The Scholarship of Teaching and Learning at EMU

Volume 3 *Spiraling Upwards: EMU Faculty Transform Through the Scholarship of Teaching and Learning*

Article 10

2010

Chapter 6 - Humanizing Academic Accountability: Embedded Assessment Gets at the Heart of Teaching and Learning

Peggy Liggit
Eastern Michgian University, pliggit@emich.edu

6

HUMANIZING ACADEMIC ACCOUNTABILITY: EMBEDDED ASSESSMENT GETS AT THE HEART OF TEACHING AND LEARNING

Peggy Liggit
Department of Biology
Eastern Michigan University

"The connections made by good teachers are held not in their methods but in their hearts – meaning heart in its ancient sense, as a place where intellect and emotion and spirit and will converge in the human self."
--Parker Palmer, *Courage to Teach*.

"You should never worry about your good ideas being stolen in educational reform, because even when people are sincerely motivated to learn from you, they have a devil of a time doing so."
-- Michael Fullan, *Change Forces: The Sequel*.

"One important feature of embedded assessment is that it 'blurs the lines' between teaching and assessment."
-- James J. Gallagher, *Improving Science Teaching and Student Achievement through Embedded Assessment.*

INTRODUCTION

Like many academic departments across the nation, ours was recently faced with a request to formally document student learning in our undergraduate and master's programs. The first two quotes presented above well illustrate the big question about creating a type of accountability system that will appease both internal and external stakeholders.

What method or model is best to document the teaching/learning process for program improvement?

As Parker Palmer's statement (1998) suggests, it is difficult to identify the exact methodology used by good teachers because their ability to teach from the heart is unique to each individual, and how do you begin to describe the process in which one brings together his or her "intellect, and emotion, and spirit, and will?" Accountability at institutions of higher learning, however, asks that we try to do just that. As a Higher Learning Commission (HLC) affiliated university, Eastern Michigan University (EMU) is required to demonstrate that students learn what we claim they can. This is a high-stakes endeavor. If the university is not able to document the student-learning process well enough, there is concern that institutional accreditation will be at risk. For faculty, the concern lies more with being labeled as poor teachers; not necessarily about instructional practices, but that good documentation about instructional practice is lacking.

Recent reports on the status of where institutions of higher learning are in their approaches toward accountability show that despite years of work there is still no best way to address institutional effectiveness (Engelmann, 2007; Shupe, 2008; Jaschik, 2009). Additional difficulties include the lack of assessment models and formal training in assessment for programs not accredited by professional organizations. As an example, in our department the certification program for secondary biology teachers must follow rigorous standards required by the National Council for Accreditation of Teacher Education (NCATE) and the National Science Teachers Association (NSTA). Faculty who teach in this program know and practice the pedagogy and assessment strategies related to teacher preparation. However, for our general undergraduate and master's programs, there are no specialized accreditation bodies (like NSTA) providing standards, and most faculty who teach in these programs have had little or no formal training in assessment practices. On a national level, there are several good examples of departments effectively documenting the teaching/learning process at their university (Maki, 2004), but sharing this process externally is hindered by the complex nature of trying to transfer educational reform practices from one institution to another (Fullan, 1999). Even if good, working assessment models are available and faculty are on board in wanting to implement them, there is often great difficulty in getting external models to work in exactly the same way at another institution.

What is unique about our journey is we started out by addressing program assessment, and, along the way, faculty had also discovered new ways to improve their teaching methodologies. Surprisingly, the methodology that indicated we were making the right kinds of steps to move our assessment efforts in a positive direction was using the reflective practice of embedded assessment. James Gallagher's statement (1999) above reminds us that embedded assessment blurs the lines between documenting what and how we teach with how well we teach. By writing down the ideas and decisions made during the *teach* (or *implement*), *assess*, *analyze*, and *adjust* steps of embedded assessment, we were able to provide a rich picture of the student learning and continuous improvement processes required for HLC reporting. An added benefit to this model is that it also fostered faculty development toward building an assessment culture, rather than a testing culture (Treagust et al., 2003). Thus, at the program level, we were able to use embedded assessment strategies to: 1) inform about the extent to which students were learning; 2) inform about the extent in which faculty developed better instructional strategies; and 3) enhance our documentation for capturing our methodology used during the decision-making process for program improvement. We found this method to be a more humanized approach to program assessment, because faculty were less resistant to the work, and the process aligned more naturally with what they were used to experiencing in their own classrooms.

Knowing that other departments would soon be following in our footsteps down the same pathway to assessment, I wanted to document and share the processes that our department followed as we began to create a formal system to evaluate student learning in biology. As a participant observer, my documentation procedures included taking detailed notes and saving all handouts from each faculty meeting and work session, analyzing course syllabi and curriculum maps (which is an exercise in aligning courses with learning outcomes), and processing hours of audio-recorded transcripts from faculty interviews. My analysis revealed that, yes, we too had struggled in trying to apply assessment models from outside sources, and we also had difficulty in capturing the teaching/learning process that showed our hearts were in right places.

In this chapter, I describe the ways in which embedded assessment served within our program documentation system, and I provide examples of what this type of assessment looks like from the work our department had accomplished in the last year. My intention here is to pose the idea of using embedded assessment as an alternative model, or a supplement to the traditional models, for addressing academic assessment. I also offer a discussion about the validity of this assessment methodology. Current literature on educational reform tells us that the most influential change forces are those that can remain balanced in times of greatest uncertainty and flux, such as the harsh economic climate and competitive global market we are experiencing now. Frameworks for reform include a combination of theoretical and applied practices, for

instance, those that: 1) support working environments that function on the edge between chaos and structure (complexity theory), Fullan (1999); 2) involve people working together in interdisciplinary groups (evolution theory of relationships), Fullan (1999); and 3) rely on working through discrepancies in professional perceptions in order to make conceptual changes (Piaget's theory of intellectual development), Wadsworth (1971) and Stepans et al., (2001). The embedded assessment indicators in this study show the approach aligns well with these change theories, thus providing further evidence that this alternative model provides a robust solution to the difficulties of implementing and sustaining institutional accountability.

