UNDERGRADUATE CRITICAL THINKING ASSESSMENT PROCESSES AND EFFECTS IN A PUBLIC UNIVERSITY CASE STUDY

by

Mary E. Zamon
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
the Requirements for the Degree
of
Doctor of Philosophy
Education

Committee:

Martha Ford _____ Chair

Gary R. Gallegos _____

S. David Brazer _____

Tim E. Evans _____ Program Director

_____ Dean, College of Education and Human Development

Date: November 24, 2008 _____ Fall Semester 2008
George Mason University
Fairfax, VA

Undergraduate Critical Thinking Assessment Processes and Effects in a Public University Case Study

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

By

Mary E. Zamon
Master of Arts
Webster University, 1983
Master of Science
Georgetown University, 1967

Director: Dr. Martin Ford, Professor
Senior Associate Dean, College of Education and Human Development

Fall Semester 2008
George Mason University
Fairfax, VA

DEDICATION

This work about thinking is dedicated to the hearts and minds of Frederick and Jen Zamon in appreciation for all their efforts to support my education, projects, and ideas. Another dedication goes to Carla, Judy, and Andrew for providing listening ears and hearts when I needed them. Truly this was an example of the extent to which family matters and to the Zamon family motto: “Zamons never quit.”

ACKNOWLEDGEMENTS

I want to thank a number of people whose support and aid enabled me to complete this endeavor. My husband, Fred, whose famous patience was put to the test, is the key to my being able to work fulltime and, for six years, follow this additional path. Inspiration for starting came from my daughter Jen, whose own Ph.D. work and academic achievements opened new ideas and horizons for me. My whole family did without attention and support I missed giving during coursework, writing and more writing. My appreciation for their tolerance is wide and deep. Carla, Judy, and Andy thank you again and again.

My thanks for guidance, inspiration, support and patience go also to many professors at George Mason. Not only have they been teachers, but the many friends I have worked with have often and kindly asked how it was going and offered their encouragement.

My two supervisors at work, Dee Holisky and Karen Gentemann have always been supportive and encouraging, allowing a good deal of flexibility so I could attend classes or write. Karen especially supported me as a member of my first advisory committee during my course work. Members of the assessment office past and present were unfailingly tolerant and supportive of the effects that dissertation work had on me. Erin, Rawa, Ying and Karen M, thank you! I also had support from Marymount University employers when I worked there during the first year of the program.

To friends and colleagues in the Ph.D. program, especially James Young, my mentor for two years and Sam Steen and Wendy Matthews who constantly offered collegial and strong support, I offer thanks.

Last but not least, I owe a debt of thanks to my dissertation committee, Drs. Martin Ford, Gary Galluzzo and David Brazer. Their comments, suggestions, and constructive critiques make this a better product than it would have been. Thanks to Dr. Brazer who first interviewed me for the program and has continued on my committee ever since. Dr. Ford's quiet patience, detailed commentary, and encouragement contributed to the success of this work and those same qualities also were evident in his service on my advisory committee. Dr. Galluzzo could always be counted on for a challenge that made me better.

To the many people who expressed interest, who wondered why I started such a project but encouraged me anyway, I say thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

UNDERGRADUATE CRITICAL THINKING ASSESSMENT PROCESSES AND EFFECTS IN A PUBLIC UNIVERSITY CASE STUDY

Mary E. Zamon, Ph.D.

George Mason University, 2008

Dissertation Director: Dr. Martin Ford

This case study follows the work of a large state university tasked with designing and completing state-mandated assessment of undergraduate student critical thinking competency. A participant researcher case study examines assessment records, field notes, and data collected from faculty and administrators through interviews and questionnaires. The research followed how this assessment was carried out, what processes and decisions were involved, the effects of the process of assessment on the institution, and used assessment, critical thinking, group work and institutional renewal concept models to propose future research and best practice directions. Findings and conclusions include a new Process Self-Renewal Model which can be applied in future research and suggestions for specific best practices in assessment.

1. Introduction

Eye on Assessment

2006 may well be viewed as a watershed year for change in higher education as universities engaged issues associated with student learning more closely than ever before. The higher education atmosphere is changing and chief among change issues is the question: How do faculty and administrators actually know if and what students are learning? Administrators and institutional leaders are not the only ones interested in such questions. Students, faculty, various educational boards, accrediting organizations, state bodies, national organizations, and, most recently, the federal government are all focusing on what institutions know or should know about what their students know. The effects of the K-12 No Child Left Behind federal policies are drawing attention to ideas about similar federally mandated standards for higher education. Institutions are becoming ever more entangled in the web of oversight described by Lane (2007) as having both latent and manifest forms, including various levels of influence and oversight from states, the national government, agencies, organizations, and the general public.

Higher education administrators are in a spotlight focusing on questions about student success and, thus, institutional success. In order to be able to make decisions that will contribute to both, administrators need to be able to know how institutional

learning goals for students are met and develop institutional processes to provide that information.

One of the most important sources of such information is *assessment of student learning conducted according to strong professional standards*. As discussions of national standards for higher education multiply, and the report, commonly called the Spellings Report, of the U.S. Department of Education Commission on the Future of Higher Education (2006) generates daily news and concerns, university administrators are focusing ever more intently on how their institutions will be perceived both now and in the future.

A close watch is necessary on what happens as a result of this report and consequent efforts to influence accrediting agencies and Congressional legislation. The Commission report (United States Department of Education, 2006) focused on accountability, perceived imbalance of attention to research and student learning by professors and administrators, the need for better data, and shortcomings in accreditation. The report called for "a robust culture of accountability and transparency throughout higher education" and states that all other goals "will be more easily achieved if higher education embraces and implements serious accountability measures" (p. 21). A large portion of the report presented the argument for standard, comparable tests, and addressed perceived serious issues with accrediting agencies. Some recent signs suggest that the department leaders are toning down the argument. Secretary of Education Spellings spoke to the National Advisory Committee on Institutional Quality and Integrity, explaining: "Let me repeat: no one-size-fits-all measures. No standardized tests.

All I ask is that institutions be more clear [sic] about the benefits they offer to students. Through the accrediting process, we can help bring this about." (http://www.ed.gov/news/pressreleases/2007/12/12182007.html)

However, it is too soon to assume that the spotlight has dimmed by much. It seems that if universities are seen as not willing to address issues related to learning accountability, future federal action is increasingly likely, either from the Department of Education or from Congress. The persistence of the Department's attempts to force change on accreditation bodies through rule making caught the attention of the federal legislature, which slowed efforts to impose federal administrative control of what accrediting agencies must require their institutions to do. In spite of the fact that some specific policy changes were deflected, attention is not likely to wane. Expectations that a new federal administration will not return to questions of higher education accountability are probably misplaced.

Accrediting associations and others are responding reactively, with traditional lobbying efforts and other initiatives. Regional accreditors like the southern regional body are asking for specific student learning outcome measures to be included in on-going institutional planning and evaluation (Southern Association of Colleges and Schools [SACS], 2005). The SACS Resource Manual not only asks for data on assessment, but also asks if budget and other support for assessment are present, and how results of assessments are used for institutional improvement. Revisions to reporting requirements were voted on at their December 2007 meeting, and include strengthened requirements for self-assessment of program-level student learning outcomes, plus double

scrutiny through off-site and on-site review teams (http://www.sacscoc.org/principles.asp). In general, regional bodies like SACS want to keep the way in which accountability is demonstrated in their hands, agreed upon by their constituent members, and do not want to give up those choices to federal agencies.

It remains to be seen how accountability will be defined by the bodies that make up the many strands of the assessment web, which will no doubt remain sticky for higher education institutions. Accountability issues, once raised, are not likely to disappear. The stakes are high and apparently if accreditors and institutions do not prepare their own accountability procedures, the federal legislative and executive branches stand ready to prescribe cures. Institutional leaders are alarmed. For example, Lawrence University's president warned:

> A fundamental strength of American higher education is its remarkable institutional diversity. Rather than embracing diversity, the Commission's proposals and rhetoric - in areas as varied as tracking academic progress through a centralized database to improving 'productivity' by imposing misguided benchmarks - have the effect of trying to homogenize American higher education at the expense of its ability to help all of our citizens realize their potential (The Annapolis Group, 2006).

National attention to higher education accountability is also increasing through organizational studies such as *Student Success in College: Creating Conditions that Matter* (Kuh, Kinzie, Schuh, Whitt, & Associates, 2005), *Liberal Education Outcomes: A Preliminary Report on Student Achievement in College* (Association of American

Colleges and Universities, 2005), and *Understanding University Success: A project of the Association of American Universities(AAU) and the Pew Charitable Trust* (Conley, 2003), all of which reflect growing concern by research and professional associations. Some associations are very serious about assisting members with seminars, workshops and other tools such as those offered by the AAU whose report on liberal education even includes a CD offering implementation guides. A new survey of employers (Peter J. Heart Research Associates, 2008) makes clear that although recent graduates have skills to begin employment, they are less prepared for advancement. The report suggests that employers clearly desire more. For example, in the area of critical thinking, employers gave a rating suggesting that preparation left much to be desired. As for the question of how to determine if students in fact have the desired skills, the authors report that employers "have the most confidence in assessments that demonstrate graduates' ability to apply their college learning to complex, real-world challenges, as well as projects or tests that integrate problem-solving, writing and analytical reasoning skills" (p. 4).

Other organizations raise the volume of discussion considerably. One of these is the American Council of Trustees and Alumni, which commissioned a report (Latzer, 2004) called *The Hollow Core: Failure of the General Education Curriculum* which condemns current university undergraduate general education programs for lack of rigor in course requirements. According to the report, only one of the fifty institutions surveyed required six or seven core courses. At first glance, the title alone suggests educational disaster; however, this report did not investigate what students actually know but only what course offerings are required. Significantly, this is an association of

trustees turning its eye on assessment of course offerings and adding one more line to the web. Associations and their efforts form a kind of mini-industry, and the enticement of institutional administrators to conferences and seminars raises yet another flag to institutions already inundated with warnings.

One of the newest association efforts is a pilot effort currently underway by The National Association of State Universities and Land-Grant Colleges (NASULGC) to produce a voluntary, common reporting method which includes both testing and institution-specific measures (2007). The common template includes many measures already being reported by universities and colleges as well as space for new sections. At the 2007 meetings of three associations for professional assessment officers all of which I attended, this effort was a topic of cautious optimism. The Indianapolis Assessment Institute, for example, sponsored a lunch discussion with experienced assessment professionals to present and explain the results of the Lumina Foundation grant that enabled development and use of the template, dubbed College Portrait. The latest information on this project is available from http://www.voluntarysystem.org.

In addition to national attention, regional accreditation mandates, and concerns of university administrators, some states are already requiring reporting on assessment of specific student learning outcomes. In South Carolina, the state has tied such assessments to funding (Cook, Johnson, Moore, Pauly, Pednarvis, Prus, & Ulmer-Sottong, 1996). Another example is Virginia, where reports on undergraduate student learning competencies are now required in submission of Reports on Institutional Effectiveness (http://research.schev.edu/roie/?from=reportstats). However, at this time,

Virginia's competencies are not tied to funding. If a direction for institutions can be discovered from these environmental pressures, then that direction leads straight to institutional assessment and the ways in which it is carried out.

Significance of Investigation of Assessment

The oversight web beginning to envelope institutions serves to focus attention on student learning outcomes and to suggest that institutions will need to understand how to design, implement, evaluate, and use results of measurable and meaningful student learning outcomes. It is disappointing to see that a new publication offering methods and advice on higher education organizational performance (Miller, 2007) does not address student learning assessment in terms other than student surveys and counts of alumni employment. Deeper knowledge of assessment processes based on faculty involvement can provide guidance to administrators looking for accountability response, and can have institutional impact as courses and programs respond to assessment results

The question remains: "But what are students learning?" Looking at a process of assessment can provide an answer to this question and can make significant contributions to three areas, (1) knowledge of how to direct assessment to improve students learning; (2) insight and suggestions relevant to assessment best practice; and (3) potential for modeling useful future research efforts.

In addition to the efforts previously described, the Association of American Colleges and Universities provides a guide (Leskes & Wright, 2005) for assessing undergraduate general education through direct measures. The guide includes a primer on assessment and practical advice that describes the kind of assessment effort chosen for

this study. The report offers key advice to build on what is an already occurring process, share responsibility, and keep assessment on-going, with sustainable institutional management. As the report states:

> Yet too often assessment is regarded as a matter of gathering some data (which may be more or less related to an actual learning goal), writing a report and then forgetting about the entire exercise until the next request years later (p. 22).

In a particularly insightful statement, the report places responsibility firmly. "For faculty and institutions, on-going assessment may be a fractal image of the lifelong learning they expect students to manifest" (p.23).

These many outside environmental pressures form one set of factors motivating change in institutional assessment practices. There is another set that is internal. Institutions of higher education are sometimes thought to work in mysterious ways. Decisions are made, studies are carried out, and reports are written and shelved. Frequently there is a high level of uncertainty about how things work. Work in some sense is done, but that work may not lead to either conscious validation or change. Intentionality and application of knowledge may be missing. Birnbaum (1998) and Kezar (2001) present several models of ways higher education institutions actually carry out their work. Their discussions dispel some myths, such as the universal friendly and collegial decision-making that Birnbaum finds inaccurate. Both authors propose, in slightly differing terms, that a successful institution will be self-renewing or using Birnbaum's term, cybernetic. An institution that aspires to continually replicate its successes should develop reflective self-knowledge, the will to change, and intentionally

supportive operating processes. Institutions themselves can become the lifelong learners the American Association of Colleges and Universities (2005) report suggests. Prior models like those from business, Birnbaum and Kezar report, have not worked. If attention is now on student learning, the development and use of information about student learning outcomes can be an even larger contributing factor to that self-renewing state.

Administrators tasked with moving to the next level of self-renewal need information that provides useful direction for action. Kezar (2005) again offers insight on collaboration as a key prerequisite - faculty and administration need to work together. Bauman (2005) adds recognition of the link between organizational learning and equity in student outcomes. In a special issue of *New Directions for Higher Education*, Lieberman (2005) provides a concrete look at faculty development and the potential for change in centers for teaching and learning. Close observations by these researchers have led them to call for improving institutional practices related to change. The impetus for change from the external environment is meeting with recognition of needs from within institutions, multiplying the significant uses to which information and results from the case study undertaken here might be applied. Certainly, researchers and administrators can conclude that there are strong indicators that outcomes assessments are important, that reliable information will contribute to institutional success, and that well conceived assessment can answer questions of accountability.

Purpose

The previous description of external and internal environments suggests that deep understanding of a particular case will be helpful in contributing relevant knowledge. Ample internal and external impetus for examining how an institution approaches assessment of student learning outcomes now exists, and influenced the decision to undertake this project. The project aim is to investigate processes in development, implementation, and use of the results of one university-wide assessment of undergraduate student learning. The purpose of this study is to shine a strong light on one case, to present and analyze one assessment process, and to identify its effects in a higher education institution. Additionally, information and ideas generated by the study can offer preliminary insight and direction to institutional leaders and the professional staff who will be facing those sticky accountability issues both today and in the near future. As is common with many case studies, close examination of a case can also result in identifying future research directions.

To that end, this study makes initial steps toward a conceptual model of process that can be useful for further studies. Although full model development is asking a great deal from a single case study, it should be possible to initiate construction of a model that provides support for cultivating a culture of accountability. In this respect, the case intends both practice-oriented and research-oriented results.

Research Questions

What are the organizing questions of this study? This case study looked for answers to the following questions:

1. How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?

2. What were the processes and decisions that contributed to the development of an assessment of this student critical thinking learning outcome?

3. What effects did these processes and decisions have on the institution during the time of the study?

4. What models of assessment, critical thinking, group work, and institutional self-renewal help us understand this case?

These four questions were chosen after considering the implications of the accountability story that introduces this chapter. If one measure of a successful university is the success of its students in the paths set out for them, then how that achievement is measured becomes very important. The College Portrait project (National Association of State Universities and Land-Grant Colleges, 2007), for example, recognizes and includes institution-specific sections as well as standardized comparable sections. Of the many ways to measure student success, assessments of student learning outcomes are gaining importance and becoming institutionally required.

In a strong institution, there should be a thread from university mission statements, to program goals, to expressed student learning outcomes and assessment which develops usable data that traces back along the same path in order to provide actionable evidence of the level of student success. These are the goals of new reforms in accreditation. This circular pattern is common to analysts (Birnbaum, 1998; Kezar, 2001, 2005, 2006, 2005a) as they review ways for universities to achieve ongoing, active self-

renewal. For this study, the focus is on critical thinking assessment of undergraduates as demonstrated in the synthesis course that is the final course in a general education program required of every student. The institution is Mid Atlantic State University, a pseudonym used for this study. Mid Atlantic State University (MASU or Mid Atlantic) enrolls about 18,000 undergraduates and has a doctoral level, research intensive Carnegie classification.

Is critical thinking an important element to measure? Several recent analyses indicate that it is. The national studies by the Association of American Universities (2003) and the Association of American Colleges and Universities (2005) as well as books like Donald's *Learning to Think* (2002) and Hersh and Merrow's *Declining by Degrees* (2005) point to critical thinking as a key to the cultivation of successful graduates.

Consistent with these analyses, it seems evident that the graduate desired by national and regional bodies and institutions is the critical thinker. Critical thinking is often expressed as an explicit educational goal for undergraduate students. For example, Harvard University, the University of Virginia, and Mid Atlantic all include fostering critical thinking by students in the university mission or goal statement.

Mid Atlantic includes the following in its mission statement. Mid Atlantic "will be an institution of international academic reputation providing superior education for students to develop critical, analytical, and imaginative thinking and to make well-founded ethical decisions (1991)" (MASU *Factbook*, 2006). [As will be the case

throughout this report, when the complete citation will identify the institution, the citation will be referred to by a general note in the text but not included in the bibliography.]

Students also report in MASU surveys on their opinions of critical thinking as a part of their degree programs. The evidence is that they too consider critical thinking to be a valued part of their education both as they graduate and five years after graduation. For example, the survey of graduating seniors in 2004 included the request to indicate the extent to which Mid Atlantic contributed to growth in critical thinking (web report on institutional assessment site). Students' mean response was 3.53 on a scale ranging from 1 (not at all) to 4 (very much). Alumni surveyed five years after graduation marked critical thinking ability 3.59 in importance and MASU's contribution to their ability as 3.32 on a four point scale, indicating a continuing value placed on critical thinking on the part of graduates.

Specific programs and courses at Mid Atlantic also include critical thinking components in their goals. For example, the Communication Department sets as a goal "critical thinking about the nature of theory as well as the validity and utility of communication theories." A recent review of university syllabi for the culminating course in general education showed that all those course goals included, at least implicitly, critical thinking aims (MASU Assessment Office Records, Synthesis syllabi 2005).

Looking at critical thinking from another direction, officials find that recommended goals for student success such as the AAU&C report (Conley, 2003) include critical thinking as an important factor. Professors themselves are also asking

questions related to critical thinking. In a review of critical thinking in practice, Phillips and Bond (2004) are concerned that what universities view as critical thinking and what is needed for graduates to be successful may not match. Their study of student experiences in a New Zealand second-tier course cast doubt on how effective the development of student critical thinking can be in traditional settings, and they ask for the university to provide a "proper account" (p. 293).

Both practitioners and researchers can profitably attend to critical thinking as an undergraduate goal and should realize that information about processes that address and track this student learning outcome responds to calls for accountability. An aggregated assessment of student levels of critical thinking might also be a proxy contributing to gauging the competency of their undergraduates and, thus, success for their institution.

Determining how to conduct such measures and how to make use of resultant information is a necessary goal for today's institutions. Understanding how such processes work and what results were obtained can be investigated through the research outlook of a focused case study.

General Setting and Researcher Statement

This case study takes place in a state university that began as a branch of an older institution and was founded as a separate institution in the early 1970's. The following information is from the MASU *Factbook* for 2006-2007. Presently there are 30,000 students, with about 18,000 undergraduates. As a suburban school, Mid Atlantic currently has more commuting undergraduates than residential students and includes students from 112 countries. In-state students compose 84% of the student body.

Additional information about the institution will be included in the section in chapter three on the research setting.

Knowledge of the researcher forms an important part of the understanding of the design, methods and interpretations in qualitative research methods like this case study. Merriam (2002) points out that "the researcher is the primary instrument for data collection and analysis" (p. 5), and thus the researcher should clarify assumptions, views, and theories at the outset (Merriam, 1998). Maxwell (2005) recommends that the research relationship should be discussed openly to clarify any ethical points or to address validity threats.

An open statement about point of view and involvement in the research is doubly important in the case of participant research where the investigator is directly involved in the activity under study, as is the case here. In addition, some background can illustrate point of view and shed additional light for readers. Sound advice from Glesne (1999) states that "awareness of your subjectivities can guide you to strategies to monitor those perspectives that might, as you analyze and write up your data, shape, skew, distort, construe, and misconstrue what you make of what you see and hear" (p. 109). She concludes with "A reflective section on who you are as a researcher and the lenses through which you view your work is now an expected part of qualitative research" (p.109).

The story which follows tells how I came to this study at this time, and is offered in order to provide the openness and reflection necessary for qualitative work. The lenses through which I see my work include belief in teaching and learning as a partnership -

students have a hard time learning if not taught. The second view I bring to my work is that of an organizer, one who believes that some order and rationale for initiating and working toward a goal will enhance success. Other operational outlooks include persistence and the desire to discover connections and explore new territories. The following mini-biography illustrates my way of working and knowing.

My first work was in the U.S. Agency for International Development as a program officer helping to manage foreign aid projects. I returned to this type of work later as director of training for the Peace Corps in the Czech Republic, and as leader of close-of service-seminars. This interest in how to manage and support projects has remained and also informed my work as a school team leader and department head during my ten years of K- 12 experience as a social studies, history, geography, and psychology teacher. I maintain a strong interest in making sure components are in place to allow my work or that of others to function smoothly toward successful goals. This principal interest can be applied to both managing projects and to the educational processes of teaching and learning. These interests continued as I shifted to eight years of adjunct history teaching in both a community college and a liberal arts university.

Some of the earliest conceptual and thought provoking ideas that came my way also have remained with me. For example, two books, *Metamagical Themas*, (Hofstadter, 1985) and *Consilience* (Wilson, 1998) both intrigued and challenged me. The common threads of both Wilson and Hofstadter were efforts to understand integration, the nature of thinking, creativity, and change. Both remain challenging works, spurring me forward as I developed skills and thinking abilities as an educator.

During time as a college instructor, I also worked as a learning specialist, helping college students discover what difficulties they were encountering and then finding a way to resolve problems so students reached success in class. As a teacher and instructor, I have always felt a duty to know and try many ways of teaching in order to reach many kinds of learners. One constant theme from these two books and my work with students was that learning takes place when ideas are integrated with the already known, when informed by multiple sources of information, and when relationships among the seemingly disparate parts fit together.

At present, my work is in the assessment office of Mid Atlantic State University and I have primary responsibility for the process of academic program review. However, prior to that work assignment, I was Director of Undergraduate Academic Programs for the arts and sciences college, and participated in the university's first efforts in discussion and faculty development related to critical thinking and undergraduate learning. In fact, as a K-12 teacher I had made use of critical thinking ideas and materials produced by a professor at Mid Atlantic.

When I was hired by the assessment office, my main duties were for academic program review. However, I also worked on the critical thinking assessment due to the previous work during a Teaching and Learning Center (TLC) internship. The assessment office practiced teamwork on all projects, so that all professionals were involved at some level in all projects. At first, I prepared materials and agendas for the faculty working group meetings, and worked with the director and associate director preparing for implementation. As time passed, my efforts increased to include preparing training for

raters, recruiting and meeting with professors, setting schedules for the observations, and collecting the data and starting the analysis. Just before the first pilot I asked the director how I should describe my role, and the response was that I was the "lead team member." I realized too that I was deeply involved, and that keeping a detached view for research might be very difficult. So it was at this time (early spring 2006) that I focused on following qualitative researcher practice to keep researcher memos, notes on meetings and on proposed ideas, as well as collecting as much data from meetings and planning sessions as I could so that later findings and interpretations would be as free from researcher bias as possible.

As I approached the decision point for the final stages of preparation in my Ph.D. program, Mid Atlantic State was also beginning a mandated assessment of critical thinking. It seemed natural to focus on critical thinking and its assessment in undergraduates for this research and dissertation. From the time I chose this topic and approach, I knew that a participant researcher focus offered both benefits and liabilities and I committed to search for both with equal strength.

However, I did not want to focus on the student results but on the process that a university goes through in carrying out an assessment of student learning. What elements are present, how do they fit together, to what degree is the process itself successful or beneficial? I felt that knowledge of these ingredients would make it easier to improve assessments, or replicate good processes.

There are advantages to association with the institution and people who became the participants. Knowledge of the people and organization forms a solid background

and enables a starting point closer to the research topic than if one begins as an outsider. In fact, when the focus is on a process, *post facto* investigation may not be able to retrieve the data in the same way as it developed. Dunleavy (2003) encourages research that departs from a strict pattern of what has gone before. Consistent with this orientation, I was very involved as a participant, and was able to plan, attend meetings, recruit the necessary faculty, offer workshops and training, compile and write up draft reports, and talk on a continuing basis with the administrator and faculty groups that became the source of my participants. In Glesne's (1999) terms, the rapport with participants was largely built during my first assignments at the institution, and contact and interaction with faculty and administrators continued and grew in my new assignment. Further discussion of participant research and how I monitored subjectivity will be in chapter three.

A participant research project faces disadvantages too, as discussed in every handbook on qualitative research in sections on ethics and validity (Creswell, 2003; Glesne, 1999; Maxwell, 2005; Merriam, 1998). The temptations of bias, of interpretation without data, and of pre-disposition to conclusions that are personally satisfactory are well known and are addressed again in the methods section of this paper. This project is one of immersion, observation, data collection, and reflection and all efforts are directed at integration which responds to needs for information related to accountability in higher education. It is my belief that by following well recognized principles of qualitative research, providing open statements about involvement, relying on data of many kinds,

and making careful observations and conclusions, I have completed useful work that can inspire further research.

Summary

In summary, there is a very strong push from governments and organizations outside higher education institutions for accountability, particularly in the area of student learning outcomes. There is also an internal interest in becoming and remaining successful as an institution, with a concomitant need for authentic information. Students themselves are interested in and value the education they have received. A case study approach to student learning assessment has both potential practical and research significance, and a deep participant research case will be informative for both research and practice paths. The next chapter will establish the conceptual framework for the study and supporting literature background.

2. Conceptual Framework and Background Literature

Introduction

This case study of the processes of a critical thinking assessment for undergraduates at Mid Atlantic State University brings together four conceptual areas: assessment, critical thinking, self-renewing organizations, and ideas concerning group work. The following discussion includes relevant concepts needed to unify the study of this particular assessment effort, not only in terms of what did happen, but also in terms of what it means to those concerned with responding to accountability demands from many sides. Potential outlines of both practical and conceptual models pointing to both research and applied results emerged. A background literature component is presented below for each of the four areas.

The themes expressed by Maxwell (2006) that literature reviews should focus on relevance to the specific topic at hand direct this discussion of concepts and literature. There are concerns among researchers about the purposes and requirements for literature reviews, and some advocate extensive reviews that synthesize a broad range of previous writing (Boote & Beile, 2005) and who also call for explicit criteria for inclusion or exclusion. The authors also address use of a review in topic based dissertations, such as this case study, concluding that the literature can successfully be interspersed throughout the dissertation. That is the case here, as additional chapters will also include relevant

work at appropriate times. Combining views of how to integrate literature into the study enables a stronger structure and is consistent with Maxwell's argument that "the primary purpose is not to summarize and synthesize some body of literature, but to use this literature to inform and support some decision or argument external to the review itself" (p. 29). The following discussion is organized around the four major sets of concepts supporting the case study: principles of strong assessments, critical thinking development as an institutional goal for undergraduates, group work processes, and the self-renewing university

Conceptual Foundation-Assessment

Assessment in higher education is in the spotlight, with a multitude of calls for accountability built on data from assessments, particularly of student learning outcomes. Principles of good assessment in higher education include clarification of purpose, direct measures of student learning, involvement of faculty, communication and use of results, and feasibility, which in many cases is related to the culture of the institution. Increased attention to principles of good assessment in higher education is evident from the increasing numbers of publications about assessments, such as those included in Banta's annotated bibliography (2007).

Exploration of the effects of an institution's culture is the focus of work by Khademien (2002), who addresses how culture affects what is possible and what does not happen as planned or ordered. Her conclusion is that the most important factor in success or non-success is often related to implicit culture, what "everyone knows" about how an institution operates. An implication is that administrators developing an assessment in

higher education will obtain better results if knowledge of the culture is considered. For example, Mid Atlantic's use of faculty working groups to develop the assessment may prove more beneficial than top down directives prepared by a staff office when the culture is heavily weighted toward faculty independence.

Terenzini (1989) advocates *Assessment with Open Eyes* as the tool to help scholars and professionals address the Who? What? and Why? questions associated with assessment. His diagram below reflects the many issues and purposes of assessment. Having clear aims for assessment is an important factor in its effects.

