriculum was defined as a developed series or set of courses that all students had to take.

Modern approaches to general education have typically synthesized both traditional liberal arts and practical applications into one program (Penn, 2011). Usually, the general education requirements comprise, on average, 30% of the undergraduate curriculum (Brint et al., 2009), which makes general education a significant part of the college experience. No matter its incarnation, general education is one aspect of higher education that all students partake in, and few higher education professionals dispute its value. Over two decades ago the importance of general education was affirmed in an important national study conducted by Ernest Boyer (1987), who found that approximately 75% of undergraduates in American colleges and universities felt that general education courses “added to the enrichment of other courses” and “helped prepare [them] for lifelong learning” (p. 85). Pascarella and Terenzini (1991) supported this claim when they discovered that the greatest gains in students’ ability to think critically were found at institutions with courses specifically designed to meet general education requirements. Given the importance of general education for baccalaureate students, it becomes important to know what students are learning in these courses.

Assessment

Assessment of student learning in higher education is necessitated by multiple stakeholders requiring data that serves multiple purposes. For example, the student needs to know her competency in an area, a department needs to know if learning outcomes are being achieved, universities need to know whether programs and certifications are effective, and
legislators need to be assured that their investments in state institutions are being used wisely. This is not an exhaustive list; rather it alludes to the complex nature of the assessment process.

Assessment as a Way to Collect Data on Teaching and Learning

Focusing on teaching and learning in his seminal essay, Scriven (1967) said that assessment was a single process that should serve multiple purposes. Bloom et al. (1971) suggested that assessment has two functions, firstly to systematically collect “evidence to determine whether in fact certain changes are taking place in the learners” and secondly to assess “the amount or degree of change in individual students” (p. 8). A decade later, Ewell (1988) reaffirmed this idea, sharing that assessment is linked inextricably to learning and development and is fundamental to education. In defining assessment as “the systematic basis for making inferences about the learning and development of students,” Erwin (1991) clarified assessment’s purpose. He continued:

Assessment is the process of defining, selecting, designing, collecting, analyzing, interpreting, and using information to increase students’ learning and development. It includes discussions about what should be assessed and how information will be used, not just the hands-on testing of students. (p. 15)

Overall, it is important to recognize that assessment is ultimately about assisting students in the learning process rather than any summative score.

Over time, assessment was recognized in three distinct ways concerning student learning. Assessments of learning are part of institutional quality assurance and accountability processes to validate the award of credentials as well as the quality of disciplinary offerings, professional training, and the student experience (Kuh, Kinzie, Shuh, & Whitt, 2005; Maki 2004; Ramsden 2003). Assessment as learning puts the onus on students by asking them to reflect on how well
they are doing and how they can improve their learning (Carless et al. 2007; Gibbs 2006; Gibbs and Simpson 2004). Assessment for learning suggests that assessment ultimately informs teaching by highlighting what is effective and ineffective in students achieving academic success and is most helpful if used for meaningful change or reform (Penn, 2011; Banta, Jones, & Black, 2009; Huba & Freed, 2000). Brown (2006) enhanced the categorization with a fourth perspective: irrelevant to learning.

Current data collection on student learning is widespread and thorough. Typically, accrediting bodies like Schools and Middle States Commission on Higher Education require data beyond reported grades, such as evidence of students meeting specific learning outcomes (http://www.msche.org/). In their national report on the current state of data collecting and use, Kuh and Ickenbery (2009) noted that most colleges and universities collected multiple types of data. As previously noted, Blaich and Wise (2011) also showed a trend of evidence being collected. Many of the results from these large studies lead researchers (Iqbal & Azam, 2009; Ewell, Paulson, & Kinzie, 2011; Black, Harrison, Lee, Marshall, & Wiliam, 2004) into exploring how student learning data affected the institution, the program, and the individual instructor’s course planning.

**Data Collection Tools for Assessment**

Part of the reason institutions and educators have so much data is that collecting it is easier than ever before. Testing tools are available from summative computer scoring programs similar to the SAT, GRE, or CLA; formative tools like MySkillLab, developed and distributed by Pearson publishing, that offer a myriad of exercises; in-class assignments; in-class tools; and institutional surveys all contribute to the abundance of data. Angelo and Cross (1993) collected an assortment of over 50 assessment tools. Other organizations have also made data collection
tools widely available. The Buros Institute of Mental Measurements summarizes and analyzes hundreds of newly released instruments in every biannual volume, while the Valid Assessment of Learning in Undergraduate Education (VALUE) project and the AACU brought together experts to create meta-rubrics that could be adopted or adapted (Penn, 2011). However, it is important to note that while there is a plethora of published instruments, for many of them, evidence of their quality and value remains unknown (Suskie, 2010).

There is a great deal of research concerning the liability of choosing and using particular assessment tools. Ratcliff et al. noted that even when outcomes are specified, one cannot merely collect data through traditional testing methods that are easily administered and scored; instead, they must be designed according to general education outcomes as applied in the students’ major (as cited in Aloi, 2003). A decade later, the Council on Higher Education Accreditation (2003) reaffirmed that view, stating it was imperative to avoid collecting data that narrowly defines student learning or using excessive standardized measures of student achievement. Though the difficulty many educators find is that “as the performance task increases in complexity and authenticity, which serves to increase validity, the lack of standardization serves to decrease reliability” (Polomba and Banta, 1999, p. 89). In other words, increasing the validity of student learning assessment data makes both the tool used and the data collected more unique and specialized; however, decreasing standardization has a negative affect on the data’s reliability and usefulness.

**Assessment in General Education**

Assessment of general education outcomes poses some unique issues for universities. First, assessing general education is extremely challenging because as Janzow, Hinni, and Johnson state, “General education is almost always the largest academic program, but, unlike the
The smallest department, has not had a head, a faculty, or a budget” (as quoted in Aloi et al., 2003, p. 243). In other words, assessment is a complex undertaking because it must be coordinated across several different stakeholders. Further, assessment of general education is difficult because there is a lack of clarity in institutions regarding general education learning outcomes. (Stone & Friedman, 2002, p. 200). This is not difficult to imagine since there is a lack of consistent language concerning general education assessment and outcomes. Yin and Volkwein (2010) examined general education and outcomes from six accreditation agencies and found that they not only defined general education and outcomes differently, but also none of them aligned with Association of American Colleges and Universities (AACU) principles and standards. A lack of consistent language makes it difficult for administrators, teachers, and other stakeholders to communicate effectively and address common issues surrounding assessment.

Gillmore (2004) expands on this idea, articulating the difficulty of assessing general education curricula because the shared goals of general education are “largely illusory” (p. 15). Consider how different general education can look depending on one’s major and how difference between students might affect the applicability of particular learning outcomes to an entire student body. Complicating this issue, students and faculty do not recognize the importance of general education, but often view it as something to be completed as fast as possible so students can focus on their chosen majors (Furman, 2013; Wehburg, 2010; Miller & Sundre, 2008).

Findings from Assessments in General Education

These difficulties in assessing general education learning outcomes and student proficiency did not stop researchers from attempting to identify and assess student achievement in general education. Many researchers (Laird & Garver, 2009; Hart Research Associates, 2009; Laird, Niskode-Dossett, & Kuh, 2009) focused a great amount of attention on the subject, most
often via descriptive surveys. In one of the earliest studies on the subject, Gaff and Wasecha (2001) analyzed a study they completed in 1990. They surveyed over 200 chief academic officers and found some data is infused in the existing courses that alters the way content is delivered and learned; for example, 93% of deans reported some kind of writing across the curriculum, and 71% cited critical thinking, and assessing student learning in those areas is difficult because they are not taught in one course, but across many courses.

Eleven years later, the American Association of Colleges and Universities (AACU) published results of a survey of its members which affirmed Gaff and Wasecha’s (2001) 1990 study’s findings concerning assessing learning outcomes and general education. While the survey was self-reported so its results may be skewed (for example, members may want to make themselves look good or there may have been a disproportionate number of reformers who responded), the results are still worth noting. 89% of the 433 Chief Executive Officers stated their institutions were in the process of evaluating or modifying their general education programs with most focusing on common learning outcomes (Hart Research and Associates, 2009). The survey discovered that more institutions were using common learning outcomes to identify proficiency, competency, or growth in those outcomes via a wide range of introductory courses. Emphasizing undergraduate research, first year experience, and learning communities were other major findings.

**Challenges in General Education Assessment**

The 2001 AACU survey also points to information that has not been captured by general education assessment. Focusing on specific learning outcomes like critical thinking, rather than only content knowledge, data collection, and subsequent analysis in general education, is more challenging. The survey also depicts assessment of general education along multiple courses as
remaining a concern with just over half (52%) of the respondents collecting and using that data. Blaich and Wise (2011) noted that part of the problem was that many institutions may have a hard time analyzing what the data might mean and how to use it effectively. As Vincent Tinto, professor and chair of the Higher Education Program at Syracuse University, claimed, there was plenty of information on student success, but what was missing was the ability to transform it into practical knowledge (as cited in Malandra 2008).

There are other significant difficulties associated with the types of data collected. For instance, even if a common grading scale is used in multiple sections in a general education course, individual instructors may “weigh course elements differently” (Penn, 2011, p. 10). In response to this phenomenon, many accreditation agencies avoid using grades or GPAs as sole confirmation of student learning or achievement (Penn, 2011). Another significant obstacle in assessing general education courses revolves around the instruction itself. Allen (2006) indicates the reliance on adjunct faculty is significant. Due to their unique status in academe as part-time faculty, they may not be active in the development of pedagogy or assessment.

Assessment: Challenges in Humanities Courses

Those who teach humanities courses in general education have unique issues with which to contend. Just as getting a clear, agreed upon definition of general education is difficult, many humanists struggle to define what humanities are and what kind of knowledge is desired. Consider Mann (2000) who relates the issue thusly, “Even the National Endowment for the Humanities tends to rely on a list of ‘humanities disciplines’ as the best method for explaining its purview” (p. 82). Helm (2000) agrees, stating, “humanities were best characterized as a collection of disciplinary contents” (p. 90). While Helm develops a working definition that focuses on a “peculiar kind of activity” (p. 92) which includes reading, inquiring, evaluating,
judging, and articulating meaning, Mann (2000) highlights the overall concern of relying on ambiguous terms stating that “[W]ithout a good definition and its related goals, assessment of students’ learning and the likely improvement of teaching are not possible” (p. 90). Defining learning outcomes in general education is more challenging for the arts and humanities than for other academic areas because of the variety of disciplines and its abstract nature (Anderson, 2002; Bers, 2000).

A significant issue concerning assessment of humanities learning outcomes is that as Chanan (1974) argued, it coopted the language of the natural sciences where results were predictable. Chanan wondered how a teacher would judge a moment when an insight occurred or individual growth took place. Building on Chanan’s argument, Weldhen (1981) described the difficulties in assessing humanities specifically due to the types of knowledge desired, cogently arguing that humanities should be looked at as experience (being or states of consciousness) instead of knowledge, exploration instead of explanation.

As Heiland and Rosenthal (2011) contend, the concept of accountability grounded in measurable outcomes generally identified and assessed through standardized testing misses the kinds of knowledge valued by many humanists. The broad categories would be better served by exploring disciplinary learning and assessment. This is particularly appropriate because often the assessment methods in the humanities are qualitative in nature. Heiland and Rosenthal (2011) summarize the issue facing the humanities and assessment extremely well, stating that ultimately humanities’ learning outcomes are nebulous and can only be recognized in the “minute particulars,” and to attempt to generalize them is to reduce them into unusable, ineffectual concepts. In other words, the learning outcome itself lacks its intended significance and becomes something that bears little weight or meaning for the student, the teacher, and the administrator.
It should be noted that from some domestic surveys, one might infer perceptions of assessments via reports on how faculty used the collected data (Townsend, 2010; Blaich & Wise, 2012), but none specifically address faculty perceptions in a direct, methodological way. One study highlighted how humanities departments claimed they used assessments for curriculum improvement (53%) and teaching improvement (62%), but affirmed that actual changes were made in curriculum (80%) and teaching (44%) (Ewell, Powell, & Kinzie, 2011). What we might infer from this data is that humanities’ faculty might believe assessment should improve teaching, but in implementation assessment results affect curriculum.

To summarize, general education and assessment serve multiple purposes for multiple stakeholders, and parsing out the implications of each is a complex undertaking. While Blaich and Wise (2011) showed a trend of evidence being collected, assessment of general education outcomes poses some unique issues for universities. It is difficult to assess general education because it is large (Aloi et al., 2003) and more difficult to identify specific goals (Stone & Friedman, 2002). In an attempt to address these challenges, most colleges are now using common learning outcomes (Hart Research and Associates, 2009), but this does not make general education assessment “easier” (Gillmore, 2004). General education humanities courses also embody distinct challenges compared to their higher education counterparts. Specifically, many humanists see the learning outcomes as too reductionist to truly capture any real “deep” learning in their courses.

The focus on learning outcomes does not give a complete picture, however. The range of courses in the humanities, with their bevy of unique course goals, creates a dilemma in understanding how humanists approach assessment and use it. One possible way a researcher may categorize or explore possible common thematic elements between humanists’ views of
assessment is with Browns’ Conception of Assessment. Brown (2004) developed his Conceptions of Assessment framework to try to quantify how secondary school teachers perceived assessment. Brown (2008) claims conceptions “express human purposes and intentions, by categorizing in an organized manner the world around us…how groups of people with shared experiences understand, represent, and intend to act within a domain” (p. 7-8) and he identified four general areas: school accountability, student accountability, student improvement, and irrelevancy. In the broader assessment field these areas can be seen as of, as, for, and irrelevant to learning. These four areas will be differentiated and discussed in the next section.

**Conceptions of Assessment**

One useful method to classify how assessment is perceived and used is Brown’s (2004, 2008) Conceptions of Assessment. Brown’s Conceptions of Assessment attempts to quantify how teachers view assessment by placing it in one of four thematic areas including assessment for, as, of, and irrelevant to learning. As previously noted, Brown (2008) claims conceptions express a person’s purpose and intention and represent “how groups of people with shared experiences understand, represent, and intend to act within a domain” (p. 7-8). Conceptions are malleable, so a professor’s perspective on assessment may change depending on one’s experience and current reality. Gilles, Detroz, and Blais (2011) concisely explain why understanding the conceptions of assessment is important:

In a context where public action must demonstrate its effectiveness and efficiency, and where the links between teaching and the quality of learning are regularly highlighted, it seems relevant to identify the trends and logic that govern university professors’ decisions with respect to the modes of learning assessment favored within the framework of their delivery of teaching services. Moreover, given that university teaching practices
are changing rapidly due to the introduction of different views of the learning process, one might conclude that the same holds for assessment practices. (p. 720)

In other words, when internal and external accountability needs meet internal community’s pedagogical needs and practice needs, it is important to understand the professors’ perspectives that drive the teaching and learning decision-making.

**Contemporary examinations of conception of assessment.** It is essential to discuss the current literature on conceptions of assessment in higher education. Brown’s (2004; 2006) work is not included since he examined secondary education. Faculty and student perceptions of assessment data are not well addressed in the assessment field. One international study (Samuelowicz and Bain, 2002) considered faculty’s orientation of assessment orientation, which they defined as a “coherent pattern of beliefs inferred from, and grounded in, academics’ assessment practices and their explanations for those practices” (p. 176). Other international research (Iqbal & Azam, 2009; Fletcher et al., 2011; Carless, 2006; Gilles, Detroz, & Blais, 2011) used Likert surveys to identify conceptions of assessment and its purposes. All four studies had large sample sizes adding to their reliability. Iqbal and Azam (2009) collected data from 297 students and 37 faculty while Fletcher et al. (2011) collected surveys from 877 faculty and 1220 students. Carless (2006) used results from a 2004 questionnaire solicited from 460 staff and 1740 students then followed up with a 52 person qualitative study. Gilles, Detroz, and Blais (2011) identified 643 university professors and lecturers associated with the Université de Montréal and its affiliated schools (École Polytechnique de Montréal and HEC-Montréal).

Two elements were very similar in all the studies that should be highlighted. First, all of them are international studies. Iqbal and Azam’s study was completed in Pakistan, Fletcher et al. was completed in New Zealand, Carless completed his in Hong Kong, and Gilles, Detroz, and
Blais was in Montreal. Further, when discussing their work, Gilles, Detroz, and Blais (2011) recognize their difficulty in finding research on the topic saying, “…we have not found anything equivalent in the English-speaking research community” (p. 722).

Second, all but Gilles, Detroz, and Blais’ (2011) study were fairly homogenous populations in terms of disciplines represented. Even with the largest sample, Fletcher et al. (2011), acknowledged this concern in their paper. In light of those similarities, it is noteworthy that all three came to the same conclusion: faculty and students were not in sync concerning assessment data and how it is used to facilitate learning. Carless’ (2006) findings are representative of all three. He indicated that tutors perceived their feedback more positively than students did. Specifically, he found the difference in perceptions were extreme. 38.4% of teachers thought students were often given detailed feedback which helped them improve their next assignment, while only 10.6% of students responded in the same way. Worse, 37.8% of students felt that feedback was rarely followed by actions to improve student learning, as opposed to 16.1% of tutors. Although all of the studies directly address the problem of how professors perceive and use assessment data, all of them are international sites, so one does question the generalizability to American higher education institutions.

From some domestic surveys, one might infer perceptions of assessments via reports on how faculty used the collected data (Townsend, 2010; Blaich & Wise, 2012), but none specifically address faculty perceptions in a direct, methodological way.

For Learning. Assessment for learning suggests that assessment ultimately informs teaching by highlighting what is effective and ineffective in students achieving academic success and is most helpful if used for meaningful change or reform (Penn, 2011; Banta, Jones, & Black, 2009; Huba & Freed 2000). This type of assessment is considered formative. Scriven (1967) first
introduced the concept of formative assessment as distinct from summative assessment, though both types of assessment could be completed in a single assignment. Taras (2005) noted that the power of formative assessment, i.e. assessment for learning, continues to be eroded because Scriven’s advice has not been heeded and a false separation has been created between summative and formative assessments. For example, a student’s essay may be used as both a summative assessment in determining a student’s grade while offering feedback for teacher and learner improvement.

After Scriven introduced the concept of formative assessment, multiple researchers developed and refined it. Black and Wiliam (1998) defined formative assessment as a practice where evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction that are likely to be better, or better founded, than the decisions they would have taken in the absence of the evidence that was elicited. In other words, as Dunn and Mulvenon (2009) explain, formative assessment includes, “all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (p. 10). Using data in such a way can be described as “closing the feedback loop” (Wehlburg, 2007; Watson 2003). After being exposed to a concept or skill and assessed on their competency, the learner uses the feedback to improve his or her proficiency on that concept or skill. Equally important, the teacher uses the information garnered from those assessments to review his or her instruction. The teacher also identifies weaknesses or areas of improvement in instruction to improve his or her pedagogy or practice in the classroom.

comprehensive review of about 250 articles reported significant positive student learning gains using formative assessment strategies. Black and Wiliam (1998) demonstrated gains with low-achieving students making the largest increase. Qualitative studies (Borish, Kachum, and Lyon-Jenness, 2009) also depict the worth of formative assessments for student learning. While the data analysis portends positive results, some researchers express concerns about the power of formative assessment. In their review of Black and Wiliam (1998) and nine other articles, Dunn and Mulvenon (2009) found a lack of a consistent lexicon across the field and discussion of formative assessment. Second, they claimed there are methodological concerns in all of the studies they examined. While Dunn and Mulveson (2009) claim there is some evidence of improvement of student learning due to formative assessment, it is not as complete or thorough as presented in the field.

Other issues related to formative assessment focus on the data and how it is used, or not being used, after being collected. In multiple quantitative studies (Blaich & Wise, 2011; Wang & Hurley, 2012; Kuh & Ickenberry, 2009) it was clear that the lack of time, resources, and support for using assessment for learning was widespread and disconcerting. In those same studies, they also found that using assessment for learning was worthwhile, but participants thought it too time consuming to warrant pursuing.

A second issue identified was that it would not benefit faculty in professional goals (like tenure, etc.). This theme was identified in qualitative case studies like Gaff and Wasecha (2001) and Gillmore (2004). Overall, these studies saw faculty answer that there was little time or resources to use data effectively. A number of descriptive case studies reached a similar conclusion (Hachtmann, 2012; Borish, Kachum, and Lyon-Jennesse, 2009; Savagian, 2009). All the studies identified faculty members’ views of assessment as something worthwhile, something
they would more fully support if it were recognized as a scholarly activity by their respective institutions.

**As Learning.** When examining assessment *as* learning, students are individually held “accountable” for their own learning (Brown, 2008). Guthrie (2002) claimed the purpose of students being accountable is to certify that they have the requisite knowledge, skill, and ability necessary to meet the next level of education, training, or employment opportunities (as cited in Brown, 2008). Brown (2008) identifies any assessment where a professional assigns a score, grade, etc. based upon how a particular student did according to specific criteria, and uses that as a way to place the student in a class, group, etc. Many high-stakes tests would fall under this umbrella. Brown (2008) also notes certifications (for example, driver’s licenses, lawyers, dentists, doctors, etc.) are high-stakes assessments that demand students be held accountable for their education.

Another purpose in higher education is that assessment *as* learning puts the onus on students by asking them to reflect on how well they are doing and how they can improve their learning (Carless et al. 2007; Gibbs 2006; Gibbs and Simpson 2004). Faculty members who view self-assessment as a significant aspect of education might subscribe to the idea highlighted by Boud and Falchikov (1986) who state that the capacity to evaluate one’s own work is the primary purpose of education. Archer (2010) was more specific, stating “Self monitoring is the ability to respond to situations shaped by one’s own capability at the moment in that set of circumstances, rather than being governed by an overall perception of ability” (as quoted in Evans, 2013, p. 87-88). For a student to be able to do self-assessment, Sadler (1983) said s/he needs knowledge of the standards, the ability to compare one’s own worth to those standards, and the ability to “close the gap” (as quoted in Boud, Lawson, & Thompson, 2013, p. 4). The importance of such
evaluation is that it allows students to identify their own weaknesses and to develop their skills in those areas.

The third area of assessment as learning emerges from student self assessment encouraging lifelong learning. Evans (2013) underscores that fact by relating Boud and Lawson’s (2011) work. They show it takes time for students to develop such self-assessment skills. Strategies to develop those skills include: student self-assess and have a rationale for their assessment (Sendzuik, 2010) or students self-revise essays based on all sources of feedback gained throughout the course (Graziano-King, 2007).

Of learning. Assessment of learning directly relates to the accountability issue. Generally it pertains to school accountability and submits that assessment can be “used to account for a teacher’s, a school’s, or a system’s use of society’s resources” (Brown, 2008, p. 18). Assessments of learning are a part of institutional quality assurance and accountability processes to validate the award of credentials as well as the quality of disciplinary offerings, professional training, and the student experience (Kuh et al. 2005; Maki 2004; Ramsden 2003). In fact, to meet these requirements, certain techniques were implemented. For example, there was an increase in common assessment instruments (Steele & Lutz, 1995, as cited in Chun, 2002), and the introduction of performance-based funding approaches (Ewell, 1991, 1997).

For higher education, the accountability push seemed to begin after the publication of reports critical of the quality of higher education (Bennett, 1984; National Institute of Education, 1984), and linking the provision of federal funding for financial aid eligibility to institutions’ student assessment efforts (Banta, 1991). As Leveille (2006) observed “Control of costs, elimination of duplication…and evidence of other efficiencies are the focus for legislatures and higher education regulating and coordinating agencies. Similarly, demands for greater
productivity in higher education will continue to be heard with greater frequency than at any
time in the past” (p. 5). He further claims this push forces people to institute a “business model”
and be concerned with “the bottom line.” For many in higher education, the bottom line is
accountability to accreditors. In a National Institute for Learning Outcomes Assessment report,
provosts and academic officers from over 1500 institutions reported that assessment at their
institutions was most commonly used for accreditation (Kuh & Ickenberry, 2009).

**Not relevant to learning.** The final conception is that assessment is *not relevant to*
learning. Brown’s (2008) original conceptions were developed for the K-12 audience, however,
the underlying premise of the conception stands in higher education as well. The literature is
fairly extensive in terms of specific critiques of the assessment movement and what might cause
professors to believe it not relevant to learning.

