Spring 2010

Volume 5, Issue 2

Illinois State University

Progressive Measures

Mission Statement:

"The University Assessment Office is responsible for conducting a variety of assessment activities related to student learning outcomes using qualitative and quantitative research techniques, providing support services to other units engaged in such assessment, and sharing best practices for and results of assessment activities."

Inside this issue:

From the Director

At the 2010 Teaching and Learning Symposium in January, the University Assessment Office team provided, "UAO and U," a panel presentation detailing the services we offer to the university community. I think the title aptly reflects the relationship the UAO is promoting between campus units, including ours, and the work of assessment itself.

Since the last newsletter, we have continued to assist departments/schools with survey development and preparation for program review. We have also recently introduced new tools to help campus units initiate assessment activities, including the Assessment Plan Tutorial and the Online Survey Guide using SelectSurvey software. These tools and services are aimed at encouraging members of the university community to recognize the value of assessment, facilitate its processes, and integrate it into our usual educational practices.

Ideally, assessment should facilitate thoughtful educational practices by informing curricular planning and allowing us to identify the "active ingredients" in our educational efforts. As assessment becomes more familiar and integrated into our systems, it will more readily enhance programs. Our goal is to help campus units develop plans to move in that direction because it is not just "UAO and U," but also assessment and you.

What excites me about this issue of *Progressive Measures* is the participation of contributors from different areas of campus who are pursuing assessment. Check out what students think about using technology in the classroom in one article. Read what two associate deans who are leading the assessment efforts in their respective colleges have to say about their vision of assessment in a Q&A with Amee Adkins and Todd McLoda.

In other articles, learn about the importance of tying program goals to outcomes in "Embedded Assessment – Closing the Loop," and learn how to determine when a research study needs IRB approval from the experts in the Office of Research Ethics and Compliance. Also learn some of the outcomes of the General Education assessment of two shared learning outcomes: Public Opportunity and Critical Inquiry and Problem Solving.

Along with the sense of renewal that comes with Spring, I hope this publication reinvigorates and informs readers about the different ways assessment can influence students and programs. Remember, the UAO is here to facilitate quality assessment. Please contact us at 438-2135 or at uao@ilstu.edu with your questions. Happy Spring!

Renée M. Hoi

Acting Director - University Assessment Office

& Associate Professor of Psychology

Assessing General Education at ISU: Public Opportunity and Critical Inquiry and Problem Solving

Alycia Hund, Ph.D., Associate Professor, Psychology

on behalf of the Council for General Education and in collaboration with the UAO staff

General Education at ISU provides students with a broad, common foundation of study upon which to build an undergraduate education. ISU students complete 14 courses (42 credit hours) as part of the program, including courses in the inner, middle, and outer cores. The inner core focuses on foundational skills and includes five courses: Composition as Critical Inquiry, Communication as Critical Inquiry, one mathematics course, and two natural science courses. The middle core focuses on interdisciplinary perspectives and includes five courses, one from each category: Quantitative Reasoning, Language in the Humanities, United States Traditions, Individuals and Civic Life, and Individuals and Society. The outer core focuses on varieties of disciplinary knowledge and includes four courses, one from each discipline group: Science, Mathematics, and Technology; Social Sciences; Fine Arts; and Humanities. In total, the General Education Program develops students' capacity to (1) critically think and solve problems, (2) comprehend and contribute to diverse and global perspectives, (3) be stewards of life-long learning, and (4) advance public opportunity. These represent the four shared learning outcomes of General Education.

The Institutional Artifact Portfolio (IAP) process provides a comprehensive method to evaluate our progress in accomplishing the four shared learning outcomes of General Education. General Education assessment also supports teaching and learning in the disciplines by providing an accurate representation of student abilities. Lastly, General Education assessment realizes many of the requirements of systematic review of educational outcomes required by external constituencies such as state government and accrediting agencies. Each shared learning outcome is assessed one semester every two years as part of this ongoing assessment process.

As part of the first wave of IAP reviews, the University Assessment Office invited instructors of General Education courses related to public opportunity (Fall 2008) and critical inquiry and problem solving (Spring 2009) shared learning outcomes to participate in the review. An artifact is any form of tangible student work. It is a product of the students' learning experience that addresses the identified traits of the four shared learning outcomes. Possible examples of artifacts include essays/papers, written assignments, exams, speeches, presentations, posters, artwork, performances, or musical recitals.

For each shared learning outcome, 300 artifacts (selected randomly from submitted artifacts for each course) were assessed using rubrics developed by the General Education Assessment Task Force. Three interdisciplinary review teams (each consisting of two faculty members) carried out blind reviews in which members were asked to come to a consensus regarding the extent to which each primary trait (and self-reflection and disciplinary knowledge) was *developing*, *established*, or *advanced* using the established rubrics. Reviewers also had the option to note that primary traits were *not present* in a given artifact. Not present ratings should be interpreted broadly because it is possible that the assignment for which the artifact was created did not incorporate particular aspects included in the rubric. It is also possible that students did not show evidence of an aspect when requested by the assignment. Inter-rater reliability was acceptable, given the applied nature of this project.

Continued on page 3...

Assessing General Education at ISU (cont'd)

Table 1

Artifact rating distribution for Public Opportunity.

