

A PROPOSAL TO USE CLASSROOM OBSERVATIONS AS ASSESSMENT DATA TO MEASURE AND EVALUATE EFFECTIVE TEACHING

Jace Hargis, Ph.D.

Director, Center for Teaching and Learning, University of California, San Diego
La Jolla, CA 92093 USA
jace.hargis@gmail.com

Melissa Soto, Ph.D
Assistant Professor

San Diego State University, San Diego, CA 92182 USA
melissoto@gmail.com

ABSTRACT

Effective teaching has historically been evaluated by using student perception surveys. This paper proposes a method to supplement the student surveys. The method would empower faculty to gather additional assessment data aligned with effective teaching and subsequently provide them with measurement tools to self-evaluate their effective teaching. For the theoretical framework, we integrate the three major stages of assessment; measurement; and evaluation. The assessment stage addresses three further steps, which include a/n pre-observation; observation; and post-observation. Within the observation step, there is also three parts, a quantitative checklist; qualitative field narrative; and faculty flow diagram. In the measurement stage, an analytical rubric has been created and proposed to assist faculty in measuring their assessment data and enable a more complete final evaluation step. The ultimate goal is to provide faculty with the ability to gather broader representations of their teaching, which will enable them to monitor, update and continuously improve their personalized teaching philosophy.

Keywords: Effective Teaching, Student Perception Survey, Observations, Assessment, Measurement, Evaluation

INTRODUCTION

In nearly all U.S. university classrooms, student evaluations, or more accurately, student perception surveys, are used to evaluate teaching effectiveness (Murray, 2005). This study addresses effective teaching by recommending the collection of additional data, to supplement student perception data. The intent of the new process is to help provide a more holistic and therefore accurate representation of the teaching and learning experience through the well-established process of assessment, measurement and evaluation. First, it is critical to operationally define these three key terms, used as the foundation of this study:

1. **Assessment** is a process of gathering information from multiple and diverse sources to understand learner knowledge, skills and dispositions (Huba & Freed, 2000).
2. **Measurement** is an assignment of marks based on an explicit set of criteria (rubrics) (Sadler, 2005).
3. **Evaluation** is the process of making judgments on the level of performance - based on assessment and the set of criteria (Macquarie University Learning & Teaching Centre, 2010).

All three of these concepts are necessary to create an accurate account of teaching and learning effectiveness. Therefore, this study will recommend three distinctive steps, which will allow the efficient collection of this data and empower faculty to receive personal data on the methods and approaches they offer in their classrooms. Subsequently, faculty are provided guides to assist them in interpreting and reflecting on the data to create their own evaluation, teaching philosophy/statement and action steps for continuous improvement.

Author (2014) described an observation process in which he used a teaching attribute instrument, modified from Chism (1999); qualitative notes of behaviors; and a faculty flow diagram. Author (2014) observed one hundred and ninety-six faculty members from three different institutions in two countries over a ten-year period. Faculty came from many different disciplines including the arts, sciences, engineering, information technology, business, law, education, nursing, pharmacy, physical therapy and dentistry disciplines. The purpose of this current study is to describe how observations can be integrated into a larger evaluation process and to illustrate how this process has been scaled to include larger classroom settings and a more diverse student population. In this study, data is incorporated from two universities (a small private liberal arts and a large research intensive), which adds 40 faculty members (for a total of 236 faculty) and 3,165 students to the original data set.

LITERATURE REVIEW

Effective Teaching

There are several interpretations of effective teaching. Through a fifteen-year study of nearly one hundred college teachers in a wide variety of fields and universities, Ken Bain (2004) summarizes effective teaching attributes in his book entitled, "What the Best College Teachers Do." The result of Bain's research shows that effective teachers know how we learn, process information and make conceptual connections; teach as serious intellectual endeavor; assess their own efforts and make continuous adjustments; create an environment where diverse learners explore, analyze, synthesize and construct meaning for themselves; and integrate the Scholarship of Teaching and Learning (SoTL) as an essential contribution of effective teaching.

Furthermore, Bain (2004) connects effective teaching attributes to research. Effective learning is supported when students are actively engaged in the learning process (Chickering & Gamson, 1987). This can be accomplished by creating a climate of inquiry where students feel appropriately challenged and activities are linked to research and scholarship (Jenkins, Blackman, Lindsay, & Paton-Salzberg, 1998) and students' experiences are acknowledged, valued, and drawn on in learning and teaching approaches (Bransford, Brown, & Cocking, 1999; Humphreys, 1998). Another important aspect of effective teaching is providing students with ample opportunities to engage in formative assessments. These include structured occasions for students to reflect on their experiences to challenge current beliefs and develop new practices and understandings (Gibbs, 1981), as well as reflect on and implement meaningful and timely feedback (Isaacs, 2001). When students are encouraged to take responsibility for their own learning, they are more likely to develop higher-order thinking skills such as analysis, synthesis, and evaluation (Zimmerman, 1998). Effective teaching can also be enhanced when students become part of a community of learners in which they can draw on each others' ideas. If dialogue is encouraged between students and teachers and among students, student motivation and engagement can be increased, thus creating a community of learners (Stefanou & Salisbury-Glennon, 2002). Ultimately, instructors must clearly articulate course expectations, goals, learning outcomes, and requirements to increase student motivation and improve learning (Race, 2001). By integrating all of these aspects, effective learning is facilitated and support the achievement of desired learning outcomes (Biggs, 2002).

