

Using Online Survey Tools to Improve Professor Assessment

Christopher Gennari

Camden County College, Blackwood New Jersey.

Once again assessment is a hot button issue. Teacher quality, at all levels, has become a political issue as much as an educational issue. It has also become a blame issue; with “bad” teachers being blamed for everything from the failure of the American educational system to the failure of the American economy. 2010 was a bad year to be a teacher. Of course, the attacks on educators was always prefaced with a differentiation between “good” and “bad” teachers yet a reading of the statements from a host of editorials, speeches and newspaper columns seems to accuse the education system of willfully harboring bad teachers in order to fleece the public and ruin our children.

Problematically, the definition of a “good” teacher used in these editorials is always left in a vague limbo akin to Justice Potter Stewart’s “I know it when I see it” definition of obscenity and pornography.¹ A Google search for the term “Qualities of a Good Teacher” does not bring up any authoritative reference. Instead the top websites are fluffy articles and blog posts which define the qualities of a good teacher in much the same way the children describe the qualities of a good nanny in Mary Poppins. These articles and posts emphasize qualities such as confidence, compassion, patience, fairness, control, guidance, entertainment and creativity.² Nowhere on the dozen sites I searched did “Knowledge of Material” factor into the qualities of a good teacher. This puts educators in a terrible conundrum; how can one be a good professional when the only standard used by society is a combination of fictional movie saints and the nostalgic remembrance of halcyon days gone by.

Everyone, including teachers and professors, want teachers and professors to be excellent in their field. Assessment in the modern classroom is based on a multi-method approach: peer evaluations, administrative evaluations and student evaluations. Other methods such as a professional portfolio and classroom management schooling may also be part of the evaluative package. Yet even these parts can be problematic. This paper discusses a new way to treat student evaluations, in the college setting, by using online survey sites in order to obtain a better assessment result.

¹ *Jacobellis v. Ohio*, 378 U.S. 184, 197 (1964)

² Examples of websites which were highly placed are http://www.school-survival.net/kit/How_to_be_a_good_teacher.php, <http://www.ripplesofimprovement.com/the-top-10-qualities-of-a-good-teacher/>, <http://www.school-teacher-student-motivation-resources-courses.com/goodteacher.html>.

Problems with Student Evaluations

Student evaluations of a professor make sense on the surface; the students are in the classroom with the teacher and have the longitudinal knowledge necessary to evaluate a teacher's professional success. In 1980 Gordon Greenwood wrote "student evaluations have become an important, if not the most important, kind of systematically collected data" concerning teachers in the classroom.³ James O'Hanlon agreed when he wrote student ratings "constitute the primary source of data" concerning teacher effectiveness.⁴ Suzanne Young, in 1999, commented that institutions spent "a lot of energy" on evaluating the characteristics of good teachers and student evaluations were a major direction of that energy.⁵ This systematic reliance on student evaluations has begun to affect the salary of professors. A few universities and colleges in Oklahoma and Texas began adding bonus pay to "good professors." The most important aspect in the decision-making process, a decision which could add ten thousand dollars to one's income, was a professor's student evaluation numbers. There is one small business school which terminates professors based on their student evaluations.⁶

This reality is problematic for educators for a number of reasons. First, the idea that one's employment or income might be dependent on an eighteen year old student who wears pajama bottoms to class understandably makes many educators queasy. Second, a reliance on student evaluations is mathematically problematic given that the validity of such evaluations is still "far from settled" according to Greenwood.⁷ A reliance on a poor methodology, even if the method sounds like it should be the correct course, has the problem of creating unintended negative consequences without accomplishing the original goals of the action.

³Gordon Greenwood, "Alternatives to Student Rating of College Teaching," *Journal of Higher Education* 51, no. 6 (1980): 673.

⁴ James O'Hanlon, "Making Teacher Evaluations Work," *Journal of Higher Education* 51, no. 6 (1980): 664

⁵ Suzanne Young, "Profiles of Effective College and University Teachers," *Journal of Higher Education* 70, no. 6 (1999): 671.

⁶ Katherine Mangan, "Bonus Pay Based on Student Evaluations Evokes Skepticism at Texas A&M," *Chronicle of Higher Education*, January 20, 2009.

⁷ Greenwood, 673.