	Not Present		Developing		Established		Advanced		
Primary Trait	#	%	#	%	#	%	#	%	Total
Critically informed position	86	28.7	24	8.0	25	8.3	165	55.0	300
Knowledge	146	48.7	55	18.3	61	20.3	38	12.7	300
Individual life	106	35.3	80	26.7	80	26.7	34	11.3	300
Social & community life	72	24.0	97	32.3	84	28.0	47	15.7	300
Resources	177	59.0	47	15.7	59	19.7	17	5.7	300
Civic participation	141	47.0	79	26.3	31	10.3	49	16.3	300
Composite Measure	728	40.4	382	21.2	340	18.9	350	19.4	1800
Self-reflection	181	60.3	50	16.7	55	18.3	14	4.7	300
Disciplinary knowledge	62	20.7	126	42.0	86	28.7	26	8.7	300

Public Opportunity

In the IAP process, public opportunity is described with the following statement, "Students will identify the resources and articulate the subsequent value of civic and community engagement." It includes six primary traits: (1) critically informed position on civic life, (2) influence of civic participation on the social and collaborative nature of knowledge, (3) contributions to the public affecting individual life aspects such as family, religion, business, and/or the state, (4) contributions to the public life affecting social and community life such as family, religion, business, and/or the state, (5) resources for civic engagement, and (6) civic participation in the social, economic, technological, and/or political dimensions of community development. In Fall 2008, 196 individual instructors were invited to submit artifacts from 279 classes offering 20,570 seats in courses in 16 departments/schools. Forty-three instructors volunteered artifacts from 47 courses (16 inner core, 19 middle core, 12 outer core; 7 in the College of Applied Science and Technology, 35 in the College of Arts and Sciences, 1 in the College of Fine Arts, 4 in Interdisciplinary Studies) related to public opportunity, yielding a 24% instructor response rate and representing 30% of total enrollments in public opportunity courses. Among participating instructors, 14.0% were graduate students, 6.9% were Administrative Professional staff with teaching responsibilities, 23.3% were Non-tenure Track Faculty, 23.3% were Assistant Professors, 18.6% were Associate Professors, and 13.9% were Professors.

In an effort to simplify the presentation of data, an overall public opportunity composite score based on these six dimensions was created, evincing adequate internal consistency (Cronbach's $\alpha = .76$). Finally, the assessment of all artifacts includes details about self-reflection and discipline knowledge. General trends for each public opportunity primary trait, the composite measure, self-reflection, and discipline knowledge can be found in Table 1. Initial examination of overall patterns for native and transfer students revealed broad similarities across student samples, so this factor is not considered in the sections that follow. Moreover, examination of patterns based on student designation (freshman, sophomore, junior, or senior) was similar to the overall trends described below, so this factor is not considered further. In general, ratings of critically informed position on civic life, contributions to the public life affecting social and community life, contributions to the public affecting individual life, and disciplinary knowledge were very positive, with civic participation in the social, economic, technological, and/or political dimensions of community development, influence of civic participation on the social and collaborative nature of knowledge, resources for civic engage-

Continued on page 4...

Assessing General Education at ISU (cont'd)

Table 2

Artifact rating distribution across cores for Public Opportunity and Critical Inquiry and Problem Solving Composite Measures.

Core	% Not Present	% Developing	% Established	% Advanced	Number of Reviews
		Public Opportunity			300
Inner	14	40	38	8	100
Middle	19	48	28	5	100
Outer	29	38	20	13	100
		Critical Inquiry & Problem Solving			300
Inner	73	9	6	12	100
Middle	52	18	15	15	100
Outer	72	9	10	9	100

ment, and self-reflection yielding somewhat less positive ratings. Table 2 includes the percentages of artifacts rated as developing, established, advanced, and not present for the three cores.

Critical Inquiry and Problem Solving

In the IAP process, critical inquiry and problem solving is described with the following statement, "Students will develop and communicate a range of interests and curiosities, engaging those interests and curiosities through critical thinking, reasoning, and problem solving." It includes nine primary traits: (1) variety of ideas evaluated, (2) quantitative reasoning used to address problem, (3) critical analysis expressed through writing or speaking, (4) the context of other viewpoints in developing arguments, (5) consideration of potential moral and ethical issues, (6) theories to resolve moral issues, (7) forces and consequences that influence life, (8) development and use of technology as it relates to society/environment, and (9) uses information from outside resources responsibly. In Spring 2009, 286 individual instructors were invited to submit artifacts from 417 classes offering 17,509 seats in courses in 27 departments/schools. Thirty-seven instructors volunteered artifacts from 41 courses (17 inner core, 15 middle core, 9 outer core; 3 in the College of Applied Science and Technology, 28 in the College of Arts and Sciences, 4 in the College of Business, 1 in the College of Fine Arts, 1 in Interdisciplinary Studies) related to critical inquiry and problem solving, yielding a 13% instructor response rate and representing 24% of total enrollments in critical inquiry and problem solving courses. Among participating instructors, 5.4% were graduate students, 13.4% were Administrative Professional staff with teaching responsibilities, 27.0% were Non-tenure Track Faculty, 5.4% were Assistant Professors, 18.9% were Associate Professors, and 30.0% were Professors.

In an effort to simplify the presentation of data, an overall critical inquiry and problem solving composite score based on these nine dimensions was created, evincing adequate internal consistency (Cronbach's $\alpha = .75$). Finally, the assessment of all artifacts includes details about self-reflection and discipline knowledge. General trends for each critical inquiry and problem solving primary trait, the composite measure, self-reflection, and discipline knowledge can be found in Table 3. In general, ratings of the variety of ideas evaluated, forces and consequences that influence life, and disciplinary knowledge were somewhat positive, with quantitative reasoning used to address problems, critical analysis expressed through writing or speaking, the context of other viewpoints in developing arguments, consideration of potential moral and ethical issues, theories to resolve moral issues, development and use of technology as it relates to society/environment, uses of information from outside resources responsibly, and self-reflection yielding less positive ratings. Table 2 includes

Continued on page 5...

Assessing General Education at ISU (cont'd)

Table 3

Artifact rating distribution for Critical Inquiry and Problem Solving.

	Not Present		Developing		Established		Advanced		
Primary Trait	#	%	#	%	#	%	#	%	Total
Variety of ideas	133	44.3	86	28.7	43	14.3	38	12.7	300
Quantitative reasoning	200	66.7	15	5.0	14	4.7	71	23.7	300
Critical analysis	176	58.7	30	10.0	39	13.0	55	18.3	300
Context of other viewpoints	199	66.3	40	13.3	31	10.3	30	10.0	300
Moral & ethical issues	211	70.3	46	15.3	27	9.0	16	5.3	300
Theories	264	88.0	18	6.0	7	2.3	11	3.7	300
Forces & consequences	128	42.7	23	7.7	93	31.0	56	18.7	300
Technology	269	89.7	9	3.0	13	4.3	9	3.0	300
Resources	192	64.0	54	18.0	16	5.3	38	12.7	300
Composite Measure	1772	65.6	321	11.9	283	10.5	324	12.0	2700
Self-reflection	246	82.0	24	8.0	15	5.0	15	5.0	300
Disciplinary knowledge	94	31.3	118	39.3	54	18.0	34	11.3	300

the percentages of artifacts rated as developing, established, advanced, and not present for the three cores. Overall, ratings of middle core courses were consistently more positive than were ratings for inner and outer core courses, with the exception that outer core courses received quite favorable ratings with regard to theories to resolve moral issues and forces and consequences that influence life.