CHALLENGES IN FORMALIZING PROGRAM ASSESSMENT

ASSESSMENT WORK TRANSFORMS DISCIPLINARY EXPERTS INTO INTERDISCIPLINARY NOVICES

In Pat Hutchings's *Ethics in Inquiry* (2002), she states, "The very idea of documenting and sharing work of teaching and learning – a core principle of the scholarship of teaching and learning – is new to most faculty" (p.1). We were confronted with this dilemma, as well, when our department discovered that the latest round of Program Review was very different from the reporting process we had experienced previously. In 1981, EMU initiated Program Review as a tool to evaluate the quality and effectiveness of programs in order to maintain institutional accreditation through the Higher Learning Commission (Eastern Michigan University, 2002). Since then, Program Review has, and still remains, a means for documenting the extent to which each program contributes to the greater good of the University's mission. Program Review has gone through several iterations over the years with the most recent including a major overhaul as a result of revisions in HLC accreditation criteria. In 2006-2007, EMU moved Program Review from a retrospective, hardcopy process to a dynamic Internet-based system called the Integrated Program Review and Continuous Improvement Cycle (IPR) process. In the IPR process, program faculty complete online templates which are aligned with HLC accreditation criteria.

Our department was one of four who piloted this new version. It was apparent that this round of Program Review, compared to the report we completed four years earlier, required much more data about our programs, particularly about student learning. In one section of the IPR templates, faculty were asked to respond to this statement, "*The program faculty reviews whether student learning outcomes are being met; assesses how well students are progressing through the program; and uses the information as a basis for programmatic changes*." (III. Criterion 3a) (Eastern Michigan University, 2007). At that time, the Biology Department had a sophisticated assessment system for the secondary teaching program (for NCATE accreditation), but nothing formal was in place for the undergraduate or master's programs. Previously, our methods and processes for assessing student learning were deemed effective in earlier rounds of Program Review. Several of our faculty had also received prestigious awards for their teaching and almost all faculty had participated in professional development to improve their teaching methodologies. Documenting instructional effectiveness is critical for our tenure and promotion process, and all faculty who had applied for tenure and promotion in our department have received it. To address HLC's III. *Criterion 3a*, we had only our past documentation practices to refer to, and our unstructured methodologies used for the general undergraduate and master's programs were now insufficient for meeting this particular HLC standard. Clearly, we had to create a new assessment system for these programs.

To determine our readiness for this work, I asked Biology faculty a series of confidence questions and asked them to rate their level of confidence (based on a 5-point Likert scale) about knowing and applying assessment terms and methodologies. On these questions (Table 1), confidence levels ranged from 25-50% for positive responses of agree or strongly agree. This survey indicated that more than half of us felt we were out of our comfort zone and feeling more like novices, rather than disciplinary experts.

TABLE 1: CONFIDENCE QUESTIONAIRE ABOUT PROGRAM AND STUDENT ASSESSMENT

Biology faculty (16 responders) with a rating of Agree or Strongly Agree to eight questions beginning with "I am confident…"		
Confidence Survey Questions		% Agree or Strongly Agree
1	"that I know the difference between a program goal and a program student learning outcome."	43.8 (n = 7)
2	"in my understanding of the term program goal."	50.1 (n = 8)
3	"in my understanding of the term student learning outcome."	43.8 (n = 7)
4	"in my ability to identify components that align with a particular student learning outcome."	43.8 (n = 7)
5	"in my ability to create a rubric to assess student work based on a particular student learning outcome."	50.6 (n = 8)
6	"in my understanding of Bloom's Taxonomy."	31.3 (n = 5)
7	"in my ability to apply Bloom's Taxonomy to student learning outcomes."	25 (n = 4)
8	"in my ability to create a curriculum map for our undergraduate program."	28.8 (n = 3)

WHERE TO START? EXPERT MODELS ABOUND, YET EXPERTS SAY MODELS DO NOT TRANSFER WELL

Increased emphasis on student learning has been a trend in higher education since the mid-1990s and numerous books are available to provide guidance throughout the process[1]. These books are useful in explaining assessment vocabulary and walking you step-by-step through a basic assessment plan. For instance, one should start first by writing outcomes, then proceed onto curriculum mapping, collecting and analyzing data, and finally making recommendations for program improvements. Understanding the concepts related to program assessment in higher education is not difficult. There are also chapters which include identifying reasons why assessment efforts are prevented or delayed and suggest tips for overcoming these roadblocks. Some of these tips suggest providing training and assistance to faculty, supporting leadership from those who understand the formative nature of assessment, and integrating assessment into campus-wide operations.

[1] (e.g., *Assessing for Learning: Building a Sustainable Commitment across the Institution*, *Assessing Academic Programs in Higher Education*, *Assessment in Practice*, *Assessment in Student Affairs*, *Policy on Assessment of the First College Year*, *Outcomes-Based Academic and Co-Curricular Program Review*, *Assessment Essentials* and *Taking Ownership of Accreditation*)

One of the struggles with academic assessment in higher education today is that assessment plans are often designed and written to appease external accreditation bodies and faculty see this work as an add-on to their workload (Shupe, 2008). Over time these efforts can fall by the wayside as faculty, staff, and administration are overwhelmed by overambitious data gathering or a decline in institutional support (Braskamp & Schomberg, 2006; Nichols, 2008). Engelmann (2007) portrays a very graphic description, "Under pressure from accreditors and others, just about every college and university has declared that it has some form of measuring learning. But we also know that assessment data are gathering dust in file cabinets around the country, and that learning outcomes have gone into syllabi and quietly died" (para. 2). Even worse, faculty raise questions about who reads these reports after they are turned in (Shupe, 2008).

Despite EMU's good intentions to improve Program Review, faculty have voiced numerous complaints about the system and process. The two greatest concerns were the faculty work load inherent in Program Review, including its impact on time for teaching and scholarly work, and faculty members' difficulty in identifying value derived from Program Review. Faculty also asserted (Eastern Michigan University, 2009) that assessment reporting across campus (e.g., Program Review, General Education assessment, external accreditation by professional organizations, and AQIP Action Projects,) appeared to be dissociated activities and redundant processes.

My desk is piled high with books and journal articles that read like Do-It-Yourself instructions on how to make assessment work at your institution. Many resources provide common sense and grounded suggestions for engaging faculty as they teach and assess learning outcomes, and attempt to renew the public's trust that higher education is responsive to its concerns to be more transparent and accountable. These suggestions include:

> "Draw on the expertise of professors who are already… doing effective work in teaching" (Engelmann, 2007, para. 5).
>
> "Leadership…both formal and informal, is critical" (Engelmann, 2007, para. 9).
>
> "Move toward reward structures that encourage and recognize… faculty collaboration" (Engelmann, 2006, para. 6).
>
> "Emphasize that collaboration to improve the teaching and assessing of student learning need not violate academic freedom or faculty autonomy" (Engelmann, 2006, para. 8).
>
> "Create communities of practice around teaching and learning issues that faculty themselves see as critical to their work" (Engelmann, 2007, para. 7).
>
> "Assessment can serve both those within the academy and those outside it" (Braskamp & Schomberg, 2006, para. 8).
>
> "Focus on creating a culture of evidence as opposed to a culture of outcomes… a rigidity of ends (teaching to the test) vs. the dynamic nature of learning, student development and solution making" (Braskamp & Schomberg, 2006, para. 9).
>
> "Connect assessment with development and change." (Braskamp & Schomberg, 2006, para. 13).