Student Perception Surveys

There are numerous studies which have shown the inadequacies of student perceptions on effective teaching and learning (Benton & Cashin, 2012; Carrell, & West, 2010; Clayson, 2009). A recent meta-study by Uttl, White, and Gonzalez (2016) found zero correlation between student evaluations and learning. The authors suggest that institutions focused on student learning and career success may want to abandon student evaluation ratings as a measure of faculty's teaching effectiveness. Stark (2016) also found strong evidence that student responses to questions of effectiveness do not measure teaching effectiveness. Stark's recommendation is that "the common practice of relying on averages of student teaching evaluation scores as the primary measure of teaching effectiveness for promotion and tenure decisions should be abandoned for substantive and statistical reasons (p. 2)."

Additional studies have been conducted to specifically identify biases in student perception data. A variable explored by Walsh, Millar, and Westfall, (2016) was the "attractiveness" of a professor. The authors asked 131 randomly assigned college students to listen to a recording of a 20-minute physics lecture. While the lecture was playing, a computer displayed an attractive person in some cases and not attractive in others. Using a 25-item quiz, there was a statistically significant difference in favor of the students viewing the attractive instructor. In addition, Boring, Ottoboni, and Stark (2016) recently found that student evaluations of teaching are biased against female instructors by an amount that is statistically significant; the bias varied by discipline and by student gender; and biases could be large enough to cause more effective instructors to get lower evaluations than less effective instructors. Effective evaluation requires reliable and valid data collection instruments and well aligned outcomes from which to be assessed and measured. Student perceptions of faculty performance are very important, however, that data alone is insufficient to describe the full extent of effective teaching.

Assessment (Observations), Measurement and Evaluation

Although every university teaches students the importance of assessment, measurement and evaluation, ironically, using these critical attributes to evaluate our own teaching effectiveness is often overlooked. Using multiple assessments and measures is essential to determining an accurate evaluation of the instructor (Goe & Croft, 2009). In addition, the lack of complete assessment data fails to recognize the wide range of ways in which teachers contribute to student learning. In our study we operationally define assessment as a process of gathering information from multiple and diverse sources to understand learner knowledge, skills and dispositions (Huba & Freed, 2000). The primary method for gathering information is through classroom observations and the diverse sources are through quantitative, qualitative and graphical representations (Author, 2014). Danielson and

McGreal (2000) cite multiple assessments, such as reflection; collaboration; self-directed inquiry; and participation in a community of learners as assessment data to measure and use for subsequent evaluation.

Measurement is an assignment of marks based on an explicit set of criteria (Sadler, 2005). A common approach to measuring assessment data is to create a set of criteria that align with key aspects of the intended goals and outcomes. To accomplish this approach of measurement, an analytical rubric is often used. There are many types of rubrics which can provide a clear set of criteria associated with various levels of success, with gradations of points or quality descriptive terms (Hargis, Cavanaugh, Kamali, & Soto, 2014). Rubrics create a clear bridge between the outcome and assessment (Sebastiá, et al, 2011). They define the criteria, especially with processes or abstract, qualitative concepts. Rubrics are objective, consistent and an inter-rater reliability coefficient can be generated, which determines the rigor. Rubrics are measurement instruments that articulates the expectations for an artifact by stating the criteria and describing levels of quality (Maki, 2012). Rubrics can be a useful way to communicating expectations, and providing focused feedback on works in progress (Brown, Race, & Smith, 2005). Rubrics can provide clear prompts to monitor and continuously improve effective teaching. When used as part of a formative assessment and measurement, rubrics have the potential to help instructors develop knowledge, skills, and dispositions as well as make dependable, consistent judgments about the quality of their own work (Davies, Welch, & Hargis, 2008).

This study adds to the literature by providing research-based methods to integrate the observation assessment data as part of a larger measurement and evaluation process.

APPROACH

This data collection was conducted at two universities, a small private Catholic liberal arts and a large research intensive, both in the Pacific Rim of the United States. For the updated data, 16 faculty members (eight female) were from the small private liberal arts university; and 24 faculty members (12 female) work at the large research intensive university. Data were collected over a two-year period including the 2014 to 2016 academic years. Participants from the prior study included 196 faculty members (102 female) from three different universities in two different countries. Seventy-six faculty (47 male) were from a mid-size comprehensive public university; eighty-three (52 female) from a private liberal arts university and thirty-seven (21 female) from a federal middle eastern college (Author, 2014).

A Three Stage Process for Evaluating Effective Teaching was developed and deployed.

- I. Assessment
 - a. Pre-Observation Meeting
 - b. Observation
 - i. Quantitative Checklist
 - ii. Qualitative Field Notes
 - iii. Faculty Flow Diagram
 - c. Post-Observation Meeting
- II. Measurement
- III. Evaluation

Assessment (Collect Observational Data in Three Stages)

To initiate the classroom observation process, faculty voluntarily contact the Center for Teaching and Learning to request a classroom observation. The Center responds with a request to schedule a pre-observation visit to begin the process.