Conclusions

The overall pattern of results for public opportunity evinces very strong student learning outcomes related to public opportunity and disciplinary knowledge. Thus, we identify public opportunity as an area of strength in our General Education program and on campus in general. Although the program-level focus of the IAP does not provide details about the relation between particular sets of courses or campus initiatives and learning outcomes, we speculate that the strong outcomes related to public opportunity indeed relate to the campus-wide focus on civic and community engagement.

The overall pattern of results for critical inquiry and problem solving evinces somewhat limited evidence of student learning outcomes related to critical inquiry and problem solving. It is possible that this limited evidence is, in part, a result of the divergent courses that address critical inquiry and problem solving from unique perspectives. That is, only a subset of courses focuses primarily on quantitative reasoning, another subset on technology, another subset on argumentation, and so on. Given that the IAP review focuses on the General Education program as a whole and does not link specific assessments of learning outcomes to sets of courses, we are unable to determine the locus of these results. Additional complementary assessment of student and instructor responses linked to the 12 General Education goals for particular course categories within each core would provide helpful details regarding possible strengths and limitations of General Education. Overall, the IAP review suggests some interesting trends that merit further study in a more focused manner. The design of the current assessment did not permit the CGE to make more than general observations of possible patterns.

Please visit the General Education website for additional information: http://gened.illinoisstate.edu/.

Program Evaluation, Quality Assurance, and Research: Shedding Light on the Gray Areas

Gary Creasey, Ph.D., Chair, ISU Institutional Review Board

S. Kathleen Spence, J.D., Assistant Director, ISU Research, Ethics, Compliance Office

The term "Institutional Review Board" (IRB) does not evoke the same images as "Caribbean vacation," "French truffles," or "University is awarded $24 million National Science Foundation grant." It would not be surprising to many readers that these regulatory oversight committees are often viewed as mechanisms to obstruct valuable scientific inquiry, are prone to nebulous decision-making processes, and are eager to step into domains that are not viewed within their purview (Fost & Levine, 2007; Lederman, 2006). Such perceived IRB encroachment has led to scathing commentaries on national IRB operations, such as the *Illinois White Paper* (Gunsalus et al., 2007).

Indeed, major debates have erupted locally and nationally as to what types of activities an IRB should regulate. These domains range from oral history work to student class research projects to graduate student-training mechanisms, such as clinical practicum. One context that has received considerable attention recently concerns quality assurance or program evaluation activities, perhaps due to the considerable national media attention (Gawande, 2007) on work reported in the *New England Journal of Medicine* (Pronovost et al., 2007) concerning the use of a routine checklist completed by nurses (e.g., *Does physician wash hands?*) that was used within hospital settings to reduce catheter-related infections. The work significantly reduced infections and probably saved a number of lives; yet, the primary university responsible for the work ran afoul with the federal Office for Human Research Protections (the Public Health Agency that oversees and provides guidance to all IRB operations) because they did not pursue IRB review for their work. The difficulty was that the investigators viewed the work as "program intervention and evaluation" and the OHRP defined the work as "human subjects research." The work was subsequently halted by OHRP, but has since resumed after agreements were worked out between the federal agency and the participating research sites.

Such attention is not welcomed by OHRP, and certainly not by institutions in which individual research projects are halted and institutional-wide federal research dollars held up as OHRP investigates the inquiry in question as well as the local IRB operations that were supposed to have regulated the work in the first place. Such actions probably also explain (at least partially) somewhat overzealous IRB operations at some institutions as these bodies attempt to protect the human participants, the institution and maybe even themselves in the process. It also leads to the important question that is routinely asked in hospitals, corporations and academic institutions — "Is it not our right to assess the quality of our programs, services and operations to ultimately benefit our patients, customers or students without the need for IRB oversight?"

Our position is that this question boils down to whether the work reaches the threshold of human subjects research as defined by the policies and procedures outlined in the Code of Federal Regulations, Part 46: *Protection of Human Subjects* (also known as the Common Rule). To start, the initial question that should be asked is whether the work involves *human subjects*, or any living individual *about whom* an investigator obtains data through interaction or intervention with the individual or obtains identifiable private information (Protection of Human Subjects, 2009). If one interviews someone or gives them a survey then they have interacted with them. Or, if one is analyzing data sets with identifiers

Continued on page 7...

Program Evaluation, Quality Assurance, and Research (cont'd)

then they have identified someone. Put simply, almost all of the work that constitutes program evaluation or quality assurance work involves human subjects, unless, the gathered data is *about* the institution and not the individual. For example, asking students to identify the types of computer support on campus does not make them a human subject because the line of questioning is about the institution. Such a simple line of questioning though is quite rare, because it is only natural to next ask, "What's your degree of satisfaction about this support?" In a single stroke, the respondent is now transformed into a human subject according to the federal guidelines.

Work that involves human subjects though does not always reach the threshold of *research*, defined as a systematic investigation, including research development, testing and evaluation, designed to contribute to *generalizable knowledge* (Protection of Human Subjects, 2009). This language contains the key criteria to determine if program evaluation or quality assurance work reaches the threshold for regulation in the eyes of the IRB. Basically, the work constitutes research if the parties who conduct the work intend to present or publish the work, intend to replicate the results in other settings and/or intend to generalize the results to a larger population. The data dissemination plan is a key ingredient in this picture because most of us who present at conferences or plan to publish our results usually do so to convince others our findings are valid, replicable, and should inform the decisions of others. Put simply, many times the very reason program evaluation or quality assurance results are presented at conventions or submitted for publication is to demonstrate to other parties that the institution has good programs or services that could be used by other like-minded parties.