Michael Fullan (1999), one of the leaders in the educational reform movement, discusses why innovative ideas are "difficult to disseminate and replicate" from one organization to the next (p. 63). These efforts are not easily transferable, because it is difficult to capture in words all the "subtleties of the reform in

practice" (p. 63). The inability to replicate another's model can also come from replicating the wrong thing – "the reform itself, instead of the conditions which spawned the success" (p. 64). There can also be problems with scale: a reform method that was successful on a small -scale may not work well on a wider-scale. For instance, an accountability system that works well at a small private university of several thousand students may not work well at a public institution with twenty-five thousand students coming from non-traditional backgrounds.

CHALLENGES WITH IMPLEMENTING EXTERNAL ACCOUNTABILITY MODELS

Traditional program assessment models (Maki, 2004) recommend that faculty collect aggregate data on student performances that align with stated student learning outcomes. Typically, selected assignments or exam items that reflect these performances are identified in core courses that all students in the program take, and data on student scores and work samples are analyzed. The final step of this process involves reporting what improvements to the program will be made based on the evidence of the data; this step is often referred to as closing the loop. Using best-practice methods provided from the literature and assessment consultants, the Biology Department began to document program assessment using these similar approaches. Some of the problems we encountered in implementing these models included the following:

Vocabulary problems when writing outcomes – In various assessment resources, the term *objective* is used synonymously with outcome, competency, standard, and goal. There had also been confusion between the meaning of course and program outcomes. Since our department was one of the first to address *III Criterion 3a* (see above) in the Program Review documents, we initiated discussions toward a university-wide practice for adopting one set of vocabulary. EMU now uses the terms *student learning outcomes* as a description of what students will be achieving by the end of a program and the terms *program goals* and *program objectives* only refer to those activities that will be put in place to improve student learning as part of a department's annual plan.

Application of Bloom's taxonomy – We know that student learning should be at its highest level possible at the culmination of a particular program. We identified three major areas of learning in our undergraduate program – ability to conduct scientific investigations (Outcome 1), ability to communicate about scientific investigations (Outcome 2), and use of scientific knowledge (Outcome 3). Our understanding and ability to write coherent outcomes and performance standards for Outcomes 1 and 2 had been straightforward. However, it was difficult to clearly articulate what the learning looks like, for Outcome 3: "Students will *integrate* knowledge in higher-level thinking processes and problem sets." We have oscillated back and forth between using the verb *integrate* or *evaluate* (the highest level of Bloom's).

What should change if data doesn't imply making changes? Table 2 summarizes the comparison of student scores for Lab Reports I and II for the Winter 2009 and Fall 2009 semesters in the Cell and Molecular and Genetics Laboratory (Cell Lab) course. This course has been identified for collecting assessment data to indicate how well students are meeting Outcome 1 (scientific investigation) and Outcome 2 (communication) in the undergraduate program. The average student scores improved between Lab Reports I and II (a gain of 7.5 percentage points in WI09; and 4.7 for FA09) and students improved between semesters (a gain of 3.3 percentage points for Lab I; and 0.5 for Lab II). The number of students who received scores below 70% (C-) also declined between labs (dropping from 5 to 3 for WI09, and from 5 to 0 for FA09) and between semesters (dropping from 3 to 0 for Lab II). From this analysis, we concluded that students were performing relatively well over multiple semesters in the Cell Lab course, and no substantive change in teaching strategy was called for.

TABLE 2: PROGRAM-LEVEL SUMMATIVE ASSESSMENT INFORMATION FOR CELL AND MOLECULAR GENETICS LABORATORY

Winter 2009 and Fall 2009 Summary of Lab Reports I & II						
Program-Level Outcome 2: Students will communicate scientific knowledge, concepts, experimental results, and conclusions in written form.						
Assessment and Semester	No. Reports Turned In	No. Reports Not Turned In	Scores for Students Who Turned in Lab Reports			
			High	Average	Low	# of Students Scoring Below 70%
Lab Report I WI09	30	1	102%	77.3%	47%	5
Lab Report I FA09	28	4	102%	80.6%	62%	5
Lab Report II WI09	30	1	97%	84.8%	50%	3
Lab Report II FA09	30	2	101%	85.3%	71.5%	0

EMBEDDED ASSESSMENT – A HUMANIZED ALTERNATIVE TO TRADITIONAL MODELS

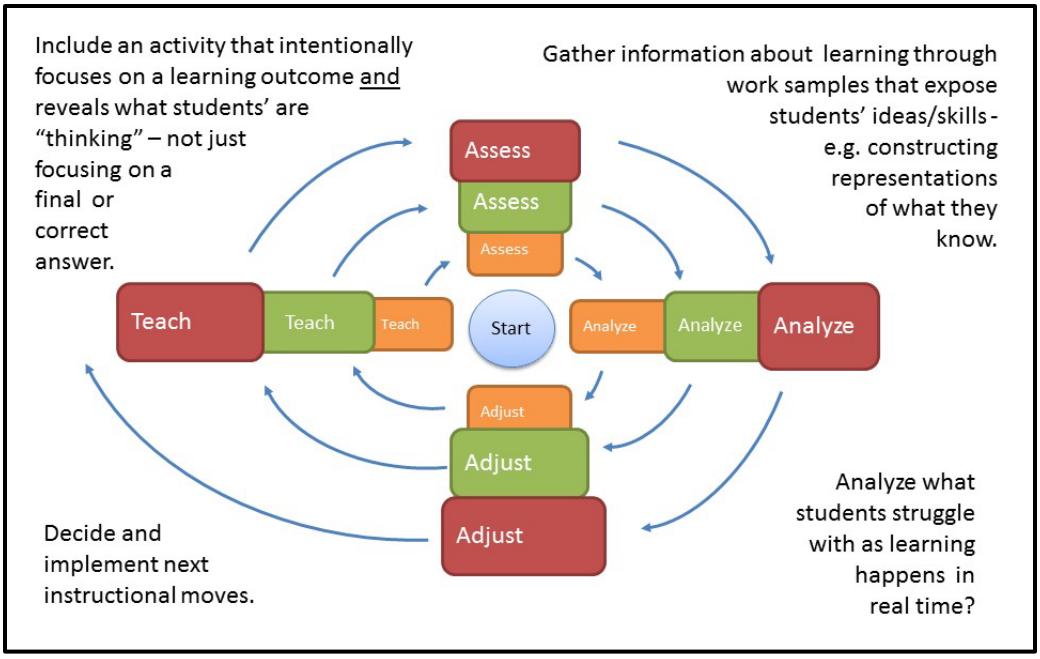
Embedded assessment (also known as ongoing, continuous, or classroom assessment), is a reflective practice performed during the instructional period that helps gain insight into a students' ideas and reasoning about the subject matter as instruction happens in real time.