Pre-Observation Visit

There are several reasons why a faculty member might request a classroom observation. They might be addressing one or more student perception comments. They may wish to obtain documentation of their teaching for an upcoming promotion and tenure review or to add to their annual report as evidence for effective teaching. Most frequently, they are simply interested in continuous improvement and engaging in a lifelong learning model, similar to the one that they encourage for their students (Malmberg, Hagger, Burn, Mutton, & Colls, 2010).

Angelo and Cross (1993); and Davis (2009) suggest that an observer schedule a 15-30 minute session with the faculty member to be observed to share the process. For this study, a process that included three data instruments was used:

- A quantitative quality teaching checklist;
- Qualitative field-notes on teaching and learning activities; and

- A graphical faculty flow diagram diagramming noteworthy actions.

Prior to the observation, the observer and faculty member meet to discuss the process using the following Pre-Observation prompts.

1. What are the student learning outcomes (knowledge, skills & dispositions)?
2. What teaching approaches will be offered (active learning, discussion/Socratic, handouts, videos, lecture, demonstrations/simulations, project-based learning, etc.)?
3. What are the instructor goals for this session?
4. What can the observer expect the instructor to be doing in class?
5. What do you expect students to be doing to reach your goals?
6. How will you know that students have met your goals?

Observation

Using the three data collection instruments, the observer visits the classroom and collects observational data.

i. Quantitative Effective Teaching Checklist (Chism, 1999 reduced to 73 items)

The observer uses a quantitative teaching attribute checklist as a guide to identify which of the attributes are observed during the session. The observer uses notations other than a check indicating the behavior was observed. For example, the observer indicates NA or Unobserved; or a “Y” indicating “Yes” this was easily and often observed; likewise for an N for “No”, etc. The goal for this more open notation is to provide the observer space to document what they observe in all of the nuances that we know happens in the art and science of teaching. It is understood that this list is limiting, and only captures a snapshot in time, indicating that the instructor may or may not have exhibited the ability to attend to an attribute once (or twice). The purpose of this checklist is not to use the number of attributes checked and conclude a level of instructor performance. The purpose is to note which of the attributes are observed, expand upon these in the Remarks section and if significant, expand upon what was observed in the Qualitative notes. Upon debriefing, the observer summarizes and shares the data in a conversation, where the faculty member being observed can ask questions and determine the value and/or intent of the attributes. Ultimately, it is up to the faculty member to place a value and evaluate each attribute as per their goals, class level, and individual instructional philosophy.

ii. Qualitative Field Narrative

The observer collects qualitative field notes of student behavior and specifics on their questions/responses; dis/engagement; student note-taking; work on their computers or other mobile devices (on/off task); talking or other distracting behavior; arriving late or leaving early.

iii. Faculty Flow Diagram

A visual faculty flow diagram, which graphically displays the movement of faculty and students who respond aloud during the session is created (Author, 2014). The observer diagrams the room configuration, complete with the location of doors, windows, whiteboards, etc. (or downloads the room schematic from the university website) as the location of these physical parameters may affect learner interaction. The observer also identifies the location of male/female students, and as they participate, make notations on when they respond. The observer documents which students are working on laptop computers, the movement of faculty around the class, and any other notable activity (i.e., classroom disruption, group activity movement, students arriving late, departing early, technology used or not working, students off/on task, etc.). Following the three-part data gathering, the observer creates a one-page summary of the events, similar to a field guide.

Post-Observation Debrief

To be most helpful, a post-observation debriefing occurred within one week of the observation. It was beneficial for the observer to begin the debrief by reiterating the purpose of the observation and maintain a collegial approach. The goal is not to be punitive, policing or “find” problems. The goal is to provide useful and specific information about teaching and learning dynamics so that the instructor can make informed evaluations. Throughout the entire debriefing session, the observer offers open-ended, non-assuming prompts.

Plans for the Future

Of the three stage process, we have collected extensive data on the first stage of Assessment, which includes the Pre-Observation Meeting; Observation (Quantitative Effective Teaching, Qualitative Field Narrative, and Faculty Flow Diagram) and Post-Observation Debrief. Combining this research with recent conversations with over two hundred faculty members, who have consistently requested more effective ways to evaluate teaching, we have created a Measurement rubric (Appendix A), which can be used in concert with the Observation data and Student Perception surveys for faculty to prepare their own data-informed self-evaluation of their teaching.

II. Measurement

An analytical rubric was created to assist faculty while interpreting and evaluating their assessment data. Portions of the rubric were derived from the Western Association of Schools and Colleges (WASC) Senior College and University Commission (WSCUC) (of which some portions were aligned with the Association of American Colleges and Universities (AAC&U) rubrics) and address an Initial, Emerging, Developed, and Highly Development indicators. For the fullest picture of an instructor's accomplishments, reviews of student survey materials (quantitative and comments) should be augmented with:

- Pre-observation discussion notes;
- Syllabus (aligned, measurable learning outcomes; assessments with rubrics; clear policies, which help create an engaging learning environment; sensible schedules, which provide complete information, perhaps links to external resources, etc.);
- Quantitative quality teaching checklist with remarks;
- Qualitative field-notes on teaching and learning activities;
- Graphical faculty flow diagram diagramming noteworthy actions;
- Post-observation debriefing notes; and
- One-page observation summary.

