

Lawrence University
Proposal:
Researching Assessment Methods in
Tutorial Education

Abstract

The purpose of this two-year project (September 1, 2008-August 31, 2010) is to develop and test a method for the formative and summative assessment of tutorial education in a consortium of three liberal arts colleges: Lawrence University; College of Wooster; and Williams College. Tutorials are an intensive form of one-on-one or very small group instruction using essay writing and discussion as primary educational activities. Although there is general agreement that tutorials in liberal arts subjects, including the visual and performing arts, support students' habits of mind and abilities to think independently and other self-regulated learning behaviors, the participating colleges lack methods for assessing student work for these outcomes. Faculty and administrative representatives of the colleges will develop a shared assessment model (SAM) in a face-to-face workshop in fall of 2008. The SAM will define relevant learning outcomes, evidence and scoring for student work at all stages of the tutorial - - research, writing or project progress, critical discussion, and presentation or performance. Six instructors (two from each school) will: apply SAM in Tutorial Assessment Studies across multiple subjects in winter/spring 2009; refine SAM in a conference in fall 2009; and apply the refined SAM-2 in Tutorial Assessment Studies to both tutorials and non-tutorials (3 students each) in winter/spring of 2010. This second round of studies will provide iterative data for evaluating the usefulness of the SAM both in tutorials and other liberal arts courses. The expected outcomes of this project are: Between 30-42 research cases, depending on the number of tutorials with one student and those with two students; a Shared Assessment Model for evaluating individual student work in tutorials; a Tutorial Assessment Guide including samples of work and grading for application to other liberal arts institutions; recommendations for tutorial instruction in different disciplines; and a proposal for the modification and application of SAM-3 to assess four-year college outcomes.

Assessment of College Tutorials to Improve Student Learning and Accountability

The present project responds to principles and commitments to action communicated in the "New Leadership for Student Learning and Accountability" by the Association of American Colleges and Universities (2008). The participating colleges will *develop quality standards themselves* and not rely on external agencies. These quality standards refer to learning goals and assessments *anchored in the tutorial and liberal arts course curricula*. Perhaps because tutorials represent the epitome of the liberal arts approach, liberal arts educational philosophies stipulate anticipated outcomes, such as learning independently and thinking for themselves, that are very similar to outcomes of tutorials. However, there are currently no assessment strategies for measuring the extent to which such liberal arts outcomes are achieved. Do liberal arts colleges, in fact, achieve the goals that we advertise as integral to our educational missions? Is it justifiable for us to make the learning claims that are the primary source of our high tuition costs? Our research project posits that if we develop a successful method of tutorial assessment, it may also prove useful in measuring the same learning outcomes in other liberal arts courses.

Thus, the project will address two major questions regarding the development of an assessment instrument: 1) Can a standardized assessment be developed for liberal arts outcomes in tutorials? 2) Can the assessment inform the design of assessments of other courses and, by extension, potentially, to the assessment of a four-year liberal arts education?

Having referred to tutorials as the "epitome of the liberal arts approach", we need to justify this claim. In a recent review of the role of dialogue and tutoring, Hacker and Graesser (2007) wrote: "The superiority of one-on-one tutoring over traditional classroom instruction has been well documented". In an experimental K-12 study, Benjamin Bloom (1984) demonstrated that tutorial instruction had a two standard deviation superiority over all other forms of instruction. Cohen et al. (1981), in an excellent meta-analysis, came to the same conclusion. Yet, Hacker and Graesser continue: "However, the identification of exactly what aspects of one-to-one tutoring contribute to this superiority has received little attention".

We concede that no comparable studies have been conducted in liberal arts education, but believe that most colleges in the Annapolis Group, for example, have some form of one-on-one instruction (one-on-one is the same as tutorial education) covered in our taxonomy: tutorials used to teach disciplinary subjects; tutorials used to conduct independent studies; tutorials used to conduct research. We have not done a survey of these colleges to assess the extent of one-on-one instruction, but the volume of student research, the majority of which is individual work, has increased dramatically in recent times as an important transition to graduate work. The number of colleges offering independent studies is comparably large. Add to this the increasing emphasis on "Senior Capstone-type" projects as culminating work in undergraduate education. We think, therefore, that we must pay attention to "what aspects of one-on-one tutoring contribute to this superiority".