Although embedded assessment is a continuous process, Gallagher (1999) breaks a cycle of embedded assessment down into four distinct parts as diagrammed in Figure 1. Part of my research for this work included analyzing 53 hours of audio-recordings and journal notes taken from interacting with seventeen Biology Department faculty during scheduled meetings, workshops, and one-on-one or small group interviews. I asked faculty to tell me about one particular question they had about student learning in at least one course they taught. As faculty became more engaged in discussing student performance in their courses, their passion and personal investment in the teaching and learning process became more and more evident. As much as faculty were excited when their students did well, they were clearly disappointed and concerned when discussing times when students were struggling with content or techniques. They were also very willing to voice concerns with problems outside of their courses that affected the program, such as: issues with advising, teaching environment, prerequisites, or course scheduling.

These interviews were opportunities for me to hear about the reflective practices of our faculty. In these conversations, I was able to identify all four steps of the embedded assessment process as they spoke. As I listened to faculty discuss the ups and downs of the teaching and learning process, I was reminded of Palmer's definition of "good teachers": those who teach with heart and make inspirational connections with their students. Although our primary mission as a department was focusing on ways to document student learning, in the act of discussing the student learning process, faculty brought into their awareness the continuous improvements they perform as a natural and almost unconscious practice of their teaching.

I wanted to test this idea further. The following sections describe three ways in which embedded assessment has been used to capture our reflective practices to inform about student learning in courses, to document our questions and methodology for decision making for program improvement, and inform about professional development.

FIGURE 1: EMBEDDED ASSESSMENT CYCLE: BLURRING THE LINES BETWEEN INSTRUCTION AND ASSESSMENT



1. EMBEDDED ASSESSMENT INFORMS ABOUT STUDENT LEARNING IN INDIVIDUAL COURSES

In the previous section, I mentioned the difficulties in reporting program improvements when assessment scores reveal students are already meeting course and program outcomes. When I interviewed Andy, the faculty member who teaches the Cell Lab course, he shared his concerns about those students who struggled with writing laboratory reports and the actions he took to help them. As Andy was talking, Gallagher's (2007) four-part embedded assessment model came to mind. The ideas of the model are summarized in Figure 1 and the steps are outlined below:

- Step 1. Teach or Implement – deliver an instructional plan based on an intended understanding for student learning and include an activity that intentionally reveals student ideas.
- Step 2. Assess – collect real-time information through student work to capture what students are thinking as they are engaged in learning; the intent is to seek out their reasoning, including naïve concepts and misconceptions.
- Step 3. Analyze - interpret the meaning and significance of students' current understanding.
- Step 4. Adjust - determine the next instructional steps needed to advance students toward the intended instructional outcome.

Figure 2, illustrates the use of the embedded assessment model to capture the continuous improvement mind-set that Andy reflected upon during the Winter 2009 and Fall 2009 semesters in the Cell Lab course. As a supplement to the assessment data presented in Table 2, the embedded assessment narrative (Figure 2) illustrates a richer picture of the teaching/learning process. In Andy's description, there is a sense that he is truly invested in how to best help students. Andy's initial question about student learning led him to ask more

questions. This question/answer process set up positive tension between wanting to know how students were doing with how to better assist them in the learning process. Andy's case study reveals a more authentic picture of what is really going on in the Cell Lab classroom compared to the traditional presentation of assessment data in Table 2. As an aside, when I asked Andy to help me report this information, he told me, "I did this in a couple hours on a Saturday night, after the kids went to bed," because he was so interested to see what the actual data looked like. He was actually excited to review and report his assessment data.

FIGURE 2: SAMPLE CASE STUDY: EMBEDDED ASSESSMENT APPLIED TO STUDENT LEARNING

Winter Term 2009 – Fall Term2009

TEACH (IMPLEMENT): Andy requires students to complete two scientific investigations and communicate their findings in Lab Report I and Lab Report II in Cell Lab (results found in Table 2).

ASSESS: The instructors of Cell Lab encourage all students to turn in drafts to receive feedback about their writing prior to turning in the final version of their lab reports. Andy wants to collect data from his WI09 courses to quantify how much better students who participate in the revision process perform. He poses the following questions: To what extent does feedback on draft reports influence the final lab report grade? Are students who turn in drafts only the higher achieving students, or do students from a range of performance levels benefit from the process? Andy gathers additional data about his students. He compares the data from students who turned in drafts and compares it to those that did not turn in drafts:

- Number of drafts - 12 students (39%) turned in drafts: 7 students turned in drafts for Lab I, 9 students turned in drafts for Lab II, 4 of the 12 students turned in drafts for both Lab I and Lab II.
- Comparison of Lab Report Scores: Students who turned in drafts to get feedback received higher lab scores than students who did not; on average 12 percentage points higher for Lab I and 6.7 percentage points for Lab II.
- GPA comparison: Students who turned in drafts were not all high achieving students. Of the students who turned in drafts, cumulative GPA's ranged from 2.28 – 3.74 (ave. 3.02).
- Course Grade comparison: Not all students who turned in drafts received a high final course grade (B or better) in BIO306W. The breakdown of the final course grades is as follows: 5 = A; 2 = A-; 3 = B-; 1 = C+; 1 = C-.

ANALYZE: Turning in drafts is a method that appears to supports all students; high - low achievers benefited of those who choose to take advantage of this voluntary system. Andy reflects, "Some students commented to me that this course has changed them, that they are better students because of all the writing they had to do."

ADJUST: Andy wanted more students to "buy into" the process of turning in drafts. He feels that students in his FA09 Cell Lab course would be more receptive to this idea if he shows them the quantified results of the data he analyzed from the WI09 semester.

TEACH: Andy presents his student learning study to his students in the FA09 Cell Lab course.

ASSESS: He repeats the study with his new crop of students. In Andy's words, "The trends are similar to the previous semester, although the numbers are somewhat less striking. Approximately 1/3rd of the students participated in the draft process for each lab report. A total of 13 of the 32 students (40.6%) who completed the course participated in the draft process - 6 students submitted drafts for both reports, and the other students submitted drafts for only the 1st or 2nd report.