III. [Self] Evaluation (of the faculty member being observed)

The final aspect of this on-going cycle is evaluation. In this process, each faculty is provided the observation forms and rubric, which they use to evaluate themselves. Instructor outcomes for this step can vary widely and include

- descriptive narratives of the process culminating in a self-evaluation report;
- tables/charts/figures which summarize their systematic reflections on teaching and/or identifying challenges with associated actions to enhance;
- on-going development of teaching statements/philosophies, demonstrating an evolving instructional repertoire; or
- perhaps digital audio or video files, where faculty represent their teaching with voice or text overs in documents or [info]graphic representations (PowerPoints) to describe their thinking and intent.

The evaluation can be formative (obtained and reviewed while the course is being taught) and/or summative (at the the end of the term/year as part of an annual performance dossier). Above all, the evaluation is a "SELF" evaluation, which is performed by the faculty member and not the observer. This approach can empower the faculty member to play an active role in the evaluation of their career direction and progress. The main goal of this entire process is to provide useful and consistent data to help faculty evaluate their own teaching in individual and meaningful ways. The Center is available to discuss ways in which to evaluate the data and suggest methods, mechanisms and formats if the faculty member would like to initiate this conversation. However, the Center is a resource and not an evaluative entity (Yee, & Hargis, 2012).

Assessment Results

Over the twelve-year period, classroom observation data was collected from five different universities with highly varied demographics and mission. Universities included a Comprehensive with an enrollment of 16,000 students; a Liberal Arts with 6,000; a Middle East Women's College with 2,500; a small private Catholic university with 2,500; and a large research intensive with 33,000 students. Faculty member demographic data from the prior study is included in table 1. Table 2 includes observational data from the two additional universities.

Table 1. Faculty Member Demographics from Prior Study

Institution	Total	Female	AS	EIT	Law	BUS	Ed	Hlth
Comprehensive	76	29	36	8	0	10	13	9
Liberal Arts	83	52	34	6	10	0	9	24
Middle East	37	21	9	5	0	19	0	4
Total	196	102	79	19	10	29	22	37

AS – Arts and Sciences

EIT – Engineering and Information Technology (which includes Computer Science)

Bus – Business
Ed – Education
Hlth – Health (which includes Nursing, Pharmacy, Physical Therapy and Dentistry)

Table 2. Discipline and number of students for two additional universities.

Small Private Liberal Arts University
(Student Enrollment of 2,500)

Faculty Discipline	# Students
English	10
Literature	15
Psychology	25
Environmental Studies	18
Marine Ecology	15
Business	12
Philosophy -1	18
Philosophy -2	16
Religion	21
Sociology	22
Media	17
Sustainability	12
Criminal Justice	18
Education	11
Rhetoric	15
Marketing	19
Total = 16 Faculty	265

Large Research Intensive University
(Student Enrollment of 33,000)

Faculty Discipline	# Students
Political Science	110
History	90
Computer Science	240
Ecology Lab	35
Biology Education	80
Computer Engineering	170
Economics	250
Education	40
Anthropology	150
Visual Arts	20
Calculus	170
Anatomy	60
Japanese Language -1	60
Japanese Language -2	40
Physics I	100
History	270
Computers	160
Medicine	6
Electrical Engineering	135
Physics II	300
Philosophy	35
Endocrinology	100
Mathematics	125
Total = 24 Faculty	2900

DISCUSSION

This purpose of this paper is to share a method to gather useful data in addition to student surveys to help faculty, department chairpersons, and universities evaluate effective teaching. The method empowers faculty to

gather specific data, which can assist in their reflection and analysis on how they are and/or will become a more effective educator. The proposal requires three major stages, which include assessment; measurement; and evaluation. The assessment stage includes pre-observation; observation; and post-observation. Within the observation step, there is also three parts, a quantitative checklist; qualitative field narrative; and faculty flow diagram. For the measurement stage, we created an analytical rubric to help faculty make sense of the and prepare an evaluation of their teaching. One important aspect to keep in mind is that this proposal is a supplement to student surveys and not a replacement. We believe that students' perception of our instruction is critical, although it is only one piece of a much more complex storyline. If we are serious about improving our teaching and offering a high quality instructional experience to our students, then - as with most things in life - our approach will require careful consideration, resources and most of all, our time. This proposal is not a magic bullet that will quickly and perfectly determine teaching effectiveness. Although we can and do gather student perceptions quickly, through multiple choice and open response surveys at the end of the term, we have shared the evidence that clearly demonstrates that this approach is severely lacking in evaluating teaching effectiveness. We also believe that additional information, such as peer evaluations from colleagues can be useful, although produces an entirely different type of data. Peer evaluations (observations) are typically self-selected and can be biased due to the relationship between the observer and the faculty being observed. In addition, the peer evaluator frequently becomes focused on the content and not the pedagogy, therefore, we recommend that peer evaluators should be from outside one's own discipline. Finally, the peer evaluator most likely does not have experience in observing effective teaching, nor do they have the instruments, such as the one's in this proposal, especially the Faculty Flow Diagram, which has been seen to be the most helpful for faculty members.