Thus, the work is not human subjects research if the primary purpose of the program or service evaluation is for internal use to guide improvement of existing programs or services. This is not to say that institutions can "do whatever they want" in such cases to collect data — the institution may have policies and procedures for such work in itself or at least pursue basic ethical steps that human subjects researchers follow anyway. For example, when assessing a general education program or conducting a campus safety survey, the institution may not require that people participate as a result of their standing (e.g., student or employee) and/or follow the normal procedures in the informed consent process (that again reinforces the voluntary nature of participation and basic rights for human research participation).

What are some best practices to take away from this article? If the data collection is just to improve basic internal services and is for internal use only, then an IRB protocol is probably not necessary. However, in all other cases, we recommend that the very best practice is to consult with the IRB about program evaluation or quality assurance work before any data are collected. It is our belief that there are many cases, particularly in academic institutions, that the data collected for program evaluation or quality assurance work *is* intended for both internal use and eventual dissemination to inform a wider audience. The latter communicates the idea that we are doing a good job, are using data to inform our decision-making processes, and that officials at other institutions should think about implementing what we are doing at our institution.

To sum, it is hoped that this article has somewhat clarified the distinctions between program evaluation, quality assurance work, and human subjects research. When in doubt, it's probably always a best practice to submit an IRB protocol (even if we later determine that the work is not human subjects research). There are many advantages in going through a prospective IRB process, even for work that does not seem at first glance to constitute human subjects research. How one views the intention to present or publish their work can change over time, and it is very difficult (and sometimes impossible) for the IRB to grant retrospective review for the use of data that have already been collected.

References

Fost, N. (2007). The dysregulation of human subjects research. *Journal of the American Medical Society, 298,* 2196-2198.

Continued on page 8...

Program Evaluation, Quality Assurance, and Research (cont'd)

Gawande, A. (2007, December 30). A lifesaving checklist. *The New York Times*. Retrieved from http://www.nytimes.com

Gunsalus, C., Bruner, E., Burbules, N., Dash, L., Finkin, M., Goldberg, J., Greenbough, W., & Aronson, D. (2007). The Illinois White Paper. *Qualitative Inquiry, 13*, 617-649.

Lederman, R. (2006). The perils of working at home: IRB "mission creep" as context and content for ethnography of disciplinary knowledges. *American Ethnologist, 33*, 482-491.

Protection of Human Subjects (2009). *45 C.F.R.*, pt. 46.

Provenost, P., Needham, D., Berenholtz, S., Sinopoli, D., Haitao, C., Cosgrove, S., & Goeschel, C. (2006). An intervention to decrease catheter-related bloodstream infections. *New England Journal of Medicine, 355*, 2725-2732.

Embedded Assessment — Closing the Loop

Linda L. Miles, Ph.D., Assistant Professor, Finance, Insurance & Law

The goal of an assessment program, as commonly understood, is to measure outcome-based learning. A review of the process is discussed by Chang (2009, Fall). Implicit in the assessment process is the understanding that measuring outcomes will lead to an identification of areas for improvement. Much of the literature on assessment focuses on developing measurable outcomes and gathering valid data. But we must remember that obtaining and analyzing data is perhaps the easier part of the process. Considerably more difficult is using the feedback provided through the data analysis to improve learning outcomes.

A key tenet of learning assessment is that the outcomes are measured based not on mastery of material from a single course, but rather on individuals' having acquired important capabilities while completing an entire program of study. Addressing assessment in this comprehensive manner surely is a noble and valid concept. However, we must acknowledge that key foundational components taught in required introductory courses often are the building blocks on which material in subsequent courses expands: Students are expected to master basic concepts and then apply that knowledge while progressing through the program's sequence of requirements. Therefore, while program assessment certainly measures overall outcomes from numerous courses, those of us who teach required foundational courses are on the front line when areas for improvement are identified from outcome assessments. We should strive to address any identified problems in an effective and timely manner.

Trefzger and Newgren (2006, Fall) point out that too frequently we tend to equate student grades with evidence of mastery of material; someone who scores an "A" on an exam (or for a semester) usually is deemed to have learned well, and to know more of the most relevant material than peers with "B" or "C" results. And if a reasonable number of students are earning "A" grades, then the desirable result of meaningful overall learning is deemed to be taking place. Yet these relationships need not always hold true. Fortunately, the grade-generating data embedded in the routine course examination process also provides the basis for an instructor to identify gaps in student understanding — thus the name *embedded* assessment.

Trefzger and Newgren provide a scenario in which we cannot conclude that a class overall has mastered one of the most important concepts being tested on, despite the fact that numerous students have earned scores in the "A" range. The authors offer a hypothetical comprehensive final exam containing five essay questions or computational problems, each worth 20 points, which they decompose by question and student to unlock information about mastery of specific concepts. An "A" turns out to be computationally possible for a student scoring high on four questions and failing the fifth. Hence, many students could possibly earn the highest letter grade, while they and others enrolled generally show a serious lack of learning in an area that the instructor has identified as highly important.

The described approach is valuable in a small class

Continued on page 9...

Embedded Assessment (cont'd)

setting, as illustrated by Trefzger and Newgren's 26-member class example, but becomes more difficult in the form they present when class size grows. Because many required introductory courses — the foundation on which program assessment rests — are taught in large sections, and by necessity exams tend to become based on numerous multiple-choice questions rather than a few comprehensive questions, other approaches are needed. Two methods of embedded assessment for measuring the mastery of course content in large foundational class settings are discussed below. These strategies might be used very effectively during the semester, in addition to collecting data solely at the final exam, letting the instructor address areas of weakness in student understanding while the term is still under way. The feedback-to-proactive-change cycle is dramatically shortened, allowing for more timely strengthening of foundation material on which students build mastery of concepts across the entire program.