Learning outcomes un-assessable. One of the primary critiques, especially concerning the
humanities, is that learning outcomes are just not assessable. Here, assessment does not have a
“legitimate place” for teaching and learning because teachers’ “deep understanding of
curriculum and pedagogy” automatically allows them to assess students in normal interactions
with students (Brown, 2008, p. 25). Hamman (2009) states, “… not everything is quantifiable.”
For example, “engaged citizens” or “productive members of society” as aspirations are
admirable and may get at the heart of an education; however, trying to quantify success in these
areas actually distracts from the goal. Hussey and Smith (2008) agree with this line of thinking,
stating that the farther away one gets from the unit or course where the learning outcome
originated, the more abstract and meaningless learning outcomes become. More directly stated,
they claim the farther from student and teacher the more “remote, generalized, and irrelevant
statements of learning outcomes” are (p. 114). Glenn (2010) defines learning outcomes as
“jargon-riddled” and states they have “little to do with what they (teachers) want to do in the classroom (para. 5). Brottman (2009) agrees, calling them “superficial projections and assumptions” (para. 1) and saying they have “no bearing” on what actually takes place in the classroom. Brottman (2009) clarifies what many humanists feel, stating, “[N]one of these ‘objectives’ are ‘assessable,’ nor would I ever want to ‘assess’ them” (para. 13). She continues by rhetorically asking how one would assess “habits of mind, nuances of thinking, an appreciation for subtleties and ambiguities of argument, and an appreciation of the capacity for empathy, as well as the need, on certain occasions, to resist this capacity” (para. 13).

Consequently, some researchers question the very purpose of learning outcomes. If someone accepts the premise that learning outcomes in some courses are just not assessable, then what is their purpose? Bennett and Brady (2012) argue that learning outcomes and subsequent assessment is actually “a device for monitoring and auditing rather than a tool for teaching and learning” (p. 35) and has to do with “administrative and regulatory necessity rather than students’ deep engagement with curriculum” (Hussey & Smith, p. 357). They continue to call it a disciplinary mechanism for college administration, government entities, and accrediting agencies.

**Deprofessionalization.** Another argument in the literature that makes assessment irrelevant for some professors is that it challenges them as professionals. It may be perceived as deprofessionalizing them by negatively affecting their autonomy and actually distracting them from education (Brown, 2008). The fact that assessment is subjective in many areas, and there is a degree of “inaccuracy,” undermines the teacher’s confidence in their efficacy (Brown, 2008). Faculty may feel that if they are hired as an expert and professional in their field, they should be trusted enough to fairly and accurately evaluate student learning. Carpenter and Bach (2010)
explain that an expert within his or her field should have the legitimacy to evaluate students on what is learned and it should be demonstrated in the grade given.

Essentially, they are not treated as experts and professionals able to complete their duties as they see fit. For example, many point to the standardization of materials and the undermining effects it has on the teachers. Wilson (2010) notes, “Because professors prize their autonomy, they are leery of any efforts to standardize classroom teaching” (para. 17). Consider the use of a standardized syllabus used across sections in a particular course. Bennett and Brady (2012) believe that this is essentially losing control of one’s classroom, which is ultimately required to generate data for externally dictated learning outcomes.

Many faculty may contextualize this standardization in a larger framework. Denecke, Kent, and Wiener (2011) suggest that faculty resent what is perceived to be the bureaucratization of US higher education and they fear that “if standards for content and teaching are defined from outside the institution, department, or classroom, they will result in nothing more than a legitimization of compulsory mediocrity in US higher education curricula” (p. 13). Fritschler (2010) recognizes the concern of applying some of the standardization presented in K-12 education to higher education stating, “Standardized testing does not fit well with the idea that higher education should foster creativity, innovation and critical thinking” (para. 12).

Finally, Bennett and Brady (2012) relate some of the more ideological concerns some faculty may have which would cause them to see assessment as irrelevant. They state that the motivations for learning outcomes assessment stems from conservative movements for industrial scientific management; it is merely adoption of corporate language, and it acts to bolster the cottage industry of assessment and support mechanisms.
Unworthy of investment. Lastly, there are other issues in higher education that might cause professors to view assessment as irrelevant. Since teaching is not a priority in their department, discipline, or institution, assessment is not worth the time or resources (Hamman, 2009). Consider, “Faculty rewards have nothing to do with the ability to assess student learning,” says one professor. “I get promoted for writing lots of articles, not for demonstrating learning outcomes” (Wilson, 2010, para. 3). Academic rewards need to be reevaluated and refocused to recognize teaching as a valuable product of professors’ time and effort.

General Education, Assessment, and Resistance to Change

A major complicating factor in using assessment data to enact change in general education is resistance to change. In the Wabash Study, Blaich and Wise (2011) offered an essential insight to keep in mind considering how stakeholders resist reforming general education for various reasons: they claim it may be just as important for assessment leaders to draw on the literature of facilitating institutional change, as it is to know how to collect, analyze, and use data.

Institutions may have the data, and recognize the importance of making changes in their general education programs or courses; however, there may be a number of reasons why they may not take advantage of that data in any significant way. Organizational change and resistance literature might contextualize the difficulties they face. As noted by Gess-Newsome et al. (2003), reforms that seek to change fundamental structures, cultures, and pedagogies in school are inherently difficult to implement and sustain. When institutions attempt to revise general education, it usually is episodic in nature. Weick and Quinn (1999) described episodic change as “infrequent, discontinuous, and intentional” (p. 365). Episodic change, they continued, occurs in periods when organizations move away from a state of equilibrium and includes fundamental
shifts in the organization. It was found in the literature that when an organization moves away from its standard operating procedure, everyone in the organization is affected. Szabla and Sanders (2012) highlighted that emotions are key elements of change and can be feelings that make change “difficult and painful.”

Research indicates that university administrators and faculty hold a myriad of concerns regarding assessment that may hinder efforts to effectively use it for student learning. Kuh and Ikenberry (2009) shared an evaluation of a Teagle Foundation-funded project to foster a culture of evidence on liberal arts campuses and found that both faculty and administrators “expressed reservations about the power of assessment data to change teaching and learning” (p. 27). Other research indicates faculty rarely have a strong sense of ownership concerning the general education program compared to their disciplinary home and, “except in rare cases,” general education is not constructed as a program in the same ways majors are (Wehlburg, 2010, p. 90). As Penn (2011) notes, faculty may also not know what to measure when considering the complex general education learning outcomes. For instance, one piece that is a fairly typical example in the literature is Savagian (2009), who gives an anecdotal account of working through the process of attempting to redesign assessment of general education history courses at Alverno College.

Another aspect of resistance comes from their professionalism and expertise. Specifically, they may not see themselves as a part of a “team.” Betters-Reed et al. (2008) indicate that faculty members generally think of themselves as individual contributors with the sole purview over the courses they are responsible for. At most, they continue, faculty may think of themselves as one part of the department. Rice says this mindset is “deeply rooted in
professional assumptions,” and it must change for successful implementation of a new culture of assessment (as quoted in Gano-Phillips et al., p. 12).

Consequently, when trying to create a more unified front, departments may infringe on what faculty see as their autonomy. Andrade (2011) discusses this difficulty, explaining that instead of the solitary process of developing assignments and grading criteria, faculty members are now being asked to identify learning outcomes and assessment measures across programs that may challenge faculty autonomy. Quite a few authors (Hammon, 2009; Brottman, 2009; Fendrich, 2007) share Andrade’s (2011) concern and fear of the possible deprofessionalization of the faculty.

Departments may resist general education data collection and reform for other reasons as well. Citing work by Martinson and Cole (2002) and Martell (2005), Betters-Reed et al. (2008) believe that, because there is a lack of “regularization,” the development of an assessment plan can be a “painful and arduous process” for departments (p. 227). Some studies show that department chairs are concerned about the faculty and department time these efforts take (Pederson & White, 2011). Others corroborate those concerns by observing the expense that would be incurred in any type of general education reform that involved a new assessment strategy (Hersh, 2005). Spalter-Roth and Scelza (2009) report that department chairs worry since assessment is an activity that generally comes with no additional resources provided for performing it and, when complete, no resources or support to revise the curriculum are provided. Lastly, Kuh and Ickesenberry (2009) also found that departments at less selective schools might feel pressured to demonstrate their value through assessment but may feel threatened by the possible results.

Chapter Summary
General education and assessment are complex, multi-faceted concepts that work on multiple levels in the university. While general education may take different forms depending on the institutional goals, ultimately the program serves as a foundation for the student’s educational experience and developing the “whole person.” Assessment is used to communicate messages to multiple stakeholders such as students, teachers, administrators, parents, the public, etc. Scriven (1967) said that assessment was a single process that should serve multiple purposes and Ewell (1988) noted that assessment is linked inextricably to learning and development and is fundamental to education.

There are four broad conceptual frameworks where assessment may reside. Assessments of learning are part of institutional quality assurance and accountability processes to validate the award of credentials as well as the quality of disciplinary offerings, professional training, and the student experience (Kuh, Kinzie, Shuh, & Whitt, 2005; Maki 2004; Ramsden 2003). Students are individually held “accountable” for their own learning (Brown, 2008) in the thematic category of assessment as learning; it puts the onus on students by asking them to reflect on how well they are doing and how they can improve their learning (Carless et al. 2007; Gibbs 2006; Gibbs and Simpson 2004). Assessment for learning suggests that assessment ultimately informs teaching by highlighting what is effective and ineffective in students’ achieving academic success and is most helpful if used for meaningful change or reform (Penn, 2011; Banta, Jones, & Black, 2009; Huba & Freed, 2000). Assessment that is irrelevant to learning suggests that assessment does not have a “legitimate place” for teaching and learning because teachers “deep understanding of curriculum and pedagogy” automatically allows them to assess students in normal interactions with students (Brown, 2008, p. 25). Assessment may also be viewed as irrelevant because it may be perceived as deprofessionalizing teachers (Brown, 2008). Hamman (2009) claims not
everything is quantifiable and Bennett and Brady (2012) relate some of the more ideological concerns some faculty may have which make assessment irrelevant.

Black and McComick (2010) succinctly capture a major concern, stating, “[T]here appears to be a lack in general discourse on linking learning with assessment, within a pedagogical framework” (p. 499). Three important notes can be made about the literature concerning assessment data and analysis in general education humanities courses. First, there is a lack of comprehensive research on what role the collected data play in the institution’s general education classes. Second, while humanities professors have expressed multiple perspectives and strong opinions in various venues, there is little research in identifying and systematically exploring humanities professors’ conceptions of assessment, particularly in their general education courses. Lastly, there is a gap in the literature concerning how humanities professors actually use collected student data in general education courses.

This case study is situated to identify, capture, and explore how humanities professors in general education conceive and use assessment. If we can identify common themes or perception of assessment across multiple humanities disciplines, we might be able to better address some of the confusion and ambiguity that revolves around humanities learning outcomes in general education.
CHAPTER 3: METHODS

Introduction

The purpose of this study was to understand the common conceptions of general education assessment by humanities professors and how they use that data. Since general education courses command approximately 46 credit hours (Hart & Associates, 2009) or 30% of a student’s collegiate experience (Brint et al., 2009), it is important to validate student learning and competency in the general education goals. Typically the goal of general education is to develop a student’s critical thinking, communication, social responsibility, and global citizenship skills, among other things (AAC&U, 2005; 2007).

Currently, institutions gather a great amount of data concerning general education, but that data is normally not being used to assist in teaching and learning (Blaich & Wise, 2011); instead, the data is used as an accountability tool (Blaich & Wise 2011; Reeves 2006) or for institutional reporting (Blaich & Wise 2011; Kuh & Ickenberry, 2009). As highlighted in the literature, data collection needs to be tied to reform for student learning for it to be helpful (Banta, Jones, & Black 2009; Huba & Freed 2000). If universities do not use data to inform teaching and learning practices, there may be little student improvement in general education learning outcomes (Gaff, 1980). How humanities general education assessment data is perceived and used to improve student learning may have significant effects on both student and institutional success. This led to the research questions:

• How do humanities professors typically conceive of general education assessment?
• In what ways do they use the assessment data?
• What is the rationale behind their use of assessment data?
The following chapter will include a discussion of qualitative research, the constructivist approach, the researcher’s role in the study, and the case study tradition. Sampling procedures, data collection, and data analysis follow a description of the site and the participants. The chapter concludes with an explanation of the study’s trustworthiness and protection of human subjects.

**Methodology**

**Research Paradigm**

I applied the constructivist paradigm for this study. Researchers (Stake, 1995; Yin, 2003) often subscribe to the constructivist paradigm when considering case studies as it involves a meaningful negotiation and interpretation of one’s reality that is influenced by one’s prior knowledge (Cobern, 1993). Guba and Lincoln (1994) state, “realities are apprehend-able in the form of multiple, intangible mental constructions, socially and experientially based, local and specific in nature…and dependent for their form and content on the individual persons or groups holding the construction” (p. 110-111). In this case, the humanities professors’ experiences lead them to hold particular ideas concerning assessment. Constructivism characterizes the conception of knowledge of the field of humanities. As such, a constructivist approach is important for this exploration of how humanities professors understand and use assessment in their general education courses.

Constructivists believe there are multiple, constructed realities rather than a single true reality (Ponterotto, 2005). More specifically, constructivists like Hanson believe reality is constructed in the mind of the individual, rather than it being an externally singular entity (as cited in Ponterotto, 2005). Ponterotto (2005) puts it succinctly, “Reality, according to the constructivist position, is subjective and influenced by the context of the situation, namely the
individual’s experience and perceptions, the social environment, and the interaction between the individual and the researcher” (p. 130). Humanities professors have had many experiences (intellectually, academically, etc.), have been involved with many groups (professional, departmental, university, etc.), and have been a part of many cultures (disciplinary, university, etc.), which influenced their constructed reality concerning assessment. To deconstruct and identify how humanities professors perceive general education assessment and use it, they must be given an opportunity to reflect and consider those influences. This qualitative study allowed them that opportunity.

In this study, the participants and I discussed assessment at length in order to clarify each participant’s view of assessment in general education humanities courses. I was not seeking objectivity, but instead I was uncovering each professor’s subjective understanding of assessment that has been created by personal as well as professional experiences (Guba & Lincoln, 1994). Guba and Lincoln (1994) claim “the aim of the inquiry is understanding and reconstruction of the construction that people (including the inquirer) initially hold, aiming toward consensus but still open to new interpretations as information and sophistication improve” (p. 113). The goal of using a constructivist approach in this study was to understand how assessment is constructed from the “lived experiences” (Ponterotto, 2005) of faculty members who have a responsibility to assess student learning.

**Role of the Researcher**

In the constructivist paradigm, the role of the researcher is complex. As Guba and Lincoln (1994) put it, “the researcher and the object of investigation are assumed to be intellectually linked so that the findings are literally created as the investigation proceeds” (p. 113). Ponterotto (2005) confirms this view stating, “The researcher and her or his participants
jointly create (co-construct) findings from their interactive dialogue and interpretation” (p. 129).

With this in mind, one should recognize that just as the participants hold values, beliefs, and assumptions concerning the subject, so does the researcher. In this view, the researcher’s values and lived experience cannot be divorced from the research process (Ponterotto, 2005). The researcher should acknowledge his or her values, experiences, biases, etc. and clearly indicate them within the study (Ponterotto, 2005; Creswell, 2013; Yin, 1994). The study is often in the first person and is often personalized (Ponterotto, 2005) where the inquirer’s voice is presented as the “passionate participant” (Lincoln, 1991, as quoted. in Guba and Lincoln, 1994, p. 115). In terms of the case study methodology, Creswell (2013) highlights the importance of the researcher spending a great deal of time in the field, observing, collecting documents, interviewing, etc., so the researcher’s role is highly interactive with the participants.

**Research Design**

In qualitative studies, researchers believe reality is a social construction, and people make meaning through subjective experiences (Creswell, 2012). Researchers generally do not know what they are looking for; instead, they allow for meanings to present themselves as details emerge from the study (Creswell, 2012). One could argue these studies attempt to get at how people think or what makes them “tick.” The initial research question attempted to explore the participant’s conceptions, and by doing so attempted to identify the assumptions s/he may have concerning what assessment is and its purpose(s). Participant’s values and beliefs concerning assessment and education, including disciplinary realities, may be intertwined with those conceptions of assessment.

A quantitative study can identify the participants’ broad conceptions of assessment but not the complexities of how humanities professors came to those conceptions, or why they use
assessment data in particular ways based on those conceptions. Understanding those conceptions is crucial because evidence exists that teachers’ conceptions of teaching, learning, and curricula influence strongly how they teach and what students learn or achieve (Clark & Peterson, 1986; Calderhead, 1996). Further, as Tang and Chow (2007) note, attitudes 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.

Nor can a quantitative study examine and truly explore the specific ways participants use the data in their work life. The difficulty, as McLellan (2004) and Iqbal and Azam (2009) found in their quantitative studies concerning conceptions of assessment, was when the faculty espoused values concerning assessment that differed from their actual use. This type of discrepancy is not easily identified; however, in a qualitative study it could be identified and investigated through document analysis and interviews.

Currently, there is a great deal of diversity of views concerning assessment in the humanities. While there are those in the humanities who argue in support of new assessment initiatives (Graf, 2010), it seems that there is still considerable resistance to assessment, especially when seemingly dictated by external forces or attempting to meet some mandated criteria (Scobey, 2009; Brottman, 2009). These complexities within the humanities highlight the importance of using a study method that can identify an instructor’s conceptions of assessment and attempt to contextualize those conceptions.

Research Tradition: Case Study

The methodological approach for this purpose was the single case study with two embedded units. Typically, in the case study, the researcher investigates “a contemporary
phenomenon within its real-life context, especially when the boundaries between the phenomenon and context are not clearly evident (Yin, 1994, p.13). Yin (2009) states the case study has a distinct advantage over other methods when a “how” question is being asked about a “contemporary set of events” of which the “researcher has little or no control over” (Chapter One). One fundamental problem of case study research is actually defining the “case” or the unit of analysis (Yin, 2009). The conception of assessment was examined in the history departments at two separate institutions. Thus, the units of analysis were the departments of a single humanities discipline at two four-year public institutions from the same southern state. State University Red (pseudonym) is a high research university while State University Blue is a doctoral research university.

Finally, the case study uses multiple sources of information and further explains that it reports a case description and case themes (Yin, 2012; Creswell, 2013). By getting to know the professors through interviews and by reviewing various course and university artifacts, I explored the subtle nuances of conception that humanities faculty hold concerning assessment and get a broader, more unambiguous understanding on how they conceive of and use their assessments.

**Site and Participants**

Originally the study was designed to examine multiple disciplines at one institution; however, due to lackluster participation it was modified to include a second institution. The participants of the study were humanities faculty from one humanities discipline at two state universities located in the South, State University Red and State University Blue. The researcher is a professor at State University Blue. In this particular state, all state institutions are under a single administrative umbrella giving the study consistency, as each institution was required to
address the same broad general education goals and outcomes; however, each site has autonomy to tackle those goals in the best way they see fit.

Site

State University Red is a public, coeducational, doctoral-granting, residential university. It has more than 17 thousand students with a little more than 30% recognized as an ethnic minority. State University Red offers more than 100 degrees with more than a dozen humanities offerings with many concentrations within each. For example, in the area of art State University offers majors in art education, new media and design, and studio art, among others.

General education at State University Red can be described as a distribution model. Its most recent modifications were seen from 2009-2012 when student learning outcomes and categories were reexamined and approved. According to the institution’s website, students must complete a minimum of 36 hours from the following categories: Fine Arts; Historical Perspectives; Natural Sciences; Mathematics; Literature; Reasoning and Discourse; Philosophical, Religious, and Ethical Principles; and Social and Behavioral Sciences. Their learning goals included: 1) Foundational Skills: Think critically, communicate effectively, and develop appropriate fundamental skills in quantitative and information literacies 2) The Physical and Natural World: Understand fundamental principles of mathematics and science, and recognize their relevance in the world 3) Knowledge of Human Histories, Cultures, and the Self: Describe, interpret, and evaluate the ideas, events, and expressive traditions that have shaped collective and individual human experience through inquiry and analysis in the diverse disciplines of the humanities, religions, languages, histories, and the arts 4) Knowledge of Social and Human Behavior: Describe and explain findings derived from the application of fundamental principles of empirical scientific inquiry to illuminate and analyze social and human conditions,
and 5) Personal, Civic, and Professional Development: Develop a capacity for active citizenship, ethics, social responsibility, personal growth, and skills for lifelong learning in a global society.

State University Blue is also a public, coeducational, doctoral-granting, residential university. State University Blue is an Historically Black College or University (HBCU). It has more than 10 thousand students, of which 80% self-identify as black or African-American. State University Blue offers more than 55 degrees and more than 6 humanities offerings with many concentrations within each. For example, the area of visual & performing arts offer majors in music education, music performance, and music (general).

State University Blue has a similar general education plan to State University Red. It is important to note the first general education reform took place in 2006. This is where the university moved to an interdisciplinary general education core. The second reform was six years later in 2012 and returned the university to the distribution model described here. The general education curriculum includes 33 semester hours of credit distributed across six broad categories of student learning: written communication; mathematical, logical, and analytical reasoning; scientific reasoning; social/behavioral sciences; humanities/fine arts; and student success. Each area had specific learning outcomes which were taken from the institution’s website: Written communication included applying writing practices appropriate to specific tasks and audiences and integrating the use of appropriate information technology tools throughout the writing process; Mathematical, logical, and analytical reasoning included: 1a. Apply quantitative and mathematical reasoning to solve problems in diverse contexts, 2a. Evaluate quantitative information using a variety of methods, and 3a. Communicate quantitative or mathematical information in multiple formats. For courses emphasizing logical/analytical reasoning: 1b. Apply logical reasoning to solve problems in diverse contexts, 2b. Evaluate claims using a variety of
methods, and 3b. Communicate logical reasoning in multiple formats; Scientific learning outcome was to analyze real-world phenomena, issues, and problems using principles and processes of scientific inquiry; Social and behavioral sciences was to apply methods of analysis used in the social and behavioral sciences in the examination of individual and group behavior; Humanities and fine arts learning outcomes included examining human experience through the interpretation of artistic, intellectual, or cultural expression.

General education requirements and general education assessment were extremely similar at both institutions. All participants documented the process by which courses were approved for general education, how reports were requested, and detailed the steps their respective departments took to satisfy those requirements. As one participant at State University Red explained, each of the learning outcomes categories might have multiple different disciplinary courses that satisfy that particular outcome. This was also true at State University Blue.

Both institutions had a General Education Council or Committee that oversaw the general education program. The purpose of this body was to bring, as one participant said, “unanimity and uniformity to those proposals.” Each year the respective Council or Committee would identify specific outcomes to review, typically one or two out of the six, and randomly select general education courses that were addressing that specific outcome. Instructors of these courses received notice from their chairs and received the rubric by which the outcome would be assessed. The faculty members were able to choose certain assignments, or questions within those assignments, that would be associated with those learning outcomes and send that data to the chair who compiled and sent them on to the council. From there, faculty also had a form to fill out to submit results to the general education committee.

Participants
There were a few criteria used to identify participants. Faculty had to have taught at least one general education course in the previous academic year. The faculty included full-time, tenured faculty who taught general education in a single humanities discipline. Because chairs of departments are in a unique position of being both faculty and administrators, they were included to help contextualize the leader’s view of assessment. Other full-time institutional administrators were excluded, as the focus of the study was on faculty and faculty conceptions. A faculty member’s academic rank, experience, or other demographic identifiers did not affect his/her ability to participate in the study. Data for this study was conducted through interviews, two per participant. These 7 participants were identified via snowball sampling and included Ken, Dale, and Jen (pseudonyms) from State University Red and Rob, Scott, John, and Brian (pseudonyms) from State University Blue.

Ken is the chair of his department. He holds a doctorate in his discipline and has been teaching for over 20 years. He typically teaches in the general education program. Dale is a tenured faculty member at State University Red. He is a white male who teaches 2 to 3 general education courses per year. He’s been teaching for close to 20 years and is tenured. Jen is also a tenured faculty member at State University Red. She has been teaching since 1994. She described her background with general education this way: “One of the main ways I’ve been involved in Gen Ed—I was the first Chair of the University’s General Education Council when the faculty senate created that job. I was the chair for the first three years of its operation when we were trying to figure out some reform initiatives because our Gen Ed program got a very scathing review from external reviewers.”

Scott, John, Brian, and Rob are all faculty members at State University Blue; Scott, John, and Brian are tenured. Scott and John identify as Hispanic. Scott is a tenured faculty member
who is close to retirement. He was actively involved in the general education reform (both 2006 and 2012) at the institution. On average he teaches 2 general education courses per semester. John is recently tenured. He has been at the institution since 2007 and has taught general education every year since his hire. Brian has been teaching in higher education for 12 years and is currently tenured. He is very animated when discussing general education and assessment, obviously invested in the process. Rob has taught in higher education since 1989. He was tenure track and taught at State University Blue through the Spring 2014 semester. He, too, was very concerned about general education and its impact on students.