ANALYZE: "In terms of perspectives about this data, I have a number: 1) It is clear that drafts improve the final written products (and therefore grades) of participants in the draft process and forces students to participate in a more realistic writing process. Students who participate also report higher satisfaction with the products of their efforts. 2) Students who participated in the draft process earned extra credit for their efforts, but the extra credit points themselves only partially explain the grade increase relative to the non-drafters. 3) I believe that more students would likely participate in the evaluation draft process if they weren't as pressed for time."

ADJUST: "It will be interesting to see if the percentage of students who participate in this process increases once Cell Lab becomes a 3 credit course. [Andy was a leader in making this curriculum change for the undergraduate program based on his previous observations of the teaching/learning process in this course]. Since Cell Lab is currently a 2 credit course, I believe that many students register for it without fully appreciating the workload, and therefore opt to take it on top of an already heavy workload. Perhaps this trend will change once the course is worth 3 credits."

The result of Andy's project generated additional questions about student learning in other laboratory courses. In the undergraduate program, students have a choice to take either the Cell Lab or Laboratory in Ecology. In the Laboratory in Ecology course, Julie the faculty coordinator, has a policy that requires all students to turn in drafts of their laboratory reports. Both Andy and Julie have had extensive training through Writing Across the Curriculum workshops, and they agree that turning in drafts is important. Andy wants to support student responsibility in context to encouraging students to work with a mentor – that in real life individuals should seek out assistance by their own initiative. However, due to the time- and labor-intensive nature of the evaluation draft process, Andy is protective about how he spends his time editing drafts. About 30-60 minutes per draft was spent providing formative feedback, and some students submitted multiple evaluation drafts. Andy also feels that it is important to invest time and energy to assist those students who take initiative and buy into the draft process rather than diluting the time and effort over the entire student population. Although turning in drafts is mandatory in the Laboratory in Ecology course, Julie cautions that some drafts are not of high quality even though she places a substantial number of points onto this first draft assignment.

After interviewing Andy and Julie, as a department we are now asking: Do different course practices (in regard to turning in draft reports), similarly support how well students achieve Outcome 2? We want to support student performance as well as academic freedom in the department. Our next steps are to develop a study comparing the draft policies between the two writing intensive laboratory courses to answer the question about student learning and faculty freedom in regard to setting course policy. The embedded assessment process will be used to capture the findings of this study too. We found the teach (implement), assess, analyze, and adjust steps used in capturing student learning in courses, for instance the Cell Lab, can also be used to describe continuous improvement activities at the program level as described in the next section.

TABLE 3: BIOLOGY DEPARTMENT SYLLABI EVALUATION IN 2003-2008

Syllabi Item Evaluated	Faculty Results: (n) = 19 people = 112 syllabi	Lecturer Results: (n) = 21 people = 62 syllabi
Course is referenced in context to the Biology UG Program		
▪ An individual always referred to the program	36%	52%
▪ An individual referred to the program half the time	32%	14%
▪ An individual referred to the program < half the time	32%	34%
Reference to Program Outcomes 1 & 2*		
▪ Language refers to inquiry (Outcome 1)	44%	40%
▪ Language indirectly implies inquiry (Outcome 1)	13%	8%
▪ Language refers to communication (Outcome 2)	37%	39%
▪ Language indirectly implies to communication (Outcome 2)	45%	24%
Inclusion of specific courses outcomes (objectives or goals)		
▪ Yes	47%	61%
▪ No	53%	39%
If outcomes were listed, how well were outcomes described?		
▪ High rating (descriptive language distinct learning level, such as Bloom's Taxonomy)	25%	21%
▪ Good rating (somewhat descriptive language)	28%	20%
▪ Poor rating (poorly described outcome)	47%	59%

2. EMBEDDED ASSESSMENT INFORMS ABOUT MEANINGFUL PROGRAM DECISION PROCESSES

O'Brien et al. (2008) discusses the importance of writing 'learning-centered' syllabi to articulate clear expectations and meaningful information to students. Using the components of a 'learning-centered' syllabus as our criteria, we evaluated 174 syllabi submitted by Biology instructors over a five-year period (Table 3). Mid-way through this analysis, one faculty member asked if there were differences in the levels of quality between those syllabi written by lecturers compared to those written by faculty. An embedded assessment case study is provided below (Figure 3) to provide a clear picture of the data analysis, key findings, and future actions proposed by our faculty based on these findings.

FIGURE 3: SAMPLE CASE STUDY USING THE EMBEDDED ASSESSMENT MODEL FOR PROGRAM IMPROVEMENT

IMPLEMENT: The intended understanding in this study was to establish a baseline of how our faculty and lecturers articulated information about the undergraduate program in their syllabi. Carmen, an undergraduate student in the Biology Department, conducted the syllabi evaluation as part of work she completed in a 2-semester independent study. She reviewed 174 syllabi for courses in the undergraduate program submitted to our department secretary between the years 2003-2008.

ASSESS: Syllabi were evaluated by the degree to which faculty and lecturers: 1) referenced their course to the Biology undergraduate program; 2) included program-level learning outcomes (or objectives or goals); 3) included course-level learning outcomes (or objectives or goals); and 4) described the learning activities presented in the course (learning objects or artifacts).

ANALYZE: The extent to which outcomes were included in syllabi and the quality of the written outcomes ranged from 20% - 61% for faculty and lecturers. Although lecturers outperformed faculty in referencing courses to the undergraduate program (61% vs. 36%) and inclusion of specific course outcomes (61% vs. 47%), faculty outperformed lectures in referencing program outcomes (44% vs. 40%) and writing outcomes using Bloom's Taxonomy (25% vs. 21%). From this analysis, we concluded that syllabi for courses in the undergraduate program did not consistently nor comprehensively articulate meaningful outcomes for student learning. Both faculty and lecturers require professional development in writing more program-based syllabi.

ADJUST: During our Fall 2009 Biology faculty retreat, we conducted a work session on improving department syllabi. Copies of our department syllabi from the 2008-09 academic year were spread across one table and this 'paper collage' made it quickly evident the difference in our syllabi format and language use. Faculty were split into groups according to those who taught common core curriculum classes and those who taught electives. As these working groups discussed the important biological content and student learning outcomes to be covered in each course, they were also asked to consider a handout called "Components of a Learning-Centered Syllabus."(Bauer, 2007). The conclusion of this work session resulted in numerous suggestions for creating a master syllabus for our department:

- Create a departmental syllabi template with boilerplate statements (University policies, student resources and conduct codes.)
- List program student learning outcomes that are assessed in the course and describe how the course "fits into" the structure of the program.
- List course outcomes and content topics that are agreed upon by all faculty who teach the course.
- When adjunct faculty and lecturers are hired, insist that they follow the same course and program outcomes specified in the master syllabi.

