COMPARATIVE MODALITIES
Assessment Project Synthesis 2010/2011

Background

In the Fall of 2010, Vice-President of Academic and Student Affairs Sonya Christian asked the Lane Community College Assessment Team to address an accreditation comment: that the college should assess relative effectiveness of online, online/hybrid, and traditional classes. The team convened a group of twenty faculty in February, 2010 to discuss ways of comparing instructional modalities.

Over the following eighteen months, more than thirty faculty across the college, from General Education as well as Career and Technical programs, joined this conversation. Eventually the group's question morphed from 1) how to compare modalities to 2) what factors can be used in developing "a rubric describing the fundamentals of quality instruction across modalities," to 3) how can we devise "a rubric to inform and improve instruction across modalities?"

During this developmental period, faculty statistician Ben Hill analyzed broad college data (retention, success, etc.) surrounding online, online/hybrid, and traditional classes. The Assessment Team sought to extend Ben's quantitative analysis to consider how modalities of instruction differ and converge *qualitatively*, what strategies are most effective in each, and how to improve pedagogy across the college by calling to mind the institutional and national standards.

By January, 2011, a rubric had emerged, drawn from best practices measured by the CCSSE survey,[1] Quality Matters,[2] and the AAC&U's LEAP core learning outcomes.[3] The Assessment Team initiated a Spring, 2011, trial project, inviting faculty across the college to test the rubric in their classes. The rubric's stated purpose: *to compare within a curriculum the same course offered in different modalities*[4] in order to glean best instructional practices from each and to optimize quality of instruction.

The team awarded curriculum development funding to eight faculty applicants[5] to compare for instructional effectiveness separate modalities of the same courses: in developmental math, college math, science, and writing. The Team requested that each project engage self and peer assessment with the modalities rubric and provide a synthesis of their evaluations. At least one

---

1 Community College Survey of Student Engagement, http://www.ccsse.org/.

2 Quality Matters Program, http://www.qmprogram.org/.

3 Association of American Colleges and Universities: Liberal Education and America's Promise, http://www.aacu.org/leap/vision.cfm.

4 It is becoming increasingly difficult to separate out distinct instructional modalities in this era of increasing online presence both in the classroom and outside it. Typical instructional modalities include online, online/hybrid, course-management-system-enhanced (using, for instance, Moodle or Blackboard not as a stand alone course but as a traditional classroom enhancement), web-enhanced, and traditional (face-to-face only). We submit that many if not most courses may use a blend of more than one modality.

5 Applicants were encouraged to apply in teams so as to reduce subjectivity in evaluation. Hill and Selph applied together, while Taylor, Mitchell, and Jensen devised their projects and then requested peer participation in discussing course design and reviewing rubric ratings.

of the projects added a student assessment to the self and peer findings. This document intends to collate individual faculty syntheses into a meta-synthesis of the four projects, providing a snapshot of instructional practices and effectiveness in these courses at this point in time.

Pilot project background[6]

A. The four projects included

1. Chemistry (Brooke Taylor)
Brooke's goal was to assess and compare three chemistry classes using an online homework system (OWL) to traditional instructional methods (classroom tutorials, visualizations, and simulations) used for those classes over six previous trimesters. The project compared OWL-enhanced to Moodle (M)-enhanced courses. An ancillary goal of the project was to assess and recommend improvements for using Lane's course management system, Moodle, in these classes.
2. Mathematics (Ben Hill and Stephen Selph)
Ben and Stephen sought to compare Stephen's traditional (face-to-face or F) and Ben's hybrid (H) College Algebra classes, looking for strengths, weaknesses and ideas for mutual course improvement.
3. Developmental Mathematics (Adrienne Mitchell)
Adrienne proposed comparing new online and hybrid sections of basic developmental math to build on anecdotal observations of effective instructional strategies by systematically assessing and analyzing how these promote student engagement and learning.
4. Writing (Sandy Jensen)
Sandy intended to test the modalities rubric by comparing online and face-to-face (in a computer classroom or Moodle-enhanced environment) Technical Writing classes.