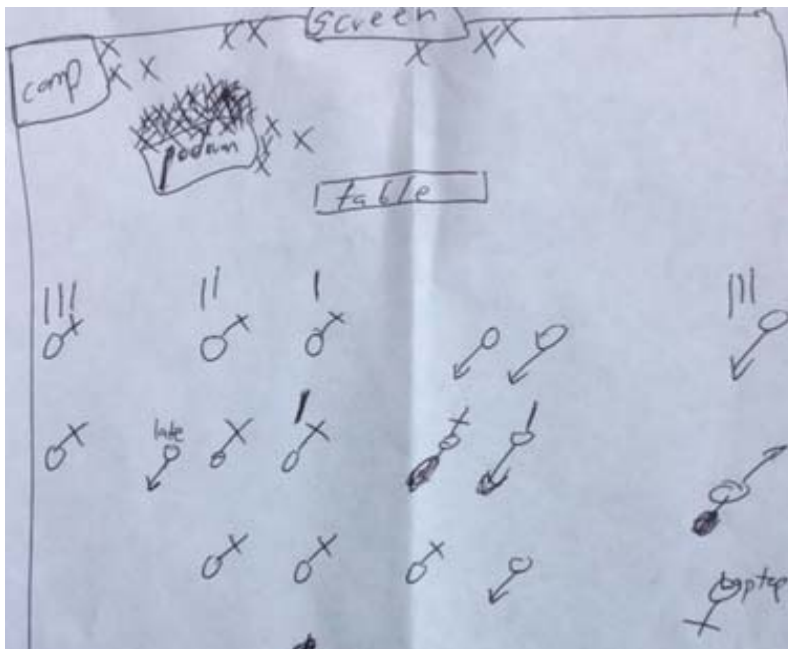


Figure 1. Example of a Faculty Flow Diagram of 18 students at a small Liberal Arts Catholic university

Faculty Flow Diagram Examples

One of the most unique and discussed aspect of the evaluating effective teaching proposal is the Faculty Flow Diagram (Figure 1). In these diagrams, such as the example in Figure 1, the "x's" represent the instructor at various times during the session, the gender symbol represent the location of each student and the marks near the student represent the number of times when the student spoke aloud (answering or asking a question). For Figure 1, you can see that the instructor remained mostly near the podium for the duration of the class session. When debriefing with the instructor, they shared the reason for this is that they felt they needed to be near the podium, which housed the computer, that operated the PowerPoint presented so they could advance the slides. When debriefing, it is key that the observer remains objective and avoid judgment on the instructor's behavior. Instead, the observer asks further questions, such as in this case, "what if we were to provide you with a remote control to advance your slides, where would be your preferred location in the class?" In this particular example, this question was asked and the instructor responded that they most likely remain near the podium because they felt that was where the students expected them to be. As an observer, the first author shared research on effective teaching, which included student engagement and how the instructor's movement could enhance engagement, especially if one of their goals was to encourage students to ask questions. Gunter, Shores, Jack, Rasmussen, and

Flowers (1995) found that when instructors increase their proximity control, or decrease the distance between them and students, that students are more likely to ask questions. When this happens, students are less anxious to talk aloud to an expert at the front of the room, and instead will quietly share their confusion when the instructor passes by them.