Young's experiments point out an uncomfortable truth: student ratings correlate highly with instructor personality traits.⁸ Consequently, high student evaluations were based upon a professor's entertainment value rather than on their pedagogy. Ms. Young illustrates that charisma granted the professor highly effective marks regardless of content.⁹ Ms. Young's experiments also bore out the conclusions of a series of experiments in the 1970s which found student achievement levels highly correlated to teacher effectiveness in evaluations.¹⁰ An "A" student evaluated a teacher as good while a "D" student evaluated the same teacher as inept. Alexander Tolor's experiments found that teacher ratings were highly dependent on the class level of the students performing the evaluations.¹¹ A sophomore was far more likely to give poor marks to an effective teacher (as defined by peers, administrators and parents) than a senior level student.¹² Tolor also found that students were far more likely to define a teacher as ineffective than a professor's peers or administrators. Consequently, according to Tolor, students were bad at revealing poor teachers and often gave them poor marks not for content but for "non-cognitive" reasons.¹³ The mathematics suggests that we should pity the Organic Chemistry professor with an eight am class.

These facts highlight the essential problem with student evaluations of teachers: what makes the students a good judge of professional ability other than their accidental location? Who is rating the raters? Kathleen Crittenden points out "student evaluation of instructors is a process of person perception."¹⁴ This means that the personality of the student directly impacts their perception of the teacher since they are active in creating their own perceptions. Problematically, the investigation of student characteristics and values has been largely ignored.¹⁵ Most evaluations, such as the ones conducted at my school of Camden County College, are treated as if they are written by a monolithic

⁸ Young, 671.

⁹ Young, 671.

¹⁰ Young, 671. and Alexander Tolor, "Evaluation of Perceived Teacher Effectiveness," *Journal of Educational Psychology* 64, no. 1 (1973): 102.

¹¹ Tolor, 102.

¹² Tolor, 103.

¹³ Tolor, 102.

¹⁴ Kathleen Crittenden, "Student Values and Teacher Evaluation," *Sociometry* 36, no. 2 (June 1973): 143.

¹⁵ Crittenden, 144.

student body instead of two hundred individuals with their own value system, criteria and perceptions. A question as innocuous as “Does the professor return materials in a timely manner” thus becomes a chaotic cauldron of individualized personality quirks, peevs and values. “Timely” to one student will not be “timely” to another; and these students will also put different weight into their meanings of “timely.” One student may be fine with receiving materials within two weeks while the student sitting in the next chair may be irked into clenched fist rage if the grades are handed back at the end and not the beginning of the next class. This is born out in a Letter to the Editor of the Chronicle of High Education by Ross Keiser, Professor of Psychology at the University of Louisiana. Keiser points out that even when he handed back graded materials during the next class – the earliest it could possibly be returned - he never received higher than a 4.5 average on a 5.0 scale.¹⁶ If returning materials as soon as possible still earns a professor a one out of five in “timelessness” from a student then these types of student observations clearly suffer from a deficit of objective, statistical reality.

The student evaluation procedure is itself terribly marked with problems. Student evaluations are usually only given once a year perhaps just after the first exams (as in my case) or at the end of the semester. This single evaluation robs the teacher and the administration of the longitudinal data only students, in the class nearly every session, can provide for analysis. Secondly, having an evaluation so tied to examination periods must have a substantial effect on the students’ answers, especially since other research has already proven that student performance is correlative to evaluation ratings.

Camden County College uses a standardized Scantron form produced by an external company which sells the same generic form to a multitude of other schools. The effects is that the uniqueness of the class or the college, as a whole, is reduced to conformist bubbles and a small written section for additional comments which are rarely used. Finally, since these evaluations are done during classtime, which is already chock full of material, the students only have ten or so minutes to complete their evaluations. The vast majority of students, feeling that the free time incurred by finishing before the ten minute allotment is more important than the evaluation itself, simply fill in bubbles by rote and go about

¹⁶ Ross Keiser, "Merits of Student Evaluations," *The Chronicle of Higher Education*, March 18, 2005, 47.

their business. The effect is that nearly every teacher on campus has a score around 3.5 on a 5.0 scale. Students who “like” the professor give 5’s while students who “dislike” the professor give 1’s. Without any information about the student giving the answer and without any way of filtering “real” answers from the nonsensical ones the evaluative result is more akin to a large sample version of the rate-my-professor website than a tool which is meant to help my improvement. In fact, the whole process is not meant to actually help me improve in the classroom since I do not receive the evaluations until after the semester is over – by which time it is too late to make any corrections in instruction.

Online Survey Solutions

This brings us to an evaluation of one possible solution: the use of online survey software in order to create customized, individualized yet anonymous answers which take the student’s behavior into account. I found the scantron form evaluations to be unhelpful in actually helping me improve my teaching. First, the questions were vague and thus the answers were unreliable. Why did certain people answer that I was ineffective at communication? This was especially vexing when sixty students might say I was “very effective” but then six said “very ineffective.” More vexing were the three people who answered that classroom instruction was “non-applicable.” These types of responses forces everyone to rely on the averages; which left me cold in terms of improvement and left me wary of putting too much faith in a printout which suddenly appeared at the end of the semester.