B. Modalities examined across the projects included

- Moodle as stand-alone online course (O)
- Traditional, or face-to-face (F)
- Hybrid, a class including both discrete online participation and classroom attendance (H)
- Traditional using Moodle as a course enhancement (M)
- Other online or web course enhancement within a traditional course (W)

C. Participating faculty summarized their findings within the modalities rubric. The rubric is intended to help faculty evaluate instructional design and effectiveness over six dimensions:

1. Does the course promote *civic engagement and collaborative learning*?
2. Does the course engage students by providing *multiple experiences with course content*?
3. Does the course make optimal use of *appropriate media, technology, tools, and resources*?
4. Does the course *promote student success* through instructional contact, feedback, and supportive information on available resources?
5. Does the course address or adapt to *students' preparedness*?

[6] Summaries below. Complete participant reports will be posted on the Assessment Team website.

6. Does the course clearly articulate and achieve targeted *learning outcomes*?

Project summaries

Preparatory and General Chemistry (CH221, 222, and 150)

Taylor requested peer review assistance from colleague Christine Andrews, who reviewed Taylor's Moodle site to assess her classes according to the rubric dimensions. Because Taylor's course was Moodle-enhanced (M rather than O), some rubric dimensions were not relevant. Although Taylor and Andrews discussed their evaluations, they submitted separate rather than conjoined ratings.

The three classes Taylor assessed differ primarily in depth of content coverage. She wanted to examine the effectiveness of the mathematics department's new OWL-enhanced classes (online homework system or W).[7] Taylor recognized that OWL did not make learning outcomes explicit even though they improved students' opportunity to achieve learning outcomes in varied ways. Taylor and Andrews found the student support dimension of all classes (M and W) relatively weaker than other dimensions, particularly in the lack of computer support for a course requiring computer homework assignments. And they found that both courses could be improved by offering more varied experiences with content, such as forum discussions to increase peer interaction and student reflection.

College Algebra (MTH111)

The two courses both scored high on all dimensions of the rubric, demonstrating their mutually strong design. Since both face-to-face (F) and hybrid (H) share the characteristic of live instructor contact, the closeness in ratings makes sense.

- On civic engagement, F scored higher than H; the evaluators expressing the superiority of an instructor to software in contextualizing course content. Collaboration and interactivity were strong in both F and H but for different reasons; each course used effective strategies to promote peer interaction and student engagement.
- F and H used the same variety of learning experiences. F, however, linked learning outcomes more explicitly as a part of test preparation.
- Both F and H used rich and meaningful tools. H, however, had the edge in terms of state-of-the-art software and video.
- Student support elements were likewise strong in both F and H, and again there were trade-offs: F providing more face time and office hours but H distinguishing itself it its immediate feedback capabilities.
- Student preparedness in math ability is equivalently addressed in F and H, thanks to college testing and enforcement of prerequisites. The rigorously structured H holds students more accountable for attendance and deadlines, suggesting more flexibility and less firm accountability in F, a possible modality-dependent factor. As for student background, both instructors rely on personal socio-cultural sensitivity, a characteristic that may be less likely to depend on modality of instruction than on the nature of an instructor. This does suggest, however, that an instructor-*independent* course, i.e. stand-alone online, will be less likely to adapt to different backgrounds and learning styles.

[7] Andrews was not able to access this OWL program; rubric evaluations on this element are Taylor's alone.

• In terms of learning outcomes, although both courses provide exercise in critical thinking, assessing this core college outcome is hampered in H by computer grading.

Hill and Selph noted a design flaw in the rubric's point values and suggested a revision. Their observation sounds a cautionary note on drawing statistical conclusions from this year's projects.

Whole Numbers, Fractions, and Decimals (MTH10)

Mitchell partnered with colleague Judy McKenzie to complete and discuss rubric evaluations. Mitchell echoed the Hill/Selph caution on rubric point values, suggesting a different correction from theirs. Mitchell's O and H classes scored closely on most dimensions and most evaluations fell in the robust range. Since both the O and H class include fundamental online elements, most prominently an open educational resource textbook, we find pronounced similarities in evaluations. Mitchell took care to enhance instructor accessibility by developing screen casts and to promote student engagement with interactive online activities. This instructional foresight may have prevented more of a gap between O and H students.

While both F and H provided opportunities for student collaboration and interaction, Mitchell and McKenzie found that student-student interactivity was stronger in H through face-to-face class participation. However, they recognized more variety of learning activities in O; this advantage recurred in two rubric dimensions.

Elements targeted for improvement include civic engagement and adaptation to diverse levels of student preparedness. One issue of student preparation arose from college registration practices; students were unaware they had signed up for a hybrid class, indicating they did not understand what that meant. Finally, Mitchell and McKenzie recognized that MTH10 could more explicitly link instructional content to learning outcomes.