Sampling Strategy

The sampling strategy for the study was modified from a two-part purposeful sampling with maximum variation to take into consideration basic demographic considerations including age, gender, and race and professional characteristics including status (tenure-track faculty, instructor, adjunct), teaching experience, and course load, to a snowball sampling strategy. At first the study attempted to use a qualitative on-line survey with three departments: English, history, and philosophy. One department declined to participate. Qualitative researchers have used surveys within their case study designs, both in an integrated (Stecher and Borko, 2002) and sequential (Spillane and Zueli, 1999) fashion. The survey was to gather demographic data and ask a series of open-ended questions loosely based upon Brown’s (2008) Conceptions of Assessment model. This survey was confidential and provided general perspectives of assessment so that the researcher could use maximum variation sampling to identify six possible candidates for follow-up interviews. Using maximum variation, the researcher attempts to find diverse cases and multiple perspectives (Cresswell, 2013; Yin, 2003). There were few responses to the on-line survey so a modification was necessary.
I moved to another purposeful sampling strategy: the snowball strategy (Merriam, 1998). Snowball strategy involves beginning with the chairs and getting references to other possible participants. As Merriam (1998) highlighted, this strategy involves identifying people who would be information rich in the subject area. At site Red, three individuals were identified, with one being on sabbatical. At site Blue, four participants were identified. It should be noted that one limitation of snowball sampling is that it is a group of people who have an interest in or knowledge of general education assessment; likewise, those who might have no interest in assessment or strong views against assessment may not have been recommended to participate.

**Recruitment and Access**

I recruited faculty from one humanities discipline, history, at two state universities. Originally I contacted the Dean from the College of Arts and Sciences at State University Red through an informal email to gauge his interest in hosting the study. The Dean indicated his willingness to review the application. Once IRB Approval was granted by NEU (Appendix A), I sent the packet to State University Red’s IRB for review and approval. After State University Red’s IRB approved the study proposal, I sent the permission letter to NEU showing State University Red’s consent to the study and received final confirmation to move forward with the study. I sent the final approved packet and all supplementary material including the recruitment emails for the survey (Appendix C) and interview (Appendix D) and unsigned consent forms for the survey (Appendix E) and informed consent for the interview (Appendix F) to the Dean.

Once the study was approved by both institutions’ IRBs and the Dean at State University Red, I contacted the three identified departments: History, English, and philosophy. I sent each chair a formal letter describing the study and the faculty consent forms. The description included why the site was chosen, who would be able to participate in the study, how the results would be
used and reported, and what would be gained from the research (Cresswell, 2012). I requested permission to reach out to their department members individually via email, and for permission to attend one of their department’s meetings to discuss the study. One department declined to participate. One department forwarded my emails to faculty, but declined to have me visit the department meeting and the third department both shared the emails and allowed me to visit their departmental meeting.

In the departmental meeting, I detailed the proposed study, shared the confidentiality information including the Consent for Participation letter (Appendix C and E), and answered any questions the faculty had. The consent form included a full description of the research problem and purpose, the data to be collected, protection of confidentiality, what the individual might gain from participating in the study, and the right to withdraw at any time (Appendix D and F).

I wrote a follow-up email and included the link to the survey. Even after multiple attempts to contact faculty, participation in the survey was thin. Subsequently, I modified the study by removing the survey instrument and changing the sampling strategy to just completing interviews (Appendix B). I emailed the two chairs and included a $25 Visa gift card for participating. One chair declined to participate. The second chair agreed to participate and recommended some faculty in his department that might be interested in the study. Two faculty members agreed to participate.

Due to the lack of participants, another modification was necessary. In this case, the researcher included State University Blue. Since only one discipline was represented at State University Red, the researcher decided to recruit participants from the same discipline at State University Blue. The participants were recruited from a different department and the researcher had limited previous contact with them.
I received permission to modify the study from NEU and contacted State University Blue’s IRB. Once approval was given, I contacted the chair of the department and included a complete description of the study and sample consent forms (Appendix F). After that, I made contact with five faculty members. Out of the five members contacted, four agreed to participate in the interview process. The challenge of recruitment was significant and cannot be understated, as it resulted in the necessity for two modifications to the study.

Data

Data Collection

Collecting multiple types of evidence is a standard practice with the case study (Yin, 2003; Creswell, 2013; Patton, 1987). Multiple types of evidence assist in triangulation and adds validity to the study’s conclusions (Yin, 2003; Creswell, 2013; Patton, 1987). The researcher collected documents concerning general education and assessment from both institutions and specific instruments from individual participants. The researcher also conducted two separate semi-structured interviews with each of the seven participants at a time and site of their choice. Five faculty chose to meet at their respective offices. Two participants chose to meet off-campus.

Document Analysis

The researcher gathered assessment documents from each institution as they specifically related to general education including assessment committee products and general education annual report(s). These were all freely accessed on the institution’s websites. Internal documents like copies of course syllabi, and sample course exercises (both formative and summative), were collected from individual participants and institutional websites. Yin (1994) noted that the primary purpose of the documents should be to “corroborate and augment evidence of other sources” (p. 81). Rubin and Rubin (2012) highlight that the documents should be treated in a
“manner similar to transcripts” and should not be treated as “literal facts,” but rather “interpretations” (Chapter 3, documentary analysis). The documents allowed me to discern the stated assessment goals, the types of assessment strategies and tools employed, and the self-identified strengths and weaknesses of their assessment program.

**Interviews**

Much of the data for the case study came from one-on-one semi-structured interviews, which lasted about 60 minutes each. The purpose of the interviews, as Seidman (1998) suggested, was not to evaluate; rather it was to understand other people’s experiences and how they make sense of their experience. The interviews were semi-structured and the questions “evolved” based on participants answers; and the questions were designed to elicit depth (Rubin & Rubin, 2012). I conducted each interview and used a digital recorder. I took some notes during the interview to identify important passages and possible supplemental questions. Immediately after the interview, Rev.com, a professional transcription service, transcribed the recordings.

During the first interview (Appendix G), I established a relationship with the participants by asking general demographic questions and open-ended questions concerning their views of assessment and general education. I also asked participants to sign the consent form. The second set of questions were individualized follow-up questions based on their first interview. Many questions specifically related to the participants experiences and explored their conceptions in more detail. During both interviews, clarifying questions were asked. After the interview, the participants were encouraged to share any other thoughts or insights they had concerning general education assessment that might not have been directly inquired about.

**Data Storage**
Confidentiality was kept through each stage of the study. All participants and their places of employment were given pseudonyms to maintain confidentiality. All data was stored on the password-protected computer of the researcher. This included survey results, audio files of interviews, transcripts of interviews, copies of provided documents, and analyzed data. A backup of all the evidence was stored on a secondary drive which was also password protected and remained in the researcher’s private desk. Transcripts and the researchers’ analysis were also uploaded into Dedoose, a qualitative researchers software program. The site was password protected and only the researcher and his committee had access. All physical documents were filed and stored in a locked file cabinet. The only people with access to the data were the researcher and the members of his doctoral committee. Transcripts, fieldwork notes, informed consent, and other documents will be stored for a period of no more than two years after the study’s completion. The audio files were deleted after the transcription had been completed and reviewed by the participant for accuracy.

Data Analysis

Transcripts were coded at the completion of each round of interviews. I used analyzing data software, the on-line service DeDoose, to help code, rearrange and view the data. The strategy I employed included multiple steps of within case and across unit analysis that provided specific products. First, data from each of the seven participants was coded using a descriptive process. Descriptive coding assigns basic labels to data to provide inventory of their topics (Saldaña, 2009). Tesch (as cited in Saldaña, 2009) explains that descriptive coding are labels on the topic, not the content or substance of the data. More succinctly, it is what is talked about. These codes can have more detailed sub-codes (Miles & Huberman, 1994). This coding process is particularly useful for a new researcher with a large amount of data forms. This process
involved identifying categories, themes, patterns, and ideas using the participants own language. For example, when asked how he would describe his general education program, one participant said, “I guess the term that comes to mind is kind of a cafeteria style. We have certain areas that a student needs to take classes from: math, science, and the Global Studies and Humanities.” I coded that excerpt as “Cafeteria / distribution model”. This code was directly based on the participant’s response. Through the first coding phase that included the interview transcripts and documents, I created a descriptive summary for each.

Table 1

<table>
<thead>
<tr>
<th>Data Source</th>
<th>Strategy</th>
<th>Analytic Focus</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interviews, Documentation</td>
<td>Immersion in each survey, interview, and document and complete initial coding (Descriptive, Structural)</td>
<td>Within each case</td>
<td>Identification of significant statements; identification of themes</td>
</tr>
<tr>
<td>Interviews</td>
<td>Immersion in each survey, interview, and document and complete initial coding (Descriptive, Structural)</td>
<td>Within each and across all cases</td>
<td>Re-categorizing codes/themes</td>
</tr>
<tr>
<td>Interviews, Documentation,</td>
<td>Pattern matching</td>
<td>Across cases</td>
<td>Identify “deep structure”; themes common to all participants</td>
</tr>
<tr>
<td>Interviews, Documentation</td>
<td>Critical reflection</td>
<td>Within and across all cases</td>
<td>Initial exploration of themes and their relation to the research questions</td>
</tr>
<tr>
<td></td>
<td>Organize themes</td>
<td>Within and across all cases</td>
<td>Essential structure</td>
</tr>
</tbody>
</table>

The second coding process was structural coding. It is most effective with multiple participants, and especially with semi-structured interviews (Saldana, 2009). Namey et al. (2008) state that structural coding “acts as a labeling and indexing device, allowing researchers to quickly access data” (as quoted in Saldaña, 2009) and both codes and categorizes the data.
according to specific frames (Saldaña, 2009). Using Brown’s (2008) Conceptions of Assessment (CoA) as a model, there were four codes that acted as the labeling device: for learning (student and instructor improvement); as learning (making students accountable), of learning (accountability of schools and teachers), and irrelevant to learning (it can be safely ignored even if it must be used, and/or it is inaccurate). Each of these categories assisted me in identifying and organizing the various perspectives the participants held in various settings. For instance, when discussing assessment one participant said, “I think it helped us talk to each other about what was happening in our classes. We established an environment where we could talk about our teaching techniques.” I coded this with Brown’s CoA as “for learning,” specifically assessment used to improve instruction. By using the structural code, I was able to clarify how participants’ thoughts, ideas, and insights worked within Brown’s (2008) theory, which helped in organization of major categories and themes. After the descriptive and structural coding was complete, five categories emerged. Each category included themes and sub-themes that described the faculty members’ conceptions of assessment of their general education courses as they are currently being described and implemented at their institutions.

Throughout both coding phases, memos were completed to document the process and organize my initial thoughts, reactions, and analysis (Saldaña, 2009). Moving between across and within-case comparisons facilitated the process of intuiting, which as Swanson-Kauffman and Schonwald state, is “the critical reflection on and identification of themes as they are found in the accounts of the multiple respondents” (as quoted in Ayres et al., 2003, p. 875).

**Presentation of Data**

Even though this is a multiple case study, the findings for each school will be presented together. The researcher organized the findings from both institutions into these overarching
categories for a few reasons. First, both state institutions work under the same administrative umbrella. Second, both institutions are using the distribution model and have similar learning outcomes and goals. Third, all the faculty came from the same discipline within the humanities and the majority of them shared comparable professional qualities, such as years of service in higher education and tenure status. Lastly, the researcher did not identify any idiosyncrasies between the institutions’ general education programs or any process that would justify separating the two institutions. In fact, the participants’ expressed a great deal of commonality in their conceptions.

It is important to note that assessment is a messy process and there is a great deal of overlap between these categories. For example, one assessment may serve multiple purposes, such as acting as a summative grade and acting as a form of feedback so the student can improve performance which can then be used to quantify learning for some outside body or organization. It may also be used to inform instruction and change teaching habits or the curriculum. Even so, the categories and themes that emerged are guides that will help make sense of some of the complexities of general education assessment.

**Trustworthiness**

Establishing reliability and validity was important in establishing the stability and quality of the data in the case study. Yin (1994) identifies four recognized tests to establish the quality of case study research including internal and external validity, construct validity, and reliability.

**Validity**

Creswell (2007) claims validation is the process of assessing the accuracy of the data. He continues by stating that time in the field, thick description, and working closely with the participants are among the techniques to validate one’s findings. To establish construct validity,
multiple sources of evidence were gathered including: documents, from the universities, on
general education and assessment; documents from some of the participants; and interviews. In
this study, the researcher had two formal interviews with each participant. Further, I collected
assessment documents, specifically focusing on how the participants present assessment tools
and subsequently use assessment data.

Allowing participants to review transcripts, parts of analysis, and the findings to make
necessary changes is also a technique that will assist in the validity (Yin, 1994; Patton, 2002).
Lincoln and Guba think it is one of the most significant techniques to establish validity (as cited
in Cresswell, 2007). Participants reviewed transcripts of the interviews immediately after
transcription. They also had an opportunity to review a draft of the dissertation to confirm its
accuracy.

Internal validity “lies in establishing phenomena in a credible way” (Riege, 2003, p. 80).
Yin (1994) states that the researcher should gain internal coherence of the data through
crosschecking the results. Multiple types of data, like interviews and documents, were collected
and cross-referenced in an attempt to identify irregularities and inconsistencies in the data. For
example, I reviewed interview transcripts concerning general education requirements and
assessment procedures and checked that against official documents posted on the institution’s
website.

External validity deals with the study’s results and attempts to establish generalizability.
There was data from multiple perspectives via seven in-depth examinations of participants in
hopes of, as Yin (1994) says, clearly outlining contributions and generalize them within the
boundaries of the research. Comparing the evidence with the current literature also developed
external validity.


Reliability

In a case study, the goal of reliability is to minimize the errors and biases in the study (Yin, 1994). One way to enhance reliability is through a sound case study protocol, which details how research was conducted. Yin (1994) states that a good case study would allow for an auditor to go through the same process as the researcher and come to the same conclusions.

As Kitto, Chesters, and Grbich (2008) note, triangulation of data is important in a qualitative study to increase validity and reliability. Thick description and a thorough accounting of the researcher’s actions during the study will assist with the reliability. Merriam (2001) defines thick description as “the complete, literal description of the incident or entity investigated” (p. 30). A researcher should be able to follow the same protocols as I have described here and complete a similar case. Along with the steps concerning validity, reliability, and trustworthiness, I took detailed memos during the data collection and analysis phases. I ended the study with over 60 memos on data. These memos allowed me to reflect on the data collection process. They also assisted me in capturing faculty comments, meetings, problems, and initial reactions and thoughts of events as they relate to my study. Being conscious of my background as a teacher and student of the humanities, taking detailed notes during the data collection and analysis phases, triangulating data, and using thick description helped in reducing researcher bias.

Protection of Human Subjects

The role of the researcher in this qualitative case study was to collect site artifacts and documents related to the study, and conduct one-on-one interviews. The researcher was responsible for data collection, storage, analysis, and working with participants to ensure data validity and integrity.
This qualitative case study collected data from humanities faculty teaching general education courses. Data included a survey, interviews, course artifacts and university documents. No psychological, emotional, legal, or physical harm was anticipated in the proposed study. The study was completely voluntary, and participants could opt out at any time. The risk of social harm was minimal and only existed in the event of an unanticipated breach of confidentiality. In that event, participants risked potential embarrassment or future harm to professional careers if confidentiality was somehow breached or identities somehow disclosed. Steps were taken to preserve participant confidentiality. Confidentiality was kept through each stage of the study. All data was stored on the password-protected computer of the researcher. The only people with access to the data were the researcher and the members of his doctoral committee. All participants were given a pseudonym to help protect their privacy, and information regarding participants was not disclosed to people outside of the research group.

Audio recordings were transcribed by Rev.com, a reputable transcription company, and did not label participants with any personally identifiable information. Individual notes taken during the interviews were maintained in the same way on the same computer and participants will be referred to by their pseudonym. Other information that may inadvertently identify the participants (i.e. school, department, name of course they teach, etc.) were redacted or changed prior to presentation. Seidman (2006) underscores the significance of using long quotes that might allude to or be used to identify participants. To address this possibility, participants were able to review both interview transcripts and the final draft of the study prior to dissemination. In this way, risk was minimized.

In accordance with the NIH (2008) requirements, the researcher presented a copy of the Informed Consent document to prospective participants when they are invited to participate in
the study (Appendix E). Participant Signatures of Informed Consent were required prior to any written data collection. Before interview data collection, the researcher also reviewed the Informed Consent document with each participant and answered any questions the participant had.

Chapter Summary

This single embedded case study was designed to identify how humanities professors typically conceive of general education assessment, how they use that assessment data, and what their rationale is for using assessment data in those ways. Many qualitative researchers (Stake, 1995; Yin, 2003) subscribe to the constructivist paradigm and use the constructivist approach for their case studies. Constructivists suggest reality is subjective and influenced by individual experience, perceptions, and interactions (Ponterotto, 2005). Humanities professors may not have had the opportunity to reflect upon the myriad of influences on their conceptions of assessment and this study was designed to explore those issues with them. Ultimately, the goal using a constructivist approach in this study was to understand how assessment is constructed from the “lived experiences” (Ponterotto, 2005) of faculty members who have a responsibility to assess student learning.

A case study was chosen since this was “a contemporary phenomenon within its real-life context” (Yin, 1994, p.13) that needed to be explored. Further, embedding cases within a single case study focused a case study inquiry (Yin, 2009). By getting to know the professors through a survey, interviews, and by reviewing various course and departmental documents, the researcher explored the subtle nuances of conception that humanities faculty held concerning assessment to get a broader, more unambiguous understanding on how they used their assessments.
Following approval from NEU and State University’s IRB offices, this researcher interviewed seven humanities general education teachers in two departments chosen via snowball sampling and completed document analysis. The collection of data was completed over the period of four months. Interviews were transcribed and reviewed by the participants for accuracy. Coding of the data was completed using a software program DeDoose. The data was securely stored then destroyed as detailed within the chapter. A number of processes were employed to increase the validity and reliability of the study. These included using thick description, using multiple forms of data and cross-referencing the data, and comparing the findings with the current literature. To minimize errors and bias, the researcher triangulated the data, took detailed notes during the collection and analysis stages, and used thick descriptions.
Chapter 4: Findings

Common conception of general education assessment in the humanities affects higher education institutions and student learning. In order to better understand how humanities professors conceive of, and approach, assessment then use the collected data, the following questions were formulated:

• How do humanities professors typically conceive of general education assessment?
• In what ways do they use the assessment data?
• What is the rationale behind their use of assessment data?

This chapter will discuss the findings of the interviews, and document analysis will follow, including a description of the categories, themes, and sub-themes that were identified.

Data

In order to address the research questions, the findings from the participant interviews are arranged into four major sections, based on the theoretical framework (Brown, 2008): assessment for learning, assessment as learning, assessment of learning, and assessment as irrelevant to learning. Within each section, themes and sub-themes are described, but first, general education and its connection to assessment at the two institutions will be described.

General Education

The participants discussed general education specifically. Not only how they spoke of general education, but its relation to the major was noteworthy. Jen discussed her perspective of the reform on her campus saying, “We were basically trying to figure out what Gen Ed was. We all operate in our disciplinary silos. . . . Their main purpose was to protect the area.” This protection of the disciplinary area is not unique to State University Red; it was also identified in the minutes concerning general education reform (2012) at State University Blue.
While a majority of the participants were tenured, they all taught general education, which may be atypical nationally speaking. Rob ventured to explore why tenured faculty do not usually teach general education.

“It’s a lot harder to teach the introductory level class than it is the senior seminar. There’s different challenges there; reading the larger papers is more time demanding, but in terms of reaching the students, by the time the reach their senior seminar, they’re schooled in what it means to be in school. And you don’t have that at the freshman level.”

Students in general education need to be taught how to be university students and how to approach university courses. This is a time consuming prospect. Added to this fact, general education courses are typically larger, which affects teaching strategies and outside class time investment. Thus, the size of the courses affected the types of assessments that could be effectively administered and analyzed. For example, Scott said, “You can only really measure content in a 100 level course, unless of course it’s capped at 25 instead of 50. Then there are other things that you can do.” Relating a story about general education assignments, Dale noted one faculty member’s incredulous reaction concerning a general education course, the amount of students he would be teaching, and the subsequent grading that would be required. “How am I supposed to get any work done?” When putting general education courses in context with other responsibilities like teaching majors or graduate students, research, and service, the time commitment is significant.

John discussed the rationale for teaching general education and some faculty’s propensity to do so. “Generally speaking, instructors like smaller classes, and they like to teach upper level classes.” Rob developed that idea, explaining, “I think traditionally many academics kind of cleave to the classes that are upper level and in their own expertise, and that’s kind of where
they’re most competent, that’s where they’re most comfortable, and classes also tend to be smaller.” Essentially, the participants are preferring the more specialized content and smaller classes, where students are typically naturally engaged in the material because it is their major.

Participants discussed the unique teaching situation of general education. Rob described the students who are relatively young and not sure what it means to be a student, and the teachers who typically get those courses—adjuncts or lecturers who are rarely acknowledged. “So respecting and understanding what it was we did day to day to day as workers, as educators, God, I never heard anybody above a dean even acknowledge that difficulty of our work, or the challenges of our work, so how do they recognize the successes? They don’t really even have the capacity or the perspective to see what it was we were actually achieving.” John, though, noted that there are some professors who enjoy the challenge of general education courses. “There are plenty of instructors who really thrive on teaching the freshman survey or the general education course. The challenge of it.” Others, like Dale, noted that there was typically no choice about teaching general education: “...we’ve written into policy that you should expect to teach at least one and possibly two 200-level courses per year.”

One aspect of teaching general education was the curriculum itself and the discussion concerning how it should be coordinated. For example, at State University Blue there was a shared text, and some shared assessments across sections. Rob identified the strengths of such of a uniform program in this way:

So we had basically all of our different sections under different instructors from different disciplines charged with this task. And through the assessment, everybody talked about it, found certain ways of presentation worked and others didn’t, and certain approaches worked. Some of our instructors were more comfortable with it than others, but the point
is we were talking about it. . . . So the assessment served as a common issue that we could set upon, and then actually go into deeper issues.

The shared syllabus or shared assessments across sections allowed for a great deal of collaboration between instructors. They could share resources, brainstorm solutions to problems, and share successes.

While many discussed the benefit of having a common syllabus and assessment, Scott relayed his concern, “You can’t force history.” Scott was arguing for more flexibility in what content was taught and how it was presented. At State University Red each faculty member is in charge of both content and assessments in each individual general education course. Dale said there is no uniformity and he was not sure how faculty would react to the prospect of a lack of flexibility.

Another aspect of general education that the participants discussed was its relationship to the major. A few participants specifically discussed how those general education classes should set up students in major specific courses. Ken framed it by highlighting the skills students needed to be competent in, prior to 300- or 400-level courses. He acknowledged, “A lot of it is individuals, professors, saying, ‘Well, we just really need to spend more time laying the groundwork, because it really is like scaffolding.’ ” Dale shared the same sentiment, explaining how students in his major courses influenced how he approached his general education courses. “. . . I have taken small tweaks to my Gen Ed level class, where I’ve actively tried to make the Gen Ed class more about a [discipline] a method and a profession and less about specific content.” He then detailed an assignment where he would “walk” a student through the process and let them practice applying the skills in the hopes they would transfer the knowledge to later major courses. Scott had an analogous situation in his general education courses where he
showed his general education students how to be students. He then envisioned a more purposeful assessment program where content knowledge would be assessed prior to major courses.

“Instead of having these general assessments that are worthless across the board, why not empower the department to do the kind of assessment that will build a good major, and in the process of building a good major will also help the student learn how to extract material from a text, to analyze content, and to organize and to write clearly.” Again, he was suggesting scaffolding the knowledge and skills in general education to be meaningful to the majors.

**Summary**

Both institutions recently reformed their general education programs and both faced a similar problem in departments attempting to protect their areas. Typically, faculty do not feel ownership of general education in the same way that they do for their major courses. This may affect faculty’s interest in investing time and effort into the assessment process. While there are some senior faculty who enjoy the challenge of general education, normally general education courses are taught by lecturers and adjuncts. Also, general education courses have higher enrollment caps, thus more students per section, and those students are typically not well versed in how to be a student. In turn, faculty find general education courses to be more time consuming than courses within the majors, even without the added work of assessment.

Uniformity (whether in syllabi, content, or assessments) across sections, versus having complete instructor flexibility, was a tension felt at both institutions. While uniformity may assist in instructor collaboration, data collection, data analysis, etc., professors need the ability to meet changing student needs during the term and feel they will lose that ability if everything is codified.
Ultimately, faculty’s tension between general education and the desire to focus on the major creates an environment that makes it difficult to find general education assessment practices that are worthwhile to all the stakeholders. Making general education meaningful to majors may at times be in conflict with making it inherently meaningful for general education students. The distribution model may also facilitate more alienation with individual professor’s investment into general education assessment. A dedicated, professional faculty whose focus is general education seemed to benefit the multiple stakeholders (students, SACS, etc).

**Assessment for Learning**

Three themes and multiple sub-themes help describe how the faculty participants identified assessment as part of the learning process. Assessment for learning indicates that participants use assessment to improve learning in some way. The three themes were: 1) Assessment as a Process for Describing Abilities 2) Assessment as a Process for Improving Student Learning and 3) Assessment as a Process for Improving Teaching. These themes and their accompanying sub-themes will be described next.

**Assessment for Describing Abilities**

Every participant described the general education assessment process as describing student abilities to some degree. Overall, the participants believed that assessment able to describe student abilities was helpful. They discussed assessment that described abilities as revealing how much students learned, what students learned, and students’ ability to use higher order thinking.