To remedy this problem I began investigating different survey measures. After several false starts, including a painfully slow alliance with the institutional research department, I came across an online survey site called surveymonkey.com. The cost was reasonable (a few hundred dollars for an infinite number of surveys and questions) and the college already had an account I was eventually allowed to use. The website was user friendly and allows for a host of different types of questions and answers; including multiple-choice – single and multiple answers – essay boxes, matrix answers, and a host of different formats. Once put up on the survey monkey website it was easy to disseminate the

location to the students who could go to the site at their own volition and take their time in answering the questions.

I wanted the students to answer the questions, and you can impose a “must answer all questions” command to your survey, but I also wanted to maintain anonymity. I did not want a method for being able to track down the students to their answers. Only in an anonymous survey did I feel I could get truthful answers. What I created was a final page the students would print out and bring to me in class. To reach that page a student had to answer all the questions. In this way the student answers were anonymous, I received my answers, and the students received credit for taking the survey. To facilitate student engagement I turned to extra-credit bribery. Students proved eager to answer the forty or so questions in exchange for even a meager extra-credit boost.

The positives of this approach are numerous. The answers are tabulated in real time and I am able to conduct multiple surveys during the semester (I usually conduct four during the semester). It allows me to have accurate data on my progress. I borrowed questions from presidential polls in order to have some well tested questions. If the number of students who say the class is on the “wrong track” markedly increases between surveys one and two, I know I have a problem.

The second advantage is that surveymonkey will disseminate the material in comma delineated form. This allows me to run the material through data mining software and find correlations between answers. In fact, since surveymonkey maintains the ISP address of the answers I can see the connections between certain students and answers. For example, students who admit to not reading the textbook also overwhelming answer that the tests are too hard. White males are more likely to admit to not taking notes white African-American males are more likely to not doing any of the preparatory tests. The number one reason why students admit to texting in class is “I Don’t Know” even when given the choice of saying the class is boring. Students who received a “C” on the test were the most likely to say the grading was unfair and students who admitted to receiving a “D” were more likely to answer that there were too many comments. In one act of discovered connection I found that the same twelve students – again I do not know who they were by name – answered that they had not read the textbook before the test, that they

disliked the essay format for the test desiring a multiple choice format, and they answered the grading was unfair. This data not only told me a tremendous amount about the students, their habits, and their educational desires but also gave me the data with which to discuss the standard scantron evaluation form with my administrators. Suddenly, my class which was listed as “harder than average” by a group of students had a rationale backed by data – it was harder than average because a large group of students did not read the textbook which was important in doing well on the exams. Without this data I would be left shrugging my shoulders to any administrator’s questions about my class.

What I also discovered is a sensitivity to tone among the students. Things I felt that were straightforward were not always perceived to be so. The essay boxes allowed students to vent but also discuss. In one essay box I asked for advice for improvement in future semesters. Much of the advice was dull simple sentences – “make the class easier,” “have less tests” types of answers – others were well written nuggets usually involving some little affair I paid little notice to at the time but left a big impression on a singular student. To be reminded of the significance of a previously insignificant event was important in my classroom behavior and not something I had previously received on the scantron evaluations.

There are of course issues with this approach. You cannot know the students are taking the survey seriously. You have no proof of which student answered the survey and one that might simply photocopy the sign-in sheet from a friend. There are students who simply will not take the surveys despite the advantageous bribe. In the end, there is still no certainty to the scientific validity of the methods. Certainly this could not function as the sole means of evaluation.

Yet, for the limitations of the methods I have found the answers to be rather honest. Since there is no reason to lie there is nothing gained by it. In fact, I try to impress upon the students the need to be honest since they gain nothing by lying and do not help me at all. Compliance is about 85% of my students and given my teaching load and the number of evaluations per semester I can honestly say I am the most heavily evaluated professor at my college. Most importantly, the evaluations create easily analyzed data.

Crittenden's research argued that student values needed to be given a weighted measurement. Thus, Crittenden concluded, by knowing the value students place on certain aspects of teaching you can come to a clearer understanding of the reality of the teacher's abilities in the classroom.¹⁷ This is the closest manner I know of to integrate that methodology. One could even apply a matrix asking about the students' value system. Using the online survey software the students can do the work on their own time. I am able to access the data from any workstation. The questions I ask can be tailored to the semester and even the class. Since every class has a distinct personality this kind of flexibility is important in understanding the classroom dynamic. In short, the online survey methodology gives the most flexibility, the most information, the easiest functionality and works as the best aid in helping me improve as an instructor by tailoring my improvements to the needs of the class.

¹⁷Kathleen Crittenden, "Student Values and Teacher Evaluation," *Sociometry* 36, no. 2 (June 1973): 143