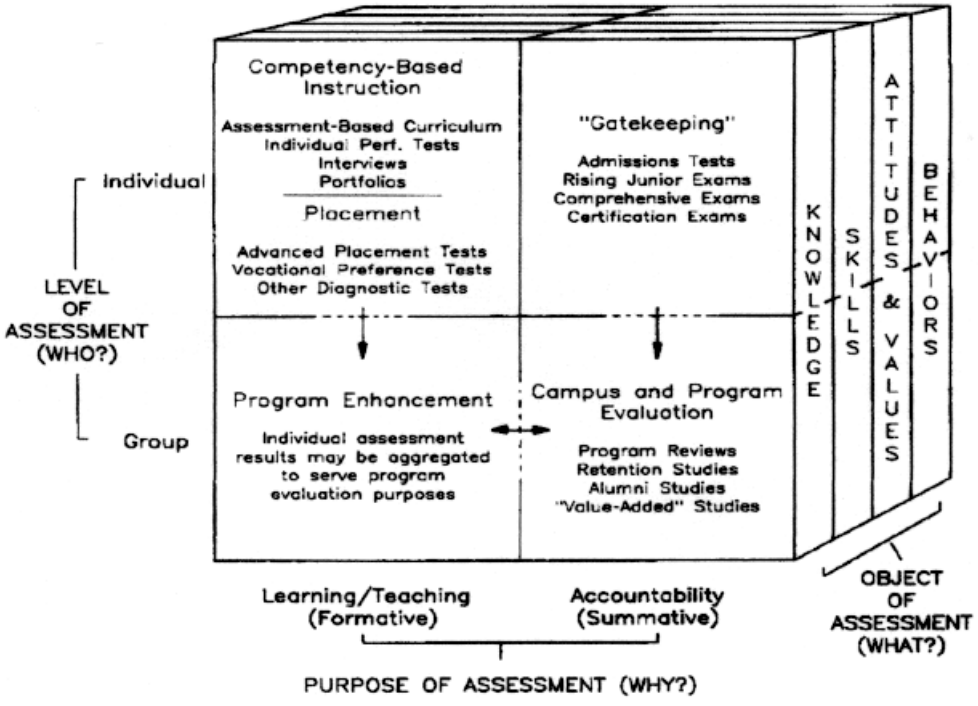


Figure 1. A taxonomy of assessment (Terenzini, 1989, p. 648).

In the case of the process under examination here, the blocks for Skills and Program Enhancement will be most relevant. The Campus and Program Evaluation blocks are also of interest, since an outcome of this assessment process might be changes to programs. As an institution is developing an assessment, it is important to note the what, who and whys, and Terenzini encourages establishing those anchors.

Hallmarks of effective higher education assessment practice offered by Banta (2004) include factors present in this study for planning, implementing, and sustaining. Banta's three groups of factors which can be used to analyze assessment are:

1. Planning: External Influences, Engaging Stakeholders, Focus on Goals, Developing a Plan, Time
2. Implementing: Methods, Faculty Development, Leadership
3. Sustaining: Interpreting Findings, Reporting Results, Using Results, Recognizing Success, Improving Assessment

These marks of good assessment will be used to evaluate the assessment effort.

Additional emphasis on the importance of assessment as a developer of usable knowledge is provided by consideration of assessment as an agent of change (Ewell, 1988; 2004). Assessment is important, but *good* assessment is crucial for positive use of knowledge about student outcomes by faculty, programs, and institutions facing the growing web of accountability. The detailed picture of assessment from the National Research Council (Pellegrino, Chudowsky, & Glaser, 2001) reinforces the need for clear goals for any particular assessment, as well as offering examples of types of assessment research, including the use of scoring rubrics (see pp. 112-120). However, the major

portion of this work relates to testing, standard quantitative methods, and establishing quantitative types of validity. The report does offer a specific guide – referred to as the assessment triangle.

> The process of reasoning from evidence can be portrayed as a triangle, referred to throughout this report as the *assessment triangle*. As shown below, the corners of the triangle represent three key elements that underlie any assessment: (1) a model of student *cognition* and learning in the domain, (2) a set of beliefs about the kinds of *observations* that will provide evidence of students' competencies, and (3) an interpretation process for making sense of the evidence. (p. 296)

These criteria will also provide a useful look at the assessment carried out in this study. Both this system and Banta's are supported by the collection of institutional good practices presented by Bresciani (2006).

Kasworm (2004) addresses some crucial issues not directly considered in the previous works. These include collaboration with faculty and commitments to understanding and valuing assessment on the part of faculty. In any project where faculty input will be a key factor, how that collaboration develops is central to achieving lasting effects. New work (Kezar, 2006) goes beyond the collaboration of faculty on any one issue, on toward the larger issue of how higher education institutions can redesign for institution-wide collaboration that affects student learning and highlights elements of the change process. This change will mean, Kezar suggests, that rather than collaboration as a result of individual leaders at individual times, a *culture of collaboration* will foster

greater improvements in student learning. The key will be to know how to foster that culture of collaboration.

Additional works of practical help for higher education include Suskie (2004), Walvoord (2004), and Angelo and Cross (1993). Important elements in all of these works are consideration of what is feasible in a given context and whether the results will be useful at the conclusion of the assessment. Angelo and Cross address classroom level assessment techniques; nonetheless, many examples can be expanded to course or program assessment. Walvoord encourages all assessment of student learning to be faculty-based and faculty-supported, and provides an outline for working with faculty on assessments of student learning through course-embedded measurement. She cautions that assessment must complete the assessment circle - set student goals, gather direct and indirect measures, and use the information for improvement. Also notable are the works of Nitko (2004) and Wiggins and McTighe (2005). Although not directed at higher education, their techniques for designing assessments for the K-12 environment have been adapted to good effect during consultation given to programs in academic program review at MASU.

Assessment is a complicated process. The choices made and the philosophy behind those choices will influence the successful creation of usable data that can lead to improvement and change. This is the kind of outlook promoted by Huba and Freed (2000), one they hope will encourage more collaboration, discussion, and preparation to teach differently. Assessment can be a powerful tool — or not. Assessment results that

sit on a shelf are not activities that should be given the scarce time and resources of an institution's faculty and staff.

Conceptual Foundations - Critical Thinking

Although the focus of this study is not primarily on critical thinking measurement *per se*, consideration of this topic as part of the case study is important. Critical thinking by students is widely considered to be a crucial part of a university education. What this particular habit of mind or skill means is less than universally defined. Dissent also exists on teachability and whether it is the same across disciplines. Since one of the tasks of the MASU working group was to develop a way to carry out the assessment, they first needed to decide on a definition, which was no easy task. There are models of critical thinking that picture a separate discipline, capable of being taught as a "subject" (Paul & Elder, 2001, 2003). Those of us who were teaching in the 1970-80's recall efforts to teach critical thinking classes in high schools as well as detailed how-to instruction from advocates like Beyer (1987). Plans for teaching a separate topic still exist, for example, in the Master Teacher Program which includes direction for teachers (http://www.masterteacherprogram.com/about/index.html) and is used at Georgia State, among other universities.

A variety of colleges have developed home-grown projects with varying definitions of critical thinking. These colleges include Tufts University (http://ase.tufts.edu/criticalThinking/), Washington State University (http://wsuctproject.wsu.edu/ph.htm), Texas State Technical College Harlingen (Bauer,

2004), and Miami University

(http://www.units.muohio.edu/led/Assessments/criticalthinking/).

At the other end of a conceptual spectrum is the model of critical thinking as being so discipline-specific that it is only visible within that discipline (Donald, 2002). Donald represents a group that believes critical thinking cannot be studied or evaluated either outside of disciplines or across them. Her detailed analysis of different disciplines and their contexts for critical thinking are directed more toward individualized programmatic examination than university-wide competency assessment.

As the California Commission on Teacher Credentialing (Paul, Elder, & Bartell, 1997) puts it, "no single definition of critical thinking will do" (p. 4). That study did find eleven common threads, comprising both the essence of critical thinking and its beneficial effects, for example:

> …that as one becomes proficient in critical thinking one becomes more proficient in using and assessing goals and purposes, questions and problems, information and data, conclusions and interpretations, concepts and theoretical constructs, assumptions and presuppositions, implications and consequences, and points of view and frames of reference (p.118).

In a more general work on critical thinking, Brookfield (1987) offers a succinct view of critical thinking as alternating phases of analysis and action.

One review of the literature on critical thinking in college students (Jones & Ratcliff, 1993) suggests that there is general agreement that critical thinking includes analysis, evaluation and inference, meta-cognition, and self-monitoring. However, that

study also highlights the debate on whether critical thinking is specific to disciplines or generalizable across disciplines and what methods measure it. Ewell (2004) suggests there is a need to get thinking about critical thinking concretized. As an institution talks about its students, it must be able to "succinctly describe exactly what they were asked to do and how well they performed" (p. 7). The difficulty of trying to decide just what is "critical thinking" as it is concretized is illustrated in a University of California San Diego's freshman course on Dimensions of Culture, designed to stimulate students' critical thinking. As described in the electronic newsletter *Inside Higher Ed* (May 3, 2007, http://insidehighered.com/news/2007/05/03/uscd), one side says the course is just a left-wing presentation, while others believe the changes made to address that concern are now too far right and doctrinaire. Another issue reported by Halx and Reybold (2006) is the perception of the faculty on defining critical thinking and the capacity for it in undergraduate students. Their study concludes that faculty believe it is a learned skill, and that they teach it implicitly according to their definitions. To teach critical thinking more explicitly, faculty should "model it, as critical reflection is a common characteristic of successful educations. How do we operationalize our definitions of critical thinking *and* critical teaching?" (p. 314).

Necessarily, one of the first steps in an institution-wide assessment of critical thinking will be selecting the operational definition and choosing a place on the continuum so that assessments across courses will be possible. The operational definition for the particular assessment must be clear, especially when there are multiple possibilities (Palomba & Banta, 1999). In order for an institution to look at critical

thinking across disciplines while necessarily examining discipline-related courses, a definition somewhere in the center is required.

This middle model, adopted by the Mid Atlantic working group, was one previously developed by a group of forty-six American and Canadian professors from multiple disciplines. This model defines critical thinking as skills that cross disciplines and have specific characteristics. This definition was one of several discussed as part of the initial workshop planning for critical thinking assessment at Mid-Atlantic State University (Assessment Records, Critical Thinking Workshop file, 2006). The original presentation of the international group has been updated, nevertheless retaining the same definition from 1998 (Facione, 2006).

> We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based. (p. 20)

Palomba and Banta (1999) used this same definition as their example of consensus needed for assessment of general education. This concept of critical thinking became the base from which the assessment at MASU developed.

Conceptual Foundations - Group Work Processes

Perhaps the most significant elements in this study are concepts of group work processes. How does such work take place, and what factors influence its success? Ideas related to beginning and sustaining group work are the conceptual heart of this study. For this case study, the term group work is used rather than committee work. Jennings

(2007) quotes Richard Long Harkness: "What is a committee? A group of the unwilling, picked from the unfit, to do the unnecessary" (p. 93). An even stronger statement of faculty attitude toward committees comes from a professor at the Merrick School of Business, University of Baltimore: "A committee is a *cul-de sac* down which ideas are lured and quietly strangled" (http://home.ubalt.edu/ntbarsh/index.html). In fairness, he continues, "The greatest things are often accomplished by individual people, not by committees. What does it mean to say that committees might have responsibility? Committees cannot have a responsibility any more than the business can. The only entities that can have responsibilities are people." The study of assessment at Mid Atlantic looks in that direction - what do people do when conducting an assessment?

Jennings distinguishes committees from task forces by noting that committees focus on on-going general tasks, whereas task forces focus on more specific charges, and are disbanded once the goal is reached. Work groups probably fall in between as they focus on a specific but continuing, if not continuous, task. Committees are very structured, often permanent, and have a wide-spread reputation of deflecting or defeating action. Any groups, by whatever name, can become stymied in their work, or can out-perform the sum of individuals in the groups, often depending on the leader's skills (Rainey, 2003). The key ingredients to success may be how processes allow "coming to agreement on the final set of goals" (p. 347) as well as the possibility for members to contribute, to operate without digression, to turn conflict into construction, to evaluate possibilities and to reach a conclusion.

Both Rainey (2003)and Khademian (2002) point out the importance of the culture in which the group operates. Culture can affect what is possible or what does not happen as planned or ordered. Khademian's case studies show that the most important factor is often related to implicit culture. For the kind of process under study, an implication is that administrators developing an assessment in higher education will obtain better results for improvement if the culture is employed rather than ignored. For example, use of faculty working groups to develop the assessment may prove more beneficial than top down directives prepared by a staff office when the culture is heavily weighted toward faculty independence, as is the case at MASU.

One issue Jennings (2007) finds with groups in schools that may not repeat in university groups is lack of collaboration skills leading to what might be called *followership*. "They are frequently too quick to resolve differences by taking a vote or simply following their assigned 'leader' " (p. 97). University faculty may not possess collaborative skills, as Birnbaum (1998) suggests in his discussion of the collegial institution, however, that lack seldom means there is a concomitant disposition to follow the leader in a group of university faculty.

For self-managing groups, such as the ones in this study, Jennings offers three sets of factors affecting success, as summarized below:

1. *Quality of the group design*: variety of skill and talent required, tangibility of outcomes, effects of outcomes on those outside group, autonomy, feedback on outcomes

2. *Competency of group members*: high levels of task relevant experience, only the number necessary for the task, at least moderate levels of collaborative skill, balance of homogeneity and heterogeneity

3. *Context provided by administrator in charge*: requirements, constraints, material resource

4. *Availability of rewards*: tangible or intangible rewards are offered for those in the group

A very common conceptual image of group work from 1965 relates to stages of group development that are forming, storming, norming and performing (Tuckman). However, this version may not prove as useful a concept because of the limited face-to-face group work in this study, as will be seen in chapter four.

A more useful set of ideas for the case at hand comes from analysis of collaboration on campuses (Kezar, 2006). Although intended to apply to the whole institution, some of the topics raised can apply downward to this case study. For example, the following summarized elements can also apply to collaboration in group work:

1. Decision on what is to be accomplished
2. Examination of processes
3. Rewards for incentives and accountability
4. Developmental and professional support

Brazer and Peters (2007) offer an additional view of collaboration and re-enforce the necessity of such skills for those working in groups. They observed that even though

the collaborative nature of the group was very apparent, the group also included recognition of categories of influence. Group members themselves recognized some were more influential than others, either for reason of official position or expertise; however, this recognition did not necessarily stop collaboration.

Kayser (1995) looked at teams and addresses many of the same concerns as Jennings (2007) and others discussed previously. Kayser presents a structured look at how teams in work situations perform depending on their collaboration and facilitation. The study, resulting from his own participant research, is directed to team leaders, and his 14 collaborative principles form concepts applicable to the processes that will be observed in this study. The following is a summary of his principles:

1. Primary facilitation comes from the convener of the group
2. Members must be secondary facilitators
3. Primary facilitator can switch hats as a content contributor
4. Command and Control will not work
5. Sessions must have purpose, outcomes, agenda, record keeping, time limits
6. Only smallest number of members necessary are invited
7. Plan a roadmap in advance
8. Conduct process check at ends of session
9. Establish process ground rules
10. Permit emotional involvement of members
11. Set stage for open expression
12. Expect support for collaborative efforts from members

13. Set rules of handling disruptions

14. Encourage constructive conflict

Because these principles also include the concepts of previously discussed writers, they will be useful in examination of work processes. What is not so clearly indicated in Kayser's model is the between-session work that may take place. That topic will need to be addressed separately.

The less than steady progress of working groups is a concept that is more fully developed in Gersick's (1998) model of group processes, which includes the concept of "punctuated equilibrium". In this view, groups work on similar categories of tasks, but not in a linear or hierarchical order, as implied by Tuckman (1965). In addition, progress is not steady and there may be periods of apparent inactivity followed by bursts of work. These are comparable, he says, to bursts of evolutionary change. The bursts may be influenced by factors that are either part of the internal environment or the external environment. The idea of punctuated equilibrium has appeal as a tool of analysis for work done in universities. Rainey (2003) follows these ideas by focusing on how organizational culture, leadership, and relationship to internal and external environments affect decision possibilities and implementation. Certainly the internal and external contexts for this assessment process will be important for why Mid Atlantic State University initiated the process in the first place and how it was carried out. Examination of the data in this case study will be more meaningful when compared to Jennings', Kaysar's and Gersick's ideas on what should and what does happen during group work. The following chart illustrates common elements emphasized by authors discussed here.

Table 1

Cross Index: Elements of Successful Group Work

Authors	Rainey	Khademian	Jennings	Gersick	Tuckman	Kayser	Kezar	Brazer & Peters
Elements								
1. Leaders skill	x					x		
2. Goal agreement	x	x	x			x	x	
3. Members' contributions	x							
4. Digression avoidance	x							
5. Conflict resolves	x				x	x		
6. Evaluation of choices	x							
7. Conclusions reached	x				x			
8. Importance of culture & environment		x		x				
9.Task success		x						
10.Collaboration			x			x		x
11.Design of the group			x		x			
12. Member competencies			x			x		x
13. Provision of context			x				x	
14. Organized processes				x		x	x	
15. Rewards		x					x	
16. Accountability			x				x	
17. Re-start if progress interrupted				x				

The table demonstrates that there are many elements that affect how groups work, and whether they successfully conclude their mission. As the case of Mid Atlantic State unfolds, this variety of ideas will prove useful in making sense of the processes. Although they focus on many different elements, the chart summary provides tools to use in examination of the Mid Atlantic State University's assessment process.

Even though the now classic look at group work in the analysis of the Cuban missile crisis by Allison and Zelikow (1999) focused primarily on decision making, a

chief element of their work was examination of recollections of participants in groups whose advice contributed to the final decision. The story of the missile crisis lends informative examples of work with groups to form action plans, then carry them out. A great deal happens between these two statements and the authors present a detailed example that emphasizes the same events looked at in multiple ways. They also took account of the pitfalls of *post-facto* recollection.

Group work concepts reviewed here not only set directions for the Mid Atlantic case, but also contributed to analysis of findings. The fourth concept area continues looking at how the assessment was conducted, using an institutional viewpoint.

Conceptual Foundations - The Self- Renewing University

In addition to assessment requirements and increasing attention to accountability, concepts of institutional self-renewal comprise the fourth part of the conceptual framework for this study. A self-renewing institution is an institution that functions so that change and improvements are continuous and contribute to viability over time. The data that come from informative assessment can become formative assessment for the institution. Change does not necessarily need to be preceded by crisis but can be intentionally built into institutional practice.

Therefore, another important reason for undertaking this study, beyond those associated with compliance with state requirements, is derived from potential intrinsic benefits to the institution as a whole. Does the process itself benefit the institution? Two concepts will assist examination of that question.

In examining the purposes of assessment, Miller (2007) reviews the background for current work on what might be called the learning university. Contributions from the areas of business and organizational learning (Drucker, 1999; Schein, 1992; Senge, 1990) are having effects in higher education. Miller says "More and more, senior leaders use assessment to ensure that organizational culture is supportive to organizational learning and innovation" and that "learning organizations are systemic, cooperative, and creative compared to traditional organizations, which are fragmented, competitive, and reactive" (p. 25). A further developed view comes from work directly related to higher education institutions particularly Birnbaum's (1998) concept of a self-renewing institution. In his terms a self-renewing or "cybernetic" institution is capable of making self-corrections through active feedback loops. The term cybernetic seems to hint at a very robotic kind of institution. However, his concept does link well to the third tool of good assessment which is making sense of evidence and using it for improvement.

Similar concepts are repeated in Kezar's *Organizational Learning in Higher Education* (2005a) although Kezar focuses on less mechanical collaboration as the driving force for organizational learning. Another call for continuous improvement comes from the Association of American Colleges and Universities (2005) asking institutions to model lifelong learning for students. Findings from the Mid Atlantic case study regarding intended and unintended effects of this assessment process were evaluated using these concepts of active feedback loops, collaboration and institutional learning.

This second support for the university as a self-renewing institution is from concepts of collaborative faculty involved in change (Kezar, 2001, 2005, 2006, 2005a; Leiberman, 2005). A concept of engaged faculty, supported by an institution that encourages collaboration and change is an important new look at institutional characteristics. Kezar advocates organizational change so that cross-divisional and group work is supported by "structure, process, people, and rewards" (2006, p. 809). She characterizes such change as being adaptive/generative, intentional, pro-and re-active, active and static, and affecting both processes and outcomes (2001).

Another study (Briggs, Stark, & Rowland-Poplawski, 2003) of university departments cited for continuous planning identifies four criteria that may extrapolate to a continuous planning university. These four are:

1. *Continuous and Frequent Circular Planning Processes*: A continuous planning department gives frequent attention to appraising the curriculum for renewal and redirection and engages in on-going planning efforts. It uses organizational structures and processes that facilitate curriculum planning as an on-going routine for renewal and redirection

2. *Awareness and Responsiveness*: A continuous planning department is attuned to and responsive to internal and external factors that may influence curriculum and is proactive with respect to future influences.

3. *Participation and Teamwork*: A continuous planning department maintains a high level of faculty involvement in curriculum issues.

4. *Use of Evaluation for Adaptive Change*: A continuous planning department gathers and uses relevant information about program successes and failures in the curriculum development process (p. 367, Table 3).

A 2007 discussion of managing for innovation (Glickman & White, 2007) includes innovation in processes, such as ones that could respond to the Spellings report's call for new ways to present student learning outcomes and measure student learning. When innovations are presented to the institutional community, there are barriers toward success. If the processes of good and effective assessment are considered to be innovation, then looking at the way innovations become adapted and adopted is informative. Rogers' (2003) classic work lists factors affecting the initiation and implementation of innovation as a process. He includes setting the agenda, matching the problem and its solution/innovation, redefining/restructuring/clarifying and then "routinizing" as the change becomes a continuing part of the organization (See Rogers, 2003, p. 421, Figure 10-3.).

How might this connection of assessment to self-renewal and innovation best be described or understood? Uses made of assessment results are one of the cornerstones of best practices as well as theory, and are comparable to Birnbaum's (1998) feedback loops, which he calls cybernetic controls, "self-correcting mechanisms that monitor organizational feedback functions and provide attention cues, or negative feedback, to participants when things are not going well" (p. 179). These feedback loops instigate responses creating course correction. This concept can work in large or small ways but is highly dependent on collected data to trigger such effects. Effective data is the element

that will contribute to non-crisis self-renewal of educational institutions. This careful look at this Mid Atlantic assessment process can reveal the kinds of circular links, collaboration, and change processes discussed here.

In the Mid Atlantic case, the multiple threads of analysis as discussed in this chapter are important, and since the study takes place in real time, the employment of a participant researcher model adds yet another tool. Certainly this effort does not duplicate high-level government and world crisis analysis, but results from this study do begin to show outlines of another model presented in chapter five that could prove useful beyond the confines of the case study. Application of the four concept elements—assessment, critical thinking, self-renewing institutions, and group work—requires a strong focus on the research questions and how answers to them might lead to new directions for higher education institutions in a warming climate of accountability.

These concepts form the backbone of the following chapters which present the context of the study and the methods used, information resulting from the case study, and ideas about future directions. The concepts described here contributed to making sense of the various data involved in case study methodology.

3. Design, Setting, and Methods

Integration through Case Study Design

This study addresses integration of knowledge about supporting student learning through assessment best practice and goals for potential modeling of future research. Careful design, attention to context, and explicit methodology support that integration. The methods used for this study are described in three parts: (1) the case study design, (2) the setting and context, and (3) the methods employed. A participant researcher case study design enabled the observation, recording, and integration of the conceptual elements described in the previous chapter as related to Mid Atlantic State University's assessment of critical thinking. Looking at student course work from this perspective allows a university to learn about student learning outcomes across programs (Palomba & Banta, 1999). Opportunities to learn how a process actually unfolds and to consider its wider implications are well supported by the methods of a case study. The design deliberately followed all three aspects of Merriam's (1998) definition:

> Case studies can be defined in terms of the process of conducting the inquiry (that is, as case study research), the bounded system or unit of analysis selected for study (that is, the case), or the product, the end report of a case investigation. (p. 44)

The research design was also informed by thinking of a process evaluation of work teams (Gersick, 1989, 1991, 1998). Observation and data collection in her studies focused on transition from discussion to action and on models that reject a simple linear progression. Gersick expands understanding of the complexity of work in groups and then identifies some commonalities. "All project groups are challenged to choose boundaries, norms and work methods, but they vary in the sequence and manner in which they settle those choices " (p.16, 1998). Her application of the theory of "punctuated equilibrium" from the field of evolutionary biology to the field of organizational behavior was useful in evaluating data from this study. Punctuated equilibrium explains work that varies from low levels of activity (equilibrium) to work activity bursts that appear (punctuate) on an irregular basis. These bursts are comparable, she says, to the bursts of evolutionary change described in the realm of natural science and may be influenced by factors that are either part of the internal environment or the external environment. The theory also proposed that groups work on similar categories of tasks, but not in a linear or hierarchical order, and was applied to groups in a face-to-face setting. These interesting concepts facilitated interpretation of results related to the two groups who worked on the critical thinking assessment at MASU.

Basic Parameters

This study was not intended to be a report on what students know about critical thinking; rather the goal was to explore evidence related to conceptual models that might lead to best practice recommendations for the processes of assessment and to potential modeling of processes. In order to reach these goals, the concepts discussed in chapter

two and as shown below were used to analyze and interpret the findings. For example, are there effects of a locally developed instrument that result in changes to teaching (Kozlowski, 2000), and was there evidence of punctuated equilibrium (Gersick, 1991)?

Time parameters are also important to establish. The study covers the time period from the fall 2005 semester through February 2008. Early information from the start of planning to initiation of the working group in December 2005 came from records of the assessment office, MASU publications, and websites. Original data were collected from participants and use of records continued between December 2006 and February 2008, the end term of the study. One change was made from the original plan for collection of data. Because the two assessment administrators also functioned as raters in the assessment process, there was only one interview. That interview near the end of the processes covered information not obtained from their rater questionnaires; therefore the protocols for the two interviews were combined. All data collection instruments and protocols are in Appendix C.

The 2002 state council's mandate for planning and conducting competency assessments set the order in which competencies were measured, so that the critical thinking assessment planning did not begin until 2005 (MASU records, Critical thinking Files, 2005-6).

Figure 2 presents a chronologic summary of data collection and research planning.

December 2005→2006 — January 2007

• Participation in working group • Collection of meeting records	• Review of records • Provost interview • Collection of related documents	• Learned software (NVivo & Digital Voice Editor) • Planned and carried out pilot	• Continued participation in working group • Collection of meeting records, feedback sheets from workshops, rater training & raters continues	• Prepared reports on pilot project • Questionnaire I to first group of participants

February 2007 — February 2008

• Participation in analysis of pilot • Collection of records continues • Began planning second assessment	• Implementation of second iteration • Analysis of research data ongoing • Questionnaire II to second group of participants	• Administrative Interviews *NOTE:* *2 planned interviews of assessment administrators combined into one because of their responses through questionnaires* • Consultation with Washington State University • Data Analysis continues	• Focus group for debriefing, and review of preliminary results • Completed data analysis	• Writing in progress • Continuing analysis • Findings and implications developed

Figure 2. Research chronology.

Organization, implementation, and data management were aspects of the case study methodology that were informed by specialized researchers. Maxwell (2005), Glesne (1999), Merriam (1998) and Yin (2003) provided direction for best practices in qualitative case study. Concepts common and important to these writers are ethics and researcher relationship; field notes or researcher memos; data collection, management and interpretation; internal validity; and development of conclusions. Among the specific techniques for this kind of case study that all three writers included are participant observations through notes, questions and interviews of participants, and the use of archival materials such as the assessment office meeting and workshop records, reports, planning notes and feedback given to that office on its activities. The protocol for use of that information is in Appendix D.

Additional guides for design include the Interactive Model of Design (Maxwell, 2005), which presents a non-linear, interconnected model of the kind of study used in this case. Maxwell's model emphasizes understanding how research design elements interact with one another and it includes inter-active elements of purpose, conceptual context, methods, validity, and research questions (p.5). Merriman's (2002) discussion of interpretative studies provided the analysis method. Interpretative studies look toward explanation and analysis of a specific instance that can inform larger views. Glesne (1999) discusses interpretive research as well, and supports the choice of a participant researcher design and the importance of context. Her work also supports the concept that participant research designs possess the potential to develop a hypothesis or theory. The design structure for this research is well supported by these three experts.

Yin (2003) included analytic techniques such as the possibility of quantifying some of the data, and determination in advance of which concepts will be used in analysis. In this study four conceptual models were directly employed in the analysis, and some information has been quantified. As do all of these authors, he emphasizes the value of well-structured and well-conducted qualitative research that results in meaningful and informative *explanatory* studies, which in turn may inform broader generalizations. As the design and work of this study progressed, it became important to attend to Stake's (2005) reminder to not overextend the reach of conclusions or findings from a case study. Where synthesis of information to something new crosses the boundary to over-extension is a difficult spot to discern and deserves close consideration. However, a main aim of many qualitative studies is to inform the course of new research direction and to suggest lessons that may be learned from the case. Perhaps the tensions between learning from a case study and concerns about over-extension are themselves the beginning of a research direction flowing from an explanatory case study on a significant topic, such as accountability of higher education institutions for student competency.

This case study used tools from assessment, critical thinking, group work analysis, and self-renewing institutions in an effort to obtain explanatory goals, while also respecting the guidance of expert research methodologists in an effort to avoid overextension of conclusions. Use of multiple tools applied to multiple types of data supported initiation of a new model and suggestions for best practice, which are presented in chapter five.

The National Research Council's assessment triangle (2001) components are evident throughout the study as overarching themes. The three components are:

1. *Cognition* - what are the models of learning and research methods in the study
2. *Observation* - collection of evidence
3. *Interpretation* - processes for making sense of the evidence

In summary, this study used a mixed-methods, primarily qualitative, participant researcher case study design, and is similar to new work from Kezar (2006) who studied collaboration in four institutions and employed the same elements used here: interviews, records examinations, and questionnaires from both faculty and administrators. Merriam (1998) and Stake (2005) also employed these types of data sources. Although the MASU study was informed by a set of guiding concepts, the study also relied on data-supported analysis and interpretation, and carefully heeded Creswell's (2003) reminder to pay close attention to assumptions that may color and shape results. The use of multiple types of data, careful attention to respondents' actual words, and reflective researcher notes and memos were some ways this advice was heeded.