Technical Writing (WR227)

Jensen collaborated with colleague Amy Beasley on this project, Beasley assessing O. To optimize validity in their comparison, Jensen had prepared her project by closely aligning the two versions of the class before the term began. Thus, all three evaluators—instructor, peer, and student—were looking at parallel course designs, each attempting to address the rubric dimensions.

Both Beasley and Jensen homed in on the issue of student preparation. Beasley found student preparation expectations implicit in the online course, while Jensen treated the issue more harshly in both M and O. The demands of the course appeared to exceed students' level of preparation. To close the assessment loop, Jensen suggests ways of addressing this challenge: by embedding First Year Experience/On Course strategies to support students' time management skills. Jensen also envisioned modularizing the course into four one-credit, self-paced components to moderate insufficient student preparedness.

This project was unique among the four in integrating student assessment of the rubric dimensions in both M and O. Naturally, while an instructor's *intent* is to achieve comparably strong results in each course, often student evaluations illuminate differences in *impact,* as Jensen found in her M and O courses.

Results

The data table below represents a collation of data from all but one project that had not submitted cumulative data. Overall, the project data are problematic in several ways:

• The Taylor/Andrews project does not reflect a complete peer review. The two instructors directly assessed M only, and rather than collaborating on a single rating, they rated the courses independently and came up with several divergent ratings. In their analysis they pointed out that the Moodle-*enhanced* course did not include some rubric dimensions, elements that were addressed face-to-face in the classroom. The data below reflect the instructor rating only.

• Hill and Selph submitted two statistical summaries, one based on a proposed revision of a flaw in the rubric. The table below resolves the flaw they cited by using Mitchell's solution of presenting a rating out of total numbers possible, thus eventually shown as a percentage.

• The headings of Robust/Moderate/Developing are difficult to translate into meaningful descriptors across projects. The table attempts to adjust placement of point totals according to the following percentages:

—Robust/16-24 = 80% or more in the modality
—Robust-Moderate/12-16 = 60-79%
—Moderate-Developing/6-12 = 40-59%
—Developing/1-6 = 0-39%

• Mitchell and McKenzie submitted dual ratings for one item in the civic engagement dimension: promoting scholarly values and breadth of knowledge. Because they were the only faculty to do so, their two scores on that item have been averaged into a single rating.

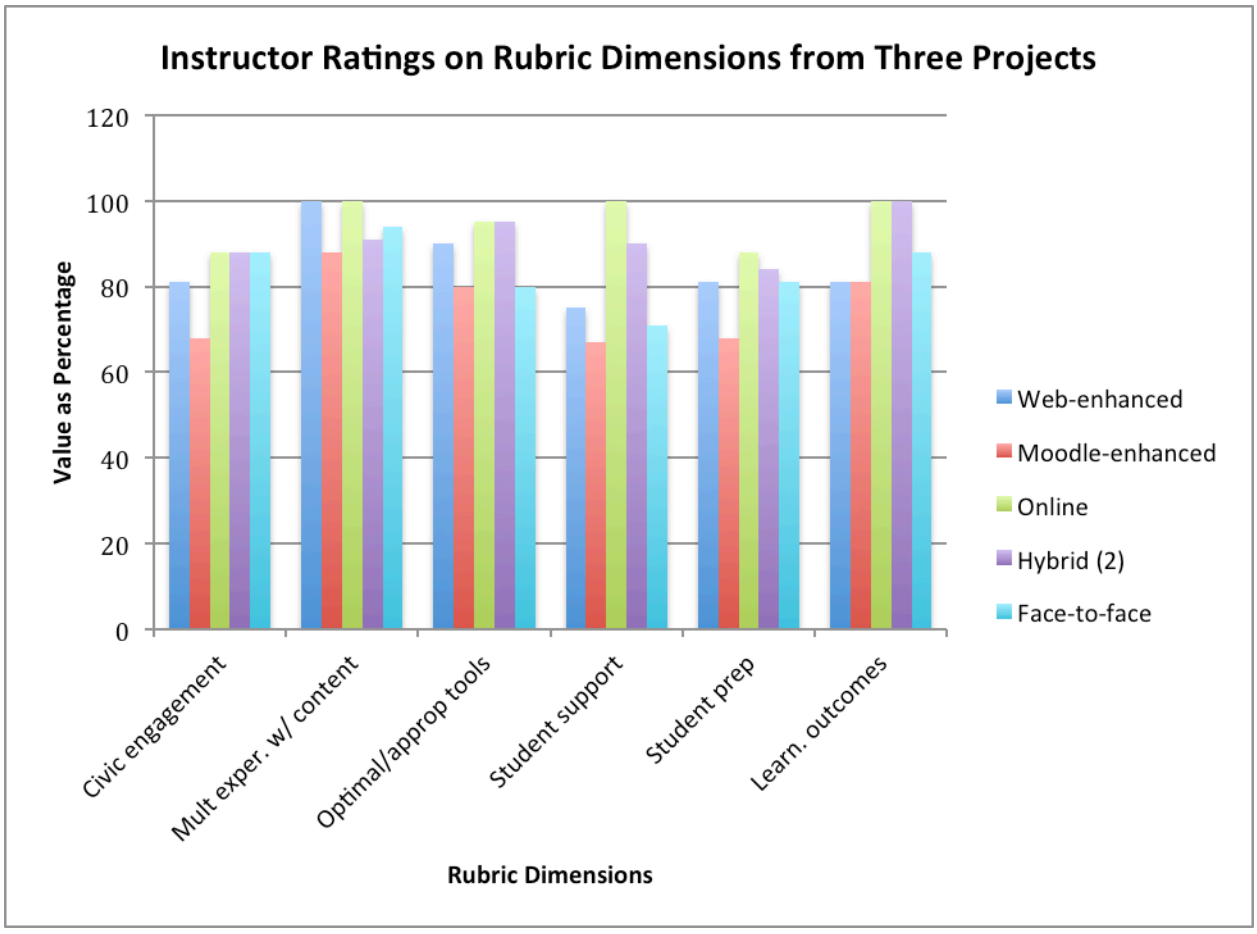
• Jensen and Beasley did not coordinate their data nor synthesize their evaluations into a single report. They did submit a detailed chart for student rubric evaluations. This source provides provocative and telling observations.