**How much students learned from teaching.** Responding to questions concerning the purposes of assessment, participants related how much of the content the students learned, or the value added due to instruction, over the course of the class. Many participants discussed the use
of pre- and post-tests, along with other summative assessments like finals and essays, to identify how much students learned. John, Rob, Scott, Ken, and Brian all used a pre- and post-test in their general education courses, for example. John shared the overall structure saying, “They’re expected to give the same pre-test and post-test which measures student’s content knowledge in the beginning to the end of the semester. Everybody does those and those numbers are crunched at the end of the semester comparing those two.” Rob expanded on the description of the process, which was fairly representative of all of the participants, saying:

We would basically give a 20-question quiz or just a test… the second or the third day of class to see if the students knew certain concepts, knew certain facts, and then I would throw in later kind of a writing sample of it. . . . Then at the end of the year, at the end of the semester, we would repeat the exact same questions in terms of the multiple-choice parts. That way you compare apples to apples.

In these assessments, the participants are attempting measure a beginning and end point in terms of student content knowledge. Essentially, how much did students learn over the course of the semester?

Instead of a pre-test and post-test, Ken used an assignment that was graded and then used a similar assignment at the end of the semester. “I have, on occasion, done an assignment of some kind that would be, again, at the very end of the semester that might measure that [learning outcome or skill].” The purpose of the assessment was specifically ascertaining how well the students learned, or did not learn, the specific skill set over the course.

**What students learned.** This sub-theme grew out of the participants understanding of assessment as measuring exactly what students learned over the term. Brian was the most succinct, stating, “I understand assessment to be the process of discerning where a student is in a
class or where an entire class is at with respect to learning the material.” John agreed, and thought there might be a little more flexibility in how those measurements were taken. “Assessment is a way of judging how well we’re able to get through, or how well our students are able to perform, to demonstrate knowledge in a particular field.” Scott focused on the importance of assessing specific content knowledge, and thought that assessing exactly what students learned should be more integral to the process. “If we had serious assessment as a department, then each faculty member, regardless of how strong he or she is, would be compelled to deliver content and then measure that student’s understanding of content with an end-of-year exam. . . .”

All of the participants shared a variety of assessment instruments and assessment techniques that they used to measure exactly what students learned. For example, John said, “You have to have figures, yeah. But how you get to that place is…there’s a lot of roads to get there.” A majority of participants specifically discussed using multiple types of summative assessments including exams, essays, quizzes, oral presentations, and clickers, among other things, to measure what students learned.

**Assessment measures higher order thinking.** One of the sub-themes attempted to capture assessments that examined higher order thinking skills. John described some of the difficulty when trying to measure intellectual growth. “[The instructors are] also struggling with getting past a student’s embedded ideas of there being a right and a wrong answer, as opposed to what we like to promote, which is that [it] is messy. Yeah, there are facts. There are also a lot of facts that are in dispute. There are debates. There are arguments that depend upon evidence.” Unlike describing what a student knows or showing how much more knowledge a student has after a course, John is alluding to the difficulty of quantifying those higher-order thinking skills
like analyzing, evaluating, and synthesizing information that does not have a right or wrong answer.

Even with this difficulty at attempting to measure a student’s higher-order thinking skills, all of the participants were engaged in attempting to capture the growth. “Assessment means tracking the progress of the student over the course of the semester, but not just tracking the student’s performance on the basis of a grade but trying to track a student’s intellectual development,” Scott said. This type of rationale underlies what many participants described as a major part of their jobs and their conceptions of assessments. John highlighted the importance of moving students from a focus on the test to a focus on the process of learning. “Preparing them for how to think, for how to think critically as opposed to how we’re going to measure them because that’s what they’re used to and that’s what they’re looking at. Sometimes that’s even what they’re most comfortable with. Thinking critically is oftentimes outside their comfort zones so they could be encouraged to think critically.” This movement from merely presenting content knowledge, to critically thinking about the material, is foreign to many general education students. Not only is the process difficult, but also students may resist it because it is unlike previous educational experiences.

Describing some of his assessments, Ken alluded to the intellectual growth that happens as skills are developed. “I’ve tried to strike more of a balance of lecturing to give them information, because a lot of times, they just don’t know anything. If you don’t know anything, you can’t do the skills anyway because you don’t have any. You can’t do them [the skills] devoid of content. Just giving them a bunch of information is not very helpful either, unless they can understand how historians come to create these narratives by looking at actual source material. . . . It’s a real skill that they have to learn how to do. It’s an important one, too, I
think.” General education students need a lot of guidance and skill development; and merely testing on content, on what they know and how much they know, is not enough. Brian was equally excited when discussing their general education course’s final exam that attempted to measure higher order thinking skills. He said,

It ties in all of the material that we’ve done in the course in a really interesting way. . . . It asks new types of questions that you would have to understand the different concepts to approach a situation. . . . Students that haven’t been learning the material come to those and are sort of, like, ‘I know what you’re asking. I can kind of do these, but now you’re asking me to put stuff together.’

Again, the focus is not just on content. While content plays a role, student thinking about the content in new ways and from different perspectives is more important. They have to apply that knowledge.

**Assessment identifies student strengths and weaknesses.** Another aspect of assessment many participants described was using the data to gauge where students were in terms of their skills and knowledge. Ken mentioned, “One thing I think the assessment data could show is things, places where students are weak. . . .” As Rob explained part of the rationale for a pre-test: “I couldn’t assume that students were all reading at a college level. The assessment gave me clear evidence that was an issue the very first day or at the very first week and so I was able to create a new practice that dealt with that problem.” In this case, he was able to identify course weaknesses and plan the curriculum accordingly. Relatedly, John defined formative assessments in his courses serving this purpose of identifying strengths and weaknesses. “You have to give work to students that are going to both prepare them for handling the materials or find out where they are with the material that you’ve assigned, okay, formative assessments.” He continued to
explain the rationale for these types of assessments: “…it gives me a window into what I can do better; at the same time it also gives me, provides a window into, the limitations my students bring into the class as far as background and also even what they’re taking out at the end of the course. That helps me figure out adjustments.” John noted that the assessment focuses what areas he needs to spend more time on to help individual students.

Brian gave a similar response concerning formative assessments, stating, “... formative assessment that we do in our classes is the type of assessment that helps a student gauge how well they’ve learned the material and what they continue to need to do. It’s also beneficial to the instructor to know where the class as a whole is or maybe individual students and what they may need more emphasis on before you give them something that actually counts as a part of their grade.” These assessments, then, help the student identify strengths and weaknesses and assist the instructor in identifying both individual student weaknesses, and also any the class as a whole might face. Rob described the purpose as an opportunity to get the teacher and the students on the same page early on. He said:

... it’s really asking questions about where the student and the teachers intersect and whether or not that connection is working because you can’t assume it is, and certainly it may not work for all students equally well, and so this is an opportunity to get some measurable data with specific questions or problems that you start with at the beginning of the semester and what we would do with it.

By identifying student strengths and weaknesses, Rob is able to meet students where they are, while also showing students where he wants them to be. This “intersection” allows them to plan and move forward together.
John speculated that the data could also be helpful for the administration, “I guess it’s not just me individually, but this is helpful for us as a department and as an institution to get those aggregate figures to figure out where our students, what are they responding to the most and whose classes are they responding to better than others. . . .” However, it was unclear if this was actually the case.

**Assessment for Improving Student Learning**

Through a series of questions related to how they use their assessments, the theme that general education assessment improves student learning emerged from the participants’ responses. Brian’s thoughts on this matter succinctly exemplify the theme, “Assessment done right, any assessment that’s done for a course, should improve student learning.” There were two sub-themes including: 1) assessment provides feedback about student performance and what they must do to improve and 2) assessment motivates students to improve learning. It should be noted that all of the participants consistently felt that their assessments served multiple purposes. For example, Brian suggested that many summative assessments like essays not only measured student knowledge of a subject or showed competency in rhetorical strategies, but also acted as formative assessment through his end comments that offered direct feedback and specific suggestions to students to improve future writing.

Identifying the way a student learns and his stake in improving that process, Scott commented that students were “ignorant of the ways in which people learn how to be students. They’re ignorant of the ways of the processes of being a student.” Students do not have the skills to be successful at the university level. Most participants agreed with this observation. Take Ken’s statement, “I guess I used to think, ‘Okay, well, I’ll show them how to do it, they’ll figure it out, they’ll do it, and they’ll move along,’ but that didn’t happen. It really was like a foreign
language to them and they really needed more practice and hand-holding to get there.” Or consider Brian’s explanation, “They even don’t understand how to approach the material, how to learn the material, how to use the material. After the first couple exams, students are really shocked and surprised at how poorly they’re doing.” Other participants conceded similar points, sharing that the assessments help students improve their learning by improving in understanding what it means to be a student at this level. Scott sums this idea up, stating:

... I showed them how to extract the information from a text, I showed them how to take notes, and I emphasized that we were doing two things here on the 200-level course. We were learning some..., we were learning some content, but in some ways, more importantly... I was introducing them to a method of study, which I think could be used throughout their university experience.

The assessment process was instrumental in identifying student needs and assisting instructors to reevaluate their process so learning could take place.

**Assessment provides feedback about performance & offers objectives.** In this sub-theme, all of the participants discussed the use of assessment to offer direct feedback on student performance and to offer future objectives. The feedback could be in the form of summative marks like grades or formative assessments like comments on student work. Brian described summative assessments as “a culmination of or an attempt at evaluating exactly what worked the best for the group, what the student knows, and what the student doesn’t.” Formative assessment were more “the type of assessment that helps a student gauge how well they’ve learned the material and what they continue to need to do.” They also assist instructors refine their pedagogy. Six out of seven participants used both types of assessments to provide students with their feedback. For example, Dale’s discussion of his grading process was typical of the
responses. This process included making specific comments on individual essays to using
Clickers to gauge student knowledge and make class adjustments in real-time.

Assessment as student motivation. Another sub-theme that emerged from the interviews was the use of assessment as a motivational tool in the classrooms. John mentioned the difficulty of using formative assessments because of the lack of meaning for students. “In some ways you can only do so much if students are, for instance, not taking the pre- and post-test as seriously because they know it’s not graded.” Brian was more direct stating, “They’re [students,] not internally motivated necessarily to do well in the class.” He continued, “There has to be some skin in the game.” Brian then shared two examples. First he described his use of a formative assessment, a worksheet in class that was worth zero points, and how students didn’t take it seriously at the beginning. “If I say, ‘I’m going to quiz you on this at the end of class,’ now they’re working on it. Now they want to know. Now they want to get it done.” After one session, Brian invited them to his office to discuss the worksheets. “I met with a bunch of them and . . . you could tell that they wanted to just know the answers because they knew it would probably be on the test. What they didn’t want to know is to actually learn and try it, learn how to do it so that they can prepare for anything on the test. That motivational aspect, these assessments are really important,” he said.

Dale came to the same conclusion about using assessment as a motivational tool. “I have also come sadly to the point where we also now have quizzes. These are content quizzes. Sometimes weekly, sometimes daily if the students aren’t doing the reading. 15 years ago, I could assign reading and 60% of the class would do it regardless of whether there was a quiz. Now that is not the case.” He, and other participants like John, Brian, and Rob, have also used Clickers with part of the purpose to encourage attendance and to take in-class quizzes.
Assessment for Improving Teaching

One of the major themes the participants discussed was the direct effect assessment had on their teaching. All the participants said assessment affected their thinking about their courses in terms of curriculum and instruction, and how they interacted with the students in them. Dale’s comment is fairly representative of the assessment movement’s effects on the instructor. “The reality is that even though I’m resistant to what I consider the whole assessment apparatus from course to the top [of the] university, the reality is that I have done a lot more thinking about how I teach and what my expectations are in teaching than I did before this came into being.” He later added how some faculty have really taken the assessment concept to heart and, “spent a little more time on their classes and more thinking about those issues.” The changes in instructor thinking can be seen in how the instructors actualized their findings in modifying their teaching or the way they approach the material. In this theme, three sub-themes emerged concerning assessment: 1) Assessment modifies teaching 2) Assessment facilitates collaboration and 3) Assessment highlights distinctions between content and skill.

Assessment modifies teaching. A sub-theme that emerged was that of the actual practices faculty implemented due to the assessment movement or the assessment results. Taking a contextual view, Jen noted, “In an ideal world the only thing that is going to make an impact is if we’re changing teaching at the grass roots level in the classroom.” If the assessment was not used to reexamine what was working or not working in terms of furthering student learning, then it is not being used in the most productive way.

There were many participants who described assessment results being used in their classes to initiate change. For example, Brian said, “I think the real utility of assessment is reflecting on what are students really getting out of my class and what are they not getting out of
my class that is really important for them to get, then going back to the drawing board and saying, ‘How could I teach this differently? Better? In a different way?’ so that students don’t consistently miss this after I thought I taught it.” Here Brian noted how the assessment sparked the reconsideration of the class. Many participants used assessment to identify student competencies and plan their classes accordingly. Rob said assessments “gave me clear evidence that was an issue the very first day or at the very first week and so I was able to create a new practice that dealt with that problem.” John agreed, stating that assessment:

- gives me a window into what I can do better. At the same time it also provides a window into the limitations my students bring into the class as far as background and also even what they’re taking out at the end of the course. That helps me figure out adjustments.

Faculty use assessments to figure out how to implement changes as the class starts, to best assist students during the semester and to refine the course before the next cohort comes through.

Other participants discussed some of the specific modifications they made in a course due to assessment. For example, Brian said:

- One thing that [assessment has] certainly done is changed the way that I practice with students in class on problems, rather than just talk about the material. Even if I practice with them in front of the class, asking questions and have a PowerPoint slide, that is not just effective with the students. . . . I have to hand out worksheets where they are independently trying to work through a problem on their own, then when they have questions, come around and answer them.

Brian moved from a lecture-based classroom where he would explore content to one with more student interaction, one where students could practice specific skills and ask questions.
Likewise, Scott went from giving a reading assignment then a quiz, to a more limited focused exercise.

I would do a PowerPoint on a chapter, but rather than doing the whole chapter, I would focus on maybe seven or eight pages in the chapter and we would discuss that, and then I would have them read it carefully. Then I would quiz them on that.

Scott used assessment to change from asking students to just complete certain tasks to modeling academic behaviors and walking through the assignment so they could complete the task.

Dale integrated assessments in class and would modify teaching on the spur of the moment.

I will have a couple of questions where I’ll stop and ask them as a way of assessing within the individual class . . . Here’s a couple possible explanations of what we just talked about. Which do you think encapsulates what we’ve done?’ If the students are all diverging from what I think we’ve done, then we stop and we go back over that material.

This type of just in time teaching uses assessment to identify student needs in real-time and addresses those concerns before moving on to new material.

Ken noted that assessment data affected his overall classroom, stating:

I think that the more I’ve seen some of the results of how difficult this is for students, I realize that I needed to take it a little bit slower and have more modeling of this for them, and give them more opportunities to try again. . . . I think one of the things I’ve done in my survey class is to have more assignments, smaller, more assignments, building towards learning some of these skills, so that hopefully, by the end of the class, they actually do understand. I’ve included more things on my exams where they actually can use these skills as well.
The assessment process highlighted student difficulties and Ken revised his entire class methodology on that data. He reconsidered his preconceptions of student competencies and helped students build those academic skills, then used assessment again near the end to see if his changes were successful.

These types of changes were not just at the individual course level. John, Scott, Ken, and Rob also discussed specific changes assessment results had on the curriculum as a whole during the 2006 reform movement at State University Blue. “We were engaged in launching a very ambitious curriculum where we wanted to demonstrate that our methods were having an impact. So the assessments were absolutely key. . . .,” said John about the program.

The general education faculty worked closely together. The assessments played a major role because the curriculum was cutting edge; therefore the assessments were needed to validate what they were doing in the classroom. Rob continued:

. . . we [faculty] were talking laterally among instructors. . . . We talked about teaching. . . . We would actually have measurable data to help us make these kind of choices and just provide a bed for this discussion and even if we kept a reading that perhaps wasn’t as successful as we wanted, we would talk about how we would deliver it differently or how we would use it differently in class.

Individual instructors were not working in isolation, but with everyone on the team, to identify strengths and weaknesses of the course and how best to assist students. Many different instructors across multiple sections of the course were having these conversations.

It should be noted that some participants disagreed with the idea that assessments were the cause for instructor reflection and implementing changes to the classroom. Ken ventured to say that it may actually be an innate aspect of teaching.
. . . faculty constantly are trying to figure out ways to make things work in their classes. I think it’s just part of teaching. Teaching is very experimental anyway. You do something and you leave class, and like ‘Boy, that didn’t work. Why didn’t that work?’ . . . Or if something did work, and go, ‘Wow, what did I do? How did that work? I didn’t understand, but write that down. . . . It varies from faculty member to faculty member, but I think most faculty have some of that in their make-up.

Thus, the changes made in a class may or may not be related to the assessments. In the same vein, Dale said, “I have changed my teaching substantially based on my own perceptions. Part of that is just what I think a teacher needs to do to be a good teacher. I don’t know that’s really a product of an assessment culture.” Jen concurred, “I see teaching in general that way so I am not sure that is tied especially to assessment. Teaching is one of the most creative things there is.” These instructors see teaching as a profession that is constantly self-monitoring, self-assessing, and “good teachers” make changes to their courses based on their classroom experiences.

Assessment facilitates collaboration. A particular sub-theme to changing instructor thinking is the collaborative and communicative aspect of assessment. Brian identified this benefit to the department in this way: “Here are some other side effects of assessment that I think are really great. I love the way assessment in General Education courses forces faculty members to talk about what they’re doing in their classes, forces them to collaborate on the effort of teaching. . . .”

Jen also highlighted assessments’ collaborative effect on the faculty by giving the faculty reasons to discuss issues: “Our challenge is to define a more intentional and coherent curriculum and assessment has, program assessment, has created opportunities for us to have those kinds of
conversations, and they didn’t exist before.” Faculty are using assessment as a reason to collaborate on both instructional and curricular issues.

Rob thought the 2006 reform integrated assessment and general education in such a way as to make a huge positive impact on the faculty.

It [assessment] became very organic for us to work through our problems and our challenges as instructors with other instructors and we of course put the data up and it went upstairs and then it got used for administrative purposes by our dean and the provost and the powers that be but at the ground level, it actually helped us get to know each other and actually have a good sentiment at the institution.

The assessments were used in multiple capacities and by a variety of stakeholders; and faculty seemed to have buy-in due to, at least in part, the teamwork that the process facilitated.

In a similar vein, John thought about the future utility of assessment to the department faculty, “I think it will bring the different units within the department closer to the others as far as sorting out best practices and how to improve while maintaining autonomy at the same time.”

The point of emphasis here is the distinction that faculty work together yet have autonomy in their work as well.

Rob offered a concrete example of how the assessment similarly affected the entire faculty saying,

Well, it [assessment] allowed us also to compare across classes because an instructor could change their midterm exams and their quizzes and their assignments, but the assessments were something that we would then be able to compare to each other and at least begin to talk because then the focus between the semesters or over the summer would be how do we need to change the assessment itself, the instrument itself and so
important, meaning elevated. It was something we could work on and it was a conduit to talk with our colleagues, a purpose and a reason to actually talk about what we felt went well in the class and what we didn’t.

The process facilitated faculty cooperation at multiple levels including curriculum development, classroom instruction, and assessment development. Ken also found faculty working together more due to assessment.

Well, I’ve mostly been involved at the department level because the Gen Ed Assessment is set at a higher level, but in the department, I see it as we’re setting goals about the things that we would like our students to learn, and hopefully, we’re designing ways to test that and figure out if in fact we were doing what we hope that we’re doing, and if not, figure out ways that we can do it better.

The focus is on goal setting and reviewing student products to see if the goals are met and what could be refined to improve the student success.

**Assessment highlights distinctions between content and skill.** This leads to another thematic element, the use of assessment to examine content versus skill. Versus does not mean in conflict with, but rather highlighting a conceptual distinction the faculty were dealing with when they discussed assessment. Ken was the most succinct, stating:

I think there’s got to be a balance between the content and the skills, is what I prefer. I think if you adopt that, then you have to realize that you’re not going to be able to maybe cover all of the contents you want to cover, so the coverage is not necessarily the only goal. If you want to try to build these skills, then you also got to just take some time to actually help students learn how to do that, which means sometimes taking baby steps on that kind of stuff.
There are two pieces that faculty recognize. One is the content of the course, the specific subject matter being explored. The other is composed of the skills involved—for example, critical reading or analysis of a text is a skill. This skill would include identifying main concepts and evaluating claims, among other things. For general education, both are important. Ken recognized that students needed both content and skill development to be successful.

Brian also commented on this distinction and how assessment has clarified the need to practice skill building. “One thing that it’s certainly done is changed the way that I practice with students in class on problems, rather than just talk about the material,” he said. Scott was specific in focus on skill-building versus content saying, “I took them through the process of, I showed them how to extract the information from a text, I showed them how to take notes. . . .”

This is not saying content is not important, but the faculty often found students unable to engage the content in complex ways without the foundational skills first.

Relatedly, Ken spoke about his general education classroom with a focus more on skill building than content, saying, “I’ve been pretty convinced by people who’ve suggested that these are transferable skills, so learning how to do these things is something that’ll help them, not just understand the past, but hopefully help them navigate life in general, by knowing how to analyze information.” Thus, while general education may be introducing students to the discipline, more importantly it is giving them ways to hone skills that they can use in other classes and, ideally, outside of the university.

**Summary**

For the assessment for learning category, three themes emerged: 1) Assessment describes student learning 2) Assessment improves student learning 3) Assessment improves the quality of teaching.
The first theme used the general education assessment process as a way to describe student abilities. This included four sub-themes: 1) Assessment measured how much students learned 2) Assessment measured what students learned 3) Assessment measured changes in higher order thinking 4) Assessment identified student strengths and weaknesses. Current general education assessment practices do a good job of allowing instructors to quantify student content knowledge and describe student learning. How much students learned was identified primarily through a pre-/post-test or having a similar assignment at the beginning of the term and at the end. In this way, instructors were attempting to identify how much students learned throughout the course. What students learned was the second sub-theme. Instructors used assessments to find out what students knew at the end of the term, and they used a wide variety of assessment instruments to gauge this.

The third sub-theme was trying to describe changes in higher order thinking. The difficulty is when instructors try to quantify student’s higher order thinking skills and or development, because there are layers of subjectivity in the assessment process that are not easily quantifiable, and instructor perception and/or description of that growth may not accurately capture student growth. For example, evaluating an argument depends on a number of factors including interpretations of data, presentation and representation of author’s ideas, and examination of rhetorical strategies, among other things. A faculty member who is coming from a certain perspective may quantify the success of that student’s endeavor based on, in part, his/her perspective. Further complicating the process, this intellectual change was the most difficult to measure and to quantify in large part due to students unfamiliarity with the processes they were being exposed to and asked to perform.
The last sub-theme in this theme was using the data to identify student strengths and weaknesses and how that affected their teaching. Assessment worked well in assisting professors in identifying both individual and classroom strengths and weaknesses through formative assessment and in offering distinct feedback to students on how they can improve.

Within the second theme of improving student learning, two sub-themes emerged. The first was that assessment provides feedback and offers the student objectives to achieve. Providing feedback could be done in a number of ways including but not limited to summative assignments and formative feedback, like responses to writing or in-class discussion, so the student may attempt to improve future performance. Most assessment instruments will likely offer the student both types of feedback.

The second sub-theme was that assessment could be used as a motivational tool. In our current numbers-focused society, assessment is a great motivator to get students engaged with the material. Some instructors found students were not internally motivated to do well or learn the material. However, if an assessment included some point value or grade, students were much more likely to practice and attempt to master the concept or skill.

The final theme in this section was that assessment improves teaching. Three sub-themes emerged. 1) Assessment modifies teaching 2) Assessment enhances collaboration, and 3) Assessment identifies the distinction between content versus skill. Most faculty were interested in using the data to improve their teaching; sometimes the assessment process just reaffirmed their “sense” of class performance and then they use it as a motivation to change. For sub-theme one, faculty discussed how the assessment movement caused them to reevaluate their classrooms and reconsider everything from goals and curriculum, to assignments and practices. This was a great use of assessments. The second sub-theme acknowledged how assessment improved
collaboration between instructors. While there were varying degrees of collaboration, most participants recognized the innate benefit of the collaboration itself. More specifically, participants liked the fact that the focus was on improving teaching and the learning environment.

The final sub-theme was the distinction participants recognized between content and skill. The content versus skill dichotomy is significant as it highlights the tension between disciplinary focus versus student skill-building. Currently faculty may not feel the same type of ownership of general education courses as they do with their majors because they may have to focus on the process of teaching students how to be students versus disciplinary thinking and content. For example, students needed to learn how to read and extract important material from a text before they could fully explore the significance of the reading. Faculty discussed refining general education coursework so students could practice skill development.

**Assessment as Learning**

Within this category, students are held accountable for their own learning. There are three major themes within this category: 1) Students taking summative assessments 2) Students meeting specific requirements or criteria, typically through some high-stakes performance and 3) Students’ metacognition and reflections on their learning.