3. EMBEDDED ASSESSMENT INFORMS ABOUT FACULTY INSTRUCTIONAL DEVELOPMENT

Piaget proposed that in order for people to acquire new information, their intellectual development must transition from initial equilibrium to disequilibrium to a new equilibrium (Wadsworth, 1971). In the initial equilibrium phase, there is a balance between current knowledge and thoughts and perceptions about particular ideas. In the disequilibrium phase, people experience cognitive conflict where there is an imbalance between assimilation, the ability to take in new information, and accommodation, the ability to make sense of new information within the constructs of current thinking. Being in a state of disequilibrium is a strong motivator to take actions for further assimilation or accommodation. New equilibrium is achieved when assimilation and accommodation are back in balance. This entire process is called equilibration. Documenting the equilibration process would be key to demonstrating that faculty are reflective and progressive practitioners.

Table 4 provides an example of how embedded assessment can be used to capture a teachable moment between two faculty members. Table 4 (right column) includes a summary narrative from my interview that revealed a misperception about student understanding and how one faculty member (Phil) assists another (Gina) in trying an alternative approach to teaching the skill of 'pattern recognition.' The left column in Table 4 highlights the various phases of Piaget's Intellectual Development (Wadsworth, 1971; Stepan et al., 2001), and the corresponding four-steps of the embedded assessment cycle (Gallagher, 1999; 2007) are listed in the middle column. During the interview process we discovered that Gina had a preconception (initial equilibrium) about students' abilities, or lack of, for learning pattern recognition. Phil's description about how he was able to train students in pattern recognition (through discovery learning) revealed to Gina that she had a discrepancy between her assumptions about what she thought students could or could not learn (disequilibrium). At the end of the conversation, she seemed genuinely pleased and relieved to have a possible solution to help students master this important scientific skill. Gina tried out Phil's suggestion in her Fall 2009 Histology course. Though the data are not quantitatively supported, Gina saw the benefits that students gained as a result of trying out a new approach to pattern recognition. She was pleased with the positive responses of the students and is going to continue with this practice (new equilibrium).

OUR EMBEDDED ASSESSMENT APPROACH ALIGNS WITH CHANGE THEORIES AND PRACTICES

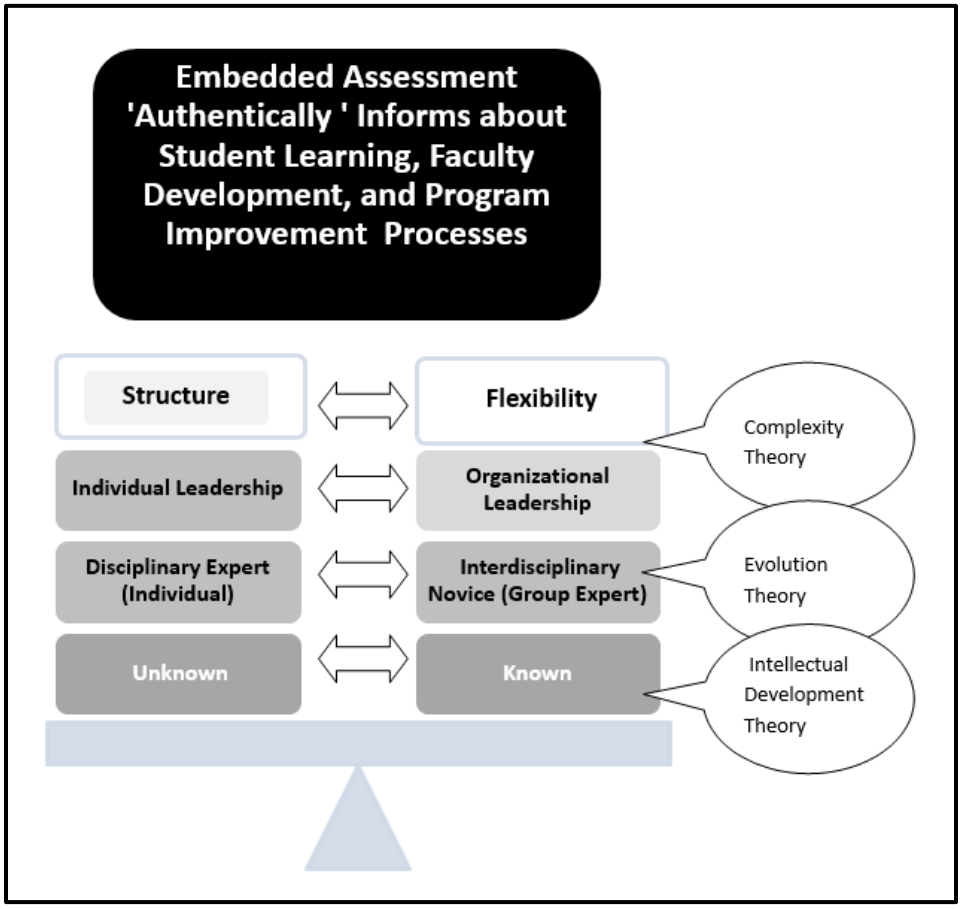
When the Biology Department first began program assessment work a year ago, we thought we would be able to follow a set model prescribed from external sources. Once we started encountering difficulties in applying a traditional assessment model, we had to append this model with our own ideas and practices. Fullan suggests that you craft your own ideas for intuitional change keeping in mind the premises of how change theories work (1999, 2006). Without consciously knowing, we applied some of the most powerful change theories described. The graphic organizer in Figure 4 summarizes these change theories and practices that apply well to our use of the embedded assessment model. These theories include: complexity theory, evolution theory of relationships, and Piaget's theory of intellectual development incorporated within the Wyoming TRIAD for professional development (Stepans et al., 2001). In the sections below, I describe each change theory or practice and provide examples of our application.

TABLE 4: EXAMPLE OF PIAGET'S THEORY OF INTELLECTUAL DEVELOPMENT APPLIED TO FACULTY DEVELOPMENT AND CAPTURED WITH A 4-STEP EMBEDDED ASSESSMENT PROCESS.