We have already argued that our work would have application to all liberal arts colleges that offer one-on-one instruction, whether these are called tutorials, independent studies, research projects, or capstones. We think that many tutorials are already being used as culminating experiences. Because the demand for the assessment of liberal arts education focuses on how well the four-year undergraduate education serves individuals, we think that tools that are developed to assess individual outcomes in tutorials are a plausible path of inquiry. Also, recall that tutorials are the only form of instruction that integrates disciplinary research, writing,

presentation and one-on-one discussion. Any four-year outcome measure would need to assess how well students combine the several skills they acquire during their liberal arts education.

Conceptual Model: Assessment of Tutorials

In a recent study of Oxford tutorials, Beck (2007) proposed that the learning outcomes of tutorials are to develop students' abilities to work independently, think for themselves, develop a skeptical orientation, acquire mental flexibility, demonstrate creativity and imagination, learn to argue, engage in continuous self-assessment, and produce a documented example of original work. Concurrently, because tutorials are conducted through literacy enhancing activities, they are expected to improve students' written, oral, and discursive literacy skills. To meet these objectives, after the setting of a question or problem, the essential student activities in tutorials comprise a period of research involving the reading and review of literature, which leads to the production of a written essay, report of science experiment, or artwork, followed by a critical discussion of the relevant work with a tutor.

While tutorials are a distinctive form of instruction, they are assessed by standards used in whole classroom instruction: quality of paper or other work, examinations, and attendance. Our purpose is not to discard these forms of assessment but to develop an assessment that measures tutorial objectives. Tutorials offer opportunities to assess the development of students' habits of mind, such as thinking independently. To assess students' independent thinking, we need to collect evidence of student work in the research, essay, presentation, and discussion phases of tutorials. And we need to develop rubrics for assessing evidence of independent thinking during these learning activities. For example, independent thinking in tutorial discussions might be assessed by the degree to which a student initiates new lines of inquiry, provides original interpretations of texts, and can articulate both sides of arguments. A similar process of collecting evidence and developing rubrics for assessing evidence would occur for learning outcomes such as skeptical orientation, mental flexibility, self-assessment, and so forth. To do so requires agreement on a *shared model*, not only on what constitutes skeptical orientation (learning outcome), for example, but what kinds of evidence (student work) would demonstrate this attribute, and how we would measure the level of excellence demonstrated (rubric). Representatives of the participating colleges will deliberate on these tutorial assessment issues and arrive at a Shared Assessment Model (SAM) for testing during the Tutorial Assessment Studies phases in the winter/spring of 2009 and 2010.

Tutorials at Participating Colleges

The participating colleges in this study are all well known for offering tutorials, but approach instruction in different ways. Williams College structures tutorial courses in which professors conduct five sets of tutorials of two students each using a common syllabus, while Lawrence University employs individual tutorials negotiated by instructors and students on unique reading or research projects, including many in the arts. The College of Wooster emphasizes independent studies for seniors. There is a range of practices of tutorials across divisions. For example, a music student at Lawrence recently composed music for a silent film and conducted a small orchestra in a public screening of the work; science students will typically conduct experiments, write research papers, present at student symposia and engage in regular discussion; humanities students (history, classics, etc.) generally do library research, write essays, present their papers, and discuss their work with tutors.

While tutorials vary in scope, length and number of tutees, they share common objectives, activities, and expected outcomes. The following characteristics distinguish tutorials from conventional classrooms:

- Tutorials are individualized instruction, involving 1-2 students, and significant student involvement in the subject matter.
- Students create the work that comprises the learning materials in tutorials.
- Tutorials involve process-based instruction in which instructor-student, and where relevant, peer-to-peer interaction, are major forms of teaching and learning.
- Tutorials feature continuous formative assessment as well as summative assessment.
- Tutorial outcomes are habits of mind/forms of thinking, not facts/skills alone.

It is obvious that our consortium believes that assessment is the best way to research tutorials, for the articulation of outcomes, evidence and rubrics in the research, writing, presentation and discussion phases of one-on-one interaction gets at the fine points of the dialogic processes that indicate the range of superior to inferior approaches. If an instructor has well-constructed assessment tools, the application to student work will reveal both to the professor and student ways the process has fallen short or triumphed. In the reflective practitioner model to which we subscribe, the results of such assessments feed back into the revision of instruction. Therefore,

we wish to develop assessment tools in order to improve tutorial education, and improvements in tutorial education may require revisions in assessment tools. The two are inextricably linked.