**Students Taking Summative Assessments**

Six out of seven participants touched on the use of grades as one assessment method that attempted to hold students accountable for their own learning, though none stopped there. For example, Scott said, “Assessment means tracking the progress of the student over the course of the semester, but not just tracking the student’s performance on the basis of a grade but trying to track a student’s intellectual development.” To him, grades are significant, but also just a marker
for intellectual growth over the semester. Dale used Clickers, a student response system device, to make students accountable for their learning. “For the purposes I just described, that is, doing a content-oriented quiz, it is very useful.” Here, the content quiz made students accountable for their own learning by ensuring they read the material.

Quite a few participants described the assessment as learning through specific assignments and how they used that grade for multiple purposes. Brian described his use of exam grades thusly:

> After the first couple exams, students are really shocked and surprised at how poorly they’re doing. I’ve certainly told them all the things that they need to do, but for whatever reason, they don’t do it until I make them and I show how it helps them.

Here the grade holds the students accountable and gets them to be more responsive to new strategies or ways of exploring future content.

**Students Meeting Qualifications**

As an important outlier, Brian identified the only high-stakes test out of all of the participants. It was not related to the general education program specifically, though it affected how he perceived general education curriculum and skills. Brian discussed how many classes lead to the Law School Admission Test, which “…is like their ultimate summative assessment.”

**Student Metacognition & Peer Assessment**

A couple of the participants were aware of the self and peer-evaluation that affected student learning. For example, John described the rubric he uses for essay assignments and how students are encouraged to review expectations before starting the assignment. “We actually print the rubric out for students to look at. As they are taking the test, they can see how they are
going to be graded.” Essentially the rubric included the range of points available and the description of what skills or knowledge was required to achieve those points. Rubrics were pretty standard practice with all the participants.

Rob described a process where students had to hold each other accountable as well. When students were charged with evaluating peers, Rob said students were typically not afraid to evaluate honestly. He also noted another type of student learning taking place. He said:

. . . I would have students who would come after class and say, “I’m on this team but clearly these members of my team aren’t pulling their weight.” And I’d say, “Well, that’s going to be exactly what you face when you go out into the real world and you have to have a job. Your first job is to figure out which of your coworkers are the ones you can rely upon, and which ones are kind of dead weight. There’s no more higher-order thinking skill than that, my friend.”

Again, this highlights the complexity of assessment. If a student did indeed learn this lesson, how would Rob quantify that for the general education committee or outside body?

Summary

All participants used summative grades as the main means to hold students accountable for their own learning. Conversely, very few mentioned using assessments as a high-stakes way to measure students’ ability to meet requirements. Briefly touched on by the participants was metacognition and peer assessment. The difficulty with that process is documenting learning, specifically how professors quantify student self-reflection of their own learning in a timely and succinct manner. While almost all of the participants discussed using summative grades, this was the least explored aspect of assessment; faculty recognized the utility of the process, but it did not excite or interest them as much as other areas.
Assessment of Learning

Within this category, all participants acknowledged the use of assessment as a tool for school accountability. The themes included 1) Assessment as university accountability via re-accreditation or other body outside of the university 2) Assessment to hold the department accountable 3) Assessment used for teacher accountability.

University Accountability

Every participant conceded the significance of outside pressure on the universities general education programs resulting in significant reform and assessment efforts. Jen, Dale, and Ken considered the significance of Southern Association of Colleges and Schools (SACS) on their general education program. Jen noted two decades of recommendations from SACS that claimed, “You all are not doing what you need to be doing with Gen Ed.” Ken was more diplomatic in his understanding of the history, but agreed with Jen’s assessment.

I guess deficiencies that State University Red had was that it didn’t have a very coherent Gen Ed program, and it wasn’t doing very much to assess that program. I think that became a particular focus of people in the administration, of trying to make sure that they didn’t get dinged for that as we did this most recent re-accreditation.

Ken was clear that the general education program was in dire straights due to previous SACS reports, and that this instigated changes in the system. Dale noted the recent progress concerning general education and assessment. “We were re-accredited this year with SACS but the previous time that SACS came through, its only major ding on State University Red was the lack of assessment of the Gen Ed program.” All three participants noted the progress made by State University Red in this area, reporting that the general education program did not receive any recommendations from SACS during the last re-accreditation. Ken sums up the overall process
by stating, “…we’re doing a lot more assessment, and I know that part of that is for external reasons, because there are people that want us to answer to them, that we’re actually accomplishing something.” Ken believes that regardless of the cause of the change (SACS), positive results were being realized.

The participants at State University Blue went through two such reforms of general education in the last decade. The first, in 2006, was in large part due to outside reviewers’ concerns about assessment, which resulted in significant reform for the general education program. The general education approach moved from the distribution model to an interdisciplinary core. Rob captured the rationale for that reform, highlighting that it focused on assessment because, “the institution’s accreditation was at stake”; therefore, “we had assessments up front. Assessment was always a priority.” John said that focus on assessment resulted in positive results. “They (SACS) gave us high marks. In fact, the highest in the university, so we must have been doing something right,” he said. The other participants from State University Blue also suggested the initial general education reform was positive as it related to general education and assessment.

The 2012 reform to general education was due to internal pressures and it returned to a standard distribution model. Brian commented that the current collection and use of general education assessment data is due to the SACS requirements, stating, “The data is currently being collected and recorded so that it can be given to SACS, not so it can change the classroom.” Beyond curricular changes, this attitude toward assessment seems to be the most significant difference between the two reforms.
It was clear that a larger, more complex system was at play in the assessment of general education. Jen was very aware of the complexities of general education assessment and the multiple stakeholders involved in general education assessment. She said:

The system wide stuff is being driven very much by the Board of Governors and what they want. Faculty would not have volunteered to do that without that external pressure. Our Gen Ed assessment is being driven at least as much by SACS and the external pressures of that kind of review and, removed from that, by the Federal Government because they are breathing down SACS’s neck.

Jen pointed out two things here. First, faculty probably would not have “bought in” to general education and assessment changes without the prodding of external forces. Second, she highlights that SACS is only one major stakeholder concerning assessment and that universities must be aware of the external pressures those stakeholders also face.

Rob concurred with that sentiment, sharing, “[Assessment] become more of an administrative tool. . . . They want to be able to have reports that they can export and then take to the Board of Trustees and the accreditation agencies and I suppose, in a handful of conditions, alumni or state legislatures for public schools and say, we are doing a good job. . . .” Rob acknowledged that the numbers generated from these courses are being used for many purposes well beyond the classroom.

Each institution has a university-wide body that creates policy, reviews courses, and compiles and evaluates the data. This is the General Education Committee at State University Blue and the General Education Council at State University Red. “I think that that’s why we have a general education council is for some kind of accountability outside the department. I think that in time that will improve,” Brian said. A few participants thought having these
committees was a positive move, because otherwise “you’d have colleagues reviewing colleagues” as Scott said, and that would cause all sorts of problems within the department. Some participants expressed an interest in allowing the departments more responsibility in the process. For example, Dale said, “I think it would work better if Gen Ed assessment was also handled by the departments. I know that the University has a stake in overseeing this process, but I think it might work better if departments were conducting the assessments (although I know that assumes all departments will take this responsibility seriously).” Dale acknowledges the inherent difficulties of having departments involved, but thinks they should have more control over the general education process.

**Department Accountability**

A second theme to emerge from the interviews was how the general education assessment data was used in relation to holding the department accountable. When asked to define assessment, Dale referenced the departmental use of assessment, stating:

> It can be assessing the way that a particular class achieves what its stated goals are. It can also be a department level where the department has goals for its majors. It’s a process therefore deciding whether the department’s mission is being met by constant process evaluation.

Essentially, Dale sees the process of assessment assisting departmental level goals, specifically as they relate to its majors. Ken concurred. “I see it as we’re setting goals about the things that we would like our students to learn, and, hopefully, we’re designing ways to test that and figure out if in fact we were doing what we hope that we’re doing, and if not, figure out ways that we can do it better.” Both Dale and Ken discussed using the data to effect change within the department to assist with student learning.
Other participants also recognized the possible utility of collecting and analyzing the general education data. John commented:

I guess it’s not just me individually, but this is helpful for us as a department and as an institution to get those aggregate figures to figure out where our students, what are they’re responding to the most and whose classes are they responding to better than others. . . .

In this situation, he is not sure if the data is actually used to enact change. Later in the interview, he again notes the discrepancy between what could be useful versus his knowledge of what actually takes place, “Again, in our area, that’s I think the biggest downside to the whole process is it’s not actually utilized at the departmental level, really, besides being collected and then individual instructors might use it maybe. Then who knows how to or is it that they use it.” This highlights the lack of communication and transparency in the process and reinforces the instructor conception that the data is not necessarily used in any meaningful way beyond reporting. It also highlights a lack of significance placed on the data by the department, as individuals may or may not use it to affect classroom teaching (it is not discussed at the departmental level).

Discussing his experience with the collection and use of general education data in the first major reform, Rob explained:

We actually had a unique situation where we knew the assessments we were creating were going to reflect our work and also be appreciated, certainly by the first level of management that hit it, and it certainly was used by the next level too because the institution was able to get re-accreditation and solve a lot of its problems.

During the 2006 reform with assessment at the forefront, the faculty knew that the work with
general education would be appreciated at all the levels of the institution. Scott was less sanguine about the current process, describing it this way:

I mean, if you’re going to assess education, it has to be done by a body that stands outside of all of the departments and can apply objective criteria for analysis and measurement and recommendations if they are in order. We don’t have anything like that right now.

What we have is a shell game. That’s basically what it is. It’s a shell game.

Scott recognizes the 2012 reform for a lack of accountability for the data, where data is collected but without recognition, support, focus, or improvement.

**Teacher Accountability**

A third theme that emerged was using assessment as a tool to hold the teacher accountable. When discussing the 2006 reform at State University Blue, John said:

We all developed more autonomy as over time to teach our classes, but we still were accountable. That process, I think it did make for a lot of extra work. It quantified our efforts in ways that I think were more detailed and more comprehensive.

John’s explanation was that there was a great deal of work involved and that he felt accountable for that work, but he did not expand on any specific consequences for not doing what was required.

Scott shared one incident when he could not complete the pre-/post-test that would be used for general education assessment. He completed a different assessment using a separate instrument and reported the findings, which were positive. He recollected, “I could demonstrate continuous improvement over the course of the semester in an aggregate way. In other words, the class in general improved considerably, as did individual students whose progress was quite notable, but that was just something that came out of the circumstances.” However, he continued,
“I got slammed because I didn’t do what I was supposed to do.” It appeared that Scott felt attacked for not doing what was required for the general education assessment. While he did some assessment, it was different from the required piece, thus the criticism leveled at him.

While Scott identified a verbal reprimand, many participants highlighted a lack of any true consequence or accountability when it came to general education and assessment. Brian, Dale, Jen, and Ken all mentioned the teacher accountability and the difficulty to hold professors accountable as it related to general education assessment. Brian noted the lack of follow-up by administration, saying:

Because, as of now, there’s nobody coming back and saying, “Hey, I notice that you guys were missing your numbers here. You want to have a workshop on this or that or anything like that to show that we are actually working on trying to develop these things?” You can miss your numbers every single time, every single time, and just say, “Well, that means we’re going to try this new thing.” That’s all they [administrators] want to hear.

It is noteworthy that Brian highlights accountability not being a negative experience, but one where faculty support is offered. For example, offering workshops to improve upon deficiencies. Dale agrees with Brian that there is little accountability, “Now, the reality is that there is no police anywhere standing over you saying, ‘Prove to me that these assignments have met those goals.’ ” Jen succinctly framed the issue saying,

It is hard for me to speak to the faculty because it is all ultimately up to individuals to do what they are going to do in their classrooms. There is no real enforcement mechanism for any of this. There is no real kind of (enforcement) other than just faculty taking responsibility to be accountable to what we decided as a team we needed to do.
The bottom line is there is no real mechanism in place and faculty must self-regulate. Speaking of self regulating, John mentioned that assessments were typically completed by faculty and most would do what they were asked just to “get it done” and not “make waves” or because they would not want to seem “uncollegial.”

Ken was not the only person to venture a rationale for this attitude, but he was the most concise, saying that “at the most basic level, faculty consider themselves to be the experts in their fields, and they are resistant to folks looking over their shoulders. They have not accepted the fact that people outside the academy are demanding accountability, and in ways that can be clearly demonstrated.” Essentially, faculty are experts with a great deal of education and experience who don’t want to be dictated to by stakeholders who may or may not have any knowledge of their discipline or even education for that matter. Ken also noted the sheer complexity of the process and took it further, stating that with so many people involved in general education, it would be very difficult to organize any kind of constructive feedback loop that would get back to the individual instructors.

Summary

Assessment of learning is essentially about accountability. Three distinct themes emerged. 1) Holding the school accountable 2) Holding the department accountable 3) Holding the teacher accountable. Typically when the participants discussed holding the school accountable they discussed outside forces, specifically SACS. Using assessment as an accountability tool drove general education reform at both institutions and forced faculty to engage general education in ways that they may not have done otherwise. Participants at both institutions also found that their efforts in assessment were positively reviewed by SACS. While getting faculty to think about, discuss, and change general education and assessment was viewed
positively, the problem was that the generated data was collected only for the reports. All the participants claimed that current general education assessment efforts were primarily for SACS and not for internal use or classroom improvement. A few participants also recognized other outside stakeholders involved in the process including the federal and state governments, the Board of Trustees, and possibly alumni or the public. All the participants also highlighted the internal committee that was in charge of collecting, analyzing, and disseminating information on the general education assessment program.

The second theme that emerged dealt with how general education assessment was used to hold the department accountable. While some saw the assessment process assisting their majors, others tried to see how the data might help students learn in the classroom. The major problem at this stage was that no one was quite certain how that data was being used or to what affect.

The last theme was that of holding the teacher accountable. There is more of a fear of using assessment data as an accountability tool for faculty than any actual instances of professional negative impact; of course, the participants were largely tenured professors in their departments. Instead, there was a general sense of responsibility as professionals and a willingness to do what was needed to be done concerning assessment. Currently, all of the participants recognized the lack of accountability or even a mechanism to address possible weaknesses or shortcomings. Participants expressed some fear about what accountability might look like or what it would entail if a system were put in place.

**Assessment as Irrelevant to Learning**

The final category looked at assessment as inconsequential or even antithetical to learning. It had six themes 1) Assessment is an imprecise process 2) Assessment interferes with
teaching 3) Assessment results are not used 4) Assessment affects accountability 5) Assessment is poorly supported, and 6) Assessment results in deprofessionalization.

**Assessment is an Imprecise Process**

Participants discussed many issues with the general education assessment process. Many participants expressed concerns about the vague language often associated with assessment and how that negatively affected the process as a whole. Dale and Brian both mentioned the broad, ambiguous nature of the general education learning outcomes. “But again, my complaint, if there is one, is the Gen Ed level student outcomes are pretty broad and that allows a pretty wide range of what happens in a . . . class to be at least plausible,” Dale said. Brian concurred, saying, “You could probably attribute it to these broad, general education outcomes that we’re trying to meet. . . . It’s hard to miss a broad target.” The problem is that the goal loses its meaningfulness when it is so broad or vague; if anything can “count,” or works, then the question becomes: How valid or reliable is the measure itself?

Dale took this idea further, identifying an assignment, which was subjective by nature, and how the assessment process does not somehow quantify it. For example:

What are the consequences of the [a discipline specific question]. I’m sure there’s many. Which is the right one? How would you assess that? Are they all right? . . . it was an extremely, time-intensive effort with dubious outcome, ultimately. That we’re ultimately going to be just as subjective as if you had left that to me in the classroom to decide whether I thought my students were meeting Gen Ed goals.

Brian was concerned that the outcome was so broad that the instructor would not know quite how to attempt to assess it. For example, “For the Humanities marker in General Education, to have that marker I think you need to have students learning about some very general ideas about
being creative, being thoughtful, critical thinkers.” Both participants touch on the vagueness of the language and then on the subjectivity involved in trying to quantify the complex intellectual development or skill being examined.

Scott wondered how one could quantify some of these learning outcomes, indeed if one would even want to, by saying, “I don’t think that there’s very deep thinking, and the reason why I think it’s not deep is because I think that there is this need, again, to quantify everything, because when you quantify, you’re providing data. It’s all data-driven, and I just don’t think it’s very constructive.” What do these numbers really mean and how can they truly represent something like “critical thinking”?

Dale commented on the same idea:

Now, we’re back to something, as a . . . humanist, that a lot of this assessment language and some of my resistance to it, on an institutional way, has to do with the fact that most of what we do in . . . classes, is extremely subjective. It’s really difficult to measure whether a piece of writing is an example of critical thinking or a example of excellence in temporal thought or whatever it is. Those things are not going to be the same to me as they might be to you. I mean, yes, so maybe on a basic level, but it’s very different from measuring math diversity or computer literacy.

Thus, there are some aspects of education that can be quantified, that can be boiled down to numbers; however, there are other areas that resist quantification and maybe should not be quantified at all. Dale also shared his experience on the university general education committee highlighting the difficulty in quantifying these subjective learning outcomes, “Then their implementation, particularly in the areas that I knew, which is humanities, Philosophy, things like that. I just didn’t see how it would be accurate or that the evaluation of the measurement
data that we assembled would be anything other than subjective.” Brian noted the difficulty of trying to quantify the learning outcome like ethics at his institution. “That’s not the kind of assessment that tells you they have satisfactorily achieved an ethics component. It’s just this class had an impact and changed them in some way. Maybe for the worse.” The point Brian is making is an important one. Can the university claim the student has been ethically trained? What would it even mean to be ethically trained?

Ken discussed the goals from a different perspective, suggesting the language used to discuss assessment, the goals themselves, and the desired actionable items where faculty were concerned, was off-putting and possibly harmful.

I guess the thing that, really, I don’t like about assessment is this notion that there’s this continuous improvement, there’s always this . . . I guess I’m not totally a believer in this progress that’s constantly upward… the goal of continual improvement, but sometimes I think people see that as that they worry that you have to show that you’re continuously improving, or there’s going to be some kind of penalty for you along the way. I think if people maybe didn’t have that understanding, they might be a little bit more comfortable with it.

The faculty should be attempting to improve their teaching because that’s what professionals do. However, “chasing rainbows,” this type of elusive, never ending cycle of assessment and possible improvement may actually turn people away from the idea of assessment. People may end up particularly resistant if they are not proving to someone that they are in fact constantly improving.

Another aspect of this imprecision deals with the various levels of learning outcomes and how they work together to make a whole. Consider Brian’s comments:
We have program-level outcomes and student-learning outcomes for disciplines and programs; then we have them for our classes. It’s always a little unclear how these things are all working together. There doesn’t seem to be anyone coordinating that, but there is this notion that these things should go together.

There is a lack of clarity and consistency over multiple areas and at the various levels of the institution. In his interview, Dale offered an example of how this was actualized in the department:

We all joke about this, or some of us joke about this, the fact that you could have, I think we have three Gen Ed learning outcomes. We have three or four department level student learning outcomes. We also have, for responsible faculty members, anywhere from three to seven student learning outcomes. If you add those all up they all have to go in the syllabus, you’ve got fifteen student learning outcomes, and it’s very difficult. Now, there is some overlap but it’s very difficult for you, as an instructor, to make sure that you’re meeting all those outcomes.

Attempting to measure the various outcomes is unrealistic. It creates a situation where they lose any weight or meaning for the instructor (and possibly the student).

Another comment Dale made indicated the difficulties of the theoretical ideas of the general education plan being realized.

Again, I admire that as a logical structure. The problem is that it was A) extremely contentious among the different departments to decide what those Gen Ed level learning outcomes were. Then, B) it was never quite clear how that interfaced with the actual teaching. In other words, how did, once you had these Gen Ed level student learning outcomes, how could you test them or measure them? Measurements, the right word.
Coming up with outcomes that everyone from all the different areas could agree upon was difficult; furthermore, there was no step between the development of those outcomes and the implementation of specific strategies in the classrooms.

The reporting mechanism at the end of the term can also be problematic. Between our initial interview and follow-up, John completed his general education assessment report for the previous semester’s general education courses and described it this way: “Last time I talked to you about what seemed like a good idea has actually turned out to be another administrative burden.” He went on to describe the instrument in this way: “It’s 29 pages of drop down menu items where you assemble the data and click on how many students performed. . . . The rest of it just seems overkill and tedious and not very well explained.” At his institution, Ken also highlighted the confusion the reporting results might cause. A colleague filled out the department forms and didn’t realize there were separate forms for the university level.

Many participants also noted the lack of substance in the current reports. Brian suggested, “Everybody’s looking for a shortcut to figure out whether or not that’s happening and that’s what we’re recording, that shortcut which glosses over everything that’s important.” John described the process as collecting a “snapshot” of a student’s learning and Dale called it the “tip of the iceberg” of the actual learning that took place. To truly get a sense of a student’s intellectual growth, one would need to review that student’s body of work, they said. Again, the quantification of complex intellectual development or skills seems to be too simplistic and too reductionist for the faculty.

**Imprecise instruments & measures.** One sub-theme that emerged from the participants’ interviews was the degree of effectiveness of assessments at the department level. One faculty member at State University Red said a problem with the current system was the lack of
uniformity with the instruments at different levels of courses. He continued, “In other words, my assessment instruments are going to be developed by me, Dale, as opposed to imposed by the department or even discussed in the department.” This was problematic because it would add a layer of complexity to collection and analysis, especially across multiple sections of the same course. Conversely, at State University Blue, Brian found the uniformity and control at the department level disconcerting: “. . . I do not like the strictness with which some departments, mine in particular, think that you have to teach all of these concepts, you have to do it in all of these ways.” The lack of autonomy in the classroom when it came to assessment was problematic.

The degree or effectiveness at the institutional level was also discussed. Brian noted that the results shared with the institution’s general education committee may be lacking enough context to make sense out of the numbers. For example, he imagines himself reading the report and having this response: “I have no idea whether or not the assessment they’re using . . . their assessment tool . . . is in fact assessing this General Education goal or outcome. Zero idea whatsoever,” because the tool is not included with the report, just the numbers and the analysis. He concludes, “The person that’s reporting on the success or failure of the assessment there, they don’t ever have to show their tool; they don’t ever have to modify their tool. All they have to report is did they make their numbers or not.” When Dale was a member of the general education committee on his campus, he had a similar type of difficulty concerning the instruments’ reliability saying, “There was no hard and fast way of evaluating the success of the measurement instruments which is what . . . was the whole purpose of this enterprise.”
Rob and Brian both discussed the instruments’ impact on the classroom. Rob said assessment should be the concern of the faculty and institution first, and outside groups or interests second. He continued, saying:

I don’t think Higher Education looks at it that way, and I think it’s even more true when you look at public schools, too; and that’s why a lot of teachers who are really good teachers based on really dull instruments of assessment can look bad according to the criteria of the institution, because it’s looking at these things wholesale; but, where things are really happening are in particular classrooms with particular students and by putting the mechanism of assessment in the hands of the instructors to the greatest degree possible, you’re going to capture some of that to work and you’re going to give them real changes. They’re real instruments and criteria by which they can gauge their own successes and failures and move forward at that level. That’s what I think needs to happen and that’s not just in Higher Ed.

The lack of control given to teachers is problematic. Just as Jen noted, the farther away the results were from the classroom, the less meaningful they were; likewise, Rob is saying the farther away from the classroom that instrument was created, the less meaning it has.

Brian verbalized another concern with an instrument and its measurement, “We do have the pre-test, so we could look and say, ‘They didn’t know some of these things, and now they kind of do.’ Again, change is not what is assessed for General Education. It is categorical. Did they meet the mark? Yes or no. Students that don’t necessarily meet the mark still pass and move on. . . .” The lack of accountability on the outcomes may affect the motivation for students to do well on the actual assessments themselves.

Lastly, Rob highlighted the use of results without context:
It (assessment) does nobody any good to do an average of a class in terms of evaluation. You have to have a T score where you can look at, ‘Okay, what is the individual level of improvement,’ right? Each student needs to be assessed as a student, so that somebody who moves from, let’s say, the 60th percentile up to the 85th percentile should be worth more than somebody moved from the 85th percentile to the 93rd percentile, right? Because that involves a lot more progress. That all gets lost in a class where you’re just homogenizing the whole and taking one temperature, class average performance on a pre/post instrument.

The value added, the movement seen in an individual student, needs to be considered also, but it is often overlooked. This devalues the progress made by the student and the instructor.

**Assessment Interferes with Teaching**

In the interviews, multiple participants touched on the idea that aspects of the assessment program may actually interfere with their teaching. Scott overviewed the assessment movement’s impact in this way:

It’s very, very difficult to quantify that [intellectual development], but we seem to be bent toward quantifying everything, which makes General Education, for somebody like me, who’s a product of an earlier time, when education was much more comprehensive, much more thoughtful in terms of trying to get at a student’s core understanding by getting at the margins of a student’s performance and then working inward toward the core. That doesn’t happen when you have to quantify everything.