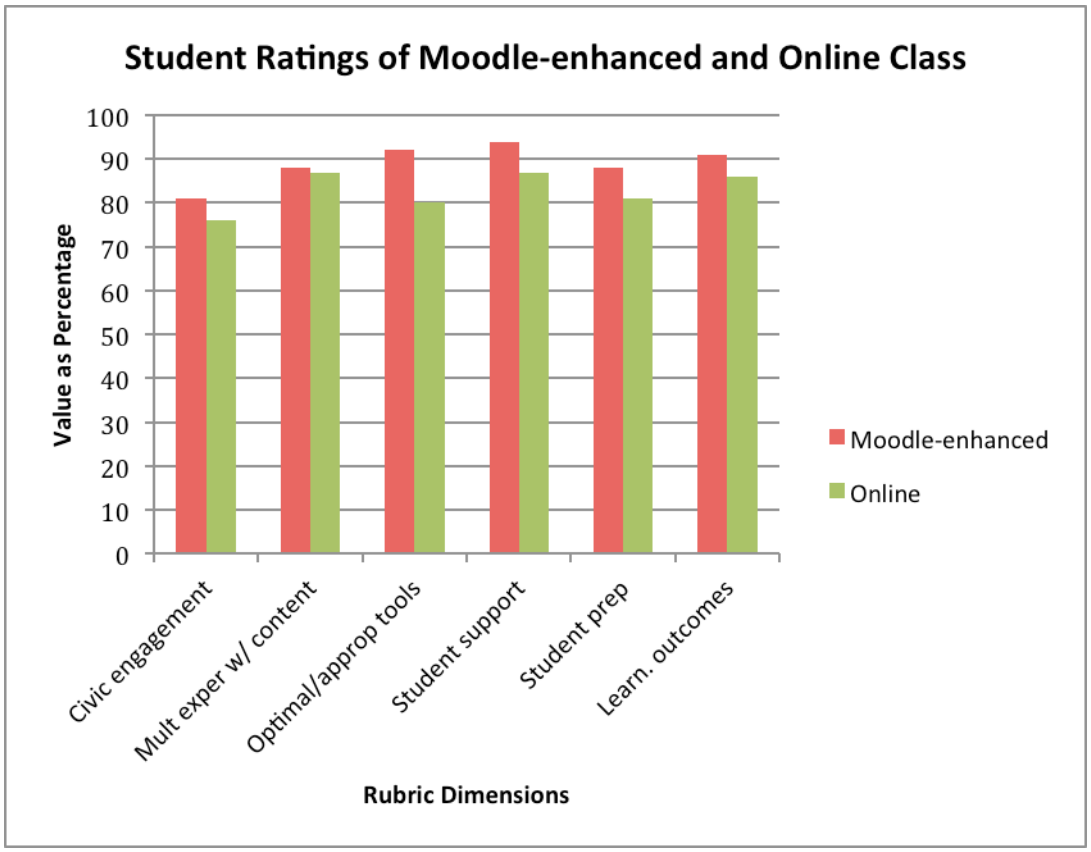
The following table and charts illustrate instructor ratings for three of the four projects (Table 1 and Figure 1) and the student ratings for one project (Figure 2). The scores in Table 1 have been computed as percentages in Figure 1. Likewise, student ratings on the rubric have been converted to percentages for ease of comparison.