The study was also informed by a variation of peer auditing or feedback recommended by Creswell (1998) and Maxwell (2005) that took place through consultation with faculty and staff at Washington State University (WSU) September 10, 2007. The opportunity to visit that university resulted from a dissertation completion grant and there was not time to add them as participants. Consequently only a general summary of those discussions is reported and no identifications are given. The staff who

directed that project and four faculty members were interviewed individually, in person, over the course of an afternoon. In a series of thirty-minute blocks, they were presented with a description of the MASU project, the research design, and a summary of what were at that time emerging themes of faculty interest in the topic of critical thinking, cooperation, and emerging consensus that critical thinking could be embedded in courses. They were asked to evaluate emerging results as compared to WSU's completed process. Responses indicated that this study matched themes concerning interest, cooperation and embedding critical thinking practice in courses. Because the WSU project was a one-on-one faculty development effort, they did not have many comments on the group work elements, although three expressed a desire for more interaction with other faculty on the topic. They were very interested in the rubric and idea of looking at student presentations, and reported that the WSU project work had been expanded to additional state higher education institutions. They stated the belief that the impact of research about assessment work in general could have wide impact (Researcher Notes, September 10, 2007). Like the MASU participants, faculty at WSU also expressed some concern about time it takes to plan and carry out new teaching ideas. No negative comments were given about the design or emerging themes. This consultation added to confidence in the design and emerging results.

The methodological rationale for the MASU study is well summarized by Schwartzmann (1989), who supports a participant research approach to studying actively functioning work groups. An anthropologic, immersive outlook, he suggests, allows a researcher to find the meaning behind group work processes and products. In the Mid

Atlantic case, the researcher was the lead member of the assessment team and work group. Consequently it was important to practice reflection during the study so that, among other dangers, overtly directive actions resulting from that role were reduced. In one such reflection, the MASU research reminded me of the tensions faced by the intrepid Captain Kirk. His prime directive to not interfere often clashed with his responsibilities for completing a task.

Best practices for designing participant researcher case studies were intentionally applied to the extent possible. These practices were enhanced by the analysis journey, which focused on describing and comprehending *processes* that might offer guidance for assessment officials looking at learning outcomes in the current environment of institutional accountability. It was anticipated that this approach would yield results of considerable significance for individuals concerned with institutional effectiveness.

Practical and Conceptual Significance

The MASU study addresses what the *Beyond Dead Reckoning* report (National Center for Postsecondary Improvement, 2002) describes as Priority Number 1: Improving Educational Quality and Institutional Performance. Particularly, administrators taking note of the growing interest in assessment of student learning outcomes will need to know how such assessment processes work in practice. A deep understanding of the MASU experience can point to elements of best practice and can support model building. Use of the methods articulated here should facilitate moving from an empirical study toward larger applications and inspiring new concepts. Taking advice from *The Craft of Research* (Booth, Colomb, & Williams, 2003), to use a logical chain to state research

aims, I use the following chain to lead to significance that is both practical and conceptual. Accordingly, my study aims can be described using the sample chain they suggest:

I am studying:

- The way a large public university assesses undergraduate critical thinking in synthesis courses at the end of the general education curriculum;
- In order to discover and interpret how the assessment was carried out, and its meaning for those developing and implementing assessment leading to accountable student learning outcomes;
- Because this study will be useful and informative to them and potentially to other institutions as requirements and desire for data on student learning outcomes increase and become necessary to support institutional change and renewal.
- key words are underlined)

Understanding the process cannot be optional. Otherwise, perceived conclusions are merely serendipity, a happy accident. Terenzini's (1989) diagram (See Figure 1 in chapter two) helps place the assessment in context. In his terms, this assessment has as its object the critical thinking skills of students, which are aggregated for potential program enhancement and evaluation. Addition of the self-renewing concepts of Birnbaum (1998) to this analysis frame yields an approach that effectively addresses major concerns of universities dealing with assessment of student learning outcomes. He envisions successful institutions as having the capacity to know where and what

improvements are needed. Careful assessments can become the "How" factor for improvement.

Context and Institutional Setting

The specific context and setting for the research effort form the world in which the investigation took place. Therefore it is important to set the context clearly and to recall the discussion of assessment issues related to critical thinking as presented in chapter two. Ennis (1993) describes the variety of ways that critical thinking might be assessed, and the Washington State University (http://wsuctproject.wsu.edu/) and Miami University (http://www.units.muohio.edu/led/Assessment/criticalthinking/) critical thinking projects provide examples of university projects that used faculty-developed assessment and standards as did MASU. Mid Atlantic State University developed its assessment informed by a review of these and other examples (Assessment Office Records, Intern Report and CT Workshop files, 2006).

Factual information and quotations in this section come from the 2006-2007 MASU *Factbook*, reports from institutional research and assessment, and the website published by the university. As stated previously, exact citations are not included in order to preserve the anonymity of the institution and confidentiality of participants. The researcher may be contacted regarding this reporting decision, based on consultation with practicing researchers.

Mid Atlantic State University is a state university, located in a suburban setting of a major metropolitan area. The university offers 64 undergraduate, 69 Masters, 25 Doctoral, and two Law degrees to approximately 18,000 undergraduates, 9,000 graduate

students and 3,000 non-degree students. About 7,000 undergraduate students are minority or international students. Out-of-state students comprise about 17 percent of the student population, approximately 55 percent of all students are fulltime, 55 percent are female and currently about 4,000 students live on campus. The six-year undergraduate graduation rate is 55 percent. MASU is a Doctoral/Research Intensive university in the Carnegie classification system.

It is important to note that both faculty and administrators see MASU, a relatively young university, as an institution characterized by a spirit ready to take advantage of opportunities and to attempt to create such opportunities in order to enhance teaching, research and growth. For a university that moved to being a separate state institution less than forty years ago, MASU presents this cultural view through administrators who encourage new ideas, attempt to provide resources whenever possible, and see their role as doing a lot with a little. The atmosphere of entrepreneurship conveys a willingness to change which may not be common among public universities in general.

The university mission begins with the statement that it will be an institution "providing a superior education enabling students to develop critical, analytical, and imaginative thinking and to make well founded ethical decisions" (*Factbook, 2007*). This mission is strongly reflected in the undergraduate general education program required of all undergraduate students, where the first goal is "to ensure that all undergraduates develop skills in information gathering, written and oral communication, and analytical and quantitative reasoning" (MASU website Catalog, retrieved December 2006). The final requirement in general education is completion of a synthesis course, designed to

demonstrate students' integration of the general education curriculum and achievement of general education goals. Many of those courses include culminating writing assignments, projects, research, or presentations.

When the State Education Council (SEC) in 2002 began requiring reports on student competency in critical thinking and five other areas, the Provost assigned that responsibility to the university assessment office. This office customarily followed a process of faculty involvement and support for developing other competency assessments, and did so for the critical thinking requirement beginning in 2005. The study follows the decision of how to proceed with the assessment, the formation of the Working Group and the work and processes of the working group and the assessment office for implementation, revision, and second implementation of the assessment. Further developments in 2007 related to the critical thinking assessment process include changes to competency assessment requirements and to the start of the Provost's initiative for Critical Thinking in the Curriculum (CTC). Both of these topics will be addressed in chapter four.

Internal and external stakeholders

There are several groups of stakeholders. Chief among them is the external instigator of the process, the State Education Council (SEC), which is charged by the state legislature with oversight of public higher education institutions. In the general climate of accountability described in the introduction, this body developed requirements for assessment of six institution-wide undergraduate competencies, among them critical thinking. The institutions were responsible for defining critical thinking and developing

both standards and methods for assessment. As an initial step, SEC required peer review of assessment plans, i.e. universities gave feedback to one another, and MASU altered their plan after peer review primarily to show more explicit alignment of the rubric and operational definition. This push by the state organization is related to national attention from the Department of Education "Spellings Report"(2006), Congress, the President, and national organizations. The external web of accountability is manifested in the SEC requirements. External stakeholders also include citizens, parents of students, and potential employers. These groups are not always as prominent, organized, or as vocal, but their concerns contribute to background interest and influence. Taken together, the external stakeholders create the focused attention on assessment described earlier.

The institutional stakeholders include the chief academic officer (Provost), the trustees, the assessment office, faculty, academic administrators, and the students. Students' responses to exit surveys conducted by the university show that they believe their skills in critical thinking have been improved during their time at MASU (Survey Results, 2006, from MASU assessment website). Obviously internal stakeholders desire results showing that the institution, its faculty, and educational processes produce graduates competent in the six key assessment areas, including critical thinking. As the processes unfold, these stakeholders commit varying degrees of attention and effort to assessment and its results. For example, during the past year, the trustees have twice asked for reports on assessment of critical thinking. All these stakeholder interests illustrate Chen's (2005) inclusion of relationships between the processes and stakeholders as an element for consideration when looking at effectiveness. The combination of

strong stakeholders and the assessment environment creates the opportunity for this study to have high significance for stakeholders, and to contribute knowledge usable by other institutions. The methods used for recording and analyzing Mid Atlantic State University's assessment of critical thinking are discussed below.

Methods

The following sections provide an overview of the methods used to select participants, and to collect and analyze data from multiple sources. Implementation of the methods was not a straight linear procedure as many of the actions described here took place simultaneously and on more than one occasion. After the narrative discussion, Table 4 in the last section summarizes relationships among the concept areas, research process and the data sources.

Participants

Although at first glance, it may be expected that participants were the students, the participants are the faculty, staff and administrators involved in the planning and implementation processes The National Center for Education Statistics (2002) included similar participants for its study of competency based assessments in higher education. Basic information concerning participants can be found in Appendix A.

All members of the original working group, professors in whose courses assessment took place, those who filled the rater function and four administrators were invited to participate. A total of thirty-two instructional and administrative faculty and staff were invited through email and, of those, twenty-eight returned signed consent letters. Roles are detailed below in Table 2, Participants, which shows that the twenty-

eight Mid Atlantic faculty and staff participated in different ways. Provision of further information about the participants violates the researcher's pledge for anonymity and confidentiality. Note that the tenured professors designated as Smithson professors are special appointments for leadership in and dedication to teaching undergraduates.

Table 2

Participants N=28

Category	**Role**	**N**	**Data Instrument**	**Additional Information**
Provost, Chief Academic Officer	Leadership	1	Interview	
Director, Teaching & Learning Center (TLC) Administrative Faculty	Leadership	1	Interview	The director position changed near the end of the study. The second director was not a participant.
Director and Associate Director, Assessment Office Administrative Faculty	Leadership & Management	2	Interview and Questionnaires I and II	Both of the participants were also raters
Professors	Working Group, Course Instructors, and Raters	24	Questionnaire I Sent after the pilot 16/17 returned Questionnaire II sent after the second iteration 10/14 returned 7 participants answered Questionnaires I and II	2 Tenured Smithson professors, • One Served on both Working and Focus Groups • One provided his course for rating of student presentations 9 Tenured • 3 working group and course • 2 working group • 1 observer • 1 working group and observer • 1 course and observer • 1 working group and course 3 Tenure Track • 2 working group and observer • 1 observer 9 Term Professors • 3 working group and course • 1 observer and course • 3 observers • 2 course 1 Librarian, Observer
	Focus Group	4	Discussed emerging themes	1 Smithson professor, 1 Tenure track, 1 term, Director of TLC

Among these participants, 14 were male. No other demographic information was collected. Return rates on questionnaires were 94% for Questionnaire I and 71% for Questionnaire II. Full questionnaires are in Appendix C. In the second round, one participant left the university and did not return Questionnaire II. Three reminders did not elicit responses from the remaining participants. The participants provided one set of data mirroring elements from Kayser's (1995) *Mining Group Gold*, which focused on the types of group members and planning. This data also covered faculty and administration collaboration issues reflected Kezar's recent work (2005; 2006), and group work concepts discussed earlier. For example, Questionnaire I, question 1, related to Kayser's attention to member selection, and Questionnaire II, question 8 addressed his attention to planning as success factors in group work. Kezar's cross disciplinary interaction and networking links to Questionnaire I, question 7 and Questionnaire II, question 6, and her attention to sense of mission is addressed in Questionnaire I, question 2 and Questionnaire II, question 7. A detailed chart of data relationships is in Appendix B. The administrator interview protocols (Appendix C) also addressed elements raised by Kezar, for example senior executive support, capitalizing on external pressures and sense of mission. The Provost protocol addressed the first topic, and the assessment office protocol includes the topics of taking advantage of external pressures and mission. The TLC director protocol addressed the wider topic of collaboration and networking. Confidentiality and identity of participants is maintained, with use of pseudonyms as needed (Tapper, 2004).

Data Collection and Processing

All issues related to data collection, confidentiality and anonymity followed approved protocols from the Human Subject Review Board. Documents with identification removed are in Appendix D, Human Subjects Documents.

Data included original data from four administrator interviews, the two questionnaires, plus archival data from assessment office records, and researcher memos and notes. Access to assessment records was obtained through a letter of permission signed by the director of assessment and the requisite protocol is also in Appendix D.

The importance of gathering data from functioning groups is emphasized by Davis and Kerr (1986) and McGrath (1989). Davis and Kerr say that data from functioning groups can replace false intuitive conclusions. An example of one such belief might be that faculty members do not want to work on assessment issues. Studying group work in context and using archival information about the group task deepens understanding. Context and documentary data are crucial, McGrath indicates, and recommends use of multiple data, multiple methods, and multiple occasions to yield a more dynamic view. Such multiple sources were given careful attention in this study.

Questionnaires, Interviews and Archival Data

The first nine-item questionnaire was answered by participants in the initial working group, professors in whose courses the first observation and writing evaluations took place, and faculty serving as raters for the first pilot implementation. The full questionnaires are in Appendix C. The questions sought to discover participants' views of the experience of working with the assessment, how they valued that experience, what

they would tell fellow faculty, and any other comments about the process and design of the assessment itself. For example,

How did you become involved in this assessment project?

Describe your participation so far in this one.

If another faculty member asked you if they should participate, how would you respond?

In what ways do you see this assessment and its results affecting the university, and your program /courses?

After the second implementation, an eleven-item questionnaire was given to the same group plus participants who joined the second iteration. There were similar questions, however, this instrument focused more on the on the process. For example

What has gone well, not so well?

Were you involved in the planning? If so, how?

How would you describe the usefulness and effects of this program to another faculty member?

In general, how might you characterize the process for carrying out this assessment?

What benefits do you see for student learning and/or teaching excellence resulting from this assessment?

Please make suggestions for how to change/improve the university's methods for carrying out assessments of student learning outcomes.

Questionnaires were sent by email along with follow-up paper copies. Results were consolidated by grouping all replies to each question. In order to preserve connection to

the respondents, a tag using participant numbers, the questionnaire number and the question number was attached to each statement. For example, Q2 22 1 identifies questionnaire 2, participant 22, question 1. In this way data could be examined, as seen in the next section, with the identification of participant and question attached. The consolidated responses were entered in the NVivo software as source documents.

Use of Software

NVivo is qualitative software that allows electronic coding, sorting and modeling. For this research I began learning how to use the sections that related to nodes, and models. "Nodes" is the name for the coding tool and the original set of 16 nodes included topics such as goals, planning, processes, and effects. The system limits choices of code names to four-character abbreviations and I found it hard to keep these in my head. Copious notes helped but were frustrating when the expectation was that the software would reduce work. Although the questionnaire data was entered, it soon became obvious that the questions themselves gave rise to the first round of codes. Analysis of the questionnaires moved from use of the software to a set of ten codes used on paper to organize the responses. Table 3 shows these initial ten codes and the final grouping after several iterations in between.

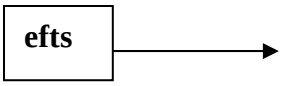
Table 3

Questionnaire Coding

Initial Codes	Final Codes
assessment group work institutional change critical thinking processes difficulties collaboration preparation communication reporting	assessment process design aims preparation critical thinking tools concepts group work processes collaboration different groups faculty administration communication and reporting institutional change assessment effects participation positives negatives

The software made more sense to me for use with the interviews and research memo data which were less structured to begin with. The interview transcripts and three researcher memos were put in the software and new codes (nodes) were created and re-grouped until there were 16. This is an example of a piece of data. The node **pcss** represented process. The sentences in italics would also appear in the code **efts** for effects of the assessment.

4-12-06/phone	M- a [political science professor-rater	I really like doing the rating *and am thinking about putting student presentations in my own classes. I see that students gain confidence and that maybe shyer ones especially; I had grad school experiences with such poor presentations that I wondered how the professor handled them. But maybe with undergraduates it will be different*



Use of the software was somewhat disappointing for this study, but may have potential in future projects.

One feature I had not originally intended to use but tried out was the model builder, a tool for creating graphics. It proved useful and stimulating as a way to entertain different ideas. Figure 3, NVivo Trial shows an early visual of the process and perhaps illustrates how much thinking and re-thinking took place during this study. The software did not produce other figures in this document because of formatting requirements.

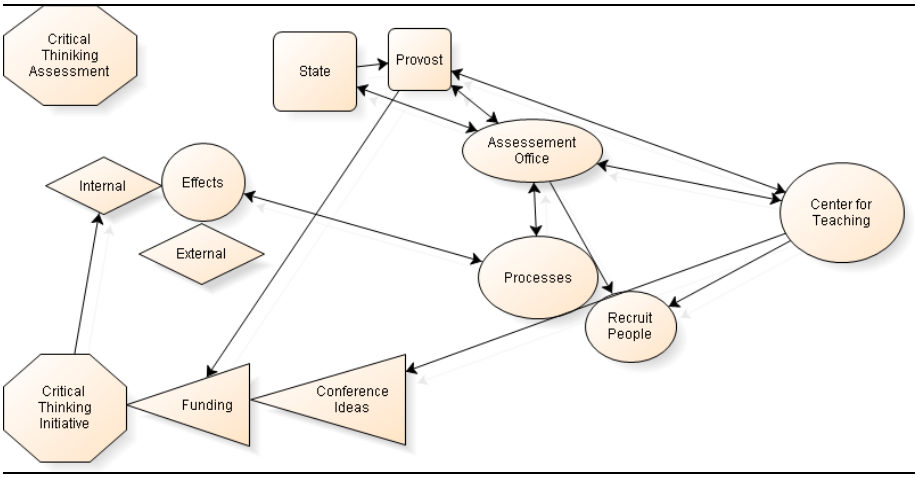


Figure 3. NVivo trial.

A combination of software coding and modeling, paper coding of the collected data and researcher documents formed important steps in data examination.

Interviews

The study includes three interviews, of the Provost, the Director and Associate Director of the Assessment Office, and the Director of the Teaching and Learning Center. The Provost was interviewed on May 1, 2007, at 11:30 am in his office. The protocol in Appendix C was used, and the interview lasted for 25 minutes. I was using my digital recorder that worked well in the past, however this time the microphone only picked up very faint sound that could not be augmented much using the electronic editor or volume controls in the digital software itself. Therefore records of the interview were primarily from my notes, which I took during the interview and reviewed immediately afterwards, adding further information. The Provost is well known for his concise speech and direct,

succinct replies to questions and this characteristic helped to assure good notes even with lack of a complete recording. During the interview I wanted to identify decisions he had made regarding the assessment and his expectations for critical thinking competency. Because the Critical Thinking in the Curriculum (CTC) initiative had just been announced, his attention was more on that subject than the assessment piece. The interview notes were entered as a source document in NVivo as described above and formed an important part of the data for research question two.

The other interviews had no technical problems and the digital editor for transcription performed beautifully, allowing me to slow the speed, to easily re-run a section, and to play the sound through the computer speakers. It was a boon to transcription that all the functions were controlled with the computer rather than by foot or hand. Controls against accidental erasures were automatic.

The Director and Associate Director of the Assessment Office were interviewed together on April 16, 2007 from 11:00am to 12:20 pm in the assessment conference room. The digital recorder worked well and I also took notes. The original plan was for two interviews, however since both interviewees had already answered the first faculty questionnaire because they were also observers, I decided to collapse the two protocols and do one interview. The protocols are in Appendix C. Both participants were very reflective, and expressed a great deal of thinking abut why the assessment was planned and carried out in a course-embedded, faculty-developed and approved design. It was during this interview that I fully realized that there were actually two different work groups—these administrators and the assessment staff who were collaborating with the

TLC director, and the separate faculty working group. It was also apparent that the first group served as the 'leader' for the second. Much of the discussion fit well with the interview of the TLC Director.

The TLC director met me for this interview in her office, from 1:00-2:20 pm on April 23, 2007. The digital recorder was used, and functioned well. I also took notes. This interviewee answered all questions in the protocol (Appendix C) and, like the assessment interviewees, the director, a former science professor, was very reflective. The themes of how to support faculty in teaching for learning and how to spread the ideas of the scholarship of teaching and learning were strong elements in the responses which the director also related to collaboration with the assessment office. The topics of critical thinking and the new CTC initiative were very important and the focus of interest, rather than the need for assessment. Transcriptions and data analysis were carried out as described earlier. Digital files from all interviews were kept on a password protected laptop in a home office and were erased from the recording device.

As a check on emerging findings and ideas, a focus group was convened December 14, 2007, from 2:00-3:00 pm in the assessment office conference room. I took notes rather than using the digital recorder. Scheduled participants included a Smithson professor, two tenure-track professors, one of them also undergraduate assistant dean for business, and, at a later time, the director of the TLC responded to the six emerging ideas. The business dean could not be present December 14 but replied the next week to the initial themes and to discussion notes. Scheduling face-to-face discussion was difficult, and accepting the information provided by the two who were not present was necessary.

Data were kept as researcher notes and email. Themes presented to the focus group included:

- Personal connections are important
- There is a very wide interest in the topic of critical thinking
- There are effects beyond the assessment
- The connection to the CTC project is important
- Faculty are willing, interested, and smart but are pulled in may directions
- There is a need for both leadership and management

The focus group function was to verify or reject these emerging themes, or add new ones. There was agreement on the accuracy of these first themes. However, discussion soon added more to be considered, for example, the value of having common vocabulary and standards to use for discussion, curriculum development effects, and student evaluation. The potential role of the general education program and its coming assessment, the idea of considering cultural differences, the many different ways critical thinking manifests in disciplines, and the strong need to follow the assessment with discussions and faculty development were also mentioned. The Focus Group discussion confirmed the emerging themes and augmented them. The focus group fulfilled the role of member checks, a common process of qualitative research. All focus group participants had roles on the working group, and two of them also served as observers and were well informed about the MASU assessment processes.

In order to achieve a deep understanding of what happens, many different types of data were collected and analyzed with the aim of leading to sound conclusions. As is

common in a participant researcher study, there was careful attention to the data aspect of the research design, to the value of field notes, and to other methodologies delineated by Maxwell (2005) during all phases of data collection. Maxwell suggests description be based on as much information as possible from researcher notes and memos as well as collected or archival data and that interpretation be based on the participants' views and the theoretical concepts of the study. Regard for data also helped alleviate concerns about validity and researcher bias by linking data to conclusions and interpretations. The steps from data collection, to analysis, to findings and to conclusions were intentionally taken and based on recommendations from the research experts cited here.

These purposeful data collection methods can in turn be connected to Birnbaum's (1998) concept of a self-renewing institution. In order to be self-renewing, an institution must have good feedback loops that can alert administrators to difficulties in a timely manner, thus enabling either proactive planning or reactive correction before crisis development. The feedback should come from acceptable data, and understanding how this feedback can be channeled from assessment processes was another purpose for data collection from multiple sources. These data were expected to demonstrate a process for calling attention to a topic (competency assessment) and to educational effects (on faculty, courses, students) which might function as the kinds of loops Birnbaum (1998) advocates. In other words, can an assessment and the processes involved do more than provide inactive, shelf-sitting reports? The concept tools used to examine the collected data will assist in that endeavor. Additional areas of data interpretation and validity related to methodology are considered below.

Data Analysis

A variety of analysis methods were applied to the data. They include use of conceptual models, coding and use of qualitative software, triangulation, and attention to connections among these tools. Laying out these connections assisted in maintaining focus on the questions and concepts. The connections were important elements of the methodology and formed the first steps in the analysis, and are detailed in Table 4, Relationships among Questions, Data and Concepts, which is at the end of this section.

Subsequent data analysis compared decisions made in the MASU case to Cohen, March and Olsen's (1972) concept of garbage can decision making. Data regarding work processes were compared with Gersick's 1998 model of punctuated equilibrium. Other concept models were applied to meeting notes and documents and field observations of implementations. Two models were particularly useful, one that combines standards and methods of high performing teams from Katzenbach and Smith (1994) and Kayser (1995), and the second from Kezar (2006) regarding collaboration. Elements of high performing teams, for example, include shared leadership, accountability, self-developed purpose, collective products, open-ended discussion, active problem-solving meetings, assessing work products, and doing real work together (Katzenbach & Smith). Kaysar adds the elements of selection of members, planning successful work sessions, and group maintenance. These concepts were used in establishing tentative coding terms for data analysis. Kezar's investigation of intra-institutional collaboration includes several applicable concepts. For example, six of her ten recommendations for promoting

collaboration also applied here. They include the mission, campus network building, cross-disciplinary structures, rewards, senior executive support, and capitalizing on external pressures. These elements played important parts in both initial coding efforts and in the development of conclusions.

Other concepts from organizational leadership and culture are also embedded in the work of these three authors. For example Katzenbach and Smith (1994) suggest that organizational leaders enhance team performance by emphasizing a performance ethic, rather than only a team-promoting environment. Kayser (1995) explicitly states that he is writing about leadership, both from the standpoint of the institution and from that of the team facilitator. Kezar's (2005) work illustrates leadership and culture change as elements of collaboration encouragement. By utilizing a variety of concepts related to group work and collaboration, the elements of leadership and culture that affect team process and work emerged. Focus Group data and administrator interviews added to the picture.

Yin (2003) suggests construction of a logic model that compares what happened to what was expected. In the case of Mid Atlantic, prior student competency assessments used the same work group models; however, those assessments did not cross disciplinary lines, as the critical thinking assessment did. Evaluating whether the work group met its charge and goals also meant examination of the additional effects of the assessment process. An important goal of this study was to spotlight institutional effects outside of the assessment processes. If change is to happen, and assessment can be an agent of change, then records of the effects of assessment can provide indicators of change.

Use of a variety of concepts to examine the various kinds of data addresses principles of triangulation. Triangulation during data interpretation is a widely recognized concept in qualitative research and was a key reason for seeking multiple sources of data. Principles of triangulation are treated in Yin (2003), Creswell (2003) and Maxwell (2005) and apply both to using diverse sources and to data comparisons. Triangulations should be visible. In this case, for example, the information from the faculty questionnaires resonates with information from meeting notes and administrative interviews. Data from the two questionnaires reflect similar themes from both participant groups. Triangulation is illustrated in Figure 4.

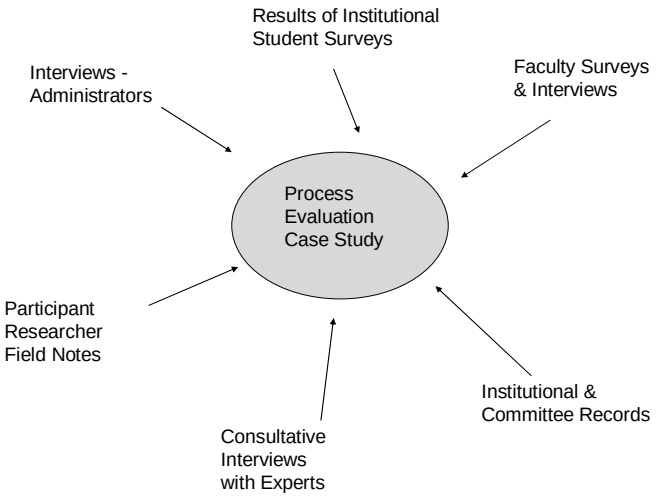


Figure 4. Triangulation of data sources.

It is in the center field that all these elements must be connected based on the methods described in this chapter. Because there were several sets of interviews of

administrators involved in the project, and one focus group session for faculty, consideration of how to integrate those data become a concern. Weiss (1994) describes how to handle the data by considering its analysis beforehand and how to set tones in interviews in order to facilitate respondents' participation. The fact that this was a participant researcher design facilitated both interviews and the focus group discussion and also helped make multiple connections among data. Since the material was so familiar and I was immersed in it on an almost daily basis, it was easier to see connections than it would have been for an investigator only looking at records. For instance, the ideas about two work groups and the connections between them, a new, wider viewpoint of "workspace", and recognition of the growing collaborative relationship between TLC and assessment are three examples of insight facilitated by participant research.

Although there was an expectation that the second questionnaire administered near the end of the study might need modifications based on initial data from the first set of questions and the interviews, this did not prove to be the case. As much as possible, information that emerged from participants was corroborated with other archival documentation from meeting notes, paper office files, draft and finished documents, reports, and web sites.

A summary of the relationships among the major concept areas, research questions, and data sources is presented in Table 4. Maintaining focus on the organizing concepts and questions was an important element of the methodology, so the first step toward analysis was laying out the tools in the following table:

Table 4

Relationships among questions, data and concepts

Concept Area: Assessment

Research Questions Addressed:

- Question 1: How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?
- Question 2: What were the processes and decisions that contributed to the development of an assessment of this student critical thinking learning outcome?
- Question 4: What models of assessment help us understand this case?
- Data addressing these concepts:
  - Administrator interviews
  - Working group questionnaire I
  - Meeting records, communications
  - Researcher field notes

References

Banta, T. W. (Ed.). (2004). *Hallmarks of effective outcomes assessment*. San Francisco: Jossey Bass

Khademian, A. (2002). *Working with culture*. Washington, D.C.: CQ Press

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment.* Washington. D.C.: National Academy Press.

United States Department of Education. (2002). *Defining and assessing learning: Exploring competency-based initiatives.* Washington, D.C.: National Center For Education Statistics, U.S. Department of Education.

Concept Area: Critical Thinking

Research Questions Addressed

- Question 1: How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?
- Question 3: What effects did these processes and decisions have on the institution during the time of the study?
- Question 4: What models of critical thinking help us understand this case?

Data addressing these concepts:

- Working group and workshop records
- Faculty questionnaire II
- Administrator interviews
- File Data - Rater feedback

Reference

Facione, P. A. (2006). Critical thinking: What it is and why it counts. *Insight Assessment/California Academic Press* http://insightassessment.com

Concept Area: Self-Renewing University

Research Questions Addressed

- Question 1: How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?
- Question 2: What were the processes and decisions that contributed to the development of an assessment of this student critical thinking outcome?
- Question 3: What effects did these processes and decisions have on the institution during the time of the study?
- Question 4: What models of self-renewing institutions help us understand this case?