To a certain degree, assessment keeps things at a surface level. He goes on to say that his classroom is affected because before the assessment movement he was able to “respond to their [students’] needs much more directly because that was the nature of what we do, we’re teachers,
as opposed to being ordered to assess their performance in some kind of arbitrary and, in many ways, empty ways.” If a student shows gaps in his or her skill levels, the teacher does not have the flexibility to truly assist the student because the instructor is worried about getting the class ready for a particular assessment. This type of attitude goes against what he sees as his mission of teaching.

Brian described how the department attempts to gain validity in its instruments through prescribing content, saying:

Too many programs, I think, are making faculty dislike this kind of assessment and this kind of process because they’re just making the classes . . . they’re erring on the side of being all exactly the same rather than erring on the side of, if we have a good assessment instrument, we wouldn’t necessarily all have to be the same.

Obviously he is against such an angle. Dale had a similar inclination. “It might be theoretically possible to imagine a . . . class where you set up a series of hard and fast facts at the beginning and then tested students about whether they did that. That sounds like high school to me and that’s not what we do [in this discipline] in the university. That’s not what [this discipline] is.” Agreeing with that sentiment, Scott described less flexibility due to the idea of teaching to the test. “There are times we used to describe as ‘teaching moments’ when everything would stop for some reason or some exogenous force would occur or some revelation from within that would just cause me, as the instructor, to basically just stop and reevaluate right there and then move on. You can’t do that when you can’t think about it. If you give a pre-test and a post-test and then something happens in the course of the semester and you don’t cover the material that you said you were going to cover in the syllabus, then what does the post-test look like?” Ultimately, the participants said that assessments should not be used to measure the student’s content knowledge
or to dictate the curriculum.

**Assessment Results Unused**

Participants shared their perception of how the general education assessment results were used. Many participants found the results had very little impact on anything of significance or worth to the faculty member or the faculty member’s classroom. “Our strategies for assessing are not always very useful. There are ways of checking off, ‘Oh, we were able to do that,’ and showing SACS a document, so it is just meaningless paperwork, and it’s expensive,” said Jen. The process was ultimately a matter of fulfilling administrative issues. Scott agreed, saying, “. . . most of us are really not assessing academic development or intellectual growth among our students. We’re just coming up with the means by which to satisfy the fact that we’re ordered to do this.” Many participants found the assessment process just something to be done and filed away. Scott captured the sentiment in this way, “People are just sick of having to jump through hoops just so we can produce the numbers that will satisfy accreditation or whatever these numbers are supposed to demonstrate.” It is an administrative burden, again, with no meaning or weight for the classroom. Brian said, “At the end of the day, what motivated this campus to do that was SACS requirement. The data is currently being collected and recorded so that it can be given to SACS, not so it can change the classroom.” Dale gives some insight into that claim, highlighting that “[The data is] so removed from the individual student and faculty member that it’s beyond the department, it’s in this star chamber of Gen Ed people that I just don’t see the value for that.” The data was not anchored to any practice and it is unclear how the university was using it.

**Lack of a feedback loop.** A sub-theme that emerged for a majority of participants was the lack of a feedback loop, an opportunity to analyze and or discuss the results either at the
classroom or department levels. Brian discussed the disconnect between how assessment “should” be used and how it “actually” was used. “Assessment feels like a way to make sure we’re doing our job and to show the administration that students are learning and accreditation that students are learning. They don’t recognize assessment as a way to actually improve student learning.” Describing the reporting mechanism, Brian highlights the lack of actual impact of the data, stating:

If you find that the percentage of students that were below the achievement line that you set for yourself, you’re supposed to say, ‘This is what we plan to do to fix this,’ which I think is a very important part, actually.

The problem is not the requirement itself, but the fact that the data moved up the university chain, and stopped. It was not used to affect the teaching or learning at the classroom level. The feedback loop was nonexistent.

It is important to note that at State University Red, there was some type of process for feedback from the general education council to the faculty, but it did not seem effective in its purpose. Ken described what State University Red does with the data in this way:

What they would do is at the beginning of the next semester, they would try to get together a focus group with some faculty, and they had some incentives to try to get . . . to talk about what the results meant, and then try to get that out to the faculty. I went to one of those and it was fairly helpful, but I don’t know that the results of that get filtered down to all the faculty in the same way that, say, our department level assessments get filtered down to the faculty.

He goes on to speculate as to why the data does not reach the faculty in general and says, “I think partly, maybe that’s just the way it is, because it is so many people involved, and how you would
organize that would be very difficult.” Thus, the process seems to be there for people who might be interested in general education assessment, but it does not seem to be done in a methodical way that would benefit the faculty who teach the general education courses.

Several participants discussed the process not being beneficial at the department level either. When asked if the results of the general education assessment ever came back to the department Jen said, “I do not remember a single departmental discussion of Gen Ed assessment results.” Scott had a similar experience, “There is no loop. It’s going into space. There is no return. It just goes into space. The faculty knows that it goes into space, so that’s why the faculty are central and just do whatever they have to do to get it done.” John mentioned, “Again in our area that’s I think the biggest downside to the whole process is it’s not actually utilized at the departmental level.” Brian was the most direct with his answer, “No.” He later added that, “To my knowledge, no one has sat down with the data at the end of the semester with colleagues who are teaching the class and discussed the results, and what we can do with those results, and how they would impact the program.” Ultimately, what the participants found was that the data was never discussed at the departmental level to refine the general education courses or think about the curriculum in any substantive way.

**Assessment Affects Accountability**

Another theme that emerged was the issues surrounding the concept of accountability in the process. Jen shared her insight into the process in this way, “Gen Ed is owned by everybody and hence, nobody knows what’s happening with it.” With so many departments that have so many competing interests combined with the sheer number of people working on general education, it is difficult to get everyone on the same page and moving in the same direction.
Jen went on to discuss the faculty specifically, saying, “There is no real enforcement mechanism for any of this….other than just faculty taking responsibility to be accountable to what we decided as a team we needed to do.” Dale concurred, “We don’t actually have very systematic ways of evaluating that and I know my department head has told me in the past that if the faculty members don’t do something, he doesn’t know what he can do to force them to do that.” Scott also highlighted the issue of accountability “We have no body, we have no group of people who are primarily responsible for taking these assessments and measuring them and then using them in some kind of constructive way to impact the future of Gen Ed at State University Blue.” The participants recognized the lack of accountability in the process, but it is also important to recognize that those adjuncts or lecturers whose jobs are renewed each year may “feel” accountable in ways not expressed here.

The participants did express a concern about how the administration may use the data. Dale touched on the faculty fears that might drive some of the resistance to assessment. “There is concern among faculty that somehow this would be used to assess, not student performance, but faculty performance.” Later he added to that idea stating:

. . . if you could imagine the howls from faculty members if the Gen Ed council said, ‘Hey Dale, your class is not achieving its Gen Ed goals. You need to change this, and that, and the other thing.’ Oh, my God, there would be an explosion of nuclear force all over the place. . . . Of course, no one to my knowledge has done that. Then that means what is Gen Ed level assessment good for? They’re collecting it to satisfy, in my view, accreditation. . . .

As with any assessment data that is collected, the conflict between faculty autonomy and administrative oversight is not far behind.
Brian was the only faculty member who mentioned the lack of accountability when it came to student performance and general education learning outcomes. If a student does not meet the required goal, what happens to students? He responded, “Mine still pass the class and move on. Yes, we’re going to reexamine the class and do some different things and retool our assessments to make sure that they’re working right, but that student X, Y, and Z, they just made it through the class. They haven’t met that requirement, but . . . it doesn’t matter. They just move on.” If a student did not improve in a learning outcome, there was no real, definable consequence.

**Assessment is Poorly Supported**

Through the interviews, the lack of administrative support emerged as a theme. Many participants interpreted the lack of administrative support to mean that general education assessment must not be something to take seriously. There were two areas where support was visible. First was in training, workshops, and amount of time given to focus on assessment or assessment related activities. Some participants said there were plenty of opportunities for training and workshops. The problem many participants found concerning assessment was the time involvement. About half of the participants commented on the sheer amount of time it takes to complete the general education assessment. Ken shared his experience as chair stating, “You should know, there is a lot of push back a lot of times to assessment because it means extra work for faculty.” Dale also touched on the time involved. “I found this very cumbersome and adding a layer of a difficult, and onerous, and very time-consuming layer of bureaucracy to what you’re already doing.” For many participants, the time investment was a significant feature of a faculty member’s willingness to support general education assessment. More so, it affected how much more time they would want to invest by volunteering to attend training or workshops on
assessment. Scott pointed out there was “plenty of training” and John recalled a mandatory workshop. Ken said at State University Red, there were a lot of opportunities, but it was up to individual faculty to take advantage of them. Dale agreed.

Rob described a very different experience concerning general education assessment and support. During his experience with the 2006 reform, the support he received from his Dean concerning assessment, and the impact that had on him and his view of the assessment process as a whole was significant.

Mac Smith (pseudonym), who was hired to implement this new core curriculum called University Studies had people come in to talk about assessment. As a matter of fact, the university before he came in had people coming to talk seriously about assessment and the head of the Institute for Teaching and Learning . . . and a few core faculty who were bought into this new curriculum were thinking about assessment as a critical feature to make sure that this new curriculum would succeed so that we were schooled in doing pre-assessment or formative assessment baseline and then doing post-assessments with very specific measures and metrics for each class, and the whole curriculum was geared towards getting specific skill sets developed into students.

The faculty were given time and support to become proficient in general education assessment. Brian also spoke highly of that 2006 reform movement and its focus on assessment.

I would say that by far and away working with [that program] taught me so much. It was some of the best times… for what’s right about assessment to improve the classroom.

JUST I’VE never seen anything like it again. . . . As a faculty member and a department, we were given time in (that program) to do that.

John and Scott also spoke highly of the program and its supportive assessment features.
Ken brought up the second aspect of administrative support for assessment through recognition in this way: “I don’t think faculty get much recognition for work in this area, at any level, unless a faculty member plays a major University role, such as being a member of the General Education Council.” Dale agreed, saying these kinds of things just don’t mean a lot in terms of post tenure review. Brian suggested it might count, but he was unsure as to the real weight or value that would be given.

Yeah, it would either fall into service or teaching. I’m sitting on a post-tenure review right now coordinating this. You have to have so many things in those packets, like there’s no way that’s really holding that much weight. There’s just no way. It’s more important that my classes get done and are well taught than the report that I write at the end. It’s far more important that the research gets done or I attend this meeting or that meeting than I have a meeting with faculty that are teaching this class, and we have a discussion and do that regularly. The percentage in the post-tenure review or going out for tenure would be so minuscule it’s just not worth it.

This lack of recognition may lead faculty to view general education assessment as “more busy work than it is meaningful,” as Dale suggested.

**Assessment Results in Deprofessionalization**

The last theme that emerged from the data was the feeling participants felt as if the assessment movement may be deprofessionalizing the faculty. One way this was achieved was through a top-down approach to assessment. Dale explained his resignation from the general education committee saying, “…the reason I resigned, was that I just felt like it was a top-down imposition of ideas that sounded good in the theoretical stage” but broke down in
implementation. Examining the 2012 reform at his university, Rob commented on the top-down approach and its specific impact on instructors.

[I]t just showed this attitude that, well, okay, we as administrators want this instrument performed and, boom, you do it, no questions asked, rather than approaching it as a shared problem and saying, ‘Okay, here’s the goal. Here are the outcomes we need to have. Here’s the criteria we need to focus in on in terms of getting data. What’s the best way to do it?’ It just kind of showed a lack of trust and respect in the people that the entire success of the institution relies on most heavily.

Essentially administrators are placing themselves in opposition to the faculty instead of in partnership with the faculty. Responding to another question, Rob came back to this line of thought, saying:

... you have administrators who are at best heavy handed in dealing with these assessments and I know the pressure’s on them and that they have goals to meet as well. That’s true but they need to stop looking at faculty as adversaries or cattle to be managed and see them as allies and they should be honored and privileged that they’re at the head of such a cohort.

The lack of shared governance when it comes to general education assessment is a significant concern, one that leads faculty to believe the one area they hold dear, the classroom, is being weakened under the guise of quantification.

This deprofessionalization is being realized specifically in the classrooms themselves according to some participants. Scott described it this way:

I think it’s dematerializing the discipline. I think that the more we drive toward this kind of assessment, the more content and analysis that we remove from our work because it
takes flexibility away from the instructor. . . . The freedom that I think is what always made higher education enjoyable to me as a student and now as faculty is lost.

Dale had a similar thought saying:

Ultimately, it reminds me a lot of the kinds of criticisms that we hear in the public schools in high school, that is, say, teaching to the test. . . . Which means, again, might be an important thing, but it would be hard to measure the students’ results there and secondly it would completely alter the way you taught.

Other participants had similar thoughts and the dangers of allowing higher education to follow in the footsteps of k-12 where assessment is concerned.

Jen summarized some of the faculty concerns in a larger context.

Part of faculty push-back, I share the concern that assessment is just going to be another one of these expensive, it is hugely expensive at that level. These companies that are creating these majorly expensive testing systems and that sort of thing. It is a resource drain. Instead of putting those resources in the classroom to improve student learning we are sending our students through yet another round of massive, this generation of students have been tested to within an inch of their lives and here we’re going to do it again at the university level.

Essentially, the students have been tested and tested and tested and the testing companies have made a lot in taxpayer funds by creating those tests. Jen sees the resources going to assessment instead of professional development and other things that will directly benefit student learning as immensely wasteful, especially at the higher education level. Scott said the whole process was actually pushing him “into retirement.” Scott’s comments serve as a cogent wrap up of this section. He said:
I think we’re headed in a direction where the substance of Higher Education is being liquidated and the paradox is that the liquidation appears as numbers. In other words, the liquidation process appears as data, but the data is what’s liquid. The data doesn’t represent anything solid. Anything that is solid is just melting into air. Some of us of an older (generation) look at this, and basically what we see here in Higher Education based on this kind of evaluation process is that much of Higher Education is being lost for most students at most universities that have to do this.

The concern is that students are losing what it means to be college educated through a series of tests that constantly try to quantify their intellectual and skill growth at every stage of their academic careers. The numbers themselves, as they move farther and farther away from the actual classroom, become less meaningful.

### Summary

Six themes and two sub-themes emerged in the assessment as irrelevant to learning category. Themes included: 1) General education assessment is an imprecise process; this includes a sub-theme of using imprecise instruments and measures 2) General education assessment interferes with teaching 3) General Education Assessment has little use, with a sub-theme that the data does not have any meaningful purpose due to a lack of a feedback loop. 4) General education assessment lacks accountability 5) General education assessment lacks support for faculty and 6) General education assessment is deprofessionalizing higher education faculty.

The first theme and sub-theme that emerged from the data was concerned with the imprecise process and measures or instruments of general education assessment. The first area of concern was the lack of precise language in the student learning outcomes and how that affected
the entire process. Assessment is a complex, multi-layered process that lacks a common language that is clear, succinct, and meaningful to all of the stakeholders involved. The vague student learning outcomes make the process difficult to complete so it becomes a game of “what do they (administrators/SACS) want” instead of how can we use this information to improve our classrooms.

The second aspect that was discussed was the inability or difficulty in quantifying vague, nebulous, or otherwise complex intellectual skills and abilities. Most participants discussed the difficulty in quantifying these complex human experiences. There is a feeling that while it may be a worthwhile goal to aspire to, the process is often impractical and even absurd. General education students are assessed to death, and the types of skills or competencies under review resist standardized tests of all sorts, consequently the assessments created challenge students in new ways that they are not familiar or comfortable with. However, this does not mean the academic process was not beneficial for student learning. For example, sometimes the challenge of assessment is to get students out of their comfort zones which means that they may not do well, but they might be learning something about themselves that is resistant to quantification. Consider a student who fails miserably on a project, but develops her problem solving skills. That learning will not be accounted for in the final report.

Another concept the participants explored was the lack of coherence between all the levels of general education assessment like the university, departmental, and classroom levels. Lastly, while the reporting mechanism may be stringent and thorough, the entire process felt reductionist and many participants said it did not fairly capture true learning. There was some concern as to the measurement itself and the tools used and the possible negative impact it could
have on the classroom. For example, more instructors may teach to the test, so to speak, which seemed like a k-12 experience, not higher education.

Many participants thought the assessment process interfered with teaching. For example, if identical content is dictated across multiple sections or if assessments are uniform, it affects faculty autonomy. Participants wanted to retain flexibility and stressed the organic process inherent to teaching where students and the content demanded the structure, not the other way around.

Little use of assessment results included the sub-theme of a lack of a feedback loop. Every participant addressed this in his/her commentary. Each thought that the data was primarily used administratively and not to assist teaching and learning. Part of the resistance to the general education assessment process was due to the lack of impact on the teaching itself. The participants seemed frustrated by the lack of communication and transparency. Faculty perceived the process as time consuming and administrative versus having any real impact on their courses, especially with the lack of any feedback on the submitted data.

Participants discussed accountability in the process and found lack of accountability at some levels disconcerting and at others welcome. For example, students are not responsible for achieving certain levels in the general education learning goals. On the other hand, faculty liked to be responsible for their classes.

The lack of administrative support was a major theme. Participants found that general education assessment, if done right, is a time-consuming process. However, there is no reason to invest the time or effort because it counts for so little in tenure and promotion. Participants also said there were plenty of opportunities to voluntarily complete professional development in
assessment; but, again, those opportunities took away from their work which would be recognized at end of year reviews.

The final theme was the participants’ concerns about assessment’s effects on the profession. One point participants discussed was the top-down approach administrators have taken when it comes to general education and assessment. This lack of shared governance and the negative effects that lack had on instruction and faculty buy-in for the assessment process as a whole was quite significant. There was very little in terms of grassroots work on assessment. Others saw the process as hurting the discipline and possibly “dumbing down” higher education and making it more like “high school.”

Chapter Summary

Most participants recognized general education as important, though many did not express the same level of ownership as they did for their major courses. While participants found general education assessment important to improving their own class and growth as an instructor, many found the assessment process an administrative burden. The purpose of general education highlighted a tension for participants, specifically how much of the course and assessment should focus on disciplinary content and thinking versus broader student skill development.

The next four categories specifically addressed how instructors conceived of general education assessment. They overwhelmingly recognized that assessments serve many purposes and hold multiple meanings for various stakeholders. The second category examined the participants’ thoughts concerning assessment for learning. Here participants shared how assessment was used to describe student abilities and competencies and attempted to quantify more complex skills and intellectual aptitudes. Participants also discussed their use of assessment to provide students feedback on their work and set future objectives so students could improve.
At the same time, many participants recognized the inherent use of assessment as a motivation tool so students would challenge themselves or learn new concepts or skills. Further, assessment was shown to improve teaching for all of the participants as they explicated ways they modified their teaching due to the assessments. Another positive aspect of the assessment movement was that it allowed opportunities for participants to collaborate on courses, curriculum development, teaching, and student success.

Assessment as learning emerged as the third category. This was essentially how assessments held students accountable for their own learning. Mostly the participants highlighted the use of summative assessments though meeting particular requirements, and peer evaluation and metacognition were also considered.

The fourth category examined how assessment was used to hold various stakeholders accountable, specifically the institution, the department, and the faculty. Participants noted two important points. First, they recognized the inherent importance of assessment in a classroom, but challenged the process, especially as it is directed from outside organizations like SACS and as it is implemented using a top-down approach instead of a grassroots one. Second, there was a paradox of wanting accountability in the general education assessment, yet dreading how it might be executed.

The final category considered all of the rationales that made assessment irrelevant to learning. Faculty discussed the imprecise nature of assessment language and the imprecise or faulty instruments currently employed. Many participants shared how the assessment process may actually interfere with teaching, specifically highlighting how by having the student learning outcome (SLO) and the assessments dictated to them, instructors may be forced to teach to the test and/or not have the ability to respond to student needs during the course. All of the
participants discussed the lack of a feedback loop and the little practical use made of the collected results. This not only fed into the sense of general education assessment as an administrative burden placed on faculty, but also reinforced negative concepts about the movement as a whole—the only reason it is in place is to provide numbers for administrative use.

Participants said the lack of support and recognition of assessment work is problematic. Assessment is an ongoing process that takes time, but general education assessment is not supported as an ongoing process, so it is treated as a burden. There is little administrative support, and little investment or potential for positive return. Faculty would be more willing to engage in it as a process if it were valued beyond accountability. While all participants recognized the need for assessment and willingness to participate in the process, the current structure and use of it generated a sense of detachment and deprofessionalization.
Chapter 5 Discussion and Implications for Practice

General education assessment is a complex and often contentious process involving multiple stakeholders. This study’s purpose was to better understand how one group of stakeholders, instructors, conceives of and uses assessment in those general education courses. Since general education courses command approximately 46 credit hours (Hart & Associates, 2009) or 30% of a student’s collegiate experience (Brint et al., 2009), it is important to validate student learning and competency in the general education goals. Currently institutions gather a great amount of data concerning general education, but that the data is normally not being used to assist in teaching and learning (Blaich & Wise, 2011); instead, the data is used as an accountability tool (Blaich & Wise 2011; Reeves 2006) or for institutional reporting (Blaich & Wise 2011; Kuh & Ickenberg, 2009). As demonstrated in the literature, data collection needs to be tied to reform benefitting student learning for it to be helpful (Banta, Jones, & Black 2009; Huba & Freed 2000). Faculty’s conceptions and uses of assessment may have significant effects on both student and institutional success. This leads to the research questions:

1. How do humanities professors typically conceive of general education assessment?
2. In what ways do they use the assessment data?
3. What is the rationale behind their use of assessment data?

In this descriptive case study, humanities faculty from one discipline at two state institutions had an opportunity to share their diverse perspectives on general education assessment through a series of interviews. Brown’s (2008) CoA Theory was used as a lens to better understand and consider those conceptions. The analysis of the findings were also informed by constructivism, which involves negotiating and interpreting one’s reality and being
influenced by one’s prior knowledge (Cobern, 1993). Assessment is not done in a vacuum and instructors do not think about assessment in isolation from their educational, classroom, or professional experiences.

As participants discussed assessment, four categories based on Browns (2008) CoA were identified. Multiple themes and subthemes emerged in those categories using the participant’s own words as labels. Ultimately, eight distinct summary findings, which can be seen in Table 2, were identified.

<table>
<thead>
<tr>
<th>Table 2. Findings.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>For learning (improvement)</strong></td>
</tr>
<tr>
<td><strong>As learning (student accountability)</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>Of learning (institution accountability)</strong></td>
</tr>
<tr>
<td><strong>Irrelevant to learning</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

These findings within those four categories will be examined in relation to the current literature, then practical implications will be considered. Afterwards, limitations of the study will be considered. Future research opportunities in general education assessment and humanities will
follow. The study will conclude with a look at an action plan.

**Category: Assessment for Learning**

**Assessment describes learning.** The participants were overwhelmingly in favor of using assessment for learning and had positive feelings towards the process of using data to encourage changes in teaching and curriculum with the goal of improving student learning. This is significant because as Willis (2007) posited, the way an instructor approaches assessment shapes the learner’s own beliefs about learning. Most scholarly work on assessment for learning has been specifically focused on formative assessments (Black & William, 1998; 2004; Broadfoot, 2008; Gipps & Stobart, 1997; Stiggins, 2002) to improve learning and the classroom.

Participants used formative assessment to help students and themselves improve, and summative assessment to judge and classify student learning (Bloom, 1969; Scriven, 1967). For example, John said, “You have to give work to students that is going to prepare them for handling the materials or find out where they are with the material that you’ve assigned, okay, [that’s] formative assessments. The summative assessments are the ones where [the students have] had time. You’ve had time to go over them and now you’re actually testing them and grading them.” Brian had a similar understanding of the terms, though he mentioned he was not “happy” with the distinction between formative and summative. Other participants discussed the generally recognized definition of the formative and summative assessment as well.

Part of Brian’s rationale for not liking the distinction was because many assessments serve multiple purposes. This uneasiness was also identified in the literature. Bennet (2009) suggested that assessments designed for summative purposes may be used formatively and the converse as well. Along similar lines, Dunn and Mulvenon (2009) claimed formative assessment include “all those activities undertaken by teachers, and/or by their students, which provide
information to be used as feedback to modify the teaching and learning activities in which they are engaged” (p. 10). Obviously, this would include assessments that might be used in a summative way in the classroom that also assist in improving learning. The participants in this study seemed to subscribe to these nuanced definitions of assessments and found that ultimately, assessment for learning suggested that assessment inform teaching by highlighting what is effective and ineffective in students achieving academic success and is most helpful if used for reform of the classroom, curriculum, or instructor pedagogy (Penn, 2011; Banta, Jones, & Black, 2009; Huba & Freed 2000; Atkin, Black, & Coffey, 2001; Black & Wiliam, 1998, 2004; Shepard, 2000).