Tutorial Assessment Studies: Research Design

Six instructors (two from each school) will test the Shared Assessment Model (SAM) during tutorials with approximately 6-12 students in winter/spring 2009 (tutor plus 1-2 tutees each) and with 6-12 new students (tutor plus 1-2 tutees each) in tutorials and 18 randomly selected students (each professor chooses 3 students) in other liberal arts courses in winter/spring of 2010. Therefore, in total, 30-42 students will participate in the project. If SAM can be used in both tutorial and non-tutorial liberal arts courses, then it strengthens its potential applicability in assessing four-year liberal arts college outcomes.

Each instructor will prepare cases that contain:

- Portfolios of multiple samples of student work collected at phases of the tutorial, including early and final samples. The samples of student work will include transcripts of discussions and written documents. The transcripts will permit instructors to rate targeted desirable student qualities more easily and can be used as teaching materials in the Tutorial Assessment Guide for application to other institutions.
- Formative assessments of student work using SAM ratings on all portfolio materials.
- Summative assessment (Final Grades).
- Annotated explanations of formative and summative ratings and how final grades were decided.
- Commentary on how SAM interfaced with traditional instructor assessment methods.

Phases of the Proposed Project

The Working Group in Tutorial Education and Assessment (WGTEA) will engage in a six-phase study with the overall goal of developing and testing a shared model for the assessment of tutorial education in each school. These phases will occur from September 1, 2008 through August 31, 2010. The WGTEA will build upon academic relationships and a corpus of speeches and research reports on the tutorial and individualized learning in colleges, collected as part of a Conference on Tutorial Education held at Lawrence University in spring, 2007. During the

proposed project, faculty and administrative representatives of the participating colleges will collaborate by: 1) web-publishing the *Proceedings of the First Conference on Tutorial Education* as a platform for recommencing consortia work; 2) participating in a *Shared Assessment Model Workshop* in fall 2008 to develop a shared model (SAM) for assessing tutorials; 3) conducting *Tutorial Assessment Studies* during winter/spring of the 2008-2009 academic year; 4) participating in the *Conference on Tutorial Education and Assessment* in fall 2009 by presenting evidence and examples of formative and summative assessment of tutorials emerging from preliminary studies and discussing the application of SAM to other liberal arts courses ; 5) conducting a *second round of Tutorial Assessment Studies* during winter/spring of the 2009-2010 academic year; and 6) writing a *Final Report* during summer 2010 summarizing and discussing the approaches, issues, and outcomes used by the WGTEA to assess education in tutorials and liberal arts courses, and present a proposal for extending SAM to four-year outcomes in liberal arts institutions.

Phase 1: Disseminate Proceedings on First Conference on Tutorial Education

During Phase 1 (September/October, 2008) of the project, we will disseminate a web-based *Proceedings of the First Conference on Tutorial Education*. To consolidate information on tutorial practices that was shared by 20 principal participants during the First Conference, we will organize the papers (14) and discussions (2 hours) on a dedicated web site. Most of the papers and all of the discussion are now in electronic form. We will recruit editors to prepare the papers in two divisions (sciences and humanities) and to redact the discussions. The web site will be built at Lawrence University. Once the papers and discussions are posted we will invite commentary by the participants. While the *First Conference on Tutorial Education* contained some reports of research, it was dominated by historical, administrative and economic analyses. Only anecdotal accounts of tutorials in science and the humanities were presented and there was a singular lack of material on assessment.

Phase 2: Shared Assessment Model Workshop

During Phase 2 (October/November 2008) of the project, Lawrence University will hold a *Shared Assessment Model Workshop* for the purpose of discussing and developing a shared model for conducting formative and summative assessments of tutorial education. The tutorial model suggested by Beck is undoubtedly just one way of framing the process associated with the outcomes of tutorials. Other participating schools may have alternative frameworks. The purpose

of the workshop is to provide a forum where representatives from these schools can discuss different formative and summative assessment models of tutorial education. Moreover, we will invite Barbara Walvoord from Notre Dame, a prominent expert in the field of assessment, to participate in the workshop. Synergy is created by having an expert in the field of assessment participating in a workshop where WGTEA members are discussing how to develop models for assessing tutorials at their particular school. Therefore, these collaborations are expected to generate innovative and productive discussions. At the end of the workshop, the WGTEA members will decide on a Shared Assessment Model (SAM) for assessing tutorials at their colleges.