Data addressing these concepts:

- Administrator interviews
- Faculty questionnaires
- Focus group

References

Briggs, C., Stark, J., & Rowland-Poplawski, A. (2003). How do we know a "continuous planning" academic department when we see one? *The Journal of Higher Education*, *74*(4), 361-385

Birnbaum, R. (1998). *How colleges work: The cybernetics of academic organization and leadership*. San Francisco: Jossey Bass

Miller, B. (2007). *Assessing organizational performance in higher education.* San Francisco: Jossey-Bass

Rogers, E. M. (2003). *Diffusion of innovations* (5 ed.). New York: Free Press Simon & Schuster

Concept Area: Group Work Analysis

Research Questions Addressed

- Question 1: How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?
- Question 2: What were the processes and decisions that contributed to the development of an assessment of this student critical thinking outcome?
- Question 3. What effects did these processes and decisions have on the institution during the time of this study?
- Questions 4: What models of group work help us understand this case?

Data addressing these concepts

- Working group meeting notes, attendance, workshop records
- Researcher memos
- Administrator interviews
- Focus group results

References
Gersick, C. (1998). Time and transition in work teams: Toward a new model of group development. *The Academy of Management Journal, 31*(1), 9-41
Jennings, M. (2007). *Leading effective meetings, teams and work groups in districts and schools*. Alexandria, VA: Association for Supervision and Curriculum Development
Kayser, T. (1995). *Mining group gold: How to cash in on the collaborative brain power of a group* (2nd ed.). Chicago: Irwin
Kezar, A. (2006). Redesigning for collaboration in learning initiatives: An examination of four highly collaborative campuses. *The Journal of Higher Education, 77*(5), 805-838

An example of how this organizing process worked is the application from section one, Table 4 of the three components of competency assessment as put forward in the National Center for Education Statistics report (National Center for Postsecondary Improvement, 2002; United States Department of Education, 2002):

- a description of the competency (faculty chosen definition of critical thinking in students from meeting notes and assessment records);
- measuring or assessing the competency (selection of course embedded methodology in records and meeting notes, and planning of the implementation and the implementation from records, questionnaires and interviews)
- standards for judging competency (creation and use of the rating rubric)

In addition to these concepts, there are several other principles related to data interpretation that assisted in developing findings and conclusions. Arnold and Feldman (1986) and Schwartzmann (1989) address need to look back to participants for member

checks and to examine implementation for information on how the organization works, along with a caveat to be cognizant of data problems. Arnold and Feldman also suggest pros and cons for types of data. For example, observations have validity from being real-time data rather than data that may be colored by the passage of time but they also contain the threat of observer bias.

Some other suggestions for conducting analysis involve the passage of time. Since faculty support for and participation in developing, executing, and interpreting the critical thinking assessment were keys to discovery, using the two-stage questionnaires as well as records over a period of time were important steps in data collection and data analysis. The interviews and surveys elicited both current and reflective information from the participants, and mirror the time-separated collection of data used in Phillips and Bond's (2004) study of students. Two dissertations, by Roberts (1996) and Albert (2004) do not specifically address the same case methodology as outlined here, but generated additional thought for this study. Roberts includes the effects of using locally developed instruments for assessing students. He points out that the process of developing the instruments, as in the MASU case, has effects on those involved and that these are beneficial for curriculum improvement. Albert's 62-university survey of general education for undergraduates illustrated the types of competencies that undergraduates should develop in general education, many of which can be seen in the critical thinking criteria for MASU.

Data Interpretation

Data collection and analysis, and interpretation of findings took place on a continuous basis. At this point, additional theories and key studies helped clarify emerging views. The question of how to interpret data and to account for rival explanations was addressed using a logic model approach advocated by Yin (2003). His model suggests that discovering a series of cause and effects will lead to strong conclusions. For example, a primary cause of the kind of assessment that took place was the director of assessment's set of strong beliefs and a cause of faculty interest after spring 2007 probably was an effect of the stipends attached to the CTC initiative. Tracing all those links can be a bit daunting as seen previously in Figure 3. Because one intention was to begin to develop a model or guide, Yin's logic model proved valuable for developing the proposed process model and the best practice statements in chapter five. Additional help structuring the findings came from two other works. McMillian and Schumacher (2001) provide a comprehensive introduction to concepts of educational research, particularly on qualitative strategies, evaluation research and analysis. For example, they support developing research that offers decision makers usable decision-making information, and characterize one form of this research as Process Evaluation, which the authors describe as including relevant questions for participants, and different kinds of data from different sources.

The second contribution pertains to the management of data, specifically use of tags for interview transcription and coding (Gersick, 1989). This technique was useful in backtracking sources in a more detailed manner than the software data management tool,

NVivo, provided. Without the tags, attributing and tracking quotations would have been difficult.

Validity and Reliability

The extensive discussion of study design, data, and data analysis leads to consideration of both validity and reliability. Maxwell (2005) is very clear about defining what validity means for qualitative research. He refers to "the correctness or credibility of a description, conclusion explanation, interpretation or other sort of account" (p. 87). Merriam (1998) responds to the discussion of how reliability, commonly perceived as replication, fits in case study research. "The question then is not whether findings will be found again, but *whether the results are consistent with the data collected* (sic.)" (p.206). She states that ensuring dependable results can be accomplished through attention to concepts for the study, expressed relationship of the researcher to the participants, description of basis for selecting informants, and information about context. The previous sections relating to researcher background, conceptual framework, data collection and methods of analysis, and triangulation support both views of reliability and of validity. Merriam also refers to the idea of an audit trail, which is a rich explanation of the researcher's work and was addressed in this case by detailed methods narratives and charts. In addition, multiplicity of sources and application of relevant conceptual theories were employed to reduce uncertainty and support validity and reliability. Finally, Yin (2003) cautions that any rival explanations or data must be considered. For example,

although most respondents agree on the value of the assessment, those who disagree must also be given voice.

A different view of validity in assessment is offered by Boud (1995), who insists on caution regarding self-referential data and suggests the concept of consequent validity be applied. "Consequent validity is high when there is a positive backwater effect on learning and low when it encourages ways of learning which are counter to what is desired…we should develop assessment procedures of high consequent validity" (p.39). The inclusion in the study design of the assessment processes effects addresses the idea of consequent validity by tracing faculty and staff reports of change.

Intentional study design, types of data choice and collection, application of conceptual frames for analysis, detailed explanation of process, and examination of effects address the major concerns of case study and qualitative research validity and reliability. As reported in the section on research background, awareness and caution regarding the role and functioning of a participant researcher was a continuous concern and one constantly attended. The methods employed here were deliberately chosen to give both depth and breadth to the study of MASU's assessment processes. Inclusion of practice and research intentions, many data sources, multiple conceptual analyses, triangulation, and continuous reflection on methods were chosen to support strength of findings and interpretation. The following chapters relate the study's findings; answer the research questions, and offer interpretation and conclusions that address both research and practice. Keeping a research focus on answers to the four research questions was enhanced by the study's structured data collection and analysis, and by the selection and

use of conceptual frames. As seen in Table 4, the research questions are informed by a variety of data which may point to more than one question. Using conceptual organization as the methods backbone sets the frame for findings as presented in the next chapter.

4. Results and Findings

Introduction

Case study results are commonly presented in narrative form, although there is no standardized format for how that is done and how findings are presented. Because it depends on purpose and audience, presentation of a case study combines "balancing description with analysis" (Merriam, 1998). This analysis involves multiple conceptual frames; therefore, results and findings are presented as narrative answers to the original four research questions, followed by an introduction to conclusions and interpretation that are detailed in chapter five. A deliberate decision was made to report findings as they relate to each of the research questions in order to make sense of the variety and scope of data. Findings resulted from employment of the concepts and tools for analysis described in chapters two and three, and as presented in Table 4, chapter three. In preparing this chapter, a review of all data was made, and intentional and continuous efforts were made to ground ideas, themes, or interpretations in these data.

Research Questions and Findings

The four research questions differ from one another in what they seek to illuminate. Questions one and two elicit the necessary history of what happened, and are addressed using straightforward narratives. Answers to question three focus on the

assessment in terms of the original institutional rationale to effect change. The findings related to question four continue interpretation of what happened by using tools from the four organizing concepts of assessment, critical thinking, group work and institutional self-renewal. Those four questions are:

1. How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?

2. What were the processes and decisions that contributed to the development of an assessment of this student critical thinking learning outcome?

3. What effects did these processes and decisions have on the institution during the time of the study?

4. What models of assessment, critical thinking, group work, and institutional self-renewal help us understand this case?

The data from questionnaires, records, and interviews form the foundation for findings related to all the questions. A detailed picture of the relationships among the concept areas and data sources can be seen in Appendix B.

Question 1: How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?

Several preliminary activities pre-dated the start of the actual assessment activities and are summarized here from the assessment office file records and from administrator interviews. The initial decision by the State Education Council (SEC) to require competency assessments was communicated to the university in 2002, with the critical

thinking reporting requirement placed at the end of the spring 2006 semester. During the same period, initiatives from the director for the Teaching and Learning Center (TLC) related to critical thinking were being developed.

> I got interested in critical thinking because of a conference I go to every year, and one of the national experts on critical thinking presents at the conference every year. So I started to think about it. Then, I came back here and started doing workshops on critical thinking and workshops at other institutions. (Interview, TLC Director, April 23, 2007)

Although the 2006 report was the first externally required reporting on critical thinking, MASU has included critical thinking in its university mission statement since 1991 (MASU *Factbook*, 2007).

The Teaching and Learning Center's faculty development workshops related to critical thinking were well attended. Particular interest was expressed at the 2004 "Critical Thinking in the Classroom: Making It Happen" which was intended to help faculty clarify critical thinking ideas for themselves, communicate expectations to students, and then evaluate student achievement. The roundtable discussion format assisted faculty in the process of creating work for their courses that reflected principles of critical thinking as they appear in their disciplines. The presentation by Dr. William Condon, leader of Washington State University's critical thinking project, certainly raised awareness of the topic and started a more general discussion on critical thinking that was an important element of the context for the subsequent assessment and this study. Thus, conversations among faculty were opened before the assessment requirements were

announced to the university community. This workshop was also the first in a series of collaborative critical thinking activities by the Teaching and Learning Center and the assessment office. A link between teaching and assessment was being forged, and was recognized by the director of TLC, the assessment office, and faculty.

"…because it's being measured they will feel they *ought* to teach it. They won't want to look bad" (Interview, TLC Director, April 23, 2007). The director and associate director of assessment add:

> Because it is part of our mission, of the university mission, that we promote critical thinking among our students and it was appropriate to begin assessing whether we actually did it. And I think maybe we say that there is also faculty interest (Interview, Director of Assessment, April 16, 2007). I pick it up from program review when they talk about what their expectation is for their graduates. They normally talk about broad things, but eventually come towards critical thinking (Interview, Associate Director of Assessment, April 16, 2007).

Additional work on other competency assessments required by the state had already taken place and was organized around faculty working groups and course-embedded assessments. This approach was based on strongly held beliefs of the director of assessment about the purposes of assessment and MASU culture, as well as intentional efforts for program improvement.

> So everything we do is faculty-based— the faculty are the ones who really have to do this or at least have to guide it. But, so ok, everything has to be faculty based.

> We embed all of our assessments….So, course-embedded, which then implies that you have to work with the faculty because it is going to be in their class. You are going to be using their classes as the source of the data for this. And we have had some experience with other competencies as well that suggested that it would be worth at least exploring whether or not we could do this across disciplines because we have done it with scientific reasoning. (Interview, Director of Assessment, April 16, 2007)

When asked if the primary reason for including faculty is related to concerns abut effecting change, the response was very direct and expressed belief in connections of assessment and change.

> If you are really trying, if your ultimate goal is to improve programs, faculty have to be centrally involved, because otherwise we collect data that goes nowhere and the faculty don't care what the results are, you know. But if they're involved and they do care— as we've found. Many of them have started making some changes even before any results were ever in, which, frankly, is pretty typical of most assessments. Just the discussion spurs change, so that is a central reason for involving faculty. (Interview, Director of Assessment, April 16, 2007)

Although this study's collection of original data began in the fall of 2006, information on the context and events prior to that date contribute to understanding the whole process. I was given access to records of meetings, workshop materials, and interview information and I benefited as well as well from being a participant researcher directly involved in the processes. My participation in the background events included

attendance at the original Critical Thinking workshops and a Ph.D. program-required internship which I took in the TLC. The internship was specifically designed to support the upcoming assessment effort. The products of the internship included preparation of a bibliography, collection of materials related to other institutions' efforts, discussions of critical thinking definitions, and proposed materials for the first faculty working group meeting in December 2005 (Assessment Records, Internship Report, 2005). This preparation informed assessment planning and helped to reinforce the decision to propose to the faculty working group the use of embedded, faculty-based assessment methods.

> I do remember when we first started, we had a grad student doing literature review of all the currently used assessment approaches to critical thinking and I remember she brought in stacks of paper you know telling us what was available including some of the tests and some of the testing questions and some approaches by other universities. I think Washington State was one of them and other universities too and also the definitions. So we really, we sat down and had a meeting and had a kind of like you know a literature review, review of the current approaches and we do feel that, you know, the way we have chosen really fits more, you know, MASU. (Interview, Associate Director of Assessment, April 16, 2007)

These two streams, the efforts of TLC and the assessment office, came together. I joined the assessment staff in the fall of 2005, and continued to work on the critical thinking assessment. Fall 2005 was the planning period for the assessment. Responsibility was

given by the Provost to the head of the assessment office. "I put Lynne [pseudonym] in charge - she is the expert" (Interview, Provost May 1, 2007).

At first glance, it may seem like the decisions for assessment were already set; however, because faculty validation for all the steps was an important factor for faculty involvement, a particular method of invitation was set up and several topics were prepared for the first meeting of the working group. Planning records (Emails, Meeting Notes, 2005) from the assessment office showed that the following steps were taken.

A review of synthesis course syllabi identified courses in which there were student assignments and activities related to critical thinking. All undergraduates, including transfer students, are required to take a synthesis course as the last portion of the general education curriculum. Courses are often part of required major coursework and may double as a writing intensive course, or as a capstone course. However, some majors do not have synthesis courses, and there are several such courses that are open to all students, for example, geography, which was selected to participate in this study partly because of inclusion of students from multiple majors. The synthesis courses were examined because the institutional goals for such courses include making connections among different ideas, synthesis of knowledge, and demonstration of advanced levels of oral and written communication (MASU Provost Website, retrieved May 5, 2008). The communication element was important because part of the assessment plan to be presented to faculty was observation of students' culminating presentations. There had recently been an extensive five-year writing assessment in the disciplines, and therefore the assessment office wanted to look for another student product to evaluate.

A group of faculty were selected for invitation based on their synthesis course syllabi and additional faculty were invited on recommendations from the TLC director, as well as others with previously expressed interests in undergraduate education, (e.g., two science professors and a Smithson political science professor). Smithson professors have a distinguished university appointment directly connected to support for undergraduate education. In order to highlight the institutional importance of this working group, professors were invited by the Provost, and their deans were notified of their selection. An individual invitation to participate was mentioned by eight of the sixteen study participants answering the question "How did you become involved?" (Questionnaire 1, question 1, Appendix C).

The Working Group Begins

The Critical Thinking Working Group convened for the first time on December 2, 2005. The agenda was structured to include consideration of critical thinking definitions, how critical thinking is manifested in various disciplines, how to know if it occurs, and development of a measure for the selected elements of critical thinking that could be used in any discipline (Assessment Records, Critical Thinking Working Group Files, 2005). Eleven faculty members from four of the six schools offering undergraduate programs attended, as well as three assessment office professionals and the director of TLC. The meeting was lively and animated, with obvious faculty interest in the topic. More quickly than expected, a consensus was reached on using the Facione (1998) definition, which was developed by a group of universities and was one of the definitions presented in the materials for the workshop. That definition follows:

> We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference as well as explanation of the evidential, conceptual, methodological, criteriological or contextual considerations upon which that judgment is based. (p.16)

A second consensus was agreement to try assessing student presentations on culminating course projects and research reports. It was also decided to include some writing projects, with a voluntary assessment of writing from portfolios in the nursing program. This program already had a cumulative assessment of portfolio writing samples and they wanted to add critical thinking to their assessment. There were reservations about whether one would measure communication skills rather than thinking skills in the presentations. Palomba and Banta (1999) present a similar method in describing the alternatives for critical thinking assessment, particularly the rubric method for examining presentations. Faculty concerns were natural since the rubric had not been developed yet. (Assessment Records, Meeting Notes, December 2, 2005).

A bigger difficulty turned on the issue of an assessment tool that would work in all disciplines. The general model for application of a scoring guide or rubric was taken from the Washington State Study (Condon, 2004). This faculty-developed model included a set of seven common points; however, each discipline or course adapted it to specific instances. Washington State was not required to develop a university-level assessment report and therefore had more flexibility than the effort at MASU. The project at WSU was seen as more of a faculty development effort.

The first working group session concluded with members agreeing to submit ideas for inclusion in a general rubric that could be used in any discipline and that supported the chosen definition. The working group had a difficult task.

In talking about this first critical thinking meeting, the associate director of assessment commented on faculty interest. "Actually, I was very impressed that the faculty showed interest in assessing critical thinking and how they feel like, you know, we are teaching it and they want to know if students are getting it or not" (Interview, Associate Director of Assessment, April 16, 2007). Faculty responding to questionnaires viewed working on this assessment positively. Fifteen of the sixteen responses were positive to question seven on the first questionnaire (Appendix C) "If another faculty member asked you if they should participate, how would you respond?" This question was particularly included because answers to it give a good holistic indication of what the respondent himself thinks about the experience. For example, "I'd tell them it would energize their teaching to talk to other faculty about how best to incorporate critical thinking in the classroom " (Questionnaire 1, question 7, participant 17). The enthusiasm, lively discussion, proposed ideas, and interest expressed at that first working group session were typified by the suggestion that those present take part in the assessment and identify student work that could be observed. This suggestion was made early in the session, before the agenda topic was even opened. The same positive attitude showed in replies to question two on the second questionnaire (Appendix C). Fifteen of sixteen replied positively to the question "How do you feel about your participation, i.e. it was/was not worth your time?" A tenure track member of the working group responded,

"Definitely, it was worth it. It was a great learning experience. It is really affecting my own teaching philosophy in a positive way. I'll be busy updating my teaching portfolio" (Participant 8).

Development of a rubric measurement tool began in January 2006 for the planned assessment at the end of the semester. At the same time, the director of assessment first presented reports on other competency assessments to the board of trustees, who expressed particular interest in the coming critical thinking assessment (Assessment Office Calendar, 2006 and staff meeting report of the Director of Assessment). From that time on, the board wanted regular reports on critical thinking.

Despite early faculty enthusiasm, few contributions to development of the assessment tool were received, even after two additional requests. Time to prepare the assessment was running out. At this point, the assessment office decided to draft a rubric based on the selected Facione definition, and incorporate any information that had been received. This draft became the agenda for two meetings, on February 16 and March 28, 2006. The draft was also circulated by email. These two meetings were less well attended (four to six faculty), and led to some frustration. "The tool is a good product. It has been a good exercise in developing it. One frustrating thing is that some members were not able to come to every meeting, so I felt I was always explaining and dealing with new ideas" (Questionnaire 2, question 3, participant 3). Scheduling meetings that can be attended by all those interested or required is very difficult when working with a faculty group. It would not be the only time that inconsistent attendance slowed down forward progress, or that faculty did not respond between meetings with requested

opinions or information. A final version was again circulated. The full rubric is in Appendix E.

The basic criteria embedded in the rubric for critical thinking assessment were:

- Identification of important questions/problems/issue
- Identification of assumptions and consideration of alternative perspectives/solutions
- Selection of appropriate methods
- Analysis, interpretation and judgment about the relevance and quality of information
- Ability to draw conclusions and make judgments based on evidence gathered
- Integration of ideas into a coherent argument/solution/presentation, etc.
- Fresh ideas/engagement with the topic/idea
- Ability to communicate the results of critical thinking

Three levels, Highly Competent, Competent, and Not Competent, were chosen. Another choice of Not Applicable was added after the second meeting so that either rater uncertainty in this first round or student assignment parameters lacking particular elements would not penalize the student score.

The committee felt that there had not been adequate time for professors to incorporate information into assignments or to provide the rubric to students. This first round was designated as a pilot, intended to test the rubric and the concept of rating presentations.

Each possible choice on the rubric had a descriptor to help the rater and to inform professors and students. Some doubts remained, exemplified by this comment "I am skeptical about one set of rubrics which will be suitable for all courses in many disciplines. It might be good to have a comprehensive one developed and choose only a subset of the rubric for assessment in a given course" (Questionnaire 1, question 9, participant 6).

The next steps were recruiting faculty to be raters, setting the observation schedules and training the raters. At least two raters observed each presentation and the nursing essays had three raters. Observations and ratings took place in twelve classes, in six disciplines, and included 110 student works between April 20 and May 10, 2006. Assessment staff met with participating professors to present the rubric, explain the process, and to collect copies of the assignment.

As time was rapidly running out, rater training on April 19 had some difficulties. There were no real presentations to practice with, so the assessment staff created scripts designed to focus on one or another of the criteria in the rubric. These were read aloud so that raters could practice for the oral presentations. Everyone found this difficult and realized that doing the student ratings would also be difficult. Feedback sheets from the training included both expressions of confidence and concerns. Not all of the raters at the training were study participants, and some of them did not give names, so these quotes are not attributed, and were taken from a consolidated feedback sheet in assessment records. As will be the case during this report, I want to clearly identify information that

came from records and is not attributable and information that came from participants and is attributable.

The limitation of the materials to just a few sentences made it harder to do ratings.

I need more practice and training.

I feel the rubric is just adequate, but easy to read and understand.

The examples that we went through today gave me a much better idea of the rubrics and how to use them.

The short excerpts went less well than the general discussion of the rubric (Assessment Office Records, Rater Training, 2006). Both trainers and raters realized how handicapped observers would be without actual training using presentations on videos.

Raters received copies of the assignments given to students, and were assigned in teams to observe student presentations, except for the nursing writing ratings, which were organized by that department as part of a regular student portfolio review. Raters completed feedback reports at the end of their classroom observations. The expectation was that there would be many problems; however, feedback reports (Assessment Records, Rater Feedback, 2006) included some positives. Once again, not all of these raters were participants, and the same source constraints as previously apply. Of the 32 comments, seven were quite positive, for example "I learned many things from observing the class and using the rubric that will be useful in my own teaching." and "The rubric was very helpful."

Less positive expressions included difficulty with depth of understanding how to apply the categories and levels of competency as were reflected in six responses, e.g., "There should be a supplementary sheet explaining more about what was meant by the abbreviated explanations on the rating sheet." These statements were taken from feedback sheets written at the end of observations, and are not identified because not all of the responders were study participants. Additional complaints addressed the kind of assignment the instructor required, the observations of internship reports, issues related to team presentations, and whether the observer needs domain knowledge. In spite of expressed difficulties, 111 separate ratings were completed (Assessment Records, Critical Thinking Results 1, Excel file, May 2006). There was only one very negative report: "Sending me to this type of class was not a good use of my time". Only two of these rater respondents stated they did not want to continue with the project, and one of those was leaving the university anyway.

Study participants' answers to Questionnaire I, given after the first implementation, included a range of statements, such as participant 11 who answered question six, "How does this experience compare to any others?" by saying the experience "was very enlightening and positive, due to the interest and dedication of participants". Of the 12 responses to this question, two thirds expressed positive experiences, and four responses identified difficulties relating to cross disciplinary efforts and observing presentations for critical thinking. Question 7 asked respondents to say how they would advise another faculty member inquiring about participation. Sixteen replies indicated highly positive responses like the one from participant 15. "I would

encourage them to do so, especially if they are interested in reflecting on how well they teach critical thinking in their classes, and or how they assess their students' learning in the area of critical thinking." Participant 6 answered in one word, "absolutely." Less support was expressed by participant 13 who thought that "the issues about thinking have to be discipline-located and the scope of generalization is limited." Since the feedback from raters was collected at the time they finished observing, the more positive questionnaire replies may be related to time for reflection or might be a result of the more negative raters not participating in the study.

Student results from the observations were reported to participating professors, and information about the first round was published by the assessment office in *Assessment Focus*, (Mid Atlantic State University Assessment Office, 2007) a periodic series of reports on university assessments. The September 2006 state report, (Appendix G) shows the following results of the pilot assessment. (Institutional identifiers have been removed.)

Table 5

Critical Thinking Results –State Report

Criteria	Number of Presentations/ Written Papers*	*Highly Competent*	*Competent*	*Not Competent*	Total* *Competent or Highly Competent*
Identification of Problem/Issues	69	30%	58%	12%	88%
Analysis of Problem/Issue/ Investigation	69	46%	44%	10%	90%
Credibility of Sources	60	65%	28%	7%	93%
Conclusion/Problem Solution	69	26%	62%	12%	88%
Creativity/Student Ownership/ Engagement	68	41%	44%	15%	85%
Communication/ Adaptation to Audience	69	39%	46%	15%	86%

*Because group presentations were included in this assessment, the total number of actual students is higher than the number of presentations/papers (number of students=110). Percentages may not add to 100 due to rounding. (Critical Thinking Report, Retrieved from MASU Assessment website, May 4, 2007, full report in Appendix G)

The working group met December 1, 2006 to review findings and begin plans for the next round of assessment in spring 2007. Further discussion on effects of this assessment is in the section below on research question three, "What effects did these processes and decisions have on the institution at the time of the study?"

During the time between the publication of results and the working group meeting, the assessment office planned for changes. Because of rater feedback, changes to some of the rubric wording, addition of another rating level, and suggestion of wide

dissemination of the rubric were proposed and accepted at the December meeting. The Director of the TLC agreed that the revised rubric would be distributed at a new faculty orientation to raise awareness of MASU's interest in having students learn critical thinking skills. It was also posted on the assessment webpage, and formed a part of the materials for proposals for Critical Thinking in the Curriculum which began in spring 2007. More details on that initiative will be in section below on effects of the assessment.

Because most of the synthesis and senior capstone courses that were part of the second round planning were only offered in a spring semester, the group also decided that the next round would be in the spring semester, 2007. Also, upon advice of professors, taping of student presentations in real time was not chosen. Professors stated that addition of that element to culminating presentations would add concerns to students already feeling pressure from high-stakes final projects (Assessment Records, Meeting Notes, December 1, 2006).

In order to address rater training issues, the assessment office arranged for two students to tape presentations they had already given. Students taped two versions, one their best effort, and another intentionally changed to reflect varying degrees of worse performance on specific critical thinking rubric elements. For example, one student was asked to leave out references, and another to demonstrate disengagement with the topic. Students did a very good job even though one was reluctant to record a poor job in case her advisor saw it. Proper releases were obtained. These efforts were also designed to show examples of the middle range as well as the extremes of very good and poor. The

plan was that raters in training sessions would have some different levels available for practice in using the rubric. These videotapes were completed in early spring 2007 (Assessment Records, Filming File, 2007).

The ten faculty members attending the next meeting on February 9, 2007 were enthusiastic, and agreed to take the rubric to their departments for comments. The group included an assistant dean from business, a Smithson professor, and representatives from English, science, engineering, and nursing. This group was obviously strongly interested in the topic of critical thinking. All but one participated in the original Working Group, and two of the professors attending brought news regarding dissemination of the rubric and assessment plans. The representative from the business school reported a decision to have raters for their senior case competition use the rubric, even though the raters were community business persons outside the university. The rubric was used to rate student teams presenting a case, help select the winners, and information from those scores was available for the university critical thinking assessment. Students were not scored individually. The second news was from a faculty member in the systems research program in engineering who had been a rater for the first round. On his own initiative he arranged for senior presentations in that department to participate.

For this implementation round, there was at least one course from every undergraduate college at the university including education, performing arts, science, social sciences and integrative studies, engineering, nursing, and business. Assessment was scheduled for twelve classes in eleven disciplines, including 158 student presentations or writings. This time writing included the health fitness program's

culminating portfolios and the nursing writing assessment was repeated. Assessments took place between April 20 and May 10, 2007.

Time was again tight and the rest of the implementation did not go fully as planned. The rater training took place April 6 and 16th, using the videos. Meeting notes indicate that the training was more effective than the pilot training, and raters felt more prepared to work with presentations. There were five repeat raters among the thirteen. These experienced raters were assigned to courses where it was not possible to have a team when others dropped out or missed a class. The business school raters from outside the university were given direction from the assistant dean, and many of them previously had served as evaluators for the student teams. One assessment staff member also rated some of the teams, and spoke informally with the outside raters, who expressed belief that the rubric was a usable tool. No problems were reported by the assistant dean (Researcher Field Notes, April 2007).

Feedback sheets from this group of raters expressed generally more positive statements than the first group. Not all of them were participants so no attributions are made.

- I thoroughly enjoyed the process since it provided me with a rare but precious opportunity to observe students' performance in courses outside my discipline.
- Some of the projects are not well suited to the rubric, but is seems it was more a weakness in the projects than in the rubric (Assessment records, Rater Feedback File, 2007).

A theme of suggesting much more work with faculty on how to develop good projects and to help students do well emerged from this round and indicated an expansion of focus on teaching skills rather than student performance alone. All raters stated a willingness to continue to work with critical thinking projects (Assessment Records, Critical Thinking, Course Files, 2007).

Responses from Questionnaire II, given to study participants after the second round, also supported better implementation and the expansion of faculty viewpoints from assessment to teaching. Responses to question three "What went well, not so well?" typically included that the rubric was better, and for those who were doing it for the second time, relatively easy to use.

> The form continues to improve - became easier to use. (participant 12).
>
> Elements/Criteria are well defined, but the standards still need refining (participant 16).
>
> My colleague and I found it to be a useful tool for assessing the project within the synthesis course (participant 31).