Table 1. Modalities Rubric Results from Three Assessment Projects 2011

Dimension	Robust	Moderate		Developing
	16-24	12-16	6-12	1-6
Civic engagement and collaborative learning	H=15/16 O=14/16 F=14/16 W=13/16 H=13/16		M=11/16	
Multiple experiences with content	W=16/16 O=16/16 F=15/16 H=15/16 M=14/16 H=14/16			

Use of optimal/appropriate media/technology/ tools/resources	H=20/20 O=19/20 W=18/20, H=18/20, F=16/20	M=16/20		
Student support	W=18/24, O=24/24 H=19/24, H=24/24 F=17/24 M=16/24			
Student preparedness	W=13/16 F=13/16 H=13/16 O=14/16 H=14/16		M=11/16	
Learning outcomes	O=16/16 H=16/16 W=13/16 F=14/16 H=13/16 M=13/16			





Interpretation

Lane's Modalities Project processes and results require that we exercise restraint in forming conclusions.

1. The rubric is imperfect. Items within dimensions are often redundant and ambiguous. The synthesis form is numerically inconsistent with the rubric. Although developers tried to correct for such flaws, the Assessment Team needs to revise the descriptive criteria (robust, moderate, developing) and numerical ratings on the synthesis reporting form.

2. Project differences challenge our ability to draw meaningful conclusions. Each project compared different modalities. No single project compared the same two modalities. Ratings for one modality may have meaning in relation to its comparator as opposed to ratings in the same modality from different instructors.

3. Participants apply the rubric in a variety of ways. Project leads were not trained or normed in applying the rubric, although all had been involved in some capacity in rubric development. Two of the peer reviewers had had no previous contact, before being approached by colleagues, with the Comparative Modalities Rubric or its development. Naturally, students had not been normed in rubric evaluations.

4. Results reflecting diverse instructional methods are ambiguous. The highest ratings from instructors appeared in multiple experiences with content; here it appears that the presence of online instruction as part of the curriculum sets a standard above that of traditional classes.

Online instruction has brought about a new range of instructional activities. The student assessment in this dimension also was closer between modalities than in any other dimension. The caveat here is that our research has not demonstrated that multiple experiences with content promote learning more effectively than a narrower range of learning activities.

5. The College, Instructional Technology, and the Assessment Team need to reflect on and discuss the role of student learning outcomes assessment in comparing modalities. O and H courses achieved top scores in making learning outcomes known and linking them to course activities. This may be a result of the college affiliation with Quality Matters,[8] which considers the link to learning outcomes a critical course element. We cannot infer that learning outcomes are more readily achieved by O and H modalities; this is not a factor measured by the modalities rubric. To date, comparative *learning* between modalities has not been measured at Lane.

6. Some of the lowest scores (though nearly all scores were above the 70 per cent level) had to do with knowing and responding to student preparedness. This is consistent with concerns across the College about improving academic advising and student placement testing and reviewing prerequisites. These goals are at this date a part of Lane's Roadmap for Student Success, under the leadership of the Association of American Colleges and Universities.

7. Some potentially meaningful data emerges in the student rubric evaluation. Here we find some equivalencies between O and M, particularly in the multiple experiences with content dimension. Although this dimension diverges from evaluator to evaluator, the overall measure of multiple experiences with content is nearly identical.

The most pronounced difference between modalities was in civic and collaborative engagement, where the M (face-to-face with Moodle) class scored higher. This rating reflects the instructor evaluations, which expressed concern with finding ways to engage this pedagogical component in an online environment.