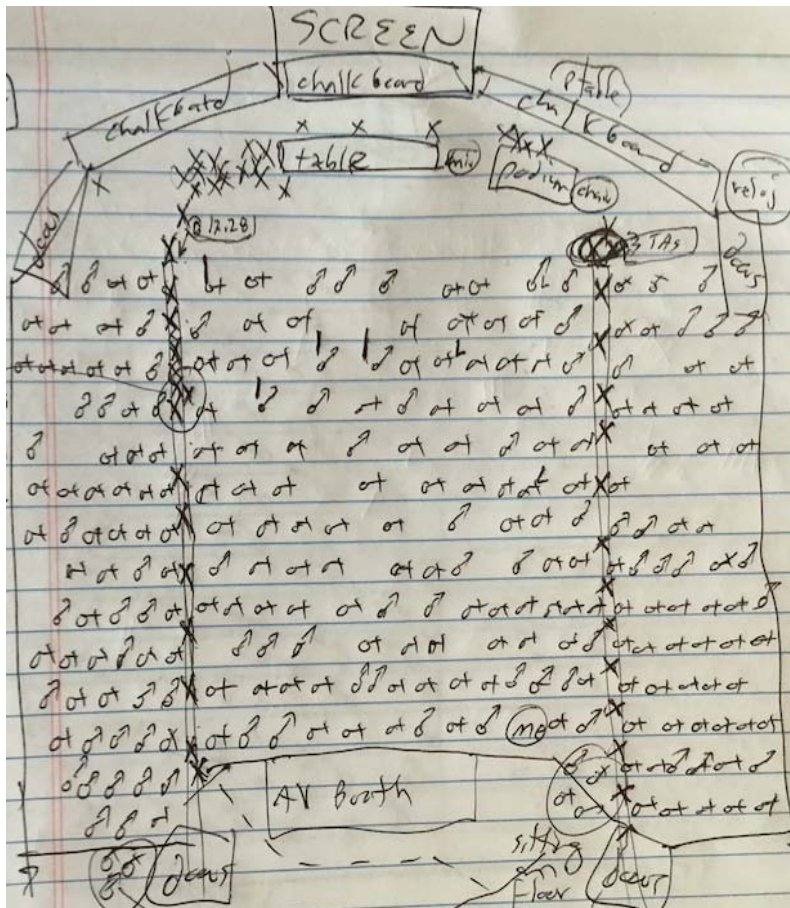


Figure 2. Example Faculty Flow Diagram from a large class at a Research Intensive university

At the other end of the class size spectrum, Figure 3 shows a Faculty Flow Diagram for a large class of approximately 300 students. Although you can see that there were not many questions from the students, that even in a very large class, an instructor can exhibit a heightened sense of proximity control by walking up the aisle. This approach allows the faculty to try and make a large class seem more personal, humanizing the dynamics as well as subtly monitoring for off-task activities. Most of the walking through the aisles occurred when the professor asked students to work on a problem and/or address an active learning method, such as Think/Share/Pair. The instructor in this example did have and use a remote to control what was being projected on the front screen.

Throughout the process of observing faculty in the classroom, the Center has identified several common requests from faculty following their observations, such as

1. Interest on how they can offer more Active Learning opportunities, which we provide and discuss a list of [228 Active Learning](#) methods to assist in student engagement;
2. How to engage students more frequently through Student Response Systems (SRS), which we share strategies including a simple four quadrant colored piece of paper (https://uminntilt.files.wordpress.com/2014/06/colored-abc_card.pdf) or electronic alternatives of Poll Everywhere (<http://www.polleverywhere.com>), Plickers (<http://www.plickers.com>), Go Formative (<https://goformative.com>), Kahoot (<https://getkahoot.com>) and Twitter (www.twitter.com);
3. How to create Electronic learning objects, which could include [Learning Glass Video](#), [Video Scribe](#), [StopMotion](#), [Screencast](#) or a [Green Screen](#);
4. How to create analytical rubrics using [Rubistar](#), an online rubric generator; and
5. How to build a more effective syllabus using our Syllabus Guide.

Limitations

There are, of course limitations to this proposal. As mentioned above, the major limitation is time, as this approach will take significantly more time than current student surveys. However, we believe that many Centers for Teaching and Learning currently offer these services and could develop ways to scale the services to accommodate the need. Possible ways could include paying a stipend to Faculty Fellows, training Adjunct Professors, identifying faculty who might be interested in this as part of their research agenda, and increasing the staff in Centers.

A second limitation and perhaps question is how large of a class is too large for the Faculty Flow Diagrams? The first author was able to draw and track a class of 300 students while at the same time monitoring the checklist and writing field notes. It has taken him several years of experience to be able to manage collecting data on three different instruments simultaneously, so class size could be a limitation.

Finally, interpreting the assessment data, while using the measurement rubric could be a challenge for faculty with limited experience in the social sciences. We believe that Centers could assist with this and ultimately (with permission) share showcase examples of faculty interpretations on their website so other faculty could use as a model.

Overall, we are hopeful that this proposal or perhaps specific data collection instruments could be helpful in enhancing university ability to evaluate effective teaching and supplement the current use of student perception surveys.

References

- Angelo, T.A., & Cross, K.P. (1993). *Classroom assessment techniques: A handbook for college teachers*. (2nd ed.). San Francisco: Jossey-Bass.
- Author, J. (2014). A ten year study of faculty classroom observations. *Transformative Dialogues: Teaching and Learning Journal*, 7(2), 1-21.
- Bain, K. (2004). *What the best college teachers do*. Cambridge, Mass: Harvard University Press.
- Benton, S.L., & Cashin, W.E. (2012). *Student ratings of teaching: A summary of research and literature*. IDEA Paper No. 50. Manhattan, KS: The IDEA Center. Retrieved from http://www.theideacenter.org/sites/default/files/idea-paper_50.pdf
- Boring, A., Ottoboni, K., & Stark, P. (2016). Student evaluations of teaching (mostly) do not measure teaching effectiveness. *ScienceOpen Research*.
- Boyer, E. (1990). *Scholarship reconsidered: Priorities of the professoriate*. San Francisco: Jossey-Bass.
- Bransford, J., Brown, A., & Cocking, R. (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.
- Brown, S., Race, P., & Smith, B. (2005). *500 tips on assessment*. (2nd ed.). New York: Routledge.
- Biggs, J. (2002). *Aligning Teaching and Assessment to Curriculum Objectives. Imaginative Curriculum Guide*. Retrieved from <http://www.itslifejimbutnotasweknowit.org.uk/files/CPLHE/Biggs%20Constructive%20alignment.rtf>
- Carrell, S.E., & West, J.E. (2010). Does professor quality matter? Evidence from random assignment of students to professors. *Journal of Political Economy*, 118(3), 409-432.
- Chickering, A., & Gamson Z. (1987). *Seven principles for good practice in undergraduate education*. Reprinted by University of Illinois, Springfield.
- Chism, N. (1999). *Peer review of teaching: A sourcebook*. Bolton, MA: Anker.
- Clayson, D.E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*, 31(1), 16-30.
- Conroy, M., Asmus, J., Ladwig, C., Sellers, J., & Valcante, G. (2004). The effects of proximity on the classroom behaviors of students. *Behavioral Disorders*, 29(2), 119-129.
- Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Davis, B.G. (2009). *Tools for teaching*. San Francisco, CA: John Wiley & Sons, Inc.
- Davies, J., Welch, C., & Hargis, J. (2008). The bridge course design: Faculty collaboration and cross-course formative assessment. *International Journal for Scholarship of Teaching and Learning*, 2(2), 6.
- Gibbs, F. (1981). *Teaching students to learn: A student-centred approach*. Milton Keynes, England: The Open University Press.
- Goe, L., & Croft, A. (2009). *Research to practice: Methods of evaluating teacher effectiveness*. Washington, D.C.: National Comprehensive Center for Teacher Quality.
- Gunter, P., Shores, R., Jack, S., Rasmussen, S., & Flowers, J. (1995). On the move: Using teacher/student proximity to improve students' behavior. *Teaching Exceptional Children*, 28(1), 12.