In response to question five, "How would you describe the usefulness and effect of this program to another faculty member?" the same participant continues, "It provides great guidelines for the student, in our case, the site supervisor, and the faculty member to follow and assess."

An additional focus on teaching in responses to question five was shown by such comments as, "It has helped me consider how to frame for students how they will show

me what they have learned. It is prompting me to revise/scale back a major project" (participant 12).

"I think it would be worthwhile for teaching faculty to see other classrooms and observe how critical thinking is taught to improve their own course" (participant 21).

One of the fifteen responses indicated they had no idea what to say (participant 27). Only one negative response was given. Although this participant indicated a willingness to work with the critical thinking project two separate times, this time there was a strong negative response stating "that SEC (State Education Council) has caused a lot of expenditure of time and faculty effort, with a request that is baseless and not useful to the higher education of the community's youth" (participant 22). Unfortunately, there were no further comments from this participant in the open comment section.

Results of this round were communicated during the summer to professors and the reports included more information, for example, every student's score was provided as well as a comparison to the preliminary university score. Preliminary data for the university indicated that 76% of the students were rated highly competent or competent on all criteria. The highest rating was for problem identification, and the lowest was for limitations of conclusions, a subsection of the conclusions criteria (Assessment Records, CT Excel Data File, 2007). An overview of the preliminary data is in Appendix H (Preliminary Second Implementation Summary Data). Final analysis of the data was not completed by the end of this study. Neither was a final university-level report prepared since the assessment office will be combining these data with additional work for the next state requirement in 2010.

As this case study came to an end in February 2008, new assessment ideas were being considered, for example, combining the critical thinking assessment with either or both writing and oral presentation assessments (Assessment Office Records, Staff Meeting Notes, 2008). The next faculty working group is scheduled for early fall 2008. The working group might be said to be on hiatus; however, actions related to critical thinking were still continuing to move forward. The choice of the working definition, identification of criteria, and production of a cross disciplinary rubric by the Working Group were concrete results that passed into requirements and processes of the Critical Thinking in the Curriculum proposals and projects. The effects of the assessment, the university initiative for critical thinking in the curriculum, and, a planned review of general education assessment form part of the discussion items in the section on effects.

Summary observations

Some general observations on implementation can be made. For the second round, there was insufficient time to have individual meetings with all cooperating professors, and direction and information were completely by email. This is connected to how and when courses are selected - for example, in order to improve the student assignment directions, professors need to know they are involved at least a semester in advance. Deans are asked to suggest courses, and instructors of synthesis courses are invited to participate. This time the requirement from the provost office was that one course in every college needed to participate in the assessment effort. Individual professors also had to agree. The assessment office spent significant time on the

communications needed to set up the agreements from all six colleges with undergraduate majors. (Assessment Records, Critical Thinking files 2006.)

In both implementations, scheduling the raters and the courses took time and when changes in location or time happened, accommodating them was very difficult. In one case, a substitute rater was the single rater, and he had no forms to use, but created a matrix from memory and succeeded in rating the students. A department chair came close to refusing to allow raters from outside the department, even though the cooperating professor clearly understood that raters were not rating the subject area. The professor reported that the next time, raters from that department would be recruited (Researcher Calendar Notes, June 12, 2006). However, for the second implementation in spring 2007, a rater from outside the department was allowed in the class, perhaps because these students scored well on most of the categories in the first round.

Researcher memos (April 2007) recorded that not all professors shared the rubric in advance with their students. Most of the reason for this was the timing issue. The revised rubric and determination of which courses would be participating did not take place far enough in advance for professors to easily change syllabus or assignment parameters. In one case, a professor telephoned to say the rubric was not given out because students were nervous enough about presentations (Researcher Phone Log, May 12, 2006).

One overall statement about assessment of critical thinking can be made. MASU faculty are interested in the topic and are supportive of efforts to improve teaching and learning. What is not so clear is what may happen as a result of the formal assessment

not repeating until 2010. Some answers to that are in the section on research question three (What effects did these processes and decisions have on the institution during the time of the study?).

Any university embarking on an assessment initiative will be making decisions and choosing processes. Information about the way decisions are made and the processes chosen is an important piece of the case study. The following section looks at decisions and process at Mid Atlantic State related to the critical thinking assessments.

Question 2: What were the processes and decisions that contributed to the development of an assessment of this student critical thinking learning outcome?

This section looks at the decision points and processes that steered the direction of this assessment. Top university leadership decisions, the participation of staff from different university units, and intentional decisions to include faculty contributed to an organized approach to the assessment task, a positive step given the atmospheres of increased scrutiny from state bodies, accreditors, and the U.S. Department of Education.

Decisions relating to this assessment and surrounding context took place at several levels of action and work. As pointed out by the director of the Teaching and Learning Center, the assessment itself calls for attention by the faculty that may not have arisen without the state requirement. "I think the assessment will foster better teaching because people will be simply more aware" (Interview, Director TLC, April 23, 2007).

The decision diagram, portrayed in Figure 5, describes the decisions taken and how they related to one another. Observations from examining the decisions using this diagrammatic base follow below.

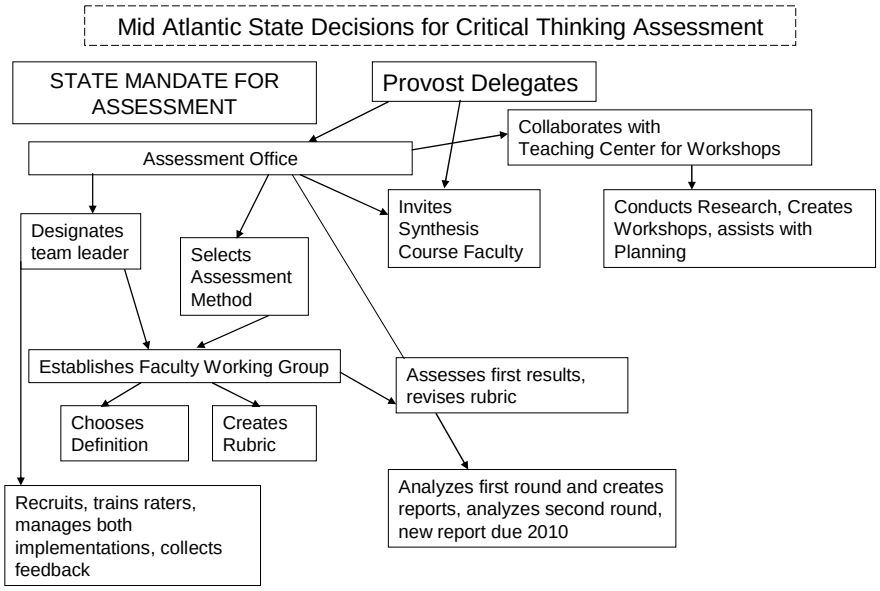


Figure 5. Decisions for assessment at Mid Atlantic State University.

The initiating mandate from the state was not in the decision control of the university and not in the scope of this study. It certainly fits with the increasing eye on assessment and student learning described in chapters one and two. The MASU Provost Office includes the assessment office and the Teaching and Learning Center, as well as units for enrollment, research, an international campus, global education, institutional research, and undergraduate and graduate education. It is not surprising, therefore, that in the case of assessment, the Provost delegated responsibility to the Director of the Assessment Office. He did express interest and expectations that effects of the critical thinking effort will be widespread and lasting. When asked about goals for the assessment and expected results, he responded, "The goal is to capture correctly the most significant contributions which vary with discipline. It is a very important outcome for

liberal education and one most valued by the community. That is in addition to the requirement." Later in the interview he added "I hope the assessment will guide the faculty better, and more systematically. I hope they will be able to use best practices and use it to inform their courses. I expect to see more student replies in exit polls that they have had more rigorous material. Current comments say they have not" (Interview, Provost, May1, 2007). Responses to questions about direct involvement indicated a decision to leave the details of plans and implementation to the assessment director. He did, however, cooperate in some details like sending original invitations to faculty and letters of appreciation from his office. The Provost's attention was given to the new initiative, Critical Thinking in the Curriculum (CTC), which "will help focus on assessment and spread information and help teaching" (Interview, Provost, May 1, 2007). More information about the CTC program is presented in subsequent sections.

Although the prime responsibility for constructing a critical thinking assessment was with the assessment office, the Teaching and Learning Center director had been participating in other assessments, and that collaboration continued. The TLC prepared materials for use of the working group and conducted joint workshops with the assessment office. The collaboration between these offices was not a specific new decision, but followed a previous pattern, going back at least to joint sponsorship of the Washington State critical thinking workshop in 2004.

The selection of the assessment method and designation of a critical thinking team leader in the assessment office helped to organize planning and anchored the administrative and management responsibilities. An agenda was prepared to present to

the faculty working group for validation. Faculty approval was important. "At that first meeting I think we asked them if we could do it across disciplines and use the synthesis courses. They really did think that was a good way to go about it. But they all did have a healthy skepticism, which is not surprising at all" (Interview, Director of Assessment, April 16, 2007).

Organizing administrative pieces for carrying out the plan were part of my work responsibilities, but efforts included all the staff in the assessment office. It was at this point of my research that I began to realize the group teamwork I was looking for actually was taking place in two groups: (1) The Assessment Team (Provost Office units of TLC and the Assessment Office), and (2) the Faculty Working Group. The discussion related to question four shows that the two groups worked very differently.

The faculty working group did make concrete decisions regarding the definition of critical thinking and the instrument for assessment. Most of them also served as raters or permitted assessment in their courses. This step (of needing the professor's agreement for the assessment in their course) was a direct influence of university culture and previous assessments that crossed disciplines, (e.g. the one in scientific reasoning). MASU is a university with an independent-minded faculty who want to be considered and to have a say in decisions about academic issues. Faculty cooperation is completely necessary to support the institution's goals for assessments that lead to improvements. "You have to start with faculty learning. They can't do what they don't know how to do." And: "That is a sort of faculty development by subterfuge. They don't think that they are doing it for their own learning, but they are" (Interview, Director of TLC, April

23, 2007). "If your ultimate goal is to improve programs, faculty have to be centrally involved" (Interview, Director of Assessment, April 16, 2007). Involvement of faculty meant that forming the working group, setting and preparing for meetings, and soliciting feedback from absentees took the efforts of all the staff of the assessment office as well as collaboration with the TLC. In answering question eight on the second questionnaire, "How might you characterize the process of carrying out this assessment?", eight of the fifteen replies specifically mentioned that the process ran smoothly, worked well, and had lots of faculty input. One participant explained.

> The planning, implementation, and reporting (yet to be finished) have been systematic and inclusive. In particular, I believe the university level committee is open to refinement and improvement, and is taking an appropriately collegial approach to this process. I believe this will be more effective in the long run due to the importance of gaining consensus among faculty (Questionnaire 2, question 8, participant 24).

One reply used only five words, "Effective, flexible and responsive to comments" (participant 12).

The availability of synthesis courses in the undergraduate programs was an advantage. These courses were designed to help students bring together the various learning goals of general education courses, and to demonstrate that learning in an integrated way. Without that option, discovering courses that were suitable venues for conducting the assessment undoubtedly would have taken much more time, and created even more difficulty for an approach involving the use of a single measurement. Another

advantage for the synthesis course venue was that any undergraduate could take any synthesis courses if the student had required pre-requisites. The assessed student population was not composed solely of majors in that subject. This potentially offered a picture closer to the make up of the whole undergraduate population. A comparison of the university population and those students assessed was not made because student ID information was not collected.

The second decision stream that interacted with the assessment of critical thinking during the processes described here is not pictured in Figure 5. That stream includes the workshops developed by TLC, one of which was the 2004 Critical Thinking in the Classroom described before. In addition, there was another workshop, February 28, 2007 offered at the same time as the working group was setting up the second assessment. This Critical Thinking Workshop was planned and presented by the TLC and the Assessment Office to guide faculty in developing assignments for teaching critical thinking. A startling number of faculty attended. There were 42 attendees from 15 different disciplines, a record for any workshop. Workshops regularly attracted between 10-20 participants. The 21 returned evaluations were highly positive. Only three of these included "not valuable" and 80% reported they were encouraged to include critical thinking in their teaching, while 70% reported now having practical strategies for doing so. One suggestion was to spend three hours, (double the workshop time) and after that have time to share examples (Assessment Records, Consolidated Feedback Sheets, February Workshop File, 2007).

The director of the Teaching and Learning Center (TLC) formed a proposal for Critical Thinking in the Curriculum (CTC) that would offer stipends to faculty to develop specific assignments for courses and programs. After the TLC director consulted with the director of assessment and presented the proposal to the Provost, the initiative was launched with stipend money allocated for a three-year span. This Provost decision gave a visible sign of the importance of critical thinking to teaching and learning. The launch was a cooperative effort of the assessment office and the TLC and one more step linking assessment, teaching, and learning. The intended model was the previous award-winning Writing in the Curriculum Program and stipend plans followed those in the technology and teaching award program. "CTC will be around for a while, and modeled on other programs - Writing, Technology and Teaching Awards" (Interview, Provost, May 1, 2007). CTC included a requirement for assessment. Requests for Proposals went out in spring 2007, with the first awards scheduled for summer. The RFP included the assessment definition and rubric and required awardees to share their projects and to assess student learning using at lest part of the rubric. Assessment and teaching were connected, and the goals of MASU's critical thinking assessment to induce change were visibly and strongly supported. This program offered potential for continuing both attention to critical thinking for undergraduates and for additional assessments, regardless of the state demand or schedule. Awards ranged from $500 for individual course projects to $2000 for team projects affecting large numbers of students. An example of these awards is the one to a team of scientists to create modules for critical thinking for all the introductory general education science courses. By January 2008, 13 projects involving

24 faculty had received stipends (Assessment Records, CTC files, 2007). The benefits for students' critical thinking abilities are still to be determined since most projects were developed in spring 2008, and then instituted in courses. First reports are due at the end of the spring 2008 semester.

It seemed that the goal of involving faculty was achieved and that effects from the assessment and from the CTC had begun and might be wide and deep. The total effects were not expected to be fully evident right away. Other researchers point out that "the apparent ability of the assessment *process* [sic] to bring about change was one of the unexpected outcomes of the assessment movement" (Banta, T. & Associates, 1993).

The director of TLC suggested that three to five years may be necessary for change to take firm root (Interview, Director TLC, April 23, 2007), and the Associate Director of Assessment said, "It takes a long time. It is really curricular management. You think about it, like information technology, it takes five years" (Interview, May 16, 2007). The conversation and action have begun, and the requirement that the working group select a definition and develop a rubric usable in many disciplines created, at the least, a common vocabulary and discussion points. "Having the CTC initiative start a year after the assessment process has given the assessment more legitimacy in the eyes of many faculty members and the assessment instrument can now be used to evaluate our CTC efforts" (Questionnaire 2, question 2, participant 4).

Processes for this assessment included difficulties, mainly centered on the use of the rubric, good training, and lack of direct interaction with the second group of instructors. As seen in previous discussion, a few faculty members reported strong

negative evaluation of the whole effort. Involvement of faculty at each step drew out the process, and provided some frustrating discussions. "The down side of faculty involvement is that it just takes a whole lot longer. You've got to have agreement across faculty, across disciplines, shoot, even just finding meeting times, you know." and "Staying on task, faculty have so many things that they have to do. Many of them do [participate] just because they really care about the students" (Interview, Director of Assessment, April 16, 2007).

The workload on the assessment staff is also a consideration, and was a concern of the assessment director. In reply to the question about difficulties of this longer, faculty oriented process she replied:

> So far we haven't seen this, but the potential for burn-out of the people working in this – because there are so many demands from so many different sides – it is important to maintain our morale too. It's not because of the work but because there is so much of it. I think it can become an issue. So it is important to keep people rejuvenated and refreshed. (Interview, Director of Assessment, April 16, 2007)

Sustaining effort in critical thinking is a factor for the faculty and the staff. The decisions for faculty connections and for collaboration with the TLC have the consequence of additional work and a need for yet more time commitments.

Another set of ways to look at the decisions involved in this assessment add deeper understanding. Comparing the MASU efforts with two points of view (1) the garbage can model (Cohen et al., 1972) and (2) ideas about rational decision making

(Allison & Zelikow, 1999) adds additional understanding of what happened. The decisions and processes used at MASU in this case better demonstrate the rational decision model, which included Allison and Zelikow's organizing concepts of selection of goals and objectives, options, consideration of consequences, and making an intentional choice. The Provost, head of assessment, and director of TLC had goals in mind; selected options they thought would lead to those goals, and then made intentional choices. It seems clear that the decisions of how to carry out assessment and the new initiative were deliberately taken.

In contrast, the garbage can model states that streams of choices, problems, solutions and energy meet, mix, and result in some output, not necessarily related to organized, rational processes. Cohen, March and Olsen (1972) suggest their model is common in education. In this model, decisions are the result of complex interrelated elements that are usually not well understood or controlled by the decision makers, and connections are made mostly through happenstance. The evidence does not show that this model describes the MASU case. The Mid Atlantic assessment case does not look unintentional or serendipitous. From one decision spot to the next, intentional choices were made, and when two sets of opportunities did appear (i.e., the outside requirement for assessment and the CTC proposal), they were intentionally linked. What neither model indicates, however, is whether the eventual results and effects will be those intended.

General Observations

The several layers of decisions regarding the critical thinking assessment at MASU were initiated by an outside decision of the state to require competency assessments. Institutional responsibility resided in the Provost Office, which included the Assessment Office. The delegation for planning and conducting the assessment was made to the Director of Assessment. The Director's decisions on how to proceed were based on firm beliefs in assessment as a tool for improvement of programs and courses, in reliance on previous faculty involvement, and on assessment that took place in courses, and was related to authentic student work. Decisions on the operational definition of critical thinking, on criteria, and on the rubric were part of the faculty-based approval and faculty-developed assessment processes. During the same time period, the Teaching and Learning Center director initiated discussions and faculty development workshops on critical thinking and proposed a curriculum-wide initiative. The provost decided to support that with funding for faculty critical thinking projects. This initiative and the assessment process merged through joint use of the definition, criteria and rubric. When the decision steps are traced and compared to two models, the garbage can model and the rational decision model, MASU's steps reflect the intentionality of the rational decision model. Indications of what effects the series of decisions may have are discussed in the next section, addressing effects of the assessment effort on the institution.

Question 3: What effects did these processes and decisions have on the institution during the time of the study?

Normally it is very difficult to find out what the effects of assessment are beyond the resulting reports. Detailed data are not usually collected to provide direct evidence, even if there are effects in courses, in programs, among the faculty, or in the overall institution.

> It sometimes can be hard to document the actual changes that have occurred. And one of the reasons for that is that frequently because we are embedding this, people don't really see that they have made changes related to assessment, they see it as just part of the way they are thinking about curriculum or reflecting on a course or whatever…I mean I feel like we need testimonies, and its hard to get them because people say, well, we were already re-thinking, we are not going to use that text, or we were going to change how we did this (Interview, Director of Assessment, April 16, 2007).

Expectations for critical thinking assessments effects at MASU were not explicit. Palomba and Banta (1999) describe a situation like MASU:

> As a result of assessment efforts, many institutions can report specific changes based on assessment findings. Courses have been modified, added, or dropped. New programs have been introduced. Resources have been provided….It pays to remember, however, that not every assessment project will result in a specific change. Many times the impact of assessment is more subtle and is felt over time (p.328).

Expectations by MASU, although not always explicit can be described as follows.

1) Expectation of student success on faculty-developed, course-embedded assessment, based on prior experiences: No student benchmarks were implied or expressed. However, previous state reports had shown competency levels for other assessments to be high. For example, students were rated as 81% satisfactory or higher in writing; 91% highly competent or competent in oral communications; and 92 % acceptable or highly acceptable in scientific reasoning. These assessments are most similar to the critical thinking process, and so an implied expectation is for critical thinking to have somewhat the same results. (MASU assessment website reports, 2007).

2) Expectation of a rise in student statements about course rigor in the exit and engagement survey: This was an explicit expectation of the provost. Those surveys are not yet processed and new engagement surveys are taking place in the spring 2008 semester. Results will be available on the assessment website.

3) Expected growth of faculty interest in critical thinking and its importance in the curriculum. No administrators or faculty described what "benchmark" or level of such interest was expected. Interest during the time of this study was primarily from attendance records at CT workshops. Unfortunately, there was no follow up with attendees to see if that interest was sustained. A sign of a downturn was the last workshop attendance which was less that 25% of the previous one. (Assessment records, Workshops, 2006)

4) Expectation of changes in courses, assignments, and perhaps even programs as a result of focusing on assessments: This expectation came from prior experiences with

other state competency assessment. As in statement one, the reports for writing, oral communications and scientific reasoning all include descriptions of course changes, and there was no reason to not expect the same in this round. When coupled with the CTC awards as incentive, one can conclude that more change was expected from the critical thinking efforts, and that it might also impact other competency areas. For example, because CTC grants were given to science and to English professors, connection to writing and science areas may be valid.

The required competency assessments are still relatively new, and collection of data about effects has not been highly developed. The state in this case has not set targets or even commented on the final reports of the assessments. One way of thinking about expectations is that it is a university determination, and one that is not so carefully spelled out.

Results and effects can also be viewed as cultural changes, perhaps the most difficult changes to identify while they are happening. Assessment can be part of this type of change and the provost at MASU refers to this more holistic view when he stated his expectations in terms of seeing student exit surveys report increased perception of rigorous courses, particularly in general education (Interview, Provost May 1, 2007). Neither the director nor associate director of the assessment office suggested any particular target for the first- or second-order changes Cuban (1999) describes. No benchmarks or guides for either first-order (quality improvements for basically sound operations) or second-order (alteration of fundamental operations because of major

dissatisfactions) institutional change are clearly offered in either the literature or data from this research.

That there are some effects is evident (see Figure 6). What we do not have is definitive yardsticks for either expectations or for when such individual effects will result in substantive or long lasting change of either the first order or second order. The two potential guides offered by Gladwell's (2002) tipping point and Rogers' (2003) routinization are analysis tools that illuminate the change process but cannot pinpoint the instant of change except *post facto*. For example, the tipping point is the moment of change, but cannot easily be observed at its dawning. Rogers' diffusion process is dependent on communication channels, time and the social system. Rate of adoption, he says, can be measured by the number of people in a time period that adopt the innovation. However, there is not a single standard period or number that applies to all situations. Using Gladwell's ideas of attention to who those adopters are and how many others they may bring along, and Rogers' routinizing, do guide us somewhat. Both ideas look at momentum and at adoption by significant entities or persons as necessary precursors to change. In the case of MASU, we do have sight of momentum from the CTC program, the change to the business capstone course, and the use of the rubric in new faculty orientation and several faculty workshops. What cannot be said at this juncture is that since we have *x* number of changes and uses, we therefore have tipped or routinized so that critical thinking is thoroughly embedded in teaching and learning at MASU. However, we can tentatively expect that in the next year or two these and other course or assignment changes should be evident in student work, in syllabi, and in those surveys

the provost looks for. Since the university is in the process of selecting an undergraduate improvement plan as required by the accreditation process, it may be that an organizational second-order change will include critical thinking and will speed up tipping and routinizing. For example, at MASU the writing program "tipped" when every undergraduate major was required to have writing intensive courses with specific amounts of feedback to students. This requirement, agreed upon by the administration and faculty, was the routinizing moment for writing. If critical thinking follows a similar process the tipping and routinizing may arrive more quickly than if it grows on its own from the seeds already planted.

Deliberate efforts were made to collect effect information. The participant questionnaires, interviews, meeting notes, and records in files were some of the means for accomplishing this objective. In addition, researcher memos and a phone log were kept so that even informal conversations could provide effect information. It must be noted, too, that effects can be expected to accrue over a period of time. For instance, it will be at least two years before any changes in student exit surveys about academic rigor are likely to be visible. The general education assessment will not be completed for two years, and that again means that student change will not be apparent until a later point in time. For these reasons, it is difficult to tell at this time what the final effects of any changes indicated here will be. It is nevertheless informative to look at what is happening even if we do not know for sure when a tipping point for demonstrated student improvements will be reached.

The Figure 6 chart, Types of Reported Effects of Critical Thinking Assessment, was a first step that collected the reported changes by types. Sources are the two participant questionnaires, administrator interviews, assessment office reports, and researcher memos and logs.

Assignment Changes	Course Changes	Program Changes	Institutional Changes	Reported Personal Change	Other Changes
11 reported *Examples*: Culminating project revised in course on non-profit management Linked IT and English course assignment modifications to include critical thinking Biology assignments for general education large classes revised	10 reported *Examples:* New synthesis course proposal for science Accounting course re structured to focus on thinking processes as well as accounting methods Music teaching course to include critical thinking elements and in specific assignments	**Business** Senior Case Competition uses rubric **Individualized Study** Senior Project content changed, & research course content improved **Academic Program Review** Includes CT information in learning outcomes training	**CTC** includes assessment piece **New Faculty Orientation** CT information and rubric distributed **More collaboration** additional workshops developed & offered **Trustee interest** in Critical Thinking continues **Community Building** through increased faculty interaction	Teaching philosophy changed 3 reports Raters gain ideas to use themselves 12 reports New ways to look at teaching 7 reports	2 requests for CT assessment in fall semester courses Request for assessment staff to participate in senior course discussion of critical thinking CT material sent in by a faculty member to assessment office Nursing professor studying CT and writing

Figure 6. Types of reported effects of critical thinking assessment.

Not every source included detailed information and intended projects may or may not have been carried out. Regardless, the interjection of critical thinking as a student

outcome into the thinking process of instructors has value, as two of the raters note effects of participating in the assessment. "It has already affected how I write assignments and ask students to think/write abut material in my class. I'm thinking about how to use it in a group project" (Questionnaire 1, question 8, participant 12).

> It has definitely affected my teaching. Since I have become interested in this project, I have been asking myself every lecture I give if I send students some clue on how to approach a topic with a critical mindset. A colleague and I are in the process of designing a synthesis course for which we want to make critical thinking the rationale that holds the course together. (Questionnaire 1, question 8, participant 8)

A problem with this information is that not enough details were collected on the assignment, course, or program changes, and no provision for further tracking was been made. At least a tentative conclusion can be made that changes recorded are broad and diverse, and may be more evident over time.

Documented effects of CTC projects will be coming in as the projects are completed. However, because the added push from assessment will not be there until 2010, how far or deep any of these changes reach is impossible to know. It is possible that the network of faculty who participated in the assessment, attended workshops, and proposed and received CTC awards will trigger a tipping point (Gladwell, 2002). Effects may only be discernable after that happens. No current way is available to tell, but Gladwell suggests that "By tinkering with the presentation of information, we can significantly improve its stickiness. Simply by finding and teaching those few special

people who hold so much social power, we can shape the course of social epidemics" (p. 259). The idea of creating an "epidemic" of critical thinking through use of better information, continued outreach, and work with key faculty is an intriguing way to look at sustaining what appears to be substantial current faculty interest and willingness to engage with the improvement of critical thinking among undergraduate students.

In contrast to this optimistic view, there are several instances pointing to factors that may lessen the impact of the assessments and slow initiation of new effects. The faculty interest evident in prior workshops did not hold up for the October 30, 2007 CT workshop, which included presentation of one of the CTC projects. Normally workshops showcasing faculty attract more participants. Only ten faculty attended, compared to the previous one with 42. Causes might be different advertising or a true decline in interest. The number of proposals for CTC declined with only two submitted for the second summer awards. This may be linked to the normal faculty view that summer is the time for their own research and time spent on course development is not a priority (unless funded up front as is the case with many outside grants). What happens in the next semesters will be important and perhaps determine whether there is in fact sustained change in concrete course assignments, in programs, in teaching, in student learning, or in institutional culture. It may be, as Khademian (2002) reminds us, that changing the culture is the most important factor, and the tipping point and routinizing will not be evident for some time.

General Observations

The tale of the effects on Mid Atlantic State from this assessment is incomplete from two points of view. First, although attempts were made to collect information about changes from questionnaires, researcher notes and logs, the data is incomplete and may represent fewer effects than reported, or since there is as yet no follow-up, reported changes may not have taken place. Faculty do report that their participation either as part of the Working Group, as raters or as course professors resulted in positive experiences and increased awareness of critical thinking as a topic of student learning. Specific change was most evident in the business school connection to critical thinking in the senior project competition, and to the launch of the Critical Thinking in the Curriculum initiative. That initiative was tied to the definition, criteria and rubric of the assessment, and demonstrated connection of assessment, teaching and to eventual student learning. However, changes in learning were not evident during the period of the study. Embedding critical thinking teaching, assessment and successful student learning as described by Rogers' (2003) concept of routinizing and Gladwell's (2002) tipping point for change takes longer than the temporal frame of this investigation. It does seem that the process has begun and is likely to continue due to the CTC projects and the upcoming state requirement for additional assessment. Further examination of the critical thinking assessment at MASU includes using four concept models to analyze what has been reported so far.

Question Four: What models of assessment, critical thinking, group work, and institutional self-renewal help us understand this case?

Models of assessment, critical thinking, group work, and institutional self-renewal shed more light on the MASU effort and lead to some conclusions about best practice and an integrated model. The previous sections described and analyzed what happened during the MASU assessment, the decisions and processes, and effects of the assessment efforts. Use of the concepts presented in chapter two makes possible further understanding. The following sections use these four concept models to look deeper into assessment, critical thinking, group work, and institutional self-renewal and result in some conclusions for practice and potential research.

Assessment Model Discussion

What can two models from assessment best practices, Banta' s (2004) hallmarks of good assessment and the National Research Council standards (2001), tell us about the MASU assessment? Banta's hallmarks of good assessment relate to the action of carrying out an assessment and include three major areas: (1) planning; (2) implementation; and (3) sustaining. Each area contains subcategories.

Planning includes:

- involvement of stakeholders, including faculty, administrators, students, and community professionals
- allocation of sufficient time
- a written plan with clear purpose
- connection to program objectives

Effective implementation hallmarks address:

- leadership
- university wide and unit responsibility
- faculty development
- use of multiple measures
- assesses processes as well as outcomes

The hallmark for improving and sustaining assessment includes:

- credible evidence
- supportive environment
- continuous communication
- use of data
- demonstrated accountability
- on-going assessments
- celebration of successes
- evaluation of the assessment process

Examination of how these hallmarks illuminate the assessment of critical thinking at MASU in the timeframe of this study presents a clearer picture of the answers to the previous research questions, and provides an "assessment of the assessment".