Summary of an audio-taped conversation between the interviewer and two department faculty in June, 2009 and a follow-up interview in January, 2010.		
Phase of Development	Embedded Assessment Step	Capturing the Conversation Highlights
Old Equilibrium	Teach	Phil asks Gina, "What do students struggle with most in your class?"
	Assess	Gina explains, in Histology class (the study of cell anatomy), some students "just don't get" pattern recognition. To her, it appeared that some students just have a natural ability, for recognizing patterns, where it seems others do not.
Disequilibrium	Analyze	Phil agrees that some of his students had the same difficulty in the courses he teaches where this skill is necessary for proper identification of animals. Phil then explains that he has discovered a method to train most students to become better at pattern recognition.
		Phil asks Gina how she typically sets up her lab to introduce the concept and skills of pattern recognition in Histology.
		She projects slides of cells on an overhead screen and gives a lecture on the important features of each cell type.
	Adjust	Phil suggests she set up a slightly modified version of this lab exercise. Rather than telling students what the cell patterns are in the projected slides, first allow students to draw and describe the various cell patterns they observe. At the end of the session, have the students compare and contrast their pattern descriptions with each other and then to those recommended by histology experts.
		Phil predicts that students will have a better understanding of cell patterns if Gina tries this approach.
		At the end of this conversation, Gina shows a sense of relief to now have a strategy to help her students in understanding this skill, and has a renewed sense of engagement and curiosity to test whether this new teaching technique will work. Gina is planned to test the following question/prediction about student learning in her Fall 2009 Histology course. If students create their own methods for distinguishing one cell type from another, they will be better at applying the skill of pattern recognition throughout the semester.
New Equilibrium	Teach Assess Analyze	January 2010 - Gina was interviewed again about her new perceptions on students' learning the skill of pattern recognition. In her own words: "I did try the method you suggested with the fall Histology class. Some of the students seemed to catch on quickly, others not so well. It's always difficult to ascertain whether these folks are just naturally more active and integrated into the courses and the others tend to be perhaps less forthcoming with answers. However, as the semester proceeded the responses improved. Patterns became easier to distinguish for most students. I'll continue on this tangent for next fall, thanks for the suggestion!"

FIGURE 4: EMBEDDED ASSESSMENT CAPTURES DECISION-MAKING AT MULTIPLE LEVELS AND IN STATES OF CONSTANT FLUX.

These indicators align with major change theories, further validating the approach.



COMPLEXITY THEORY

Complexity theory is based on the premise that all organizations are paradoxical, in that they are dynamic, are woven together by "nonlinear feedback loops," and are constantly moving in and out of states of equilibrium (Fullan, 1999). With complexity theory, decision-making happens best when organizations "operate on the edge of chaos." This is someplace in the middle between the balance of organizational structure and organizational flexibility. An organization must have enough flexibility to allow for people to figure out solutions in their own way – sometimes flexibility means allowing more time to meet deadlines; sometimes flexibility means allowing a shift in resources from one place to another and back again. Flexibility might mean allowing for differences in delivery of format. A one-size-fits-all approach is not conducive to the application of complexity theory. Application of complexity theory was evident in the following process of our assessment work:

1) At EMU, the university is in the process of formalizing program assessment through the Program Review reporting system. There is structure in the templates, but flexibility in the manner in which departments report assessment of student learning. Recall the problems we had with

applying external models. Although there were templates for Program Review, we were the first of four departments to go through the new process, and there were no internal models to follow. The environment had just enough structure to keep us on track, but enough flexibility to create our system the way we needed to (asking for vocabulary clarification and reporting in multiple ways, e.g., as seen in the summative assessment information in Table 2 and the embedded assessment case study in Figure 2.)

2) Currently, the university is flexible about due dates for providing data on all outcomes; however, this will be an expectation for the 2010-2011 academic year. We are appreciative that there is a bit more time as we are still in the process of clarifying student learning for Outcome 3.

EVOLUTION THEORY OF RELATIONSHIPS

Evolution theory of relationships, summarized by Fullan (1999) discusses the human dynamics involved around interpersonal relationships associated within an organization. In an academic setting, I would describe an individual tenure-track faculty member as a disciplinary expert in their field. Most of their experience with teaching and assessing students happens in individual courses or laboratories. Program assessment work shifts faculty out of their area of expertise and thrusts them into an interdisciplinary arena: a place where knowledge of the discipline meets the pedagogical practices of the discipline. For many faculty, this is unfamiliar territory (Hutchings, 2002). Evolution theory of relationships supports the idea that people working in groups fair better than those who work in isolation. In the realm of relationship theory, a group of people can come together as "interdisciplinary novices" (my term) pooling their ways of thinking, sensing and perceiving in such a way as to create a decision-making collective. This would function in a similar manner as a group of individual experts. Survey data (Table 1) indicated about half of our faculty, individually, felt unsure about applying assessment vocabulary and methodology. However, when we worked in pairs or groups, we were able to assist each other in making expert decisions. Evolution theory of relationships was most evident in these cases:

1) The group decisions on writing and assessing student learning Outcomes 1 and 2.
2) Phil and Gina's conversation about pattern recognition (Table 4).
3) The group decision on how to create the department master syllabi (Figure 3).

PIAGET'S THEORY OF INTELLECTUAL DEVELOPMENT: FACULTY DEVELOPMENT

The National Research Council in *College Pathways to the National Science Education Standards* (2001) reminds us that when planning professional development activities, it is through modeling and personal experience that teachers make conceptual changes in their previous perceptions and practices of teaching. This reference introduced me to the Wyoming TRIAD Professional Development Process where Piaget's theory of intellectual development is applied to professional development (PD) practices for K-12 teachers (Stepans et al., 2001). Here are some examples of our use of embedded assessment as a reflective practice for professional development:

1) Phil and Gina's conversation of pattern recognition (Table 4).
2) The syllabi analysis (Table 3) revealed about half our faculty did not understand how to apply student learning outcomes to their syllabi. At our FA09 retreat, faculty were confronted by this when they saw the results of the analysis study and were able to make meaningful decisions toward developing a departmental master syllabus.

CONCLUSIONS – SPRIRALING UPWARDS NOT CLOSING THE LOOP

It has become increasingly apparent that institutional accountability is here to stay (Engelmann, 2007; Shupe, 2008; Nichols, 2008). Although there is a wealth of books, articles and websites to assist in understanding and implementation of program assessment, accountability models that work well at one institution are not easy to replicate at others. The "teach, assess, analyze, and adjust" steps of embedded assessment assist in helping faculty confront their misperceptions and naïve assumptions about student learning (and understandings about student misconceptions and naive assumptions). This is counter to some traditional assessment practices that only evaluate students' current knowledge at a surface-level (e.g., multiple choice, true/false, fill-in exams) or do not include assessments with response items written purposefully to challenge students' conceptual thinking. The added benefit of using embedded assessment is that the practice is iterative and fosters faculty development. As faculty engage in this reflective practice, they are more prone to confront any shortfalls in the learning process and are more naturally driven to discover new and effective teaching methodologies.