In all, the student ratings of the online class in comparison to the web-enhanced class were lower, suggesting that a) online students may be harsher critics of their learning experience, or b) online students' reduced face-to-face exposure to the instructor emboldens them to express more extreme positions, or c) online students are less certain of the value of their class. Of course there are countless other explanations for this finding, including that this is a single study. The student evaluation factor calls for more intensive study and broader instances for analysis.

Closing the Loop: What have we learned about comparing instructional modalities?

Key findings in these projects include what we discovered about the modalities rubric, what different evaluators have to offer, what we can tell and cannot tell from project results, and how we might adjust the modalities rubric as we move forward.

1. Evaluators: the review process

Each project incorporated a peer-evaluation system to check and explore how to apply rubric elements to their courses. The peer review added confidence and occasional validity to project summaries. It became apparent in the Taylor and Andrews study that instructors may require

[8] Lane adopted the Quality Matters standard around 2006, as it expanded its online course offerings.

support from IT staff to access online course elements.

Faculty tended to score their own or a colleague's course at the highest levels. Jensen's student assessment results indicate that students and faculty can score quite differently. Here we are torn between dismissing a student rating on the basis of students not being in sync with educational jargon and wondering about faculty detachment in assessing the strengths and weaknesses of their own or a peer's course. What is evident to an instructor may not be so to a student and vice-versa. However, Jensen's self-evaluation produced one of the most critical assessments of student preparedness, a weakness not perceived at all by student evaluators. The Assessment Team might consider requiring or encouraging a student rubric evaluation in future runs of the modalities project.

To produce a meaningful synthesis, the project coordinators need to determine the most effective format for reporting results and to promote adherence to this standard with project participants, perhaps through more frequent check-ins throughout the term.

2. The rubric: should it be revised?

The math instructors in this set of pilot studies all pointed out computational flaws in the rubric, for which Hill/Selph and Mitchell/McKenzie proposed remedies. This project synthesis has chosen Mitchell and McKenzie's approach, but a better solution would be to group and report ratings in way that more rigorously protects statistical validity.

Mitchell and McKenzie also pointed out several redundant or ambiguous items in the rubric and altered scoring to adjust to these problems. In advance of the next Modalities Project run, the Assessment Team should correct the items cited.

3. Generalizing about modalities

Summaries of this research by dimensions is less revealing than comparisons by item analysis. Two project leads submitted full commentaries on the rubric in all its dimensions. Two submitted the summary form alone. The more practical, instruction-level insights derive from specific items under general dimension headings.

For example, whereas the Hill/Selph review found strengths in the F and H courses' instructor contact, Mitchell's O and H comparison sites strengths in the interactive online resources and activities and achieves face-to face instructor contact through screencasts. Both mathematics courses in these two projects are challenged by the goal of civic engagement and collaboration, a dimension which might score higher in a social science course. The item analyses, then, serve to inform instruction at the course level and to bring into focus areas ripe for improving instructional quality, regardless of and across modalities.

4. Usefulness of the Comparative Modalities Project

In the rubric development throughout 2010-2011, project participants suggested that the modalities rubric could be used

- as a guideline in course development
- as a checklist for an instructor to ensure that students are receiving important course elements across a variety of instructional modalities
- as an indicator in program assessment and curriculum planning to determine numbers

of sections to offer in one or another modality

To compare relative effectiveness of achieving desired learning outcomes across modalities, the current modalities rubric is not an appropriate tool. This rubric can complement other modes of assessment, such as core abilities assessment by rubric and other program assessment tools. Or the Assessment Team can revise the rubric to address learning outcome comparisons between modalities.

Participating faculty in the 2011 assessment pilot expressed universal appreciation for the modalities rubric as an aid to course development and a checklist for course revision. They felt it highlighted weaknesses and caused them to explore different instructional strategies to improve course quality.

The Assessment Team and the College were fortunate to have as project participants experienced, creative, and talented faculty alert to controversies in educational strategies and active over a period of time in college discussions on best instructional practices. All had been involved in the Comparative Modalities Rubric development; they possessed broad and detailed understanding of our challenges and goals. At this juncture, these faculty can encourage and support continued research into relative effectiveness of instructional modalities and modes of student engagement.

The Office of Academic and Student Affairs has confirmed its ongoing financial support for professional development pay for faculty in comparative modalities research. Consequently, Lane's Assessment Team will continue to coordinate the process of soliciting, selecting, monitoring, and publishing this professional activity.