Fill In the Dots ... And Then Connect Them

Multiple-choice exams in large sections typically are scored electronically, using Opscan exam sheets. The Opscan Evaluation unit in the Center for Teaching, Learning and Technology (CTLT) can provide a wealth of information in the data analyses it offers. The data an instructor can request identifies the answers selected by each student in varying degrees of detail. The short-form analysis includes data calculated for each question, including the percentage of students selecting each answer. If a large number of students choose a single wrong answer for a question, the source of the confusion might be easy to identify. The output also shows the percentage of the class answering each question correctly (of most value to those with a single-test version; more manipulation is required with multiple versions). The output also includes the discrimination index for each question, allowing the instructor to identify the questions missed by more high-scoring students than low-scoring students. While not always helpful in assessing student learning outcomes, this information aids the instructor in identifying test questions that should be evaluated for elimination from future exams. (Of course a question that relatively few answer correctly still may be worth keeping; it may be what tends to separate the "A" from the "B" students.) An instructor who asks to have the information electronically transmitted can easily construct a spreadsheet, as proposed by Trefzger and Newgren, to analyze the data and assess the mastery of concepts.

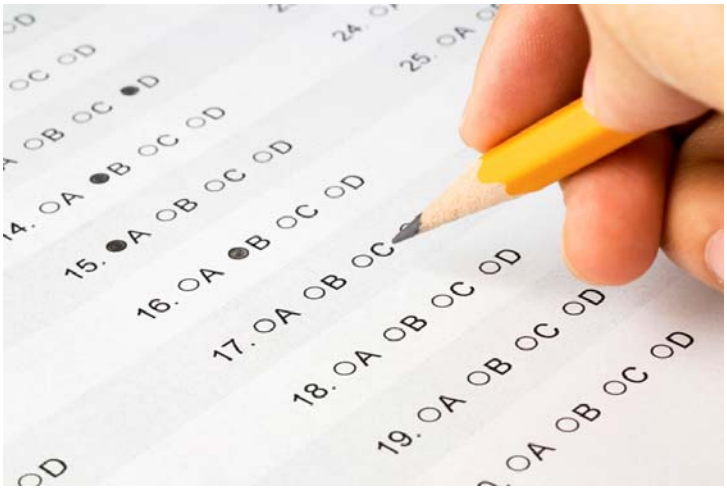
Trefzger and Newgren's example relates to an MBA finance course, in which the instructor assesses learning by examining the proportion of students who score acceptably on detailed, open-ended final exam questions that cover concepts deemed especially important to the student's skill set. This same logical approach could be used with multiple-choice testing in a large undergraduate foundation course, but with a group of questions, rather than an in-depth essay, targeting the concept of interest. Think of an exam in which the instructor can divide 50 multiple-choice questions (two points each) into groups based on the broad concepts addressed. One particular benefit of this type of analysis is that each broad concept can be further divided into smaller component parts.

The examination is constructed to assess knowledge of two broad concepts in a required undergraduate finance course: Time value of money (such as the buildup in a retirement savings plan) and bonds (large loans made to companies or governmental units), with 25 questions devoted to each concept. Some questions address general topic knowledge, while others hone in on more specific applications. The first 25 questions cover time value of money, both from a theoretical standpoint and through practical application involving calculations. In a similar fashion, questions 26 through 50 address bond theories and calculations. Assume that the examination has been constructed with 10 questions testing theory and general comprehension of

Continued on page 10...

Embedded Assessment (cont'd)

each concept, and 15 requiring some form of computation. Intentional exam design becomes imperative — there must be a clear understanding of what knowledge a given question is designed to test, which is no small feat. Likewise, if the instructor chooses to scramble the order of questions, it is necessary to create an unscrambling key so that questions on related broad concepts and related specific applications can be matched for analysis.



In addition to requesting the short-form report from the Opscan process, the instructor wants the raw data. Using the unscrambler key created when the exam was written, she creates a spreadsheet aligning the questions across the various versions of the exam. In this example, there are four versions of an early midterm exam taken by 130 students. Question 1 in Version A corresponds to question 27 in Version B, question 21 in Version C, and question 15 in Version D. Aligning the individual response data in a column allows for quickly labeling the correct responses. Following Trefzger and Newgren's format this data, embedded in the grading process, is used to assess the mastery of the tested concept. The percentage scored on the entire examination by a given individual ("row" data) no longer is the relevant statistic. Rather, what is interesting is the information revealed by the "column" analysis, which reveals the percentage of students who mastered the tested concept. In conjunction with this question/concept analysis, it is important to review the data received in the Opscan short report regarding test question validity.

With data gathered and analyzed, the instructor can "close the loop" by reviewing the graded exams with the class. An obvious benefit of assessing with midterm tests is that going over exam results allows the instructor to explain further the material relating to questions that many students answered incorrectly. Designing a course to include some midterm testing and assessment, plus a comprehensive final exam combines timely feedback with the incentive for students to retain knowledge of concepts, and motivation to understand concepts they did not successfully master for earlier exams — at least until the end of the course.

Then a similar analysis performed at the semester's conclusion, based largely on the final exam results, provides comprehensive information regarding mastery of the relevant course concepts. This analysis can provide important insights for departmental review, particularly when combined with a similar analysis conducted at the conclusion of higher level courses. Thus mastery, and utilization, of concepts important to the entire program of study can be tracked on an ongoing basis as the program progresses.

Has the Lesson Clicked?

A second method of embedded assessment arose with the adoption of Turning Technology's audience response devices (clickers) by Illinois State University. This system provides an opportunity for instructors to gather data on student understanding, in real time, through ordinary course activities. While we would all love to have students asking questions when concepts are not clear, fewer individuals seem willing to ask questions as class size grows. Many of us who teach large sections employ clicker questions to measure student understanding. When the responses are tallied, we can then proceed to the next concept if most students select the correct answer, or provide another example or further explanation if many students select the wrong answer. Clicker interaction also provides the opportunity to explain why the wrong answers are incorrect — often valuable to do when students have been asked to choose among closely related possibilities. Certainly eyeing clicker data is superior to presuming that a lack of student questions gives clear evidence that they understand, but there are obvious limits. The requisite format is multiple-choice questions, leading to the possibility of correct responses that merely are good guesses. Even when incorrect clicker responses impel the instructor to provide a more ex-

Continued on page 11...