- Hargis, J., Cavanaugh, C., Kamali, T., & Soto, M. (2014). A federal higher education iPad mobile learning initiative: Triangulation of data to determine early effectiveness. *Innovative Higher Education*, 39(1), 45-58.
- Huba, M. E., & Freed, J. E. (2000). Learner centered assessment on college campuses: Shifting the focus from teaching to learning. *Community College Journal of Research and Practice*, 24(9), 759-766.
- Humphreys, D. (1998). The impact of diversity on college students: The latest research. *Campus Diversity Initiative, American Association of State Colleges and Universities and Ford Foundation*. Retrieved from <http://www.inform.umd.edu/diversityweb/NewsRoom/studbenefit.html>
- Isaacs, G. (2001). Assessment for Learning, University of Queensland, Brisbane, pp. 10-11.
- Jenkins, A., Blackman, T., Lindsay, R., & Paton-Salzberg, R. (1998). Teaching and research: Students' perspectives and policy implications. *Studies in Higher Education*, 23(2), 132.
- Macquarie University Learning & Teaching Centre. (2010). Retrieved from <http://www.mq.edu.au/about/about-the-university/offices-and-units/learning-and-teaching-centre>
- Maki, P. L. (2012). *Assessing for learning: Building a sustainable commitment across the institution*. Stylus Publishing, LLC.
- Malmberg, E., Hagger, H., Burn, K., Mutton, T., & Colls, H. (2010). Observed classroom quality during teacher education and two years of professional practice. *Journal of Educational Psychology*, 102(4), 916-932.
- Murray, H.G. (2005). Student evaluation of teaching: Has it made a difference? *Paper presented at the Annual Meeting of the Society for Teaching and Learning in Higher Education*, Charlottetown, Prince Edward Island, June 2005.
- Race, P. (2001). *The Lecturer's Toolkit: a practical guide to Learning, Teaching & Assessment*. Farmer, London: Routledge.
- Sadler, R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment & Evaluation in Higher Education*, 30(2), 175-194.
- Sebastiá, M., Catalá-Icardo, M., Mengual, J., González, J., Meseguer-Lloret, S., Torres-Cartas, S., Llinares-Palacios, L., Gómez-Benito, C., & Lull-Noguera, C. (2011). Rubrics as a bridge between assessment and learning: An experience in a Bachelor's degree in environment sciences, *EDULEARN11 Proceedings*, pp. 1986-1994.
- Siddiqui, Z., Jonas-Dwyer, D., & Carr, S. (2007). Twelve tips for peer observation of teaching, *Medical Teaching*, 29(4), 297-300.
- Stark, P. (2016). An evaluation of course evaluations. *ScienceOpen Research*, 1-7.
- Stefanou, C., & Salisbury-Glennon, J. (2002). Developing motivation and cognitive learning strategies through an undergraduate learning community, *Learning Environments Research*, 5(1), 78.
- Uttl, B., White, C., & Gonzalez, D. (2016). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related, *Studies in Education Evaluation*.
- Yee, K., & Hargis, J. (2012). Indirect faculty development and the role of sociability. *Journal of Centers for Teaching and Learning*, 4, 61-79.
- Walsh, M., Millar, M., & Westfall, R. S. (2016). The effects of gender and cost on suspicion in initial courtship communications. *Evolutionary Psychological Science*, 2(4), 262-267.
- Zimmerman, B. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. Schunk & B. Zimmerman (Eds.). *Self-regulated learning: From teaching to self-reflective practice*. New York: Guildford Press.

APPENDIX

Appendix A. Rubric for Measuring Teaching Effectiveness.