When considering the assessment process for learning, scholars (Heritage, 2007; Wiliam and Thompson, 2007) typically noted specific steps including figuring out where the learners are, identifying the “gap”, and establishing how to best close the gap for students. Participants often described their personal approaches in comparable terms. Brian said “I think part of the process is looking at the assessment and trying to figure out what is working, what's not working, and reevaluating even the assessment tools and changing those where it's necessary, then making changes to the classroom to help them to meet those goals.” The way Brian envisioned and used assessment was to ultimately inform practice and make appropriate classroom changes.

As Brown (2008) noted, assessment for learning requires teachers to actively diagnose and ascertain knowledge and growth. The first step in this process is to identify where students are in their academics and describe that as reliably as possible. John described his department’s use of a pre- and post-test: “... the same pre-test and post-test which measures student’s content knowledge in the beginning to the end of the semester.” All the participants used a variety of testing methods, papers, quizzes, oral reports, etc. to identify and describe where students were
in their learning. Ken said, “One thing I think the assessment data could show is things, places, where students are weak.” Scott took the idea further and noted how the process was not just about “tracking students’ performance on the basis of a grade but trying to track a student’s intellectual development.” This highlights the participant’s desires to reliably determine students’ knowledge, understanding, or skill level at various points during the term.

Liu and Carless (2006) stated the term assessment is often interpreted as referring to marking, grading, measuring, and ranking rather than part of a learning process. All the participants in this study, however, used assessments to describe student knowledge as just a piece of a larger process. Just as Wehlburg (2011) wrote, “Once we know what they [students] know and what they don’t, we can more easily create an atmosphere of challenge that is appropriately rigorous,” the participants here highlighted how assessments shaped their courses. For example, many participants used a pre-test. Rob claimed, “I couldn't assume that students were all reading at ‘a college level.’ The assessment gave me clear evidence that was an issue the very first day or at the very first week and so I was able to create a new practice that dealt with that problem.” This part of the process is identifying the “gap.” Over two decades ago, Sadler (1989) claimed the essential purpose of formative assessment was to identify the gap between a student’s current status in learning and some desired educational goal. This view has not changed a great deal since then. Consider Wehlburg (2011), who said that knowing what students have learned and what they have not learned can help to guide educational decisions. The participants supported this line of thinking.

The teachers’ use of the evidence, obtained from students’ performance on assessment tasks, to adjust instruction and to guide students in adjusting their learning strategies is a significant part of the assessment for learning process (DeMeester & Jones, 2009, p. 7). Brian’s
comments were fairly representative of all of the participants’ desires to use assessment in this way. “We do take the assessment seriously in so far as improving our classes, improving how we teach a particular subject.” He continued discussing the refining process saying, “I think it's really the meat and potatoes of assessment. That's the best part.” At State University Red, Ken shared a similar sentiment concerning using data to improve the course stating, “It varies from faculty member to faculty member, but I think most faculty have some of that in their make-up.”

To guide these changes, Heritage (2007) describes the process of scaffolding, learning that builds upon individual development achievement by continuously guiding, challenging, and facilitating new opportunities to develop learning. Likewise, participants seemed to consider many assessment techniques and instruments in a way that would help them improve student learning. As Heritage (2007) suggests, the teacher’s job is to

... ensure that the student receives appropriate support so that new learning is incrementally internalized and ultimately becomes part of the student’s independent achievement. Matching the instruction to the gap cannot be done successfully without differentiating classroom instruction. (p. 144)

Differentiating instruction is supported by the use of instructors’ assessments.

For example, Dale used in-class quizzes to gauge student understanding of course materials. While the quiz was used in a summative way, the feedback it provided was used immediately. He analyzed the results in real-time, and revised the course to assist students. Other participants had similar assignments and changes to their curriculum or teaching style based on the assessment data they collected.

**Conclusions.** Overwhelmingly the participants’ thinking revolved around assessment for learning, as they understood assessment to be used essentially to improve student learning. For
humanities professors at both institutions, assessment for learning is more than just grading and evaluating students. Instead, assessment for learning means knowing who students are and what strengths and weaknesses they bring to the table and using assessment tools and practices to help students learn. Faculty use this information to help students learn both content and skills in their classes. Additionally, assessment is not just a tool to help students learn. It also is a tool for faculty learning. Professors can use assessment to strengthen their own practices as instructors. In order to use assessment for professional growth, faculty must be able to use assessment in ways that are relevant to their classes and relevant to them as individuals.

**Category: Assessment as Learning**

*Faculty use summative assessment but are apprehensive towards high-stakes testing.* Brown (2008) said that assessment as learning is essentially making students accountable for their own learning through a grade, score, etc. in an attempt to certify they have the needed knowledge or skills for the next level. Currently, in a humanities general education classroom, there are a number of assessments used to hold students accountable. While high-stakes tests are a part of the mix, participants do not rely on them. In fact, participants thought high-stakes testing on content or learning outcomes was more representative of high school assessment practices, and was anathema to good teaching and learning practices.

It is important to recognize that all the participants used multiple-choice tests. This was a decision based on the delivery system and a focus on what they wanted to measure. John recognized the stigma associated with these types of tests but said, “I’m not saying I apologize for doing multiple choice exams when I said there’s value to knowing content and being able to process no matter how it’s asked.” Affirming the purpose and utility of the test, specifically it being easier to quantify and standardize, John recognized they could be used for decent effect for
certain purposes. The literature recognizes the usefulness, reliability, and validity of the tests to measure specific aspects of student learning (Halayna, 2012; Niesser et al., 1996; Schmidt & Hunter, 1981).

Even though all of the participants used multiple-choice tests, the over-reliance on mandated tests troubled all of the participants. Dale captured the participants’ concerns when he said, “It reminds me a lot of the kinds of criticisms that we hear in the public schools in high school that is, say, teaching to the test. If you really followed the logic of all these student learning outcomes you should be teaching only to meet that student learning outcome.” This concern for high-stakes testing creeping into higher education has been voiced in the literature (Hursh, 2008; Scobey, 2009). The apprehension of using assessment via high-stakes tools offends the basic understanding of teaching and learning for the participants. The research on the most effective uses of assessment for student improvement supports them, as one size does not fit all (Shavelson & Huang, 2003). Banta (2007) reiterated the point thusly: “But there is no possibility that a single instrument given in a period of a few hours will provide a valid measure of the kinds of critical thinking, or even the kinds of writing, that faculty in engineering, history, and art, for example, expect their students to develop” (p. 2). The focus and investment is on the end product (student competency) with little to none being offered on the front end (student learning) and ultimately a focus on the prior does not benefit students or learning.

The problem of these tests, according to the participants, is that it takes one aspect of a student’s learning and puts it forth as representative of the whole person. John explicitly highlighted how participants do not rely on them to get a full picture of student learning. John called them “tip of the iceberg assessments” because “they’re screen grabs as opposed to actually going to a website and scrolling down surfing and reading the content.” Essentially, participants
thought it bad form to rely on just one type of assessment to truly get a sense of student progress and learning, especially in general education. This aligned with researchers like Aloi (2003) who found that assessments must be specifically designed for the general education outcomes.

**Conclusions.** While humanities faculty recognize the utility of high-stakes testing and believe some tests can adequately measure student understanding of content, the days of merely relying on a midterm and final multiple choice tests in higher education general education humanities courses seem long gone, if they were ever here. Participants expressed concern over a reliance on high-stakes testing and the effects it would have on student learning and engagement in higher education. Overall, faculty are concerned about high-stakes testing because it would negatively affect their classrooms by encouraging “teaching to the test” practices and providing a distorted or incomplete view of student learning in their humanities courses.

**Assessment motivates and engages students in learning.** Some participants found that students needed the extrinsic motivation provided though summative credit on many assignments to get them engaged in the learning process. Brian was the most succinct, stating, “They're not internally motivated necessarily to do well in the class, so they're not necessarily internally driven to perform well.” Later he added that students need to have some “skin in the game.” Banta (2007) said that in her experience “students may not give it their best efforts unless their performance has some consequences” (p. 3). While there may be concerns about using assessment in this way due to its possible negative effect on the depth of student learning (Kohn, 1993) and the possible deleterious effect on intrinsic motivation (Deci & Ryan, 2002), this practice is widespread and is reinforced in the literature. Elton (1996) described it this way:

... students are mostly keen to develop these [transferable skills], but in general only when they obtain some recognition for their achievements, preferably as part of their
degree assessment. Such a desire does not, of course, mean that students work only for grades; what it does mean is that they are unlikely to work hard without there being a recognition through grades. Par 16

Other scholars (Hilton et al., 2010; Pintrich & Schrauben, 1992) came to similar conclusions, finding that grades were the number one reason students of all ages work in a course.

The perceived importance, usefulness, and value of engaging in a task are motivators for student effort (Pintrich & Schrauben, 1992). While one participant described the rationale of his formative assessments to students, and met some success, other participants found students typically unwilling to put in the optimal time and effort for skill development, unless it was tied to summative credit. Discussing pre- and post-tests, one participant indicated that they are basically useless because, in part, students have no reason to do well.

One participant described a situation where students were just not completing the assigned reading for class. The formative assessment might come in the form of verbal in-class discussions on the material or in other in-class work. However, if students have not read, the formative assessments cannot be effectively employed. It appears this is not atypical, as Clump, Bauer, and Bradley’s research (2004) indicates that less than 30% of students complete assigned readings before coming to class. To address this problem, Dale modified his course and employed content quizzes. He claims this has worked well in motivating students to perform. His experience is also supported by the literature as some researchers have found success in encouraging reading by incorporating quizzes into their courses (Tomasek, 2009; Marchant, 2002). In her study, Connor-Green (2000) found that student study behavior is strongly influenced by tests, specifically that frequent planned quizzes motivated student to complete assignments. The majority of students also thought that daily quizzes maximized student learning.
As Brian conveyed, his rationale was to get students to move beyond what was going to be on the test, to “actually learn and try it, [to] learn how to do it so that they can prepare for anything on the test. That motivational aspect, these assessments are really important,” he said.

**Conclusions.** Many students are unmotivated to learn general education course content. General education courses may not stimulate or encourage intrinsic motivation in students as the major courses might. Humanities professors in general education courses use assessment, particularly summative assessment, to motivate students to do their assignments and to make them engage in the learning process. Ultimately, using summative assessment as a motivational tool keeps student accountable for their own learning.

**Category: Assessment of Learning**

**Reform, collaboration, and discussion driven by external groups.** Assessment of learning is related to institutional accountability. Participants overwhelmingly recognized the assessment process as being integrally intertwined with reaccreditation, mirroring what is demonstrated in the literature (Kuh & Ickenberry, 2009) and quality assurance and accountability process (Kuh et al., 2005; Maki, 2004). Participants also found this integration incisive in stimulating reform of general education as a whole, and assessment of student learning specifically by making them work more closely together.

For both institutions, the inceptions for general education reform were due to the reaccreditation process, which is typical in the literature (Blaich & Wise 2011; Kuh & Ickenberry, 2009). Participants at both institutions recognized the importance of SACS on their institution’s willingness to discuss and implement change. At State University Blue, Brian claimed, “. . . what motivated this campus to do that was SACS requirement,” and Rob
referenced the 2006 reform saying, “[we had to] implement a new curriculum . . . because we had to show results because the institution’s accreditation was at stake. . . .” In part these reforms were needed specifically due to lack of assessment data to improve student learning. The changes were instigated by SACS and were overall positive for both institutions. It is hard to imagine any meaningful general education reform taking place without the push from an outside agency.

More importantly, participants said general education assessment was instrumental in “forcing faculty to talk” as Brian said; or, put another way, “created opportunities for us to have these kinds of conversations,” as Jen said. All of the other participants made similar statements noting how general education assessment drove the faculty to dialogue and reform their programs and courses. This is significant because as Palumba & Banta (1999) pointed out, the involvement of those most affected by an assessment program is the most important factor for its success. It takes on further significance since national studies (Ewell, 1996; Kuh & Ikenberry, 2009) have identified lack of faculty support as the most significant barrier to effective implementation of assessment.

Conclusions. Humanities professors appreciate the rationale for general education assessment and recognize its significance to the institutional mission. Accrediting bodies and other outside stakeholders have a significant impact on assessment of learning in general education humanities courses, typically as instigators of change. Once the reform process is underway, humanities faculty view general education assessment process as a way to improve communication and collaboration between faculty, and to develop more clarity in general education and assessment.

Category: Assessment Irrelevant to Learning
Lack of ownership and responsibility related to general education. Participants seemed to exhibit some disconnectedness when it came to general education, in part due to its structure and implementation. Jen mentioned the fact that “Gen Ed is owned by everybody and, hence, nobody knows what’s happening with it” and that many of the faculty typically work within their disciplinary silos. Likewise these types of sentiments can be found in the literature. Wang and Hurley (2012) claimed faculty do not have as strong a feeling of ownership over general education as they do with their disciplinary home, and Wehlburg (2010) said the fact that no one owns general education makes it difficult to design and implement a general education assessment plan. Aloï et al. (2003) agreed and said the lack of having a group of faculty to hold accountable for its curriculum made it difficult to design and assess. Ken recognized that general education “might be a hard sell” because it was about all students and it is a bit easier when “you’re talking about we got to do this (assessment) for our students and make sure that they are learning what we want them to learn.” For Ken, he sees a difference from faculty buy-in and attitude between assessment for internal department and general education processes. This may be due to the feeling of ownership and responsibility.

Participants at both institutions found themselves agreeing on the difficulty of faculty buy-in to general education assessment. One significant stipulation was the 2006 reform at State University Blue which had a core general education faculty in place. It seemed to have a robust assessment program and an engaged faculty. This seems to be one of those rare instances where general education was designed and run like a department with its own budget, and faculty who were responsible for the bulk of the general education curriculum. Participants like Brian, John, and Rob spoke highly of the assessment program and its focus on using assessment for curricular improvement and student learning. The participants contended that the subsequent 2012
formation of general education and assessment does not have the same focus and faculty do not have the same internal motivation to get involved with the program.

One specific tension that may contribute to the disconnect faculty feel with general education assessment is the distinction between skill development and content knowledge. This tension appeared to make it difficult to create and implement general education assessment practices that were recognized as worthwhile to all the stakeholders involved. Newton (2000) described the content or skill debate this way:

Some emphasize the clear distinctions embodied in the foci and methods of the different disciplines and see the university as a loose collection of sharply defined departments drawn together under a broad institutional mission. Others stress a unity and coherence in the pursuit of knowledge that transcends departmental divisions and view the university as a common enterprise based on a coherent set of widely accepted assumptions.

Berrett (2012) explains that some see general education as a way to develop core skills like writing. He continues, “Because core skills are the goal, rather than specific knowledge, they can be taught through many disciplines, including combinations” (p. 37). For a more thorough examination of this debate see Adler-Kassner (2014). Participants reflected this dichotomy, struggling with two competing visions of general education and identifying how that affected their classrooms and assessments. Jen described State University Red’s general education as not having common content; instead, “We have to assess on the disciplinary ways of knowing. . . ,” and later recognized that according to external reviews some faculty may be too focused on content. On the other hand, Dale implemented more skill building in his general education courses because he saw the need to develop skills before the content could be fully explored. The
problem is how this tension gets translated into an understanding of general education as a whole, and general education assessment specifically.

A major problem with these competing views of the purpose and implementation of general education is its effect on the assessments. Aloi et al. (2003) found that instructors from many departments across campus teach general education courses, making it difficult for faculty members to agree upon desired outcomes, measurement tools, and administration of measurement instruments for common course requirements of general education. Jen captured the conundrum in this way, “When you're talking about something like Gen Ed you are looking at program level. Faculty are better at thinking about how to assess in their courses than they are in thinking about what is the role their courses play in a larger curriculum.” Ultimately, faculty struggle with the distinction of using assessment tools that are attempting to aggregate data across multiple sections on specific skills or knowledge competencies versus developing a tool that is more focused on disciplinary ways of knowing or skills.

**Conclusions.** General education design has a significant effect on the assessment instruments and processes. Faculty need a clear understanding of their university’s general education program and the rationale for its structure. It should go beyond protection of an area or discipline and specific courses. Once faculty have a clear understanding of general education’s goals, they should develop the tools that will measure those skills or content knowledge. A robust general education assessment program needs to have a faculty that not only support it, but are intimately involved with it and encouraged by it as meaningful stakeholders rather than finding the process just one more part of the job that needs to be completed.

**Questioned use of assessment due to its imprecise nature and subjectivity.** There is a significant amount of criticism concerning general education assessment both for its vague,
subjective language and imprecise tools. Anderson (2002) and Bers (2000) claim that defining general education learning goals in the humanities is more difficult than other academic areas because of their abstract nature and diverse disciplines. Glenn (2010) calls learning outcomes “jargon riddled” while Brottman (2009) said they were “superficial projections.” This study affirmed many of those criticisms. Participants were overwhelmingly critical of the vague language of general education learning outcomes. For example, Brian claimed that the learning outcomes for humanities like “being creative” and “critical thinking” were so wide-ranging that it became almost impossible to not achieve them, stating that “it’s really hard to miss that broad target.” Many scholars have leveled similar concerns about learning outcomes and their ineffable qualities that allow for so much interpretation that they are rendered almost meaningless.

Wehlburg (2010) specifically addresses the issue in her critique arguing that general education goals are difficult to measure, especially concepts like critical thinking or problem solving and that “Attempting to quantify some of these higher-order thinking skills is very difficult, if not impossible” (p. 90).

Faculty are put in a difficult position of having to operationalize the broad goals of general education classes into actionable and measurable student competencies in the classroom, and the frustration with that charge is clearly evident. Of course, not all of the participants shared those same concerns, at least not to that same degree. Discussing the faculty’s use of data to disengage from the assessment process, Jen shared this insight: “The results are not scientific enough so we can’t change program because we don’t have good data. We make it bad data by intention. Right? So we don’t have to change the program. We’re not stupid. Just incorrigible.” She recognizes that some faculty may be more resistant to the assessment program and may use whatever rationale they can use to discredit it or not get involved with it.
Part of the difficulty is taking the classroom assessment data and analyzing it in a way that objectively quantifies it in the reporting mechanisms. Once faculty got past the vague language and completed an assessment protocol, it is a subjective process to quantify the data. Hussey and Smith (2002) touch on this conundrum:

[Learning outcomes] are similar in that they do not lend themselves to precise definition. . . . the written learning outcomes may specify such things as ‘first year level’ and use such terms as ‘precisely’, ‘thorough understanding’ or ‘describe accurately’, but this is only pseudo-exactness. It is a vain attempt to describe in words what can be recognized by an experienced teacher--someone who knows how to judge material at the required level. (p. 111)

Here, the responsibility is on the teacher to reliably recognize and honestly assess the learning, but it is still a subjective process.

The quantification involved in assessment is suspect to humanities faculty. For example, Dale found the lack of a valid measurement instrument “demoralizing” and undercutting the entire process. Humanities scholars (Heiland & Rosenthal, 2011; Hussey & Smith, 2002) shared similar critiques. Another issue is how learning outcomes are written.

Conclusions. Humanities faculty find assessment irrelevant to learning for very specific reasons. Assessment results are broadly seen as objective. However, faculty view assessment language as vague and believe that assessment findings are the results of a subjective process. Humanities professors find that the subjective nature of assessing students is helpful, as noted in previous findings, however, when a fluid process is used in an objective way to make judgments and critical decisions, assessment becomes problematic.
Assessment is perceived as unimportant due to lack of support, value, and feedback.

To be successful, assessment programs need the faculty to be involved in the development and engaged in the process (Wehlburg, 2010; Hutchins, 2010; Walvoord, 2004); however, typically they feel disengaged from it (Kuh & Ickenberry, 2009). According to some research (MacDonald, 2014; Wehlburg, 2007) faculty are willing to support assessment programs if certain conditions (like having input at all stages of the process or clarity on why the assessment is taking place) are met.

The faculty in this study depicted a mixed result. As previously discussed, faculty generally recognized assessment as a necessary component of their professional lives. Additionally, most participants found distinct benefits of assessment to student learning and teaching improvement. While participants thought there were plenty of opportunities to engage in workshops, training, and other support mechanisms, the lack of clear recognition for the time or effort invested in the process left many questioning the rationale for doing so. At both institutions, faculty seemed to be very critical of the current administrative process and the lack of a feedback loop.

Some participants found plenty of opportunities to attend training or workshops or reach out to colleagues concerning general education assessment, but at the same time, they are unwilling to invest too much time in the assessment process without some return value or recognition. Faculty reward structures and policies are sometimes cited as key benchmarks by which faculty members allocate their time (Milem et al., 2000). Ken was pretty straightforward in his estimation, saying, “I don't think faculty get much recognition for work in this area, at any level, unless a faculty member plays a major University role, such as being a member of the General Education Council.” These impressions reinforce the current data on faculty impressions.
of the assessment process. Hutchins (2010) said faculty generally resist assessment because it does not benefit them. Others (Jones, 2010; Gaff & Wasecha, 2001; Gilmore, 2004) also found the lack of value towards tenure, etc., a major hindrance to faculty buy-in. Fairweather and Beach (2002) found that “Rewards seemed devoted first and foremost to research and scholarship” (p. 112). Wilson (2010) identified one faculty member who claimed that he would get promoted for his research not how well he assessed learning outcomes.

Even though faculty do not always see the value of assessment for themselves, they, in part, value what is recognized by the administration. Without administration investing in support up front through compensated training, workshops, time to concentrate on assessment, or through recognition given at end-of-year performance or tenure reviews, there is no reason for faculty to invest in the time or energy beyond what is minimally necessary for reporting standards.

Perhaps the most compelling message sent to the faculty concerning general education assessment comes from how the administration uses the collected data. This is essentially what Huba & Freed (2000) called the administration’s responsibility in creating policy that would encourage a culture of evidence. In the literature there is a marked concern about how the data is being used. For example, Blaich & Wise (2011) and Wehlburg (2010) identified extensive general education assessment data being collected, but they believed institutions were not aware of how best to make use of that data. At both institutions, participants detailed a lack of the use of results in any meaningful way beyond satisfying reports for the bureaucracy, and the participants described the need for assessment results to be shared with the individual faculty members, classrooms, and the department (Wehlburg, 2007; Watson 2003; Black & William, 2006). Jen concluded, “Our strategies for assessing are not always very useful. There are ways of
checking off . . . and showing SACS a document. . . .” The discrepancy between the learning outcome and the collected data is one where the fundamental issue of learning is never addressed because the result is reported and everyone moves on.

**Conclusions.** Faculty understand and recognize the importance of assessment as it pertains to the institutional mission and reaccreditation. However, humanities faculty see assessment as irrelevant to learning due to the lack of a corresponding value placed on assessment by the institution, which would be acknowledged through faculty development, funding, and/or scholarly recognition. Faculty appear to be more open to general education assessment process if their individual efforts are recognized through support and institutional recognition, and if the data resulted in meaningful, responsive change.

**Assessment movement is harmful to higher education and teaching.** Participants shared some of their concerns about the assessment movement overall. One of the most significant issues was the feeling of a top-down approach that was not clear and could lead to a loss of autonomy in the classroom. Many scholars (Kuh & Ickenberry, 2009; Whelberg, 2008; Blaich & Wise, 2011) recognized the importance of faculty involvement in the assessment process, however, the participants in this study found the assessment process actually removed from the faculty. For example, Rob found the administrators in conflict with the faculty by dictating many aspects of the assessment process. “[W]e as administrators want this instrument performed and boom, you do it.” Rob captured many participant’s thoughts stating that it should have been the administrators coming to the faculty and communicating the goals and criteria then asking the faculty how to move forward.

What makes this assessment process even more problematic is that many participants are already doing these things in the classroom and they see this process as merely adding more
busywork. This is in line with Wehlburg’s (2010) results where participants in her study said that assessment of student learning was happening regularly in the classroom, and it was a part of their jobs. Dale said that “most of the historians resent it [assessment] and find it to be a tiresome layer of the administration on top of the things that we are already doing.” Jen, Ken, Brian and others all mentioned the fact that being a teacher is inherently doing assessments and trying to improve student learning. Even if the assessments are embedded into the curriculum, the extra processes are seen as onerous and are affecting a faculty’s feeling of ownership of his/her classroom. In fact, to get reliable data more administrators are relying on standardizing the curriculum. While participants at State University Red still had significant amounts of discretion in their courses, participants at State University Blue found standardization creeping in.

Standardization of classes led to a feeling of disempowerment for faculty. This concern for the loss of classroom autonomy in general and deprofessionalization in general was prevalent in the literature (Brown, 2008; Hammon, 2009; Brottman, 2009; Fendrich, 2007; Wilson, 2010). Faculty also believe that assessment must be implemented correctly or it can be damaging to the student learning process.

While assessment is not currently being used as an enforcement or accountability tool, there is a great amount of concern that it is only a matter of time. Some critics of higher education accountability insist, for example, that accountability is a “myth,” if it does not include actions and consequences based on results (Carey, 2007). Faculty fears may be buoyed by the logical extension that, if the administration is investing money and resources into the assessment process, they would want compliance with whatever plan they put forward, or “meeting” some arbitrary goal. Wehlburg (2010) suggested that to have a successful general
education assessment program, administrators needed to assuage faculty fears by promising the data would be used to improve the program and not in any punitive way.