Embedded Assessment (cont'd)

tensive explanation, the only way to further attempt to judge understanding is with another concept-based, multiple-choice question.

A couple of semesters ago a student suggested during my office hours that a better use of clickers would be to ask class members directly to assess their understanding of the concept. To put this good suggestion into practice, I pose clicker questions along these lines (consider an example regarding the time value of money):

1. I understand time value of money, get with the program and move on.
2. I'm pretty comfortable with the concept.
3. I'm not really sure how that works.
4. I think I know what time value of money means, but that's about it.
5. Money's value at time what?

The anonymity of clicker responses (plus a little humor) allows students the opportunity to ask for further explanations implicitly, without directly having to raise their hands and self-identify as "not getting it." There also is positive feedback for the students who already have mastered the concept, when seeing that only a small percentage of their peers have progressed as far as they in understanding.

This type of clicker feedback also provides the instructor with the opportunity to "think on one's feet." In my experience, often the examples created on the spot lead to more interaction—particularly when something does not work out just right; it gives the class the chance to help "trouble-shoot" the example. Additionally, my "on-the-spot" examples rarely contain math that can be calculated in anyone's head. A plea of "someone help me here — please calculate the answer given this problem set-up; I'm beyond my mental math capabilities" seems to engage more students than does a prepared example that is carefully worked through in advance.

Who's In Charge Here?

While ultimately the process of "closing the loop" may reside with the department curriculum and assessment committees, instructors in required large-section courses are in a position to proactively monitor the mastery of the foundational concepts embedded within the framework of typical examinations and class work. With confidence gained through students' demonstrated mastery within the initial course, the department can begin to plan how to maintain that strong foundation while building the knowledge students are expected to acquire across the entire program of study. Perhaps most importantly, effective assessment initiatives in foundation courses allow for enhanced instruction to benefit students while they are still in the program (possibly even while enrolled in a given course), rather than merely making better the academic lives of those fortunate enough to enter the program after improvements based on earlier comprehensive assessment efforts have been implemented.

References

Chang, S. J. (2009, Fall). Assessment 101. *Progressive Measures*, *5*, 2-3.

Trefzger, J., & Newgren, K. (2006, Fall). Course grading is not an assessment tool: But your grade book provides data for embedded assessment. *Progressive Measures*, *2*, 13-15.

Students' Perceptions of Using 3D Computer Technology for Enhancing Learning

Jamie Perry, Ph.D., Assistant Professor, Communication Sciences and Disorders

Danielle Cunningham, Communication Sciences and Disorders

It is hard to imagine that computers became a consumer product only 35 years ago and become readily available and affordable for home offices in the early '80s. Today, we see computers immersed within every part of life. Computer technology is being routinely used in the college classroom through podcasting, classroom management systems, and online courses. Some professors are also using virtual environments and virtual reality to engage students within the classroom. Three-dimensional (3D) environments and learning provide students with the spatial-depth cues which share some similarities to that of hands-on learning. In the area of science, this is advantageous because it allows for exploration in 3D (e.g., internal anatomy, molecular biology, neuroanatomy) that might otherwise only be viewed through two-dimensional (2D) images or drawings.

Undoubtedly, traditional learning through lectures, human atlases, and 2D images has a valuable place in the educational setting. Efficacy research studies have assessed whether or not using 3D technology in the classroom actually improves student learning. Authors propose that students understand better through 3D models or spaces, can make better clinical hypotheses, provide a greater transfer of learning to the clinical environment, and allow students to see the processes to patient care (Hilbelink, 2009; Schleich, Dillenseger, Houyel, Almange, & Anderson, 2009; Khalil, Johnson, & Lamar, 2005; Lewis, 2003; Nicholson, Chalk, Funnell, & Danielle, 2006; Perry, Cunningham, Kuehn, & Gamage, 2009; Petersson, Sinkvist, Wang, & Smedby, 2009). Other studies have shown minimal to no positive effect of using 3D or multi-view images on learning (Garg, Norman, Spero, & Maheshwari, 1999; Jolicoeur & Milliken, 1989).

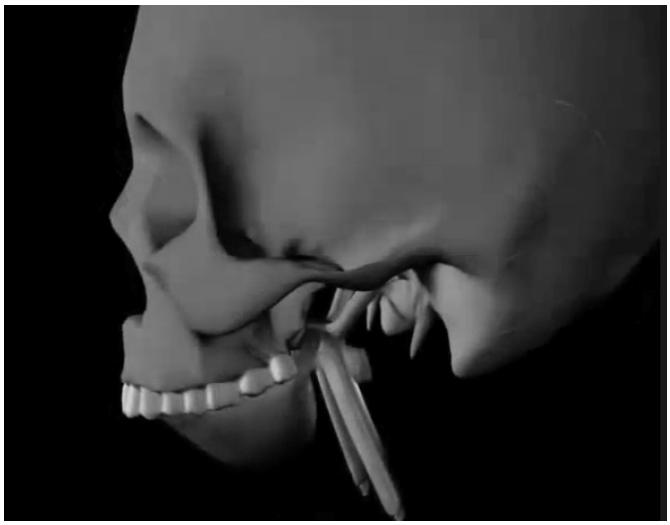
It seems evident that 3D technology for learning has a place in the education setting. The purpose of this study was to assess students' experience and familiarity with computer technology and to obtain their opinions about the use of computer technology in the classroom.

Method

Participants

Forty Caucasian students, 2 males and 38 females, were recruited to participate in the study. Participants were all enrolled in the communication sciences and disorders program at Illinois State University. The male-to-female ratio represents the typical distribution of gender in the academic classes within the department. Participants ranged in age from 19-22 (M = 21; SD = 0.58). Participants had a similar educational background in that they had all finished a minimum of three years of undergraduate study in communication sciences and disorders.

Figure 1



Apparatus

A 3D computer model was created using a visualization and imaging software program (Maya, Autodesk). The animation demonstrated the internal anatomy of the head region with specific emphasis on muscles of a region important for speech (velopharyngeal muscles, see Figure 1). Audio recordings and labels were added to the 3D animation to emphasize the major regions of interest. The animation created for this study is a typical

Continued on page 13...