Phase 3: Tutorial Assessment Studies

During Phase 3 (January-May 2009), six instructors will conduct *Tutorial Assessment Studies* using the SAM. During this phase tutorial professors will a) use the rubrics and procedures developed in the SAM workshop to collect quantitative and qualitative data on tutorial learning outcomes, b) conduct the analysis, and c) prepare a case study on the results of the assessment. Instructors will test the power of the SAM to measure change during tutorials by using the SAM in assessing student performance at baseline (start of tutorial) and as an outcome (end of tutorial). We will also hold online conferences using the NITLE system, a multipoint, desktop videoconferencing system. These videoconferences will provide the principal investigators with the opportunity to monitor and provide feedback to the instructors on the progress of the assessment studies.

Phase 4: Conference on Tutorial Education and Assessment

During Phase 4 of the project, Lawrence University will hold a *Conference on Tutorial Education and Assessment* in which all WGTEA members will present papers on their assessment projects (October/November 2009). The conference will be open to other liberal art schools (e.g., Associated Colleges of the Midwest). We envision the conference as an opportunity for participants to evaluate and critique the SAM and the student work collected and analyzed to assess the various student outcomes in Phase 3. We will invite Barbara Walvoord back to be the keynote speaker and discussion facilitator. It is expected that through the presentations and discussions, WGTEA members will arrive at some conclusions about best practices that could be used on their campuses for the second round of tutorial assessment that will occur in Phase 5, and ways in which tutorial assessment can be applied to measuring four-

year liberal arts goals in other courses. We expect that a modified shared assessment method, SAM-2, will emerge from Phase 4.

Phase 5: Second Round of Tutorial Assessment Studies

During Phase 5 of the project, a second round of *Tutorial Assessments Studies* will be conducted (January-May 2010) using SAM-2. This phase represents an important iteration of the assessment process for determining whether the learning outcomes of tutorials are being realized and are capable of measurement. The six instructors will follow the same procedures outlined in Phase 3. Video-conferencing sessions will also be held using the NITLE system. Because SAM-1 represents the first attempt to develop a tutorial assessment instrument that will undoubtedly be revised because of discussions during the phase 4 conference, it would have been premature to examine its effectiveness in non-tutorial courses during phase 2. But now that SAM-2 has been developed, we believe phase 5 is the time to use the instrument to assess non-tutorial courses. Therefore, the instructors will assess the effectiveness of the SAM-2 to other courses they will teach during the tutorial term or semester. Instructors will be asked to randomly select three students in their other classes to test whether the learning outcomes associated with tutorials are being accomplished in non-tutorial classes. For example, a tutorial professor who also teaches a literature course or an introductory science course would use SAM-2 to assess to what degree three students in those courses develop, for example, a skeptical orientation or acquire mental flexibility. There are two reasons why we have chosen three students. First, the total sample size of this comparison group equals 18 students (6 professors x 3 students each), which match up well with the total sample size of the tutorial group ($n = 12-24$). And second, given the modest stipends for professors, we could not ask them to use SAM-2 for more than three students in non-tutorial courses. The results of this assessment activity will be combined with data from the tutorials when the instructors prepare their cases.

Phase 6: Final Report

During Phase 6 of the project (June-August 2010) we will develop a *Final Report* summarizing the results of Phases 1-5, including the tutorials assessment activities conducted during the 2008-09 and 2009-2010 academic years, and the presentations and discussions of the *Conference on Tutorial Education and Assessment* (October/November 2009). This paper could be used as a blueprint for other four-year liberal arts schools that are interested in assessing tutorials and other forms of individualized learning by detailing the SAM-2 and expertise developed by the

WGTEA institutions in measuring learning outcomes. The paper also could serve as a resource for developing a new proposal for a multi-college study, whose objective would be to assess liberal arts learning goals by employing, perhaps, a further modified version of the shared assessment method, SAM-3.

References

Association of American Colleges and Universities (2008). *New Leadership for Student Learning and Accountability*.

Beck, R. J. (2007). The Oxford Tutorial System: A learning theory, *Oxford Magazine*, Eighth Week, Trinity Term.

Bloom, B. S. (1984). The 2 Sigma problem: The search for methods of group instruction as effective as one-on-one tutoring. *Educational Researcher*, 13, 4-16.

Cohen, P. A., Kulik, J. A., & Kulik, C. C. (1982). Educational outcomes of tutoring: A meta-analysis of findings. *American Educational Research Journal*, 19, 237-248.

Hacker, D.J. & Graesser, A. C. (2007). The role of dialogue in reciprocal teaching and naturalistic tutoring. In R. Horowitz (Ed.), *Talking Texts: How speech and writing interact in school learning* (pp. 253-270). Mahwah, NJ: Erlbaum.