In the case of planning, MASU exhibited all four hallmarks, although the involvement of stakeholders in planning did not include every group mentioned. Planning did include faculty and administrators, but not the students. Community members were not directly solicited, but interest in the topic was referred to by the Provost. The extension of the assessment to business senior case studies did involve

business community members who used the same rubric for judging the competition.

The assessment office followed a general plan for this assessment that had been successfully followed for other required competency assessments. These plans were very oriented to faculty involvement and course-embedded methods. The time factor listed in Banta's hallmarks did impact planning that might have been more thorough, especially regarding work with the course professors. There was a written plan, Appendix G, which was submitted to the state, and included rationale and methods, but did not include any specific target because this first plan was designated as a pilot. Because improvement of the undergraduate critical thinking goal was related to the institutional mission as well as a state competency requirement, the connection to programs was through university-level goals rather than individual programs. Planning according to these hallmarks fell short on involvement of community and students, and on clarity of a target goal for student achievement (as a result of being a pilot for testing the type of embedded assessment chosen by MASU). The written plan itself did elicit outside commendation. The state process called for review of assessment plans by peer institutions, and the plan reviewer for the pilot applauded the links to course experiences, the definition of critical thinking, the match between that and the rubric and the reflection of serious faculty commitment. (Assessment Records, Pilot Plan File, Peer Review Comments, 2006). It is possible that the next planning in fall 2008 will consider the questions of wider involvement and a stated target of achievement

Of the five Banta components of the implementation hallmark, four were met, but university-wide assessment and unit responsibility were not. Use of the Banta element of

leadership to look at decisions and indications of leadership enhances the discussion earlier in this chapter. The decisions made by the Provost, Director of Assessment and Teaching and Learning Center Director helped to involve faculty from every undergraduate college and to encourage participation in the assessment and the new Critical Thinking in the Curriculum initiative. Leadership was also evident in the links between the Assessment Team and the Faculty Working Group.

Although there was at least one course from each college teaching undergraduates in the assessment, it was not a university-wide effort. Not every synthesis course participated. Nor was the plan designed with the intent that units – individual colleges or departments – would have responsibility for the assessment implementation. Although faculty and course professors cooperated, responsibility remained with the assessment office. As seen in Figure 6, chapter three, indirect faculty development took place among those involved in the implementation and was further tied together with assessment through collaboration of the assessment office and the TLC in both the working group and workshops.

The recommended use of multiple measures for an assessment was met through use of the student exit survey questions about student opinion regarding the degree of MASU's contribution to their competence in critical thinking. Students reported a mean competence of 3.41 on a four point scale (not competent →very competent) and 90% rated MASU's contribution to that competence either contributing very much or somewhat (Mid Atlantic State University Assessment Office, 2008). This information was an indirect measure of critical thinking and complemented the direct assessment.

The next element, assessing the process and outcomes of the assessment is shown through feedback from raters, input from the Faculty Working Group, and revisions made from the first to second round. These revisions included rubric changes, improvements to rater training, increase in participating courses, and increase in communication of results. Full assessment of the second round has not taken place by the close of this study, but is planned as part of the tasks for the working group in the fall of 2008. No implementation is likely to be perfect, but MASU's use of a faculty working group and collection of feedback for evaluation were efforts that led to improvements. In addition, the assessment team and the working group will continue at the next fall meetings to consider additional improvements. The critical thinking assessment at Mid Atlantic demonstrated all but one of the implementation hallmark elements. There was a deliberate decision to place responsibility (the missing element) in the assessment office rather than in units at either college or department levels as the Banta hallmarks encourage.

It may be too early to see all of the elements in the improvement and sustaining hallmark. The evidence from assessment of student presentations is credible, but ensuring that selection of course participation is not just voluntary would improve the use of results as a university-level measure. Although what a "supportive environment" means is not well defined by Banta, support for the assessment effort was spread from the Provost to cooperating faculty. By linking assessment and a curriculum-wide focus on critical thinking through the CTC project support, the Provost has created a supportive environment by providing dedicated resources. Communications explaining and supporting critical thinking were continuous from the assessment office to the trustees,

Provost, MASU community and particularly to faculty throughout the time period examined. Presentations, published reports, workshops, and active work group meetings were parts of the assessment process. In addition, the professors received individual reports on their students. The first three elements of the improvement and sustaining hallmark were met.

Determination of data use is a bit blurred because all the changes happening in assignments, courses and programs may not have been reported. It might be helpful to think of the effect data as part of communication and discourse, rather than attempting to link specific change *x* to data element *y*. However the previous section reporting on effects of the assessment process did document some changes related to the assessment, if not formal causality (See Figure 6, chapter three). It is possible that coming work in the assessment office on general education assessment plans as well as reconvening the Working Group will result in more direct use of these data. For example, a nursing professor is working with the data for nursing to study connections between critical thinking and good writing in the discipline. Assessment of critical thinking is continuing through student surveys, in the CTC projects, and by preparing for the next required state assessment. In the periods between assessment activities, the Provost and Director of the TLC are continuing CTC project support and workshops for faculty.

Consideration of the accountability element of this hallmark requires looking at two kinds of accountability. One is the institution's accountability to the state which has been met through the report to the state which is in Appendix G. The second is accountability within the institution. The trustees are asking for reports, and the Provost

and Director of Assessment are responsible for those responses as well as the state report. A piece of accountability may be in administrator annual evaluations. Accountability was evident in the Assessment Office's openness to improvements. Determination of accountability of faculty for success of students' competency in critical thinking is part of another university area, (namely faculty evaluation) which is outside the assessment process itself. Faculty evaluators do have access to published information developed by the assessment processes.

There are no direct celebrations of successful assessment; however, there is a connection of this particular assessment to awards for critical thinking such as the Banta hallmarks call for. We could consider the CTC projects awards as a celebration. One of the large initial team awards went to a group which included one of the participants in the assessment. The only acknowledgement of participation in this assessment was in the form of an individualized letter from the Provost and these letters may be used in teaching portfolios or tenure and promotion documents. Evaluation of the process by the assessment office was described earlier, and any office staff success does figure in staff evaluations, and is also a topic of regular staff meetings. It may be that the recent budget restrictions at MASU on things like faculty lunches (which did occur in the past) reduce the opportunity for celebratory events.

Perhaps if one graded the assessment processes using the Banta hallmarks for planning, implementing, improving, and sustaining, this MASU effort might receive a B because of some missing or incomplete elements. MASU data do show that seeking

improvement is not only the topic of assessments but also part of the assessment process itself and was present in both planning and implementation.

Another, very different, look at assessments is offered by the National Research Council (2001) in its report "Knowing What Students Know." The report gives three essentials of assessment:

- a model of student cognition
- a set of beliefs about the kinds of observation that will inform
- an interpretation process

The critical thinking assessment process clearly met the first criterion, which the report explains is to "identify the set of knowledge and skills that is important to measure" (p. 44). In addition to the state mandate for assessment of critical thinking, MASU's efforts included selection a definition of critical thinking and, through rubric development, which skills and abilities to look for in student work. By connecting the Facione (1998) definition,

> We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference as well as explanation of the evidential, conceptual, methodological, criteriological or contextual considerations upon which that judgment is based. (1998, p.16)

to the rubric criteria and standards (Appendices E and F), the Assessment Team and Working Group fulfilled this first NSC standard.

The second part concerns the beliefs and choices about how to observe those cognitive elements. The decision to use student presentations of culminating projects in the synthesis courses was based on the requirement of those courses to show integration of general education coursework. In addition, some student writing was analyzed with the same rubric. Observations were carefully designed, although carrying out the observations using rubrics contained some problems traceable to rater training and student assignments. Improvement for this second NSC standard would include closer work with the course professors and some additional enhancements to rater training. As reported earlier in this chapter, some raters still express concern about the use of presentations; however, as the assessment progresses into the next round that issue will likely be addressed again, particularly as consideration of merging CT assessment with other competency assessments in writing and/or oral presentation is examined (Assessment Records, Staff Meeting notes, January 2008).

The third standard addresses interpretation. Interpretation of those classroom observations turns them into usable evidence. So far, the pilot round at MASU has not fully turned into evidence of student competency, because the pilot focused on establishing the usefulness of this kind of assessment and if it would in fact work substantively and logistically. The first steps of communicating results have taken place, and the publication of the two *Focus on Assessment* (2007, 2008) reports began interpretative steps. The assessment office and Working Group did not complete evaluation of the second round before the end of this research study, so a conclusive

result matching the third element in this model is not yet possible. The first two standards were well met.

MASU critical thinking assessment can be rated very good or good. Of the 17 elements in Banta's (2004) hallmarks of good assessment, the Mid Atlantic effort met 12 of them fully. Use of data, celebration of successes, multiple measures and connections to program objectives were partially met, and the element of unit responsibility was not met due to centralized responsibility in the assessment office. Two of three standards of National Research Council were met, and indications are that the third will be met in the next semester.

It is clear that in order to achieve high levels on the two sets of criteria connections to other university practices and policies must be made. Some examples that might assist in critical thinking assessment improvement are a focus on critical thinking in all general education courses and stronger faculty rewards for participating in assessment. A common assessment statement is that good assessment completes the circle, which moves in a continuous cycle of plan, implementation, review, and begin again. Both of the models used here address that pattern and MASU has expressed commitment to continuing the process through CTC projects and advance planning for the next state required competency assessment. Assessment of critical thinking appears to be alive, pretty healthy, and ready to grow more, but will need continued nurturing on its way to becoming a routine part of the university experience.

Critical Thinking Model Discussion

The first issues facing development of this assessment were deciding on what critical thinking means. An operational definition for an assessment process is necessary (Palomba & Banta, 1999). A clear, agreed upon statement of what is being assessed must precede any assessment, and is necessary for locally developed processes. Standard tests, for example are usually very clear about what they are testing, and what kinds of results will be produced. Without clarity for MASU's effort, no common conversation could ensue and the assessment might be lost in a quagmire of well-I-know-it-when-I-see-it opinions. As presented in chapters two and three, there is a continuum of thought on how to define critical thinking, and there are diverse opinions on whether it can be characterized in the same way in different disciplines. The work from the Washington State University project leaned toward discipline-specific definitions, but provided a common set of ideas that were adapted by the disciplines. Those common themes became a starting model for the MASU effort. Significantly, because the state required a university-wide assessment, some common definition *had* to be chosen, one that could fit the different disciplines at MASU. The task of choosing the definition and then developing a rubric fell to the Working Group of faculty. The choice of the definition from the Delphi Group's work (Facione, 1990, 1998, 2006) was acceptable to the Faculty Working Group, which was composed of eleven faculty members from multiple disciplines and colleges.

> We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference as well as explanation

of the evidential, conceptual, methodological, criteriological or contextual considerations upon which that judgment is based. (1998, p.16)

The assessment office sought faculty approval for conducting the assessment in the general education capstone synthesis course as the venue for the assessment, and to use student presentations of final synthesis projects. Initially skeptical faculty did agree to try and those already teaching synthesis courses volunteered their classes. The next step was developing the rubric to be used. The first version is in Appendix E, and the revised second - round version is in Appendix F. The development of the criteria elements for the rubric, and particularly the content of the explanatory boxes took a long time, from the middle of the fall semester 2006 to early in the spring semester 2007. During these discussions and emails, the faculty came to further agreement on how to identify critical thinking rather than presentation skills. For example, the section from the original rubric on references and sources required a great deal of discussion to get to the wording.

Resources/source materials	Appropriate and relevant (number & content)	Most resources/source materials are relevant and appropriate	Few resources/source materials are relevant or appropriate

(Appendix E)

Since there is no standard definition or evaluation process, and because the state left both of these decisions to individual institutions, MASU carried out a systematic and faculty-based development process to determine these two key assessment tools. For this assessment to proceed there had to be agreement on how to begin, and the process achieved that. Neither the assessment office nor the faculty indicated that this is the final

answer, but it is one they are continuing to examine and work with. The synergy of collaborative work in faculty development by the assessment and TLC units contributed a great deal. Multiple critical thinking workshops spread the word and developed interest even before the assessment began, and helped to sustain it. When a definition and rubric can be applied to student work in a dance seminar, a chemistry research course, and business case competition (among others), a kind of success for the tools is implied and introduces common ground for discussion. The consequent validity that Boud (1995) regards as important was present, and in Palomba and Banta's terms (1999), MASU was successful in developing and using an operational definition.

Group Work Model Discussion

Contrary to the expectation that this research would focus on one group and how it achieved the task, early review of interviews and assessment files indicated that there were two key groups at work. Although linked closely and equally necessary for completion of the assessment, they operated in different ways and had different roles. The Assessment Team (Assessment Office staff, TLC Director) operated as leader, planner, manager, recorder, analyst, and disseminator for the process, and the Faculty Working Group (professors) functioned as a collaborator, expert advisor, evaluator, and approver of the process. The first is designated as a team because it addressed continual and continuous issues of university wide student competency assessment. These two terms used by Jennings (2007) are useful in this context. Continual refers to work or issues that re-occur over time, and continuous refers work or issues that are constantly present. Jennings uses the terms to designate differences among types of group work.

The Working Group was formed to address this particular assessment for advice and consent in methods and implementation, to focus on a specific but continuing task, and then worked intermittently on that task, and did not set its own agenda. In Jennings' view, neither group is a committee which he characterizes as attending to general tasks, and not oriented to action, nor are they a task force, disbanded at the end of a very specified task.

The key links between the two groups were forged by leadership and management efforts from the Assessment Team. The work of the groups complemented one another and these complementarities fostered another insight. Research, preparation, and workshops – all activities outside the Working Group – were necessary precursors to achievements at the meetings. I was led to consider a new idea of what composes a "workspace". If the workspace is conceived as more than a physical space, more than the sum of agendas, meeting notes, and concrete records, then leadership, planning, formal and informal communication, and research, among other factors, are valued as contributing to the process. Very real work of many kinds took place outside of the hours people in either group were in a room together. Effects of that part of a workspace are seen when the face-to-face portion of the work is then able to reach conclusions and take action. Those effects were evident in data from assessment records examined for this case.

The next question is whether these were successful groups. Concepts of group work were described in chapter two. The Success Matrix in Figure 7 evolved from the Elements of Successful Group Work in Table 1, chapter two. Topics that were

mentioned more than twice were retained in the Figure 7 matrix. The first level includes those mentioned three or more times, with Goal Agreement, which occurred five times (the highest for any factor) listed first.

First level

Elements of Successful Teams	*Assessment Team* Director, Associate & Assistant Directors, Assessment Office and Director, TLC	*Faculty Working Group* Eleven faculty from undergraduate colleges
Goal Agreement	******Main goals set by State requirements, including required plan, Appendix F, central goal to effect change set by team.	******Confirmed definition, measure and methods.
Group Design & Member Competency	*****Composed of Professional Administrative Faculty, with demonstrated expertise.	Composed of Mixed senior, tenure track and term faculty.
Resolution of Conflicts	******Minimal conflicts, which were worked out quickly.	******Included open discussion, continuous revisions, regular solicitation of input.
Collaboration	******Assessment and TLC worked well together, in turn, assessment team linked to Faculty Working Group.	******Faculty remained interested in one another's views, cooperated with assessment team for implementation.

Second level

Leader's Skill	*****Director of assessment made goals related to course and program improvement very clear, delegated, worked with TLC director, valued everyone's input, example of collaborative leadership.	# The Faculty Working Group was lead by the assessment team, although different members made stronger contributions from time to time i.e. Nursing, Business, & Science faculty.
Organization & Process	******Planned and organized both the team and the working group for implementation and evaluation of results.	*****Through cooperation during the face to face meeting, and during implementation faculty were organized in carrying out and evaluating the assessment.
Reach Conclusion	******Reports given to professors, state, and published in *Assessment Focus* and on line.	******Accomplished tasks in time to conduct the assessment.
Accountability & Rewards	***, #** The assessment team was accountable through regular work evaluation, but there were no rewards other than intrinsic satisfaction.	**# , *** The Working Group was not accountable in any concrete way, participation and effort were voluntary. In addition to intrinsic rewards, faculty receive a formal letter of thanks from Provost and had the option of participating in the CTC proposals & awards.

Key: ******Highly present,***** Present, # Not Present

Figure 7. Success matrix for group work.

Two of these factors stand out in the data from this study. Goal Agreement and Collaboration were evident in interviews, questionnaires, and records. Both the Working Group and the Assessment Team agreed on what the work goals were and, worked toward completing the assessment. Collaboration within each group and between them was the glue that united the process.

Another way of looking at group work includes the concept of punctuated equilibrium (Gersick, 1989, 1991, 1998). Gersick's concept, borrowed from evolutionary biology, proposes that groups work in a non-linear fashion, and the work is not continuous, but separated by periods of inactivity or balance that may quickly be energized, only to wind down again. The Faculty Working Group seems to have operated in this fashion, as face-to-face meetings accomplished goals, but activity between meetings was minimal.

The Assessment Team did not reflect this concept because their work was continuous over time. Several factors may explain this difference. As seen in the discussion above regarding the roles of the two groups, the Working Group tasks were more focused on advancing the agenda through collaborative meetings, and the Assessment Team's roles included more of the workspace activities outside of face-to-face meetings. Joint agreement on goals and enduring collaboration formed the backbone of this assessment process. The information on group success in Figure 7 also indicates the organization and planning by the Assessment Team contributed to the Working Group's ability to reach conclusions. The concept of punctuated equilibrium does apply

to understanding of the larger workspace, and how these two groups contribute different roles and kinds of actions to one process.

In summary, there were two work groups, with different but complementary roles. Understanding how the two worked together is enhanced by thinking of a larger workspace, and considering process beyond face-to-face meetings. Both groups can be considered successful in terms of almost all of the success matrix elements. Elements marked Not Present, (Working Group- Leadership, Accountability, and Rewards) apply to the Working Group and are results of interaction between the two groups, and decisions made outside the Group's control. The question of accountability is different for each group because accountability resided solely in the assessment team. The low rewards element relates to both groups and is one that could be addressed in the future, as improvement in rewards might help to sustain work on critical thinking.

Applying the success matrix in terms of MASU's assessment shows that the idea of punctuated equilibrium helps to clarify differences in roles and work between the Assessment Team and the Working Group. Are there other views of this assessment process that inform deep understanding? If we move the focus outward there are other lessons potentially useful to the institution. The topic of how this process can inform further ideas about institutional self - renewal is examined in the following section.

Institutional Self-Renewal Model Discussion

The atmosphere of accountability in higher education today encourages institutions to consider ideas of self - renewal and to proactively develop self-accountability methods. Both ideas imply the ability to change and to know when change

is needed. Examination of the MASU process for critical thinking assessment leads to some potentially productive pathways. Accountability to the state for student competency in critical thinking coupled with internal ideas from the TLC offer three potential lessons, in culture and collaboration (Kezar, 2001, 2005, 2005a), innovation and change (Gladwell, 2002; Rogers, 2003), and corrective feedback loops (Birnbaum, 1998). The descriptive and analytical findings reported so far in the chapter focused on understanding the process itself. The attempted larger picture is not yet complete and will require another look after more time has passed. It may be too soon to see how sustainable and how deep into the culture critical thinking assessment and initiatives will grow. Nonetheless, looking at the picture now offers insight to process anchors and to examination of best practice. Shifting focus outward by looking at suggested connections to these concepts regarding effective institutions in turn contributes to the Process Self Renewal Model presented in chapter five and leads to meeting goals for an explanatory case study.

Attention to culture and collaboration (Kezar, 2001, 2005, 2006, 2005a) as factors also yields insight. In the MASU case, considerations of university culture were important elements of the Assessment Team's planning and decisions. In order to embed assessment in courses and to assess in many disciplines with the same measure, the team knew that voluntary and cooperative faculty involvement was crucial. By coupling that understanding with awareness of a faculty culture of concern and control regarding academic programs, the Assessment Team formed the Working Group with its roles for advice and consent.

As seen previously, collaboration was the necessary tool for this work. Kezar (2005) characterizes collaboration as engaged faculty, working in cross disciplinary teams, supported by structure, people, and rewards. These ingredients for a collaborative institutional outlook were present in this assessment process (even though the extrinsic rewards were few) and might be a model for other institutional activity.

The second set of thoughts from Rogers' (2003) work considers questions of innovation and change. Does this MASU case map to these ideas? Rogers suggests that innovation and change require an agenda that matches a problem to a solution, involves redefining or restructuring in some way, and ultimately leads to "routinizing" that change. If the problem is represented by the state's requirement to address student competency in critical thinking (with later trustee interest), coupled with issues expressed in student exit surveys about challenging course work, then the MASU experience offers two solutions. The first can be seen in the way MASU carried out the required assessment. The assessment was specifically designed to yield data that could be used for required reports while also providing information that would facilitate positive change in courses or programs. Reports were provided to professors, the university, the board of trustees, and the state. The second, voluntary, solution was commitment to the Critical Thinking in the Curriculum project by the Provost and provision of resource support.

Rogers' redefinition was more subtle, through communication about the topic, workshops, and other consciousness-raising efforts. A future plan for general education assessment offers another opportunity to continue redefinition of this student outcome. Nascent restructuring beyond the CTC project can be seen in the use of critical thinking

by the business school and the revision of senior project work in the individualized study program.

At this time, any additional progress of routinizing critical thinking assessment may need to wait for the next state requirement which does not re-occur until 2010. However if the CTC project and faculty development efforts continue, there could be a routinizing of faculty attention to critical thinking teaching and learning in courses and programs. Or, as Gladwell (2002) terms it, a tipping point will occur, one that results in routinizing. We may know this is occurring if we hear students saying, "Oh, yes that is like the critical thinking I was asked to do in another course." The Provost will look for student surveys indicating that students believe their courses are more challenging than reported in the past. It is simply too soon to see all of Gladwell's or Rogers' processes manifested for this case.

The third view of institution level self-renewal comes from Birnbaum (1998), who postulates that self-renewal should be through systems, including feedback loops. These loops alert responsible university centers or persons to difficulties or problems. Well collected and delivered relevant information calls attention to issues which administrators can address before some crisis point. In the MASU case, undergraduate exit surveys indicating lack of challenges served this purpose for the Provost and they are expected to continue to do so. Coupled with the requirements for state information on critical thinking, two feedback loops provide information on the same topic. By making the connections among the student survey, the state requirement for assessment, and the TLC proposed initiative, the Provost demonstrated how Birnbaum's concept might work.

Expectations are that results from the assessment and CTC initiative will turn up in the surveys, in reports, and in project results, further contributing to continuing the feedback loops. If loops continue to operate, the chance for additional revision and improvement remains. The continued flow of good and useful information, calling attention to important concerns, is the heart of the system of self-renewal described by Birnbaum and assessment efforts offer important inputs to that mechanism. When there is a confluence of assessment, other actions, collaboration between units, and usable information, potential usefulness in a culture of accountability rises.

These three models of institutional self-renewal suggest themes of recognizing and valuing institutional culture, supporting collaboration, identifying areas for change through regular feedback loops, and connecting efforts to support redefining goals and fostering change. These models do not, on the other hand, directly address the element of goal agreement so clearly seen in the data from the work groups. We saw that not only was there agreement on what had to be done, there was also agreement on the importance of the specific topic under assessment, undergraduate critical thinking, which remains part of MASU's university mission statement. It may be that findings from this case, which was connected to an academic idea considered important by virtually everyone, may not easily transfer to another topic or context and the process may not apply in the same way. It is also important to remember the element of entrepreneurship that is part of the MASU culture. This element does foster openness to change that may not be present in other institutions.

General Observations

The interpretative application of these four conceptual frames (assessment, critical thinking, group work, and institutional renewal) allows deeper understanding and organization of the findings in the MASU case study. Such interpretations are intertwined with findings presented for the first three research questions. The answers to the fourth research question promote understanding of the prior answers, and provide a pivot point to the fuller interpretation and conclusions in chapter five. The story of what happened as MASU developed and conducted critical thinking assessment begins to take on meaning, and shape itself into potentially larger stories, informed not only by what happened, but by a variety of views on meaning. Finding meaning in this story as well as telling the story were purposes behind use of case study methods, and were intentional goals of this particular case investigation.

Findings Summary

The findings of this research into an assessment process include the story of what happened, the decisions and processes that occurred, the effects of the assessment, and discussion of models that inform understanding. Models of assessment, critical thinking, work groups, and self-renewal place the assessment process in larger contexts, and inform the findings of the first three questions. In most respects the MASU collaborative, embedded assessment of critical thinking demonstrated success in terms of assessment methods that promote change, development of common references for critical thinking, and successful work in two different groups. Application of the fourth institutional renewal concept resulted in ideas beyond a single assessment, ideas that suggested forms

of self-renewal. Some larger lessons are proposed in the next chapter, including a proposed integrated model and suggested best practice.

5. Interpretation and Conclusions

Introduction

Answers to the original questions about the processes of assessment for undergraduate critical thinking competency at Mid Atlantic State University have shown what happened, what decisions were made, what institutional effects were and how the process related to four concept analyses. If an overall statement can summarize those answers, such a statement could resemble these four points

1. MASU conducted a state-mandated assessment of critical thinking using faculty to develop a course-embedded assessment based on observations of student work. The process involved a series of decisions which were taken specifically to support change in teaching and thus, learning.
2. The process was rooted in work of the assessment office team and a faculty working group. Their processes exhibited many characteristics of successful working groups; although they did not exhibit the same roles or same work methods. During the time of this study, the two groups (assessment team and faculty working group) did exhibit detailed interaction, agreement on goals, sustained collaboration, attention to assessment best practice, and commitment to improvement of the assessment method.
3. Effects of the assessment process and faculty involvement are taking place and likely to grow. Connection of the assessment to a critical thinking curriculum initiative and

communication of results to the state, trustees, faculty, staff, and community continue.

4. As the case study ended, change regarding teaching and learning directed at critical thinking competency was initiated, assessment results were produced, communicated, and are still in the process of analysis. The institution can likely learn from this history, and discover ways to improve assessment and increase its effects, which may contribute to MASU becoming a self-renewing university.

The benefits resulting from case study methodology can and should extend results further than an institution's single instance. However, the spirit and culture of entrepreneurship that is a characteristic of MASU may imply limitations to the transferability of conclusions and the proposed model which follow. As the model sources and elements are examined below, limits may not be significant, since the model is built from both the MASU experience and from the concepts applied to this study. Benefits from this case study include opening the mind to a wider view that addresses the perennial "so what". Significance of the research is limited if it can only provide ideas to a single entity. Case studies commonly result in ideas for research direction and for practice. Since the stakes for institutions regarding accountability for student learning are clearly rising, the MASU case should be able to provide some direction for research and practice or, in other words, address the "so what" question.

Two views of the findings from this case study research move out from consideration of a specific instance of assessment at an individual institution to a view useful in other instances and to other institutions. The two views show interpretations

related to research and best practice. The first view for research-oriented interpretation is a proposed model of an integrated process that can lead to self-renewal. The second practice-oriented view is an interpretation including suggestions for best practice in assessment, particularly assessments which include a goal for change.

As previous sections of this report indicate, higher education is currently operating in a web of accountability. The significance of understanding process that addresses accountability is high. In this heavy atmosphere of interest in student competency, it is an advantage to an institution to design and use processes which pro-actively recognize and enable addressing an issue before crisis stage or before outside intervention. Such processes can contribute to self- renewal goals that Birnbaum (1998) suggests can be achieved through feedback loops that trigger action. Assessment processes that collect data, interpret it, and regularly inform administrators of the state of student learning as it relates to the university mission fulfill a key role in feedback loops. Such assessments as the one described at MASU also involve the very people – the faculty – who will implement and make effective any changes suggested by feedback loops. The value of this case study includes proposing such reliable linkages through a model for self renewal built on assessment processes. Accountability/assessment can be the vehicle for recognizing issues and taking action in a systematic way that will renew the university as well as provide evidence of accountability and responsiveness. Assessment reports can no longer sit on shelves gathering dust. They must become tools taken up and used to build the kinds of university processes that Birnbaum (1998) has in mind for institutions intending to grow in achievement. A proposed model and best

practice conclusions are offered so that application and expansion of lessons learned from this Mid Atlantic State University assessment case occur.

The Process Model

A basic foundation of the model is the experience of MASU as presented in this case. The assessment process experience of MASU had three significant characteristics that inform the model's goal of presenting a path for success. First, intentionality characterized the process decisions, plans and implementations. How the process was established and carried out derived from a coherent a set of beliefs about faculty involvement and embedded assessment. In turn these beliefs were reinforced by prior experiences with similar processes. Second, the Mid Atlantic State effort included collaboration among leadership, management, units of the university, and faculty. Collaborative efforts led the assessment office to expect successful implementation of the assessment and initiation of some changes in courses, programs and faculty outlook. Again, this was based on beliefs and experiences of prior assessment activities. Third, the MASU assessment office carried out its management functions well. Nothing will stop a process faster than inattention to scheduling, lack of follow-up or poor meeting preparations. These three characteristics of the MASU effort are important pieces used to construct the model. Other pieces include new ideas and a synthesis of information from characteristics of successful assessment, work groups and principles of renewal as presented by Birnbaum (1998).

A major piece which unites the parts of this model is the concept of *workspace*, the nexus of activity and work. The Mid Atlantic State University case shows that work

of the assessment team and faculty working group both separately and together, plus the contributions from institutional leadership, took place in an arena larger than the times people were present in the same physical space. Research, thinking, communication, and more thinking were important work which affected the periods when people were physically in one space. I have named this larger arena the *process workspace*. As present in Figure 8 below, all three circles indicate the process workspace, and this larger entity takes into account institutional leadership, management, and working groups. The concept allows multiple roles of faculty, administration, and staff to manifest and come together intentionally in the *active workspace*.