Students' Perceptions of Using 3D Technology (cont'd)

animation that has been used in our department's voice and resonance course.

Procedure

After obtaining consent, students were seated in a computer lab and instructed to watch the video and complete a questionnaire. Information was collected to obtain information about their use and familiarity with computers and computer technology and to gather information about students' opinions of using computer technology in the classroom to enhance student learning.

Results

All students reported that they used computers for e-mail, word processing, web browsing, virtual environments (e.g., blogs, Facebook, MySpace), and course management on a daily basis. On average, students reported using computers for 4.2 hours per day. On a scale from 1 (extremely uncomfortable) to 5 (extremely comfortable), students reported an average comfort level of 4.3. All subjects reported that they were very comfortable with using the software program in the study and found it to be very user-friendly.

Qualitative information about the animation was obtained through the following questions: Do you wish professors would use computers more often for in and out-of-class learning? If there was one short animation (such as the one provided in the study) each week that was optional, would you watch the animations regularly? Why or why not? Do you think animations like these are helpful additions to basic textbook readings? Why or why not? and What would you add or change to the animation?

Of the 40 participants, 27 said they wished that professors would use computers more often for in and out-of-class learning. Nine students felt that professors do *not* need to use computers more often, and four reported that it would be helpful only for certain classes. When students were asked whether they would view the animation if it were provided as a supplemental material for an optional out-of-class learning experience, the majority (37 out of 40) stated that they would watch the animation. Some students indicated that they would use it most often when they were struggling with the course and if the animations were short. Some reasons students provided regarding their opinions included that animations clarified the topics discussed in class, it helped them get a better grasp of material, and it was easier to understand compared to material in textbooks.

Students were also asked whether they felt that 3D computer technology would be a helpful addition to basic textbook readings. Students responded that the 3D animation was easier to understand compared to textbook readings (16 students), provided multiple perspectives of a complex system (9 students), and was a good supplement to the readings (8 students). When asked if they had suggestions for change, 19 students felt no changes were necessary. Suggestions listed by the remaining students included slow down the animation, provide more animation/movement for interest, make it more interactive, configure it to be mobile for an iPod, more images, make the movie larger, and supplement with note print-outs/worksheets.

Discussion

As expected, students reported that they are comfortable and familiar with using computer technology. They reported regular use of computers and the Internet for both school and social interaction. Many professors use computers in the instruction and dissemination of course related material (e.g., Blackboard). This study was specifically directed at the use of 3D computer animations for providing students with supplemental learning. Sixty-seven percent of students reported that they would like for professors to use more computer technology (3D computer animations, virtual environments, cyberspace, etc.) inside and outside of the classroom for learning. Students provided feedback indicating that they feel that 3D computer animations can help them get a better grasp of the material, clarify topics discussed in class, and even provide more information than that of printed text (e.g., textbooks).

These technologies have been implemented into the instruction of two courses in the department of Communication Sciences and Disorders at Illinois State University. Assessment outcomes from using 3D computer

Continued on page 14...

Students' Perceptions of Using 3D Technology (cont'd)

animations for learning was recently reported (Perry et al., 2009). Outcome measures demonstrated that students who studied using the 3D computer animation performed better on a 10-question examination compared to that of students receiving only traditional methods (e.g., textbook and pictures) for learning.

Inevitably, as technology continues to make advancements, virtual reality will find a place in most classrooms. It appears evident that students are comfortable with this movement and are supportive of professors implementing such technologies for improving learning. As a professor of the sciences, it is important to me that the "glitziness" of technology not overshadow the importance of the basic sciences. However, our students are growing up in a visual world, and it is likely that using such technology may indeed improve learning and student engagement.

References

Garg, A., Norman, G., Spero, L., & Maheshwari, P. (1999). Do virtual computer models hinder anatomy learning? *Academy of Medicine, 74*, S87–9.

Hilbelink, A. (2009). A measure of the effectiveness of incorporating 3D human anatomy into an online undergraduate laboratory. *British Journal of Education Technology, 40*, 664-672.

Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: Effect of context of prior presentation. *Journal of Experimental Psychology of Learning Memory and Cognition, 15,* 200-210.

Khalil, M., Johnson, T., & Lamar, C. (2005). Comparison of computer-based and paper-based imagery strategies in learning anatomy. *Medical Education, 18,* 457-464.3.

Lewis, M. (2003). Computer-assisted learning for teaching anatomy and physiology in subjects allied to medicine. *Medical Teacher, 25*, 204-206.

Nicholson, D., Chalk, C., Funnell, W., & Daniel, S. Can virtual reality improve anatomy education? A randomized controlled study of a computer-generated three-dimensional anatomical ear model. *Medical Education, 40*, 1081-1087.

Perry, J., Cunningham, D., Kuehn, D., & Gamage, J. (2009). Do 3D Stereoscopic Computer Animations Improve Student Learning of Surgical Procedures? *Journal of Effective Teaching,* under review.

Petersson, H., Sinkvist, D., Wang, C., & Smedby, O. (2009). Web-based interactive 3D visualization as a tool for improved anatomy learning. *Anatomical Sciences Education, 2*, 61-8.

Schleich, J., Dillenseger, J., Houyel, L., Almange, C., & Anderson, R. (2009). A new dynamic 3D virtual methodology for teaching the mechanics of atrial septation as seen in the human heart. *Anatomical Sciences Education, 2*, 69-77.

Two Perspectives on Assessment

Matt Fuller, Assistant Director, University Assessment Office

Amee Adkins

Todd McLoda

Two colleges at Illinois State University have recently hired Associate Deans responsible for assessment activities and support in their colleges. Dr. Amee Adkins, Associate Dean for Assessment and Undergraduate Education in the College of Education, and Dr. Todd McLoda, Associate Dean in the College of Applied Science and Technology, are providing leadership and vision to their colleges in regards to assessment. I have had several opportunities to discuss assessment with Drs. Adkins and McLoda. The follow Q&A session was developed as a means to introduce their ideas and vision for assessment in their colleges.

Continued on page 15...