Criteria	Initial - Emerging - Developed
Observation (pre-discussion)	<p>The instructor has clearly articulated the ...</p> <ul style="list-style-type: none"> - concepts and learning outcomes for the session; - teaching approaches which will be offered (active learning, discussion, Socratic, handouts, videos, lecture, demonstrations/simulations, etc.); - a detailed conceptual lesson plan for the session; - their instructor goals for this session and a consistent method on how they will share these with students; - specific expectations that they have of students for the session, which will enable them to reach their goals; and - an assessment of student learning, which will provide substantive data on student abilities (knowledge, skills and/or dispositions) after the class session.
Syllabus	<p>The syllabus contains ...</p> <ul style="list-style-type: none"> - a teaching philosophy, which describes the instructional approach as it aligns with research-based effective teaching; - contact information (Instructor's name, email, phone number, office location, office hours, website, etc.); - course information (Course name/section number, credit hours, description, active, higher-level measurable Learning Outcomes well-aligned to assessments and department/university outcomes); - course policies, including attendance, academic integrity and student engagement; - ADA and Title IX compliance statements; - calendar with exams or other culminating learning opportunity date and time; - information about campus resources; - a policy on student decorum and appropriate, relevant and meaningful educational technology that encourages high-level discussions; - formative and authentic assessments with analytical rubrics; and - sensible schedules, which provide useful information to guide learners through conceptual frameworks.
Observation (quantitative)	<p>During the observation, a substantial number of effective teaching attributes were documented, such as</p> <ul style="list-style-type: none"> - Presentation skills (audible voice, varies the tone, avoids distracting mannerisms or reading from notes, maintains eye contact, spoke at effective pace, uses space well, is enthusiastic); - Rapport with students (addresses by name, attends to comprehension, provides feedback, uses reinforcement, incorporates student ideas, welcomes participation, motivates students, has appropriate sense of humor, uses effective classroom management, treats students impartially); - Clarity (defines and elaborates new concepts, uses examples, pauses for questions, responds and explains clearly, relates practical situations) - Instructor organization (arrives on time, states relation of the class, knows how to use ed tech, posts goals, is well-prepared, ensures learning activities are organized); - Content knowledge (incorporates current research, identifies sources, communicates reasoning behind operations, is knowledgeable and confident, pitches instruction to an appropriate level, demonstrates intellectual curiosity); - Variety and Pacing (uses more than one form of instruction, accepts students responses, draws nonparticipants into the discussion, helps students extend their responses, maps the direction of the discussion, provides directions for active learning tasks and how they will be evaluated); and - Instructional Strategies (teaching techniques is appropriate for goals, has a good, broad level of questioning skills, mediates discussion well, proceeds at an appropriate pace, ensures presentations are legible and organized, provides clear directions and facilitates group work well, helps students to learn from each other, effectively holds class attention).
Observation (qualitative)	<p>The faculty member was fully aware and is able to identify and assess specific student behavior which their lesson attempts to elicit. Behaviors such as the following were observed and fully reconciled:</p> <ul style="list-style-type: none"> - substantial student questions and responses; - engagement reinforced and/or lack of engagement addressed;

	<ul style="list-style-type: none"> - note-taking in various forms or other means of student interaction afforded; - working productively on mobile devices and/or distracted and addressed; - actively discussing valued and other distracting behavior addressed; and - behaviors such as arriving late or leaving early addressed.
Observation (Faculty Flow Diagram)	<p>When reviewing the Faculty Flow Diagram, the instructor is highly reflective and is able to identify and suggest strategies to address</p> <ul style="list-style-type: none"> - on if the student dynamics were what they had intended; - meaningful trends (more students responded in the front/back/T-box; gender differences, dominating students, instructor movements, etc.); - do these trends mean anything to the faculty member; - Would they like to increase or decrease certain behaviors; and - their movement and discuss how this facilitated learning and engagement.
Observation (post-discussion)	<p>Following the observation, the instructor was able to substantially describe</p> <ul style="list-style-type: none"> - their impression of how the session went, identifying key strengths and weaknesses, and rationale or hypothesis on why they believed so; - their thoughts and perspectives on learner interactions; - the quality and quantity of how well students accomplished outcomes; - potential variables for why activities may not have gone as planned; - clear areas of improvement and perhaps steps for improving; - how they can ensure students were attending, processing and applying to the instructor goals and how they knew it at the end of the session; and - different type and level of questions, as well as knowing when to ask each.
Teaching Philosophy	<p>The teaching philosophy includes</p> <ul style="list-style-type: none"> - a clear conception of teaching and learning as it connects to empirical evidence (including ways in which you integrate cultural diversity, access to underrepresented groups, attention to different ways of learning, creating an open and safe learning environment, classroom dynamics, etc.); - a complete description of how you teach and a deep background on why you selected those approaches, as well as indicators of how their effectiveness (diversity of methods, level of interaction, quality of feedback, intercultural sensitivity); - multiple, high quality examples of how you have been reflective and purposeful about your teaching; - clear aspirational goals as an instructor and corresponding actions; - rationale of your teaching values, beliefs, and goals; - a set of elevated criteria to assess, measure and evaluate the quality own teaching; - research based methods for assessing, measuring and evaluating students' learning with associated analytical rubrics; and - foundational information, such as courses you have taught, samples of course syllabi, teaching evaluations, letters of recommendation and a video of a class you have taught.