The structural frame for this process workspace derives from the success factor matrix for group work (Figure 7, chapter four); from the hallmarks of good assessment described previously; and from MASU experiences in creation of a plan, implementation of the assessment, reporting on student learning competency, and moving toward meeting goals for effective change in teaching and learning. This workspace is permeable, has flexible borders, and is inclusive of the many factors affecting the work at hand, in this case, assessment of student learning competency. The various roles for institution leadership, management, and group, are brought together in the active workspace to achieve necessary tasks.

In developing this model, I followed three rules for model building that Lave and March (1993) propose and which are paraphrased here.

1. Think process - good models address process.

2. Develop interesting implications - good models involve judgment, as well as recognition of interesting predictions that may be possible.

3. Look for generality - Good models apply to multiple situations and include language that generalizes.

This particular model is presented as a proposed model, because it derives from a single case. The aim is to present a creative way to look at assessment and perhaps other processes that will intrigue researchers and lead to testing the model in other situations. The model below was constructed with the three Lave and March principles in mind.

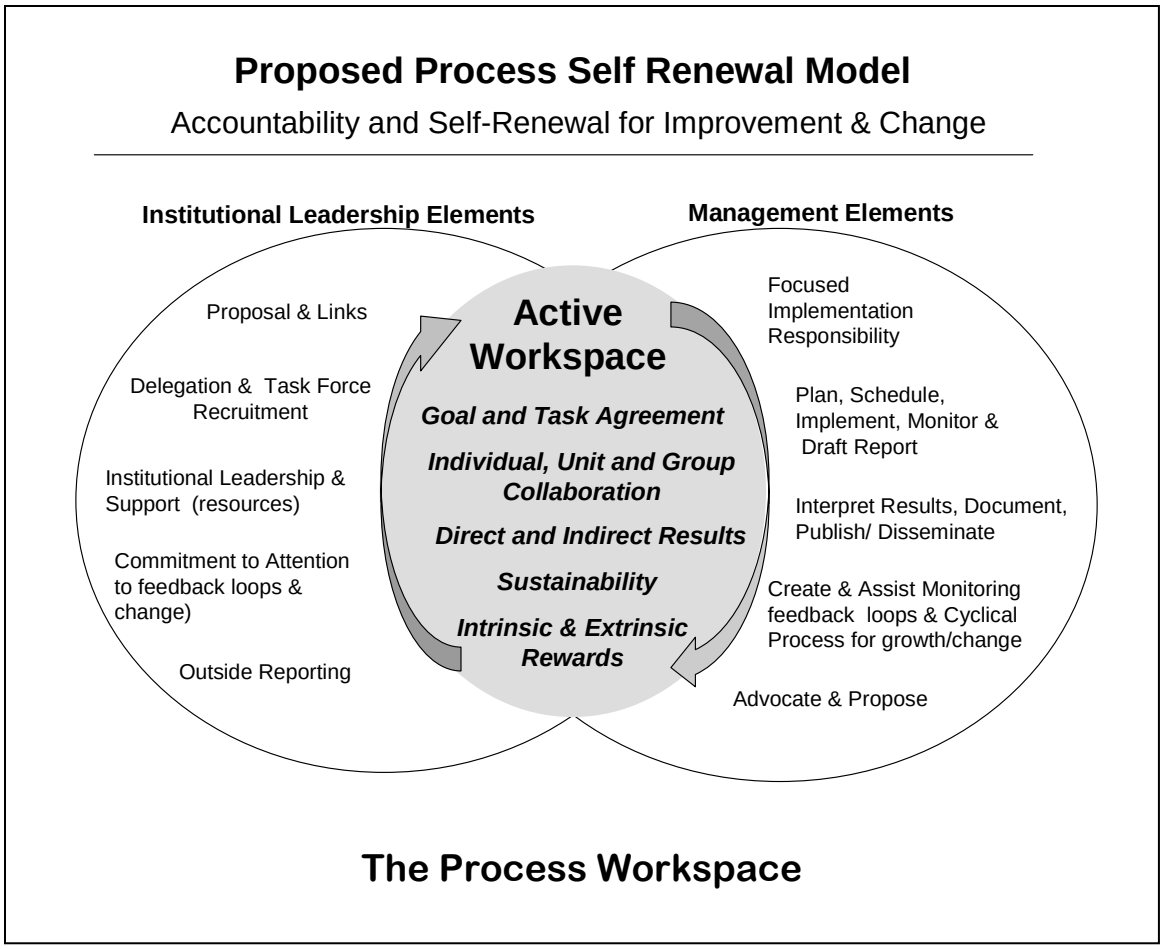


Figure 8. Process and Self-Renewal: A PSR model.

The model shows that a wide variety of roles, tasks and actions take place in the process workspace, and for a moment this might resemble the garbage can model (Cohen et al., 1972). However, Figure 8 was built on MASU experiences, and shows a more intentional and organized content in the active workspace. Goals were set and agreed upon; collaboration took place; tasks were accomplished, reported and evaluated; and results were achieved. Through interaction with the Critical Thinking in the Curriculum

initiative, a sustainable focus on critical thinking at the university is developing. Although extrinsic rewards were limited, feedback from records and collected data indicates presence of strong intrinsic rewards. Faculty consistently remarked on the value of the opportunities for interaction with other faculty, particularly faculty outside their discipline. Dissemination and communication to the university community and the state took place and appears poised to contribute to routinizing (Rogers, 2003) as opportunities open in a coming review of general education and as preparation for re-accreditation action plans takes place. In addition, communication efforts that Gladwell (2000) characterizes as opportunities to speed to tipping points are continuing. Reports were returned to individual professors for their courses, formal communication to the university board and to the state took place, and information and reports are readily available on websites. Attention to critical thinking by faculty and to moving toward a tipping point for wide and deep attention to critical thinking teaching is strongly supported by the CTC initiative. Administrators at MASU reported they think that three to five years may be needed for full effects, so it is too soon to see if there is full routinizing or if a Gladwell point has already been reached, but efforts at engaging faculty in all areas of the university appear to be headed in those directions. If the CTC project and collaboration between the assessment office and the TLC continue, there is a strong possibility that additional movement toward routinizing and tipping will take place.

In attempting to integrate all of these elements into one model, a fluid, live picture using technology might make the best presentation. However, at this point in the

research that remains a future endeavor. The presentation here may feel static but nonetheless offers a model of an interactive process workspace with wide and open boundaries which also suggests a repeatable scheme. The model is expressed in a way that attempts to show group work concepts (Gersick, 1989, 1991, 1998) indicating that work does not need to be linear or even continuous. No borders exist inside the spaces, except for the circles which imply inclusion rather than exclusion. A table arrangement was rejected as it implied linear progression and specific links. There is no numbering scheme to freeze what should come first or last. Many of these elements can take place together, at different times or in different sequences, and still retain the possibility of successful results for change. The arrows indicate that the process can be repeated either for the same task or for different tasks. For example, this model could be applied to the implementation of an undergraduate evaluation of general education, or development of strategic plans. It is expected that this model will be used for the coming second iteration of critical thinking assessment in 2010.

The model includes elements of the Success Matrix for group work matched with essentials from the MASU experience. These include delegation, responsibility, goal agreement, collaboration, faculty engagement, openness to change, effective management, communication, and connection to other institutional ideas. Once the provost delegated responsibility to the assessment office, that office began work on the other elements. Chapter four details the story of planning, implementation and reporting and recounts management efforts.

The nature of any model includes the suggestion that if some elements are not present then success is less likely to result. Future research in use of the model can help determine which pieces are essential. For example, openness to change during the process was reported by faculty who appreciated recognition of their feedback to the assessment office. How important is that? Is that an element that may not be necessary to creating successful assessment that informs change?

A short tour of the model adds details. Starting with Institutional Leadership is logical because of the importance of proposing the task and linking it to other institutional efforts. This is a model for a whole university effort, so it is extremely unlikely university-wide efforts will take place minus this crucial element. Linking to other institutional efforts can make all of them stronger and spread throughout the institution as was the case with links between the CTC program and the assessment at MASU. As the delegation of responsibility and task force creation (in the MASU case, to the assessment and then the Assessment Team and Working Group) take place, the other pieces of institutional leadership will move in and out of the active workspace in a non-linear manner. For example, commitment to action can take place at any time – either in monitoring or more directly in support through resources. Resource support can be direct in the form of allocations of personnel, money or space, or it may be indirect in the form of intangible rewards, like recognition of service. Resource inputs can take place at many times during the process, and the example of a series of provost-supported workshops and project awards were two examples from MASU.

Institutional leadership is also responsible for outside reporting and that step may be crucial to continuation of any institutional wide effort. For example if the trustees had an incomplete or poor report, they might have forced changes that would have negated the benefits of faculty involvement. At MASU that had happened in the past regarding a change to general education requirements. However, leadership efforts will not succeed alone. Management elements must be present and functioning.

The management elements of the model included the range of activities and responsibilities that were delegated. In this case, the provost clearly passed all management and implementation as well as formation of reports to the assessment officer. The many actions listed in the model for management - planning, scheduling, implementation, monitoring, data reporting, interpretation, documentation, and dissemination formed the day-to-day continuous work of the Assessment Team. Examples are research on critical thinking assessment methods and commercial tests, preparing materials, setting up training for raters, scheduling of the assessment observations, and continuous communication with all involved faculty. Through provision of these elements, as well as reporting, the assessment office helps create and monitor feedback loops, and supports cyclical change. As is the case with managers in many types of institutions, they are also in a prime spot to advocate and propose changes based on the evidence collected. A process missing good management will have little chance to develop and monitor self-renewing feedback loops, nor will poorly managed processes result in the ability to advocate or propose. Collaborative efforts of both the assessment office and the TLC on critical thinking in the curriculum enabled the proposal

that the provost funded, and joint workshops permitted advocating for improved teaching and learning as well as assessment.

In the Active Workspace, these contributions from institutional leadership and management join with the other personnel or units for development of the five central elements: Goal and Task Agreement; Individual, Unit and Group Collaborations; Direct and Indirect Results; Sustainability; and Intrinsic and Extrinsic Rewards. At MASU this happened through interaction with the Faculty Working Group, course instructors and faculty serving as assessment raters. Collaboration between the Teaching and Learning Center and the assessment office reinforced the importance and aided completion of the assessment tasks.

The MASU case and the model both support the idea that process can be as important as the product. For example, a question might be asked as follows, "What is likely to have longer effects: the required reports, or the experiences of faculty and staff working on the assessment as it was designed at Mid Atlantic?" The PSR Model which is based on faculty involvement for change presents a tool useful for future confirmatory research. If it proves sturdy, the model can offer (1) a way to improve higher education accountability for student learning competencies, (2) a potential guide for intentional choices which encourage effective change, and (3) direction for moving toward routinizing institutional self-renewal.

In the future I hope to apply this model to other activities, other institutions, and possibly to a variety of subjects beyond assessment. As this particular case study ends, one important reason for choosing case study methodology is fulfilled. The case does

point to directions for further significant research, described more fully at the end of this chapter.

Best Practices for Assessment

A second goal for this research was to discover best practice ideas that assessment officials and university administrators can look to for their own work. The MASU case clearly distills ideas useful to other institutions. In addition to considering the PSR model, practitioners and administrators can take away suggestions particularly related to assessment of student outcome competencies. What are some of those suggested lessons?

The MASU lessons for best practice in assessment for higher education can be grouped into three areas (1) goals, (2) methods, and (3) results. Some indications for sustainability of assessment can also be seen in the Mid Atlantic experiences. Success factors as well as warning flags are present in all three of these areas and around the issue of sustainability.

The first lessons relate to goals. Sometimes, as in this case, one of the goals is given from an outside requirement and serves as the initiating factor. If the institution wishes to make the requirement into an opportunity, the formal, required charge is not the end of goal discussion. The opportunity can be to reframe the goal from a short term find-out-something-and-report-on-it, to internalized goals of effecting improvements and real change in programs and courses to improve student learning and achievement. It is not enough to have a few administrators support this opportunity goal, it must become the agreed upon goal for the groups working on the assessment, and also related to other

ongoing institutional operations, for example, general education. The success matrix for work groups, (Figure 7, chapter four), shows that goal agreement was the most frequently identified factor in successful work groups. At MASU a synergy of critical thinking topic interest was also in place and augmented unity of goal agreement among the provost, the assessment office, the Teaching and Learning Center and the Faculty Working Group.

In considering working goals, institutions will find it is worthwhile to make as many connections as possible. In the Mid Atlantic case, these were to the university mission, to trustee interest, to a topic of faculty interest, and to work of the TLC. Clearly stated goals which mesh well with other institutional aims and are commonly agreed upon are the first components of best practice lessons.

Goal agreement leads directly to selection of methods for the assessment, the second group of practitioner lessons. In this case study, methods include management, make-up of the two work groups, collection and interpretation of data, reporting, and evaluating results. At MASU, selection of methods also took note of the culture of the university, of past practice that worked well, and included research. Obtaining faculty validation and collaborative cooperation for an agreed upon, course-embedded assessment was crucial. Consideration of methods also means checking back to goals. If the goal is supporting effective change to teaching and learning, that needs to be part of the methodology and was one of the chief reasons Mid Atlantic selected embedded assessment. In the instance of critical thinking, there was another piece partially related to methodology which was the connection of the assessment and its measurement tool to

the initiative for Critical Thinking in the Curriculum, a program that the Provost supported with funding for course and curriculum innovation. The rubric of elements of critical thinking was a starting point in requirements for the RFP's and requirements for awardees included assessment of critical thinking. A good assessment takes advantage of links to other opportunities.

The practice of faculty working groups developing course embedded assessment had already proven successful at MASU, and even though the critical thinking assessment crossed disciplinary lines, the same process was employed. Other institutions may have different past successes and different culture parameters, but taking notice of those past successes or cultural elements and using them to support assessment efforts is crucial.

Research and preparation for working with faculty was another important methods element and helped the working group come to task completion, as did the collaboration with TLC in recruitment, workshops and connection to Critical Thinking in the Curriculum. Management of the effort was also a part of methodology and included organizing and preparing meetings, scheduling, training, and preparation of the data for presentation back to the working group for further evaluation. These are all part of this second group of lessons for best practice.

The third group of lessons for best practice relates to results. Once assessment results were obtained, both those managing and those leading the assessment worked on preparing the reports and ensuring that results were communicated and disseminated. For example, all course professors received score reports for their students and how they

compared to university scores. Examination and discussion of the results also helps complete the assessment circle by encouraging action to make use of what the assessment demonstrated. Use could mean refining the assessment or processes, as well as noting changes in teaching and learning. At this point, looking once again at the goals is important. If one goal was to effect change, then some way to document those changes becomes necessary. Usually there are more silent effects than effects that come to notice easily. The MASU assessment process made some efforts to capture those changes in assignments, courses change, programs, and in teaching outlook. Without attention to the original goals, administrators and assessment officials may miss determining what the full results regarding change were or could be. Knowledge of all types of effects allows administrators to make informed decisions on how to go forward, to allocate resources or to offer other support for the desired change. Both direct and indirect results need to be considered and help to complete the assessment circle of design, implementation evaluation and use of results for desired improvement.

The cycle of using both direct and indirect results, re-examining that use, and re-conceiving the next steps make an effort sustainable. Madison (2006) adds that faculty become involved through a set of "progressions of steps necessary to get college or university faculty fully committed to meaningful and effective assessment of student learning. The first step is awareness, the second, acceptance, next comes engagement and finally, ownership" (p. 4). The ownership step aligns with routinizing (Rogers, 2003). This involvement then forms part of a routine of assessment which passes into common practice and establishes actionable feed back loops.

These three lessons of best practice assessment processes can be applied to other efforts at MASU and beyond. Substituting other kinds of work for assessment in these practices is another avenue to follow. For example, the practices listed here for assessment could also apply to re-structuring general education or to academic improvement plans for accreditors. Each particular institution can take these best practices and apply them to help meet their own goals in a variety of arenas.

Summary Interpretation

A final look at the Mid Atlantic assessment of critical thinking returns to some observations on their efforts. MASU met most tests of assessment best practices in its use of faculty working groups, course embedded methodology, communication, use of results, and connections to other institutional actions and goals. A less strong area was course participation. Although use of the synthesis courses fit the topic of assessment, participating courses were voluntary – there was no requirement that a professor to join the assessment. The participating courses did cover multiple disciplines and there was at least one course in each undergraduate college. Furthermore, the faculty for business and engineering programs came forward to be included even though no specific effort to bring them in was made. Deriving university level conclusions had some strength, but not as much as it would have had with all synthesis courses participating.

Mid Atlantic State's selection of a critical thinking definition and development of a rubric reflecting that were based on work in the field of critical thinking and selected after consideration of options. The definition and rubric remain in use as the assessment tool and the framing definition for a curriculum initiative and are used as information

communicated to incoming faculty. These tools are now important for both assessment and teaching.

The use of the four conceptual frames to analyze case findings points to a MASU story of good assessment, intentional decision making, and successful group work. By enlarging the view to the fourth concept of self- renewing institutions, the MASU assessment can be seen as also offering direction for research and practice.

Completion of this look at the history of Mid Atlantic's process, and institutional effects as well as the creation of explanatory pictures using conceptual tools brings satisfying and useful conclusions. As this project's analysis progressed, ideas about of what constitutes the 'work space' expanded, and led to the Process Self Renewal Model in Figure 8. The proposed model also offers practitioners and researchers useful lessons and direction for MASU and other institutions. Participants expressed some of these ideas very well, and their voices continue to be valuable at the end of this study.

- It is really fun to be in these meetings. And to hear people talking about their discipline approach to something and to hear them listen to somebody else and be surprised at how somebody else approaches something. This is why the push nationally or at the state level to just do tests as a way of assessing what students learn is so wrong - because it will do absolutely nothing to change the curriculum, to change faculty thinking about pedagogy (Interview, Director of Assessment, April 16, 2007).
- "I want to include critical thinking in course activities. How do I do that?"

  (Music Professor, Phone Log, July 20, 2007)

- "I believe critical thinking is a critical aspect of education, making learning an active process and giving students true ownership of their education."(Questionnaire 1, questions 3, "Why did you decide to participate?, participant 8).
- "I would say yes, because I think that the interdisciplinary collaboration is very worthwhile and I have gotten many new ideas on critical thinking assessment through professors from other units." (Questionnaire 1, question 7 "If another faculty member asked you if they should participate, how would you respond?", participant 3)
- "I think the combination of the CTC initiative and the CT assessment is very powerful." (Questionnaire 1, question 5, How would you describe the usefulness and effects of this program to another faculty member?", participant 4)

The future direction of research into assessment practice is advanced by ideas and insights gained from Mid Atlantic State University's assessment story. Faculty collaboration, links among assessment efforts and other university activities, longitudinal follow up to this study, and discovering applications for the PSR Model are just a few of the research paths and topics leading onward.

More explicitly, future research can take several directions, some of which have already begun. There are four areas that future research could investigate:

1. Validation and usability of the model

2. Assessment as faculty development

3. Longitudinal study of critical thinking in the curriculum and the effects of assessment at MASU

4. Assessment office collaboration with other administrative units

Model validation research can be followed in two ways, one at the same institution on another topic. For example, MASU is currently preparing for re-accreditation and one of the requirements is to form and implement a plan for improvement on an institution-wide topic that supports undergraduate education. Investigation of how the model assists that effort from the start or fits that effort after it is carried out can introduce interesting research questions that will help to validate or refine the model.

A second form of research would be to see how the model works in another institution. A first step toward finding an interested party for collaboration in that direction is a submitted proposal for presenting the model at the next American Education Research Association conference. Such a presentation route can be a way to find collaborators for additional model-based research.

The second set of research ideas involves a research path for assessment as faculty development and is also being followed through accepted regional and proposed national presentations of that aspect of the findings. For example, the regional panel discussion on faculty involvement with assessment may lead to venues and collaborators for the topic of faculty development through assessment. Presenting may also attract those interested in the scholarship of teaching and learning since well designed student assessment expects to lead to improved teaching, and thus improve learning. All of these

paths can develop from the MASU case followed over time or from investigation of new cases.

A third path is to follow the specific topic of critical thinking assessment and its effects at MASU, with a focus on students. The institutional assessment office can continue tracking effects to determine how or when permanent change in teaching and learning takes place. Since there will be another critical thinking assessment in 2010, there is a high probability of continuing this research.

The fourth branch for future research might be toward a study of how assessment offices and officers organize and conduct their units and work. Most assessment conferences offer sessions for those new to the field, and graduate schools offer assessment courses which can benefit from research on how institutions organize for assessment. Perhaps solid evidence can keep future professionals from re-inventing the wheel. Given the current large eye on assessment, all of these research paths can lead to significant contributions. Inquires from those interested in such paths are welcome!

Concluding this study and reflecting on the potential next research provides opportunities to re-energize and to take up the tools that developed from this story to see what can be built. Case study research proved intriguing, difficult, and frustrating to explain. However, it also provided the opportunity to acquire knowledge from varied sources, to discuss significant topics with many faculty and administrators, to integrate concepts, to try out new ideas, and to create a somewhat new view of how institutions work now and might work in the future. Educators must be willing to look at new ideas in order to do some of that difficult explaining of how we know what and how our

students are learning. In an era of very rapid change, we will need to be self-renewing and proactive in anticipating and dealing with change so that our students will be well educated for their own futures of change.

Appendix A
Study Participant Information

Participants are not personally identified, and no information is included here that would identify them personally. Smithson professors are tenured and have special appointments to address undergraduate student learning.

ID No.	MASU Position	Critical Thinking Assessment Role(s)
1	Tenured professor	Course professor
2	Tenured professor	Working Group
3	Term professor	Working Group, Rater
4	Director, Teaching & Learning Center	Assessment Team, Rater
5	Term professor	Rater
6	Tenure Track professor	Working Group, Rater
7	Smithson professor	Course professor
8	Tenure Track professor	Working Group, Rater
9	Smithson professor	Working Group
10	Term professor	Working Group, Course professor
11	Librarian	Rater
12	Term professor	Rater, Course professor
13	Tenured professor	Working Group, Course professor
14	Program Director	Working Group, Course professor
15	Tenured professor	Working Group
16	Associate Director, Assessment	Assessment Team, Rater
17	Director, Assessment	Assessment Team, Rater
18	Provost	Responsible for State Report
21	Term professor	Rater
22	Tenure Track professor	Rater
23	Tenure Track professor	Course professor
24	Tenure Track, Assistant Dean Business	Working Group
25	Tenure Track professor	Course professor
27	Tenured professor	Course professor
28	Term professor	Course professor
30	Tenured professor	Course Professor
31	Tenured professor	Rater
32	Term professor	Rater
	These 28 signed consent letters.	

Appendix B
Research Questions, Data, and Analysis Chart

Research Questions	**Data Source**	**Analysis Methods**
1. How did a specific institution carry out a specific assessment of undergraduate critical thinking in the culminating course of general education?	Administrator Interviews Working group records such as: · meeting agendas/ notes/attendance · workshop records · researcher notes Relevant questions on institutional surveys of students	Compared transcripts and notes, used NVivo software as appropriate to find commonalities. Constructed a descriptive narrative of "what happened" so the other research questions are context-situated Identified the group's composition, charge & goals Summarized what is known from these student opinion surveys
2. What were the processes and decisions that contributed to development of an assessment of this student critical thinking outcome?	Relevant portions of the administrator interviews, and questionnaires for the first faculty group (Working Group and pilot participants) Researcher notes and assessment records	Described the decisions made and who made them. Looked for influencing factors through answers to questions #2 on the first participant questionnaire, # 7 on the second questionnaire, # 2 and others for the assessment staff interview, and the Provost interview. The data were initially grouped & coded, using the 4 concepts as described in Chapter 2 above. Categories were then collapsed as necessary. Findings were reported as the narrative of Chapter 4.

Research Questions	Data Source	Analysis Methods
3. What effects did these process and decisions have on the institution during the time of this study?	Faculty Questionnaires I & II, (Includes members of the Working Group, faculty in whose courses assessment took place, and those serving as raters) , Assessment records & final Administrator interviews, Researcher notes & records of working group	Looked for unintended consequences and effects beyond the student information that was obtained. Data was compared to the initial charge and addressed these questions: · Did the formal reports indicate accomplishment of the goal? · What happened as a result of that formal report? · What did feedback from participants have to say? Constructing an 'after' timeline clarified effects, and helped categorized the changes i.e. instructional, procedure, institutional response etc. Unintended consequences (i.e. change and effects on programs or courses, departments, professor outlook etc) are included, and the initiator, effect and its extent are described and quantified as possible. The timeline was then translated into a descriptive statement of discovered effects.

Research Questions	Data Source	Analysis Methods
4. What models of assessment, critical thinking, group work and institutional self-renewal help us understand this case?	Working Group records: -meeting notes -communications -feedback collected from raters -researcher notes Interview & questionnaire data Focus Group results	· Identified changes made as result of pilot · Compared second implementation to the original. · Looked for successes and problems · Looked for differences among team members. Used Focus Group data for checking initial findings Used data from varied of sources for triangulation. Mapped the actual processes and implementation and effects to the concepts to organize and interpret findings
Interpretation of Findings: Looked at both practice and conceptual outcomes, for example, how does this case offer suggestions for modeling assessment processes in higher education and how might this inform administrators tasked with accountability? Consolidation of the processes described above enabled conversion to narrative form, and the start of the new conceptual model.		

Appendix C Instruments

Faculty Participant Questionnaire I

This questionnaire will be used at the start of the project and given to the committee members and the professors in whose courses the assessment takes place for the pilot project. Responses can be given either electronically or on paper, as chosen by the participant.

1. How did you become involved in this assessment project?
2. Why do you think the university initiated this assessment project?
3. Why did you decide to participate?
4. What other experiences have you had with university wide assessment projects?
5. Describe your participation so far in this one.
6. How does this experience compare to any others?
7. If another faculty member asked you if they should participate, how would you respond?
8. In what ways do you see this assessment and its results affecting the university, and your program /courses?
9. Do you have any other comments you wish to make?

Faculty Participant Questionnaire II

This will be given to the committee, the faculty involved in the pilot and the faculty involved in the implementation following the pilot. Responses can be given either electronically or on paper, as chosen by the participant.

1. What has been your role in this project?
2. How do you feel about your participation?
3. What has gone well, not so well?
4. Were you involved in the planning? If so, how?
5. How would you describe the usefulness and effects of this program to another faculty member?
6. What kind of feedback have you received from students, and/or fellow faculty?
7. Why do you think the university carried out this assessment?
8. In general, how might you characterize the process for carrying out this assessment?
9. What benefits do you see for student learning and/or teaching excellence resulting from this assessment?
10. Please make suggestions for how to change/improve the university's methods for carrying out assessments of student learning outcomes.
11. Do you have any other comments you wish to make?

Faculty Focus Group

This will be used for a group selected from the faculty participating, and take place after the questionnaires are completed. The discussion was not audio taped.

The topic will be the value of university assessment work in the area of student learning, particularly this critical thinking project, and the usefulness of methods to carry out the project. Other issues to be considered will arise from the two questionnaires. Examples might be the difficulty of cross-discipline, common assessment of student learning or the issue of raters when rubrics are used.

These next two protocols were combined in one interview.

Office of Institutional Assessment Interview Protocol I

Two administrative faculty in the Assessment Office will participate in this interview at the beginning of the project. The following are the general questions, which will be followed up with others depending on the participant replies. The discussion will be audio taped.

1. What are the goals of this assessment?
2. How was this particular way of assessing students chosen?
3. What is the reasoning behind involving faculty to plan the assessment?
4. Can you describe any issues that developed as results of the kind of plan you have described?
5. What has been easy/difficult to do in carrying out the plans?
6. Where are you in the process now; and what are the next steps?
7. Do you have any other comments you want to make?

Office of Institutional Assessment Interview Protocol II

Two administrative faculty in the Assessment Office will participate in this interview at the end of the project. The discussion will be audio taped.

1. Describe how well the project reached its goals.
2. Now that the project has been through both a pilot and a general implementation, what do you see as successful elements? What was the most successful?
3. What kinds of thing were not so successful? What was the least successful?
4. What things were unexpected?
5. What changes would you make in doing this type of assessment again?
6. In reply to another university's asking you how to assess critical thinking, what are the most important pieces of advice you would give them?
7. Can you say that this project contributes to 'best practices' or to a potential model for assessing student learning in other areas?
8. Do you have any additional comments you want to make?

Teaching and Learning Center Director and Provost Interview Protocol

This interview will take place at the end of the project. The discussion will be recorded using researcher written notes.

1. What were the goals for the university in assessing student critical thinking?
2. Do you think the goals were met?
3. What kinds of things were important to you in the ways this was planned, carried out and reported?
4. What was your level of involvement with planning and implementation?

5. What do you see resulting from the information and the processes used in this project?
6. What would you tell another Provost or Teaching Center interested in assessing critical thinking?
7. What do you expect will be lasting effects of this assessment and its processes?
8. What are some implications for the university as a whole that emerge from this type of student learning assessment?
9. Do you have any further comments you wish to make?

Appendix D Human Subjects Documents

Participant Informed Consent Letter

How a University Works: A Case Study of Assessment of Undergraduate Critical Thinking [Pseudonyms as used in the dissertation have been retained here]

RESEARCH PROCEDURES

This research is being conducted to develop a case study of a university's process for assessing undergraduate critical thinking in synthesis and capstone courses. The requirement to do so originated in the State Education Council. The knowledge from the case study of development, implementation, and evaluation of this assessment of student learning will contribute to future university efforts in assessment and may inform a model or best practice development for use of additional institutions. The case study will develop deep descriptions as well as a rich understanding of faculty participation in assessment processes. Faculty participants will be asked to complete two questionnaires, taking about 20-30 minutes each, and some will be asked to participate in an audio recorded focus group, taking about 45 minutes. Faculty teaching synthesis and capstone undergraduate courses will be invited to participate. Administrative faculty will participated in two audio taped interviews of approximately 30 minutes each, and the Provost and Director of the Teaching and Learning Center will participate separately in one non-recorded 45 minute interview. Additional materials from the committee will be used in the study. The dissertation researcher is a member of the Office of Assessment, assigned to work on critical thinking assessment.