Two Perspectives (cont'd)

***Matt Fuller*:** What is assessment to you? What is the value of assessment?

***Amee Adkins*:** Assessment is a way of examining whether what we're doing is meeting the goals, especially learning outcomes, that we set for our programs. For me, the value of assessment is first, that it reminds us to be intentional (asking us what are our goals) and second, that it's a way to see how effective we are. With assessment we either reassure ourselves that we're meeting our goals or we identify what we might do differently.

***Todd McLoda*:** The process of assessment is often diagrammed as a closed loop with steps connected by arrows. As a visual learner, I prefer to think of assessment as a coiled spring aimed at continual improvement without repetition. The process of assessment and the data collected continue to provide indicators for success and purposeful measures for targeted change. The path continues to build toward an ongoing, progressive process that meets current needs as well as future demands. I view the steps of setting goals, gathering evidence, analyzing the evidence, and using the results to set new goals as an evolution for managing programmatic stability and ensuring focused improvement.

***Fuller*:** When a faculty or staff member in your college engages in an assessment project, what do you hope they experience or come away with from the project?

***Adkins*:** I hope they finish with a sense of fulfillment and reward. I worry that people associate assessment with a "shame game" or scrutiny to justify harsh judgment. If that's all assessment could be, I sure would avoid it, too! However, when it's a process of laying out, "Here's what I think is important to accomplish. I wonder how well I'm doing," the results of the inquiry should feel satisfying, not threatening. Curiosity drives good assessment.

***McLoda*:** My sincere hope for faculty members engaged in assessment activities is that they collect high quality data and indicators of success and that they employ an active process to use that information in a way to enhance programmatic quality beyond expectations or standards. I would like to see individuals involved with assessment institute a purpose that focuses on the stakeholders. Many subscribe to the practice of assessment to meet needs related to resource allocation, external accrediting agencies, or for other process-output measures. The link to assessment that is often forgotten is the emphasis on student learning.

***Fuller*:** If you had a magic wand and could change one thing about the assessment field, what would it be?

***Adkins*:** Its momentum would not be driven by the high stakes accountability discourse because that brings a negative emotional charge. In the field of education, No Child Left Behind (NCLB) provokes a lot of anxiety, some of it rightly felt, which I think unfortunately displaces our energies — everything that we give up to NCLB-anxiety is energy we aren't directing toward good teaching and learning. With the magic wand, I would be able to engage groups of people in curiosity- and commitment-driven inquiry, the results of which would most likely be more useful to everyone than what we see coming out of anxiety-inducing demands for accountability.

***McLoda*:** I would like to see more units complete the practice of assessment by employing a plan to use the data collected as a means of enhancing curricula, programs, resource requests, and marketing efforts. All of these can be aimed at student learning. I sometimes see programs that are very efficient at collecting a variety of measures, but then fall short in applying what is learned to the ongoing process of improvement. The viewpoint seems to be that assessment is an occasional task rather than a persistent mechanism for directing change and planning for the future. Programs should be considered as fluid rather than stationary. This ensures that flexibility and foresight can be used advantageously.

***Fuller*:** Five years from now, what do you hope your college is or is not doing in terms of assessment?

***Adkins*:** Fifteen minutes from now...patience. Five years from now, I hope my colleagues are systematically assessing genuinely, on the basis of their commitments for intended learning outcomes and student benefit, instead of being driven by a motive to "comply" with a set of extraneous demands. I hope they experience assessment as routine a practice as updating a syllabus for the start of the semester: Of course this is something we do around here.

***McLoda*:** It is my aim within CAST to encourage enhanced usage of data generated both within the College

Continued on page 16...

Two Perspectives (cont'd)

and by the University Assessment Office for purposes beyond budget requests and periodic program reviews. For example, in CAST, all academic units use the IDEA Student Ratings of Instruction as a partial means of evaluating teaching. The data generated from the use of IDEA provides excellent feedback to the individual faculty member and the unit administrators related to summative teaching effectiveness and formative advice on targeted improvement. This data, when used appropriately, can guide enhancement of teaching effectiveness and can demonstrate that changes in teaching methods are resulting in the intended outcomes for the course.

***Fuller*:** Are there any other comments you'd like to include about assessment in your college, the university, or higher education?

***Adkins*:** No, not really. But I think that, at least here at ISU, we are doing a wonderful job!

***McLoda*:** Of all the instruction provided by my father during my youth, perhaps the only lesson that I consistently choose not to follow is, "if it isn't broke, don't fix it." In other words, I consistently seek opportunities to improve upon processes in the pursuit of excellence. Assessment is an invaluable tool for directing informed change. The negative consequence of such a philosophy is that collecting information and refining processes implies dissatisfaction but this is not the case. I simply encourage program personnel to not just meet standards, but seek opportunities to exceed them. This ensures that we are mindful of our collective vision while meeting the goals of our strategic plan and the ongoing, evolving needs of our primary stakeholders.

The UAO & U

Derek Herrmann, Graduate Assistant, University Assessment Office

At the 2010 CTLT Teaching and Learning Symposium, "Sustainable Teaching, Sustainable Learning, Sustainable Living," the University Assessment Office staff provided a presentation on the different services that are offered, free of charge, to ISU faculty and staff. The "UAO & U" presentation included Renée Tobin, Acting Director and Associate Professor of Psychology, as the session chair. Assistant Director Matt Fuller discussed the General Education Assessment Institutional Artifact Portfolio (IAP), the Process for Review of Academic Assessment Plans (PRAAP), and the Assessment Plan Tutorial. UAO Coordinator Jon Laird presented information about SelectSurvey and the process of developing, administering, and analyzing online surveys. UAO Graduate Assistant Derek Herrmann talked about the Alumni Survey and the *Progressive Measures* newsletter. It was a well-attended presentation with good questions and discussions. The PowerPoint presentation is available online at http://www.assessment.ilstu.edu/activities_services/index.shtml.

More information on these topics and other services can be found on the UAO's website, http://www.assessment.ilstu.edu/.

University Assessment Office

Campus Box 2500
Normal, IL 61790-2500
309.438.2135 (phone)
309.438.5602 (fax)
www.assessment.ilstu.edu