Capturing the steps of the embedded assessment process as a narrative report provides a rich picture of the teaching/learning and faculty development processes. Thus, embedded assessment reporting is a more humanized and sustainable method for faculty, while providing for a richer, more in-depth picture for reviewers and public stakeholders. The embedded assessment method can also be utilized at the program level to document the real-time decision-making processes that departments make for continuous improvement (e.g., addressing curriculum changes and use and requests for program resources).

As Fullan (1999) reiterates, educational reform models are very difficult to transfer from one institution to the next, so it is better to craft one's own. What embedded assessment can provide is a method or model to capture faculty's reflective practices, i.e., the crafting of the visionary changes that are called for in the unique environment of each department and institution. It is the embedded part of assessment that brought us closer to capturing what good teaching looks like; because the lines are blurred between teaching and assessment, the process felt more natural and was, therefore, palatable for our faculty. Educational reform theories, such as complexity theory, evolution theory of relationships, and Piaget's theory of intellectual development support our innovative use of embedded assessment. By its authentic nature, it has the potential to establish itself as a transferable model outside of our department and institution. This is especially needed in departments such as those in the College of Arts and Sciences where there are few specialized accrediting bodies providing standards or competencies.

These insights couldn't come at a better time. It appears that the previously rigid constraints for institutional accountability are shifting the balance to favor organizations that work well under the laws of complexity and evolution theory. Shupe (2008) explains, "There is a real opportunity at hand" (p. 94). He's referring to the types of data that academic institutions must collect and analyze to meet the requirements for institutional accreditation. He further states, "Academic institutions are proving their ability to evolve in their attention to educational needs. Today, the primary constraints are aspirational – to what ends are colleges and universities willing (or not willing) to aspire?" (p. 94).

Over the last year, numerous decisions have been made as we worked on formalizing assessment practices in the Biology Department. Although we are well on the way in documenting the assessment of student learning in our non-accredited programs, we still have more to complete. Outcome 3 is still a question. What does student learning look like for 'integrating knowledge' or 'evaluating knowledge?' The department continues to wordsmith the intent of this outcome, and little progress has been made in collecting assessment data without this clarity. There are also several faculty concerns to address. Faculty in our department were part of the collective that voiced their concerns about the work load issues of Program Review in a Faculty Council Report (Eastern Michigan University, 2009). In a follow-up meeting to our Fall 2009 retreat, several faculty complained that the master syllabus template is getting too long and fear the added pages will be a

waste of paper because students “don’t read the syllabus” as it is. Gina also left us with a cautionary qualifier about documenting gains made in student learning: “Variables such as 'effort' are tough to factor in. Substantial improvement may reflect substantial effort in some [students], while others may improve very little due to a lack of effort rather than a lack of ability in pattern recognition.” Regardless of the challenges we have encountered, we are proud of the progress made and the professional development we have experienced.

Now that I have a more holistic understanding of program assessment, I don’t believe the public outcry for institutional accountability is expecting perfection. I think external stakeholders want and need a better idea of what actually happens inside college and university campuses – that the people inside are caring and engaged human beings. Supplementing or substituting a traditional assessment report with an embedded assessment narrative can provide a comprehensive view, a snapshot created in real time, as learning or decisions made about learning take place. This rich picture captures the ever spiraling upwards actions and intention for improving the learning experience, and this approach goes beyond closing the loop.

REFERENCES

Braskamp, L., & Schomberg, S. (July 2006). Caring or uncaring assessment. *Inside Higher ED*. Retrieved from http://www.insidehighered.com/views/2006/07/26/braskamp

Eastern Michigan University. (2002). Program review at Eastern Michigan University 2001-2011. Retrieved from http://www.emich.edu/public/aa/pr/PDFs/2001-2002/01-02PRManual.pdf

Eastern Michigan University. (2007). Academic review and continuous improvement cycle. Retrieved from https://programreview.emich.edu/v2_adpr_II_3_3.php

Eastern Michigan University. (2009). Faculty council March 2009 report on program review process evaluation. Retrieved from http://emichportfolio2.project.mnscu.edu/

Engelmann, D. (2007). Encouraging Assessment from the Ground Up, *Inside Higher ED*. Retrieved from http://www.insidehighered.com/views/2007/09/28/engelmann

Fullan, M. (1999). *Change force: The sequel*. Philadelphia, PA: Farmer Press.

Fullan, M., Peter, H., & Crevola, C. (2006). *Breakthrough*. Thousand Oaks, CA: Corwin Press.

Gallagher, J. J. (1999). Improving science teaching and student achievement through embedded assessment. *MSTA Journal*. Retrieved from http://www.msta-mich.org/index.php/publications/journalArticle/20

Gallagher, J. J. (2007). *Teaching for Understanding: A practical guide for middle and high school teaching*. Columbus, OH. Pearson, Merrill Prentice Hall.

Hutchings, P. (2002). *Ethics of inquiry. Issues in the scholarship of teaching and learning*. Menlo Park, CA: Carnegie Publ.

Jaschik, S. (2009). The assessment gap, *Inside Higher ED*. Retrieved from http://www.insidehighered.com/layout/set/print/news/2009/12/11/assess

Maki, P. (2004). *Assessing for learning building a sustainable commitment across the institution*. Sterling, VA: Stylus.

Nichols, J. P. (2008). Liberal Education: A “Both/And” Commitment. An Association for General and Liberal Studies (AGLS). Retrieved from http://www.bsu.edu/web/agls/AGLS%20Resources.htm

O’Brien, Grunert, J., Millis, B., & Cohen, M. (2008). (2nd Eds). *The course syllabus. A learning-centered approach*. San Francisco, CA: Jossey Bass.

Palmer, P. J. (1998). *The courage to teach*. San Francisco, CA: Jossey-Bass Publ.

Shupe, D. A. (2008). Toward a higher standard: The changing organizational context of accountability for educational results. *On the Horizon*, *16*(2), 72-96. https://doi.org/10.1108/10748120810874487

Stepans, J. L., Maureen, S., Yager, R. E., & Saigo, B.W. (2001). Professional Development Standards. In E. D. Sibert & W. J. McIntosh (Eds.), *College pathways to the Science Education Standards* (pp. 25-56).

Treagust, D. F., Jacobowitz, R., Gallagher, J. J., & Parker, J. (2003). Embed assessment in your teaching. Science Scope, *26*(6), 36-39.

Wadsworth, B. J. (1971). *Piaget's theory of cognitive development.* New York, NY: David McKay Co.