**Conclusions.** Faculty are concerned that the way the assessment movement is being implemented is not conducive to improving student learning and the higher education classroom; rather, it is an extension of corporate management techniques and language being misapplied to education. Humanities faculty are concerned about losing autonomy of the general education classroom and being forced into a position where real student learning is replaced with teaching to get specific results on assessments.

**Summary**

When considering general education assessment in the humanities, faculty are largely knowledgeable and recognize assessment’s utility to the institution. They believe that assessment is important because of its positive effect on student learning and the general education curriculum. There are a variety of issues, however, that create an environment where faculty do not see the direct value or benefit to them or their specific classes, such as the lack of administrative support and recognition for assessment work, or the data being reported but not actually used to support pedagogical improvements. Further, a structural hindrance to faculty buy-in to the general education assessment is the attempt to objectify the subjective process with equally subjective instruments. Finally, for general education humanities faculty, their primary focus related to assessment is teaching their students. Therefore, assessment data that is useful to them is designed, collected, and analyzed with teaching and learning in mind.

**Research Implications**

The humanities is a complex concept that extends across multiple disciplines. Due to the difficulty in recruiting, the study was modified and focused on one specific discipline within the
humanities. Additional research should be completed on how other humanities disciplines faculty conceive of and use general education assessment, especially since the humanities area is spread over many disciplines such as art, music, and English. With a wide variety goals and competencies being explored, qualitative studies that delved into each from a humanist’s perspective may offer unique insights. It might help clarify possible goals and resistances that humanities professors hold when considering general education assessment.

The rank of a faculty member or instructor influences their professional priorities. For example, an assistant professor working towards tenure may be more invested in improving their research over their teaching. Whereas, the faculty in this study were already tenured and they all expressed a sincere interest in teaching and student learning. Therefore, it would be important to hear conceptions from those faculty at different academic ranks. It may be that the tenured participants included in this study have more training, insight, or interest in assessment in general. Further, other faculty ranks such as instructors or adjuncts who do not have as much job security may have significantly different perspectives on what assessment is, how it is being used, and its overall significance to the higher education general education classroom.

The structure of the general education program and its focus may have significant impacts on how faculty approach the general education classroom and assessment. For example, a faculty specifically hired to develop, implement, and improve a general education program may have a significantly different level of attention and focus on assessment versus a typical faculty consisting of adjuncts, lecturers, Graduate Teaching Assistants, or tenure track faculty meeting administrative requirements. A further finding that has potential for future research is how the structure of general education and a faculty’s feeling of ownership of the program affected their perspective of assessment. It appears that if there is a dedicated faculty who are
responsible for the general education program may be more open to and supportive of the overall assessment program. Likewise, exploring how a program or institution can encourage or develop a feeling of ownership and/or responsibility for general education program might be important as additional research in this area might assist both administrative and faculty stakeholders plan and implement general education structure that is beneficial to students and is supported with a robust culture of assessment.

**Practical Implications**

Faculty in this study described general education as something from which they were somewhat disconnected, even though they were responsible for teaching it. This caused faculty to see the assessment process at the program level as something to be gotten through or checked off rather than something meaningful that would assist with student learning. Therefore, universities and/or departments should support ownership of the general education program through encouraging a professional, dedicated faculty who are responsible for the process and accountable to the institution. Faculty whose focus resides in teaching and learning of general education may feel more invested in and excited about the program overall. They may be able to pedagogically specialize or innovate in ways that are not currently feasible in the present system due to administrative support and recognition being more prominent for research or work within the major.

Faculty described a lack of any kind of meaningful response, reform, or discussion due to the collected data. General education assessment will be more useful if there is a systematic feedback loop in place. The data that is collected from faculty, analyzed by the institution, and ultimately reported, needs to get back to the faculty in some way so as to improve student learning. The process should be transparent and clearly communicated (Wehlburg, 2010) and
used to assist faculty for classroom improvement (Blaich & Wise, 2009). This process would benefit all the stakeholders involved, not the least of which are the students, and would more clearly communicate assessment’s significance to the institution.

In this study, faculty discussed the lack of support, encouragement, and recognition of their work in the area of general education assessment and its negative affect on their overall willingness to engage in the institutional process. Faculty should be professionally supported and recognized for their work in general education assessment. Opportunities need to be not only offered, but also encouraged by the administration. For example, active administrative reward structures should be instituted. Things like stipends or fellowships for training or research, funding for travel for professional development, course releases or stipends to engage in Scholarship of Teaching and Learning (SoTL) work in general education courses, etc. should all be explicitly identified as essential to a strong assessment program. Further, work in general education assessment should be recognized during tenure review or performance reviews with an explicit reward structure. Lastly, faculty should be given the resources to pedagogically innovate based on the collected data and to sustain those innovations. Without allocated resources investing in the general education assessment program, faculty will continue to see the assessment process as one required for reporting, while having no real connection to their jobs or their classrooms.

Faculty’s feeling of lack of ownership of general education, their lack of a feedback loop, and the lack of recognition could be partially addressed through The Scholarship of Teaching and Learning (SoTL). SoTL should be encouraged in general education programs. Faculty who are collecting and analyzing the classroom data should be encouraged to conduct original research, record best practices, document innovative pedagogies, and otherwise engage teaching
and learning and assessment as a standard part of the general education assessment program. Not only would this encourage improved teaching and learning practices based on data rather than anecdote, it would address some of the issues currently facing higher education including documenting student learning in higher education.

**Study Limitations**

The limitation of this study is its generalizability. Although the sample included faculty at two state institutions, all the faculty participants were tenured in their discipline. This allows for a unique perspective since many faculty who typically teach in general education are not on the tenure track. More often they are adjuncts or instructors who have a different charge at the university and less voice in how the structure is developed and the assessment process implemented. Further, participants were identified via Snowball sampling, so they all had some interest in or knowledge of general education assessment.

Another possible limitation was researcher bias. The difficulty in identifying participants resulted in the researcher including his own institution as a part of of the study. However, potential bias was minimized through the use of thick description, the clear and open recognition of previous interactions with the participants, and the participants’ opportunity to review interview transcripts and analysis prior to publication.

**Action Plan**

I have a few practical applications for the results of this study. First, I plan to attempt to publish the results in related academic journals. I believe that the feeling of ownership affecting general education assessment findings are salient for researchers in the area of general education. As such, I will review *Journal of General Education* to determine appropriateness. I also believe the findings concerning faculty’s perception of a lack of feedback loop, lack of support, and lack
of recognition affect their investment or buy-in to general education assessment programs and researchers involved in assessment may find them informative. I plan to examine *Assessment Update* as a possible venue. I plan to share my results at professional conferences like the Lilly Conference on teaching and Learning, the SoTL Conference, and MLA.

Lastly, I plan to share the results with my department. It is my hope that the results may spur more action towards developing a concrete feedback loop where general education assessment data is used to improve courses and the overall program. As a member of the assessment committee, we should be able to involve faculty in ascertaining specific learning outcomes we want to examine, discuss and identify a specific assessment instrument, and implement the process over the next semester. The process will include an opportunity for the assessment committee to collect and analyze the data and, more importantly, share those findings with the faculty for their insights and judgments concerning the data. Once fully discussed, one area of weakness should be identified and a plan of action will be explored and implemented by the faculty for the following term. Ideally this process assists the students in learning, the faculty pedagogy, and the administration in delivering a quality program. After a full cycle has been completed, faculty should be consulted as to their thoughts of the process, and adjustments can be made.
Appendix A

IRB Approval

Notification of IRB Action

Date: June 25, 2014
IRB #: CPS14-05-11

Principal Investigator(s): Corliss Brown Thompson
Chad Rohrbacher

Department: Doctor of Education
College of Professional Studies

Address: 20 Belvidere
Northeastern University

Title of Project: General Education humanities Professors’ Perceptions of Assessment

Participating Sites: **forthcoming**

Informed Consent: One (1) unsigned consent for survey
One (1) signed consent for interviews

As per CFR 45 46.117(c)(2) signed consent is being waived as the research presents no more than minimal risk of harm to subjects and involves no procedures for which written consent is normally required.

DHHS Review Category: Expedited #6, #7
Monitoring Interval: 12 months

Approval Expiration Date: JUNE 24, 2015

Investigator’s Responsibilities:
1. Informed consent form bearing the IRB approval stamp must be used when recruiting participants into the study.
2. The investigator must notify IRB immediately of unexpected adverse reactions, or new information that may alter our perception of the benefit-risk ratio.
3. Study procedures and files are subject to audit any time.
4. Any modifications of the protocol or the informed consent as the study progresses must be reviewed and approved by this committee prior to being instituted.
5. Continuing Review Approval for the proposal should be requested at least one month prior to the expiration date above.
6. This approval applies to the protection of human subjects only. It does not apply to any other university approvals that may be necessary.

C. Randall Colvin, Ph.D., Chair
Northeastern University Institutional Review Board

[Signature]

Nan L. Regina, Director
Human Subject Research Protection

Northeastern University FWA #4630
Appendix B

IRB Modification

Northeastern

Notification of IRB Action
Modification

Date: November 25, 2014
IRB #: CPS14-05-11

Principal Investigator(s):
Coeliss Brown Thompson
Chad Riehler

Department:
Doctor of Education Program
College of Professional Studies

Address:
20 Belvedere
Northeastern University

Title of Project:
General Education Humanities Professors’ Perceptions of Assessment

Modification:
Add interviews via snowball sampling with an unsigned consent
participating Sites:
Permission letter on file
Informed Consent:
One (1) unsigned consent form

DHHS Review Category:
Expedited #6, #7

Monitoring Interval:
12 months

Approval Expiration Date: JUNE 19, 2015

Investigator’s Responsibilities:

1. Informed consent form bearing the IRB approval stamp must be used when recruiting participants into the study.
2. The investigator must notify IRB immediately of unexpected adverse reactions, or new information that may alter our perception of the benefits-risk ratio.
3. Study procedures and files are subject to audit at any time.
4. Any modifications of the protocol or the informed consent as the study progresses must be reviewed and approved by this committee prior to being instituted.
5. Continuing Review Approval for the proposal should be requested at least one month prior to the expiration date above.
6. This approval applies to the protection of human subjects only. It does not apply to any other university approvals that may be necessary.

C. Randall Colvin, Ph.D., Chair
Northeastern University Institutional Review Board

Nan C. Regna
Director, Human Subject Research Protection
Appendix C

Recruitment Email Survey

Dear Faculty,

My name is Chad Rohrbacher, and I am writing to you concerning the possibility of participating in a research study titled “Perceptions of Assessment by Humanities Professors in General Education.” I am conducting this research as a doctoral candidate at Northeastern University, Boston. The purpose of this study is to allow you an opportunity to describe your thoughts around what constitutes assessment and its use in a general education humanities classroom. I am attempting to better understand how higher education humanities faculty conceive and use assessment data. With your help in this study, I will better understand faculties’ reactions to calls for assessment by various stakeholders and how assessment is used in your general education courses.

This questionnaire will take approximately 30-45 minutes of your time. To ensure participant confidentiality, the questionnaire will be completed online via Qualtrics. The questionnaire is housed entirely on Qualtrics’ secure Website. If you are amenable to a follow-up interview, you will have the opportunity to indicate your interest at the end of the survey. Indicating interest in a follow up interview does not guarantee selection or commitment.

You have been asked to participate because you are a general education faculty member in one of 3 humanities departments identified across campus. This survey is open to all faculty regardless of rank.

You should have taught at least one general education humanities course the current or previous academic year. Since this study is focused on faculty perceptions and uses of assessment, if your assignment is more than 3/4 time administrative please do not participate.

Your participation in this study is voluntary at all times, and you are not required to respond to every question. You may choose to not participate, or to withdraw at any time. To do so, simply exit the survey without submitting your answers. Deciding not to participate will not result in any penalty. If you elect to exit the survey prior to submitting your answers, your answers will not be collected as part of this research. Choosing not to participate or withdrawing from participation will not affect your standing at the university.
Your information will remain confidential at all times for the duration of the study. While you will need to use your official university email to enter the questionnaire, an identification code will be used and individual responses will be confidential. To ensure participant confidentiality, the questionnaire is housed entirely on a secure Website. As a result, personal contact between the subject and researcher will not take place in reference to completion of the questionnaire.

Additionally, data will be collected from a Web-based form that will NOT collect any traceable or personally identifiable information (IP address, computer name, etc.). I will ensure that no clues to your identity appear in the thesis. Any extracts from what you say that are quoted in the thesis will be entirely confidential.

To participate, please use the following link to Qualtrics:

http://northeastern.qualtrics.com/SE/?SID=SV_3sZqyt4IbkSa7fD

I would like to thank you very much in advance for your participation and assistance in this vital research project. Please let me know if you have any questions or concerns about participating in this research.

You may contact me at: rohrbacher.c@husky.neu.edu You can also contact Dr. Corliss Brown Thompson, the Principal Investigator at co.brown@neu.edu or 617.637.6702. If you have any questions regarding your rights as a research participant, please contact Nan C. Regina, Director, Human Subject Research Protection, 960 Renaissance Park, Northeastern University, Boston, MA 02115. Tel: 617.373.4588, Email: n.regina@neu.edu. You may call anonymously if you wish.
Appendix D

Recruitment Email for Interviews

My name is Chad Rohrbacher, and I am reaching out to you about the possibility of your department's participation in a study titled “Perceptions of Assessment by Humanities Professors in General Education.” I am conducting this research as a doctoral candidate at Northeastern University, Boston.

The purpose of this study is to allow your faculty to describe their thoughts around what constitutes assessment and its use in a general education humanities classroom. I am attempting to better understand how higher education humanities faculty conceive and use assessment data. With their help in this study, I will better understand faculties’ reactions to calls for assessment by various stakeholders and how assessment is used for teaching and learning in their general education humanities courses.

The study would be in two parts. The first part would include a web-based questionnaire open to general education faculty members regardless of rank. The questionnaire is voluntary and confidential. From that data, I would identify a potential participant for a follow up interview. I have included the consent letters for your review.

Their information will remain confidential at all times for the duration of the study. While they will need to use their official university email to enter the questionnaire, an identification code will be used and individual responses will be confidential. I will ensure that no clues to their identity appear in the thesis. Any quoted extracts in the thesis will be identified via pseudonym. Participants will also have an opportunity to review interview transcripts and initial findings. Also, the institution would not be identified in the dissertation or any other publications/presentations.

I would also like to explore the possibility of reviewing any end of year reports or other documents concerning general education and assessment. They will help me contextualize the faculty responses and give me a fuller understanding of the department’s approach and goals to assessment in general education courses.
Northeastern University’s Institutional Review Board (IRB), (Redacted) IRB, and Dr. (redacted – Dean of A&S) have reviewed the study. Please let me know if you have any questions or concerns about this research.

You may contact me at: rohrbacher.c@husky.neu.edu. You can also contact Dr. Corliss Brown Thompson, the Principal Investigator at co.brown@neu.edu or 617.637.6702. Lastly, if you have any questions regarding your rights as a research participant, please contact Nan C. Regina, Director, Human Subject Research Protection, 960 Renaissance Park, Northeastern University, Boston, MA 02115. Tel: 617.373.4588, Email: n.regina@neu.edu. You may call anonymously if you wish.

Thank you,

Chad
Appendix E

Unsigned Consent Survey

Northeastern University, Department of: College of Professional Studies
Name of Investigators: Dr. Corliss Brown Thompson, Chad Rohrbacher
Title of Project: General Education Humanities Professors’ Perceptions of Assessment

Request to Participate in Research
We would like to invite you to take part in a research project. The purpose of this research is to better understand how humanities general education assessment data is perceived and used in the classroom.

You must be at least 18 years old to be in this research project.

The first interview can either be completed in-person at a location of your choice or you can complete the interview questionnaire and return it via email. The first interview should take 30 minutes to an hour. After you complete the first interview, the researcher will contact you via email to schedule the second interview. The second interview will take place, in-person, at a location of your choice and will last about an hour. We will ask you to describe your thoughts around what constitutes assessment and its use in a general education humanities classroom.

The possible risks or discomforts of the study are minimal. Risk only exists in the event of an unanticipated breach of anonymity. In that event, participants may risk potential embarrassment or future harm to professional careers if anonymity is somehow breached or identities disclosed. Specific steps will be taken to retain participant confidentiality and anonymity and the likelihood is minimal.

There are no direct benefits to you for participating in the study. However, your answers may help us to learn a more of what comprises assessment and its uses in the humanities general education classroom, which may assist humanities faculty and administrators in understanding how assessment can be used to improve the learning experiences of students.

Your part in this study will be handled in a confidential manner. Only the researchers will know that you participated in this study. Any reports or publications based on this research will use only group data and will not identify you or any individual as being of this project.

The decision to participate in this research project is up to you. You do not have to participate and you can refuse to answer any question. Even if you begin the study, you may withdraw at any time.

You will receive a $20 visa gift card at the end of the interviews.

If you have any questions about this study, please feel free to call me, Chad Rohrbacher, at 336-402-6321. You can also contact Dr. Corliss Brown-Thompson at 617-637-6702, the Principal Investigator.

If you have any questions about your rights in this research, you may contact Nan C. Regina, Director, Human Subject Research Protection, 960 Renaissance Park, Northeastern University, Boston, MA 02115. Tel: 617.373.4588, Email: n.regina@neu.edu. You may call anonymously if you wish.

You may keep this form for yourself.

Thank you.

Chad Rohrbacher
Appendix F

Informed Consent Interview

Email / Signed Consent Form Concerning Individual Faculty Interviews

Dear Sir or Madam,

You recently completed an online questionnaire concerning the study titled “Perceptions of Assessment by Humanities Professors in General Education” and have expressed interest in participating in a follow up interview. I am contacting you to follow up on your willingness to describe your thoughts around what constitutes assessment and its use in a general education humanities classroom. Please review the information below, print, and sign the document.

Northeastern University, College of Professional Studies

Name of Investigator(s): Dr. Corliss Thompson, Principal Investigator; Chad Rohrbacher, Student Researcher.

Title of Project: Perceptions of Assessment by Humanities Professors in General Education

Informed Consent to Participate in a Research Study

I am interested in exploring your thoughts around what constitutes assessment and its use in a general education humanities classroom.

This form will tell you about the study, but the researcher will explain it to you first. You may ask this person any questions that you have. When you are ready to make a decision, you may tell the researcher if you want to participate or not. You do not have to participate if you do not want to. If you decide to participate, the researcher will ask you to sign this statement and will give you a copy to keep.

Why am I being asked to take part in this research study?

I am asking you to participate in this study because you are general education faculty member in one of three humanities departments identified across campus and you expressed interest in the on-line questionnaire.

Why is this research study being done?

The purpose of this study is to allow you to describe your thoughts around what constitutes assessment and its use in a general education humanities classroom.

What will I be asked to do?

If you decide to take part in this study, I will ask you questions concerning your thoughts about assessment of general education in the humanities. The interview will be recorded, however, to ensure participant confidentiality, you will be identified by a pseudonym. Your information will remain confidential at all times for the duration of the study. I will ensure that no clues to your identity appear in the results. Any extracts from what you say that are quoted in the thesis will be entirely confidential. You will have an opportunity to review the transcribed interview and the final results to ensure accuracy.

Where will this take place and how much of my time will it take?

The initial interview will take approximately 60 minutes of your time. You will be interviewed at a location of your choice at a time that is convenient for you. If any follow up questions need to be considered, you will be contacted via email and a second interview will be scheduled. This interview will be no more than one hour and will be at a location of your choice at a time that is convenient to you.

[ signature ]

Northeastern University - Human Subject Research Protection

Date: 5-13-2011
Valid THROUGH 6-29-2011
Will there be any risk or discomfort to me?

There are some foreseeable risks or discomforts to you for taking part in this study. The risk of social harm is minimal, and only exists in the event of an unanticipated breach of confidentiality. In that event, you may risk potential embarrassment or future harm to professional careers if confidentiality is somehow breached or identities disclosed. Specific steps will be taken to retain your confidentiality and the likelihood is minimal.

Will I benefit by being in this research?

There will be no direct benefit to you for taking part in the study. However, the information learned from this study may help you gain insight into your own conception and uses of assessment data.

Who will see the information about me?

The interview will be recorded, however, to ensure participant anonymity, you will be identified by a pseudonym. Your information will remain confidential at all times for the duration of the study. I will ensure that no clues to your identity appear in the thesis. Any extracts from what you say that are quoted in the thesis will be identified by pseudonym. You will have an opportunity to review the transcribed interview and the final results to ensure accuracy.

Only the researcher and those on his doctoral committee will have access to the data.

Digital recordings of the interviews will be transcribed by Rev.com and then reviewed by you for accuracy. Once accuracy is established, the digital audio will be deleted.

All data will be stored on the password-protected computer of the researcher for three years. After three years the data will be deleted.

If I do not want to take part in the study, what choices do I have?

Your participation in this study is voluntary at all times, and you are not required to respond to every question. You may choose to not participate in the interview, or to withdraw at any time. Deciding not to participate will not result in any penalty. If you elect to exit the interview before the interview is concluded, your answers will not be collected as part of this research. Choosing not to participate or withdrawing from participation will not affect your standing at the university.

What will happen if I suffer any harm from this research?

No special arrangements will be made for compensation or for payment for treatment solely because of my participation in this research.

Can I stop my participation in this study?

Your participation in this research is completely voluntary. You do not have to participate if you do not want to and you can refuse to answer any question. Even if you begin the study, you may quit at any time. If you do not participate or if you decide to quit, you will not lose any rights, benefits, or services that you would otherwise have as a faculty member.

Who can I contact if I have questions or problems?

If you have any questions about this study, please feel free to contact me at: 336-402-9321 or rohrbacher.cj@husky.neu.edu. You can also contact Dr. Corliss Brown Thompson, the Principal Investigator at co.brown@neu.edu or 617.637.6702.
Who can I contact about my rights as a participant?
If you have any questions about your rights in this research, you may contact Nan C. Regina, Director, Human Subject Research Protection, 960 Renaissance Park, Northeastern University, Boston, MA 02115. Tel: 617.373.4588, Email: n.regina@neu.edu. You may call anonymously if you wish.

Will I be paid for my participation?
There is no payment for participation.

Will it cost me anything to participate?
There are no costs associated with participation.

Is there anything else I need to know?
You must be at least 18 years old to participate unless your parent or guardian gives written permission.

I agree to take part in this research.

Signature ___________________________ Date ____________

Printed name of person above ___________________________
Appendix G

Interview Protocol

1. In your own words, how do you define assessment?

2. How would you describe the current purpose(s) of assessment in higher education?

3. How would you describe the impact of assessment on your discipline?

4. When hearing the term “assessment” in relation to your department, what are your initial thoughts and impressions?

5. What are your initial thoughts when hearing the term “assessment” in relation to your general education courses?

6. When your department collects assessment data from general education courses, what is your impression of their utility to the department? To the faculty?

7. How would you describe your use of assessment in the classroom?

8. How would you describe the impact of assessment on your teaching?

9. Considering your general education colleagues in your discipline, how would you characterize the overall impact of assessment on their teaching?

10. How would you describe the impact of assessment on student learning and performance?

11. If you have any other thoughts concerning assessment in your general education courses, please share them here.
References


London: Routledge.


Ewell, P., Paulson, K., & Kinzie, J. (2011, June). Down and in: Assessment practices at the


http://www.learningoutcomeassessment.org/documents/PeterEwell_005.pdf


Innovative assessment in higher education (pp. 23-36). London: Routledge.


http://www.ccsso.org/Resources/Publications/Learning_Progressions_Supporting_Instruction_and_Formative_Assessment.html


http://atlantic.edu/about/research/documents/OpeningDoortoFacultyInvolvement.pdf


http://www.learningoutcomeassessment.org/NILOAsurveyresults09.htm


Lin, Yi-Guang, McKeachie, W. J. “College Student Intrinsic and/or Extrinsic


*Mortgaging our future: How financial barriers to college undercut America’s global competitiveness* [A REPORT OF THE ADVISORY COMMITTEE ON STUDENT FINANCIAL
ASSISTANCE]. (2006, September). Retrieved from Department of Education:
http://www2.ed.gov/about/bdscomm/list/acsfa/mof.pdf


*College Teaching, 43*(3), 102-103.


http://phrp.nihtraining.com


Ratcliff, J. L., Johnson, D. K., La Nasa, S. M., & Gaff, J. G. (2001). The status of general...


Spalter-Roth, R., & Scelza, J. (2009). *What's happening in your department with assessment*
[White Paper]. Retrieved from American Sociological Association:

Spellings Commission. (2006). *A test of leadership: Chartering the future of U.S. higher education* [The report of the secretary's commission on the future if higher education]. Retrieved from Education Department:
http://www2.ed.gov/about/bdscomm/list/hiedfuture/reports/pre-pub-report.pdf


Suskie, L. (2010). Why are we assessing? Retrieved from


humanities departmental survey. Retrieved from
http://www.humanitiesindicators.org/binaries/essayTownsend.pdf