RISKS

There are no foreseeable risks.

BENEFITS

There are no direct benefits; however, you may receive informal recognition for faculty service to the university.

CONFIDENTIALITY

Non- public data in this study will be confidential. Confidentiality will be maintained regarding participants' responses to questions and on the taped sessions. Real names or identifying information will not be used in the findings. Tapes and non-public, identifiable data will be stored in a locked cabinet. Audio tapes will be destroyed at the end of the study.

PARTICIPATION

Your decision to participate in the research is voluntary. You may withdraw from the study at any time and for any reason. If you decide not to participate or if you withdraw from the study, there is no penalty or loss of benefits to which you are otherwise entitled. There are no costs to you associated with the research.

CONTACT

This research is being conducted by Martin Ford and Mary Zamon in the College of Education and Human Development at George Mason University as part of Zamon's doctoral dissertation. You may reach Dr. Ford at 703 993-2004 with questions or to report a research-related problem. You may also contact the

Office of Research Subject Protections at 703-993-4121 if you have questions or comments regarding your rights as a participant in this research. This research has been reviewed according to University procedures governing your participation in this research.

CONSENT

I have read this form and agree to participate in the study as described above.

_____name _____date of signature

Some participants will be asked to take part in audio taped interviews or a focus group. I give my consent to be audio taped for the purposes described above. Please sign below if you are willing to provide your permission.

_____name _____date of signature

PROTOCOL – Involving Existing Records

1. The data for this study will include notes, minutes, communications, presentations, documents and reports of the Committee on Critical Thinking. Data will also include records of the researcher's internship related to the development of the committee. Those internship records are part of that coursework and are also the researcher's work records.

2. The attached letter [Not attached in this appendix because of identifying information] gives permission to use the records of the Institutional Assessment Office related to the critical thinking assessment. [Name] Director of the Assessments has signed the letter.

3. The confidentiality of information on personally identifiable data will be maintained for any that is not public information. An example of public information is membership on the committee.

4. Because these data are primarily qualitative data, the extraction of information is done by looking at themes and ideas developed through categorization of narrative materials by use of coding and some specialized software available for that purpose. Quantitative information might include numbers of meetings, attendance over time, numbers and types of communications or documents and other quantifiable material. The result will be primarily reported in a narrative case study, supported by numerical information as developed during the course of the project.

Appendix E Scoring Guide for Critical Thinking Pilot Version

Scoring Guide for Critical Thinking

Date: _____ Rater: _____
Course:_____ Student/Group No._____

***Score: NA = not applicable; DK = don't know**

Critical Thinking Elements	3 Highly Competent	2 Competent	1 Not Competent	SCORE **(NA/DK)***
1.Problem/Issue				
Identification	Clearly identified, explicit	Problem/issue present, but a little ambiguous	Problem/issue ambiguous or not present	
Assumptions	Clearly identified and explained	Identified, not clearly explained	Not identified, not explained	
2. Analysis of Problem/Issue/Investigation				
Method(s)	Selected and used appropriately	Some inappropriate choice/use of method(s); needs improvement	Inappropriate choice/use of method(s)	
Resources/source materials	Appropriate and relevant (number & content)	Most resources/source materials are relevant and appropriate	Few resources/source materials are relevant or appropriate	
Alternate points of view/methods	Thoroughly considered alternate points of view/method(s)	Some attention to at least one alternate view/ method	Obvious alternate views/methods ignored or superficially considered	
3. Credibility of Resources/Source Material				
References	All references are from credible resources	Most references but not all are from credible resources	Most references are from questionable resources	
4.Conclusion/Problem Solution				
Limitations	Thoroughly considered	Given some attention	Limitations ignored	
Integration	Ideas integrated into a coherent argument, solution, presentation, etc.	Ideas integrated into a somewhat coherent piece	Ideas lack coherence	
Conclusions/ Solution(s)	Conclusions/solutions based on evidence/sound methods	Most conclusions/solutions based on evidence; some unwarranted conclusions drawn, or some solution errors	Most conclusions/ solution(s) unwarranted; maintains preconceived views regardless of evidence/need for different solution	
5. Creativity/Student Ownership/Personal Engagement				
Creativity/ Ownership/ Engagement	Fresh ideas, reflecting in-depth student engagement with the topic	Some new insights/engagement; but primarily based on collection of other people's ideas, products, images, etc.	Little evidence of engagement with the topic; the work is a rehash of other people's ideas	
6. Communication with/Adaptation to Audience				
Communication /Adaptation	Highly evident	Some communication/ adaptation	No communication/ adaptation	

Appendix F Elements of Critical Thinking

Elements of Critical Thinking Assessment-*disguised to preserve institution identity*

These elements were developed by a faculty work group and are used to assess Critical Thinking for undergraduate students. Additional information on Critical Thinking at MASU is on the Critical Thinking in the Curriculum (CTC) web site xxxxxxx.

Elements	Highly Competent	Competent	Needs Improvement	Not Competent
1. Problem/Issue				
Identification	Clearly identified, explicit	Identified implicitly or incompletely described	Problems/Issues present, but a little ambiguous	Problem/issue ambiguous/not present
Assumptions	Clearly identified and explained	Clear identification, somewhat unclear explanation	Identified, not explained	Not identified, not explained
2. Analysis of Problem/Issue/Investigation				
Method(s)	Selected & used appropriately	Some inappropriate choice/use of method(s)	Choice/use of method needs improvement	Completely inappropriate choice and use of method(s)
Alternate Points of view	Thoroughly considered multiple alternate points of view	Thoroughly considered attention to one alternate view	Some attention to at least one alternate view	Obvious alternate views ignored
3. Credibility of Resources/Source Material				
Resources/Source material	All are appropriate & relevant (number, content, credibility)	Most resources/ sources relevant, appropriate, credible	Some resources/sources are not relevant, appropriate, credible	No relevant, appropriate, credible resources/ sources
4. Conclusion/Problem Solution				
Limitations	Thoroughly considered	Given some attention	Limitations mention very briefly	Limitations ignored
Integration	Ideas well integrated into a coherent argument, solution, presentation	Ideas integrated into a somewhat coherent piece	Some connections among a few ideas are made	Ideas lack connection or coherence
Conclusions/Solution(s)	Conclusions/solutions based on evidence/sound methods	Most conclusions/solutions based on evidence	Some unwarranted conclusions drawn, or solution errors	conclusions/solutions unwarranted; maintains preconceived views regardless of evidence/ need for different solution
5. Creativity				
Creativity	Reflecting in-depth student engagement with the topic; fresh ideas	Reflects some in depth consideration of topic; some new insights	Primarily based on collection & repetition of other people's ideas, products, images; very few new insights	No evidence of engagement /new ideas on topic; work simply a re-hash of other people's ideas
6. Communication with/Adaptation to Audience				
Communication/Adaptation	Highly evident	Evident most of the time	Few small considerations of communication and/or adaptation	No communication/ adaptation to audience

Summer 2007

Appendix G
State Report (Institution identifiers removed)

CRITICAL THINKING ASSESSMENT REPORT

REVISED SEPTEMBER, 2006

Definition of Competency:

As infants, we begin to develop critical thinking skills and then spend most of our lives perfecting those skills. Faculty at institutions of higher education work explicitly to develop critical thinking among students and indeed, most courses implicitly, if not explicitly, have critical thinking as a goal. This is certainly true at_____. Although the meaning of critical thinking is often interpreted differently across disciplines, these determinations of meaning are often nuanced and overlapping rather than distinct. For example, an engineering student may utilize more math and logical reasoning; the art student more critical decisions about the medium, the subject, the point of view, etc., but both students must employ critical thinking skills to succeed in their respective fields. Thus, critical thinking is a higher order thinking skill exhibited in context. At the college level, it is learned, developed and finds formal expression within contexts represented by academic disciplines. Nonetheless, because critical thinking is a transferable skill, there are core meanings of critical thinking that transcend disciplines.

The following components of critical thinking were identified by an interdisciplinary team of faculty as the essential criteria by which critical thinking should be judged at _____.

- Identify important questions/problems/issues (Rubric Item #1)
- Analyze, interpret and make judgments about the relevance and quality of information (Rubric Item #2 and #3)
- Assess assumptions and consider alternative perspectives/solutions (Rubric Item #2)
- Draw conclusions and make judgments based on evidence gathered (Rubric Item #4)
- Be engaged with one's topic/idea (Rubric Item #5)
- Integrate ideas into a coherent argument/solution/presentation, etc. (Rubric Item #4)
- Communicate the results of one's thinking (Rubric Item #6)

The mission statement of the university specifically states that _____ "will be an institution of international academic reputation providing superior education for students to develop critical, analytical, and imaginative thinking and to make well-founded ethical

decisions." The assessment of critical thinking, as defined in this report, supports that goal.

Description of Methodology Used to Gather Evidence of Competency and Standards for Critical Thinking Competence:

Critical thinking was assessed as a pilot in spring 2006 in synthesis courses, the upper division general education requirement for all _____students. Synthesis courses are available to students in one of the following ways: 1) an academic unit offers a synthesis course for their majors, 2) an academic unit encourages or requires their students to take a particular synthesis course offered by another unit, 3) a synthesis course is open and available to any student in any major, and 4) two or more academic units collaborate and offer a joint synthesis course. Students in synthesis courses "engage in the connection of meaning and the synthesis of knowledge." They are required to "demonstrate advanced skills in oral and written presentation," thus providing a venue to assess critical thinking.

A common rubric was developed by the Critical Thinking Assessment Committee, a committee comprised of faculty teaching synthesis courses. Having these faculty create the rubric was intentional and was designed to allow faculty to use the results of the assessment to effect change and enhance students' learning experiences. The rubric included criteria to address all of the above components of critical thinking. Fourteen trained faculty raters used the rubric to assess presentations or written work in six courses representing six disciplines. A total of 98 students in 57 presentations, including poster, project, and group presentations were assessed, along with the written work of 12 students in a single course. Thus, there is a total of 110 students in six courses rated by 14 trained faculty. Most presentations were rated by 2-3 raters and all written work was rated by one (of 5) faculty member. No faculty member rated student work in their own course.

Analysis and Presentation of Results

The rubric used by all raters included three explicitly defined competence categories for each criterion: highly competent, competent, and not competent. A fourth category of "don't know/not applicable" was available for the rater, but was not assigned a score and is, therefore, excluded in the following table. The results presented here include the percentage of student presentations/written work that were scored at each competency level. The data have been aggregated across all courses and presentations.

Critical Thinking Assessment Results

Criteria	Number of Presentations/ Written Papers*	*Highly Competent*	*Competent*	*Not Competent*	Total** *Competent or Highly Competent*
Identification of Problem/Issues	69	30%	58%	12%	88%
Analysis of Problem/Issue/ Investigation	69	46%	44%	10%	90%
Credibility of Sources	60	65%	28%	7%	93%
Conclusion/ Problem Solution	69	26%	62%	12%	88%
Creativity/Student Ownership/ Engagement	68	41%	44%	15%	85%
Communication/ Adaptation to Audience	69	39%	46%	15%	86%

*Because group presentations were included in this assessment, the total number of actual students is higher than the number of presentations/papers (number of students=110).

**Percentages may not add to 100 due to rounding.

In most categories, _____students exhibited good critical thinking skills in both their oral presentations and written work, but there is room for improvement. "Student creativity/ownership" and "communication" may need attention, but the small percentages of highly competent students in "identification of problem/issue" and "conclusion/problems solution" suggest further analysis as well. In the fall, the Critical Thinking Assessment Committee and the General Education Committee will consider these results and testing will be repeated in the spring. One issue will be whether the assignments are optimally designed to elicit critical thinking. Another concern is whether the same rubric can be used to assess different modes of expression, particularly oral and written presentations.

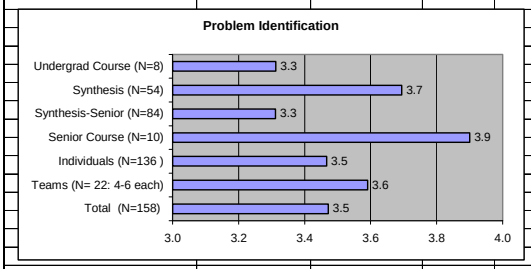
This assessment was designed as a pilot study with a convenience sample. We plan to broaden the number of synthesis courses that will be included in the next round of CT

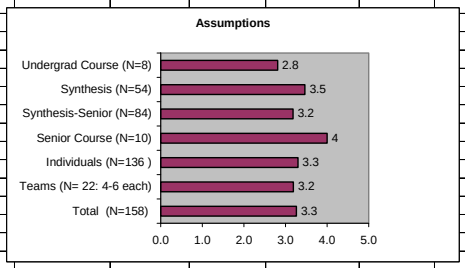
assessment. For 2006-07, we have several plans related to critical thinking. The first is to offer a fall workshop for faculty on how to deliberately include critical thinking in the curriculum. We also plan to encourage the use of the CT rubric in other classes this fall. And, finally, the Provost has begun a new initiative entitled, "Critical Thinking In the Curriculum" (CTC) which will be based on the very successful "Writing Across the Curriculum" and will include an assessment component.

Provost Signature

Appendix H
Preliminary Second Implementation Summary Data

Preliminary Second Round												
Averages												
	Problem Identification	**Assumptions**	**Method**	**Alternatives**	**Resources**	**Limitations**	**Integration**	**Conclusions**	**Creativity**	**Communication**	**N**	
Total (N=158)	3.5	3.3	3.4	3.1	3.4	3.0	3.3	3.2	3.2	3.2	158	
Teams (N= 22: 4-6 each)	3.6	3.2	3.7	3.4	3.5	3.3	3.4	3.3	3.5	3.5	22	(4-6 students each)
Individuals (N=136)	3.5	3.3	3.4	3.1	3.4	3.0	3.2	3.2	3.1	3.2	136	
Senior Course (N=10)	3.9	4	4	4	3.7	3.9	4	4	3.8	3.9	10	
Synthesis-Senior (N=84)	3.3	3.2	3.3	3.0	3.3	2.9	3.1	3.1	3.1	3.2	84	
Synthesis (N=54)	3.7	3.5	3.4	3.1	3.6	3.1	3.4	3.3	3.2	3.3	54	
Undergrad Course (N=8)	3.3	2.8	3.6	3.3	3.3	2.9	3.1	3.0	3.4	3.4	8	





		Highest			**Lowest**							
Total (N=158)		Problem ID	3.5		Limitations	3						
Teams (N= 22: 4-6 each)		Method	3.7		Assumptions	3.2						
Individuals (N=136)		Problem ID	3.5		Limitations	3						
Senior Course (N=10)		Assumtions, Method, Alternatives, Integration,Communication	4		Resources	3.7						
Synthesis-Senior (N=84)		Problem ID, Method	3.3		Limitations	2.9						
Synthesis (N=54)		Problem ID	3.7		Alternatives Limitations	3.1						
Undergrad Course (N=8)		Creativity	3.4		Assumptions	2.8						

REFERENCES

REFERENCES

Albert, A. R. (2004). *Assessment method for student learning outcomes in general education at urban and metropolitan universities.* Unpublished Ed.D., University of Central Florida, Orlando, FL.

Allison, G., & Zelikow, P. (1999). *Essence of decision: Explaining the Cuban missile crisis.* (2nd. ed.). New York: Addison-Wesley.

Angelo, T., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (second ed.). San Francisco: Jossey-Bass.

Arnold, H., & Feldman, D. (1986). *Organizational behavior*. New York: McGraw-Hill Book Company.

Association of American Colleges and Universities. (2005). *Liberal education outcomes: A preliminary report on student achievement in college.* Washington, D.C.: Association of American Colleges & Universities.

Association of American Universities. (2003). *Understanding university success*. Eugene OR: University of Oregon Center for Education Policy & Research for the Association of American Universities.

Banta, T., & Associates. (1993). *Making a difference: Outcomes of a decade of assessment in higher education.* San Francisco: Jossey-Bass.

Banta, T. W. (Ed.). (2004). *Hallmarks of effective outcomes assessment*. San Francisco: Jossey-Bass.

Banta, T. (2007). *Selected references on outcomes assessment in higher education: An annotated bibliography.* Indianapolis, IN: Office of Planning and Institutional Improvement Indiana University-Purdue University Indianapolis.

Bauer, B. (2004). *Quality enhancement plan 2004*. Harlingen,TX: Texas State Technical College Harlingen.

Bauman, G. (2005). Promoting organizational learning in higher education to achieve equity in educational outcomes. *New Directions for Higher Education, 131,* 23-35.

Beyer, B. (1987). *Practical strategies for the teaching of thinking*. Boston: Allyn and Bacon.

Birnbaum, R. (1998). *How colleges work: The cybernetics of academic organization and leadership*. San Francisco: Jossey-Bass.

Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational Researcher, 34*(6), 3-15.

Booth, W., Colomb, G., & Williams, J. (2003). *The craft of research* (2nd ed.). Chicago: University of Chicago Press.

Boud, D. (1995). Assessment and learning: Contradictory or complementary? In P. Knight (Ed.), *Assessment for learning in higher education.* (pp. 35-48). London: Kogan Page.

Brazer, S. D., & Peters, E. (2007). Deciding to change: One district's quest to improve overall student performance. *International Journal of Education Policy and Leadership, 2*(3), 1-14.

Bresciani, M. (2006). *Outcomes-based academic and co-curricular program review: A compilation of good practices.* Sterling, VA: Stylus.

Briggs, C., Stark, J., & Rowland-Poplawski, A. (2003). How do we know a "continuous planning" academic department when we see one? *The Journal of Higher Education, 74*(4), 361-385.

Brookfield, S. D. (1987). *Developing critical thinkers: Challenging adults to explore alternative ways of thinking and acting*. San Francisco: Jossey-Bass.

Chen, H.-T. (2005). *Practical program evaluation: Assessing and improving planning, implementation and effectiveness*. Thousand Oaks, CA: Sage.

Cohen, M., March, J., & Olsen, J. (1972). A garbage can model of organizational choice. *Administrative Science Quarterly, 17*(1), 1-25.

Condon, W. (2004). Critical thinking in the classroom: Making it happen. Pullman, WA: Washington State University Critical Thinking Project.

Conley, D. (2003). *Understanding university success: A project of the Association of American Universities and the Pew Charitable Trust.* Eugene OR: University of Oregon Center for Educational Policy Research.

Cook, P., Johnson, R., Moore, P. M., P., Pauly, S., Pednarvis, F., Prus, J., et al. (1996). *Critical thinking assessment: Measuring a moving target :Report & recommendations of the South Carolina higher education assessment network critical thinking task force.* Rock Hill, S.C.: South Carolina Higher Education Assessment Network, Winthrop C0llege.

Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among five traditions.* Thousand Oaks CA: Sage.

Creswell, J. (2003). *Research design: Qualitative, quantitative and mixed methods approaches.* (2nd ed.). Thousands Oaks CA: Sage.

Davis, J., & Kerr, N. (1986). Thought experiments and the problem of sparse data in small-group performance research. In P. Goodman (Ed.), *Designing effective work groups* (pp. 305-349). San Francisco: Jossey-Bass.

Donald, J. (2002). *Learning to think: Disciplinary perspectives*. San Francisco: Jossey-Bass.

Drucker, P. E. (1999). *Management challenges for the 21st century.* New York: Harper Business.

Dunleavy, P. (2003). *Authoring a PhD: How to plan, draft, write and finish a doctoral thesis or dissertation.* New York: Palgrave Macmillan.

Ennis, R. (1993). Teaching for Higher Order Thinking. *Theory into Practice, 32*(3), 179-186.

Ewell, P. (1988). Outcomes, assessment and academic improvement: In search of usable knowledge. In J. C. Smart (Ed.), *Higher education: Handbook of theory and research* (Vol. 4, pp. 83-108). New York: Agathon Press.

Ewell, P. (2004). *The academy in transition: General education and the assessment reform agenda.* Washington, D.C.: Association of American Colleges & Universities.

Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction*. Millbrae CA: California Academic Press.

Facione, P. A. (1998). Critical thinking: What it is and why it counts. *California Academic Press.*

Facione, P. A. (2006, 2007). *Critical thinking: What it is and why it counts.*, from http://insightassessment.com

Gersick, C. (1989). Marking time: Predictable transitions in task groups. *The Academy of Management Journal, 32*(2), 274-309.

Gersick, C. (1991). Revolutionary change theories: A multilevel exploration of the punctuated equilibrium paradigm. *The Academy of Management Review, 16*(1), 10-36.

Gersick, C. (1998). Time and transition in work teams: Toward a new model of group development. *The Academy of Management Journal, 31*(1), 9-41.

Gladwell, M. (2002). *The tipping point: How little things can make a big difference.* New York: Little, Brown and Co.

Glesne, C. (1999). *Becoming qualitative researchers: An introduction.* (2nd. ed.). New York: Longman.

Glickman, T. S., & White, S. C. (Eds.). (2007). *New directions for higher education: Managing for innovation* (Vol. 137). San Francisco: Jossey-Bass.

Halx, M., & Reybold, L. E. (2006). A pedagogy of force: Faculty perspectives of critical thinking capacity in undergraduate students. *The Journal of General Education, 54*(4), 293-315.

Hersh, R., & Merrow, J. (2005). *Declining by degrees: Higher education at risk*. New York: Palgrave Macmillan.

Hofstadter, D. R. (1985). *Metamagical themas: Questing for the essence of mind and pattern*. New York: Basic Books.

Huba, M., & Freed, J. (2000). *Learner-centered assessment on college campuses: Shifting the focus from teaching to learning.* Needham Heights, MA: Allyn & Bacon.

Jennings, M. (2007). *Leading effective meetings, teams and work groups in districts and schools*. Alexandria, VA: Association for Supervision and Curriculum Development.

Jones, E., & Ratcliff, G. (1993). *Critical thinking skills for college students*. Unpublished manuscript, University Park PA.

Kasworm, C. E. (2004). Outcomes assessment in adult undergraduate programs. *The Journal of Continuing Higher Education, 52*(2), 13-22.

Katzenbach, J., & Smith, D. (1994). *The wisdom of teams: creating the high performance organization.* New York: HarperBusiness.

Kayser, T. (1995). *Mining group gold: How to cash in on the collaborative brain power of a group* (2nd ed.). Chicago: Irwin.

Kezar, A. (2001). Understanding and facilitating organizational change in the 21st century. *ASHE-ERIC Higher Education Report, 28*(4), 147.

Kezar, A. (2005). Moving from I to we: Reorganizing for collaboration in higher education. *Change, 37*(6), 50-57.

Kezar, A. (2006). Redesigning for collaboration in learning initiatives: An examination of four highly collaborative campuses. *The Journal of Higher Education, 77*(5), 805-838.

Kezar, A. (Ed.). (2005a). *New directions for higher education: Organizational learning in higher education.* (Vol. 131). San Francisco: Jossey-Bass.

Khademian, A. (2002). *Working with culture*. Washington, D.C.: CQ Press.

Kozlowski, A. (2000). *The organizational effects of a locally developed graduation exam at the American school of Guatemala.* University of Alabama.

Kuh, G., Kinzie, J., Schuh, J., Whitt, E., & Associates. (2005). *Student success in college: Creating conditions that matter.* San Francisco: Jossey-Bass.

Lane, J. (2007). The spider web of oversight: An analysis of external oversight of higher education. *The Journal of Higher Education, 78*(6), 615-644.

Latzer, B. (2004). *The hollow core: Failure of the general education curriculum.* Washington, D.C.: American Council of Trustees and Alumni.

Lave, C. A., & March, J. (1993). *An introduction to models in the social sciences*. Lanham, MD: University Press of America.

Leiberman, D. (2005). Beyond faculty development: How centers for teaching and learning can be laboratories for learning. *New Directions for Higher Education, 131*, 87-98.

Leskes, A., & Wright, B. (2005). *The art and science of assessing general education outcomes: A practical guide*. Washington, D.C.: Association of American Colleges and Universities.

Madison, B. (2006). Tensions and tethers: Assessing learning in undergraduate mathematics. In L. Steen (Ed.), *Supporting assessment in undergraduate mathematics.* Washington, D.C.: The Mathematical Association of America.

Maxwell, J. (2005). *Qualitative research design: An interactive approach.* (2nd ed. Vol. 41). Thousand Oaks CA: Sage.

Maxwell, J. (2006). Literature reviews of, and for, educational research: A commentary on Boote and Beile's "Scholars before researchers." *Educational Researcher, 35*(9), 28-31.

McGrath, J. (1989). Studying groups at work: Ten critical needs for theory and practice. In P. Goodman (Ed.), *Designing effective work groups* (pp. 362-391). San Francisco: Jossey-Bass.

McMillian, J., & Schumacher, S. (2001). *Research in education: A conceptual introduction.* New York: Addison Wesley Longman.

Merriam, S. B. (1998). *Qualitative research and case study applications in education: Revised and expanded from case study research in education*. San Francisco: Jossey-Bass.

Merriam, S. B. (Ed.). (2002). *Qualitative research in practice: Examples for discussion and analysis* (1st ed.). San Francisco: Jossey-Bass.

Mid Atlantic State University Assessment Office. (2007). Assessing competence in six areas at MASU: A summary of the first cycle 2002-2006. *Assessment Focus, 12*(1).

Mid Atlantic State University Assessment Office. (2008). Learning outcomes and student competence: Results of the 2006-2007 graduating senior survey. *Assessment Focus, 13*(1).

Miller, B. (2007). *Assessing organizational performance in higher education.* San Francisco: Jossey-Bass.

National Association of State Universities and Land-Grant Colleges. (2007). *The Voluntary System of Accountability*. Washington, D.C.: The National Association of State Universities and Land-Grant Colleges.

National Center for Postsecondary Improvement. (2002). *Beyond dead reckoning: Research priorities for redirecting American higher education.* Stanford, CA: National Center for Postsecondary Improvement, Stanford University,.

National Research Council. (2001). *Knowing what students know: The science and design of educational assessment.* Washington, D.C.: National Academy Press.

Nitko, A. J. (2004). *Educational assessment of students* (4th ed.). Upper Saddle River NJ: Pearson Merrill Prentice Hall.

Palomba, C., & Banta, T. (1999). *Assessment essentials: Planning, implementing and improving assessment in higher education*. San Francisco: Jossey-Bass.

Paul, R., & Elder, L. (2001). *Critical thinking: Tools for taking charge of your learning and your life*. Upper Saddle River, NJ:
Prentice Hall.

Paul, R., & Elder, L. (2003). *The miniature guide to critical thinking concepts and tools*. Dilton Beach CA: The Foundation for Critical Thinking.

Paul, R., Elder, L., & Bartell, T. (1997). *California teacher preparation for instruction in critical thinking: Research findings and policy recommendations*. Sacramento, CA: California Commission on Teacher Credentialing.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment.* Washington, D.C.: National Research Council, Committee on the Foundations of Assessment.

Peter J. Heart Research Associates. (2008). *How should colleges assess and improve student learning? : Employers views on the accountability challenge.* Washington, D.C.: The Association of American Colleges and Universities.

Phillips, V., & Bond, C. (2004). Undergraduates' experiences of critical thinking. *Higher Education Research & Development, 23*(3), 277-294.

Rainey, H. G. (2003). *Understanding and managing public organizations.* (3rd. ed.). San Francisco: Jossey-Bass.

Roberts, M. (1996). *An assessment of the writing proficiency examination at Delta State University.* Unpublished Ed.D., Delta State University, Cleveland, MI.

Rogers, E. M. (2003). *Diffusion of innovations* (5 ed.). New York: Free Press.

Schein, E. H. (1992). *Organizational culture and leadership*. San Francisco: Jossey-Bass.

Schwartzmann, H. (1989). Research on work group effectiveness: An anthropological critique. In P. Goodman (Ed.), *Designing effective work groups*. San Francisco: Jossey-Bass.

Senge, P. (1990). *The fifth discipline: The art and practice of the learning organization.* New York: Currency/Doubleday.

Stake, R. E. (2005). Qualitative case studies. In N. K. Denzine & Y. S. Lincoln (Eds.), *The sage handbook of qualitative research.* (3rd ed.). Thousand Oaks, CA: Sage.

Suskie, L. (2004). *Assessing student learning*. Bolton: MA: Anker Publishing.

Tapper, J. (2004). Student perceptions of how critical thinking is embedded in a degree program. *Higher Education Research & Development, 23*(2), 199-222.

Terenzini, P. (1989). Assessment with open eyes. *The Journal of Higher Education, 60*(6), 644-664.

The Annapolis Group. (2006, September 26, 2006). Liberal arts college presidents speak out on Spellings commission report. *Collegenews.org*.

Tuckman, B. W. (1965). Development sequence in small groups. *Psychological Bulletin, 63*, 384-399.

United States Department of Education. (2002). *Defining and assessing learning: Exploring competency-based initiatives.* Washington, D.C.: National Center For Education Statistics, U.S. Department of Education.

United States Department of Education. (2006). *A test of leadership: Charting the future of U.S. higher education. A Report of the commission appointed by Secretary of Education Margaret Spellings*. Washington, D.C.: U.S. Department of Education.

Walvoord, B. (2004). *Assessment clear and simple*. San Francisco: Jossey-Bass.

Weiss, R. S. (1994). *Learning from strangers: The art and method of qualitative interview studies.* New York: Free Press.

Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd. ed.). Alexandria VA: Association for Supervision and Curriculum Development.

Wilson, E. O. (1998). *Consilience: The unity of knowledge*. New York: Vintage Random House.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd. ed. Vol. 5). Thousand Oaks, CA: Sage.

CURRICULUM VITAE

Mary E. Zamon has lived and worked in many places. In addition to New York, Washington, D.C., Virginia and Missouri, England, France, Africa, the Caribbean, and the Czech Republic have been her home and place of employment. She has worked for the U.S. Agency for International Development, the Peace Corps, public and private schools and community colleges and universities. She graduated from Georgetown University's School of Foreign Service and has masters' degrees from there and Webster University. She has a twin sister, top drawer husband, four children, and a growing number of grandchildren.