



Coffee Conversations Feature Article

ITLE- Teaching &
Learning Support
Oklahoma State University

Some Personal Reflections on Course Evaluations

Charles I. Abramson

As this week begins the opening of the Student Survey of Instruction site (i.e., course evaluations), I would like to offer some personal reflections on the standard university course evaluation. My reasons for offering my opinion are two-fold. First, I hope that my comments will stimulate discussions on changing our standard evaluation form. Second, I hope that the material I provide can be used by faculty who have been victimized by these evaluations as well as those who believe – as I do – that these standard evaluations are not worth “the paper they are printed on.” It is time for a change.



By way of background, I am a professor of psychology and have been at OSU for 25 years. I have been fortunate enough to have been recognized by several organizations for my contributions to teaching, and I am an ITLE Faculty Teaching Fellow. Over my 25 years, I often have been victimized by these evaluations and have been labeled twice by our student newspaper as one of the top 10 worst professors at OSU – their data, a website that included, until recently, ratings on how attractive a professor is. Compounding this abuse was the fact that several people wrote letters to the newspaper protesting this characterization that were never published. Clearly, not one of their finest moments.

In my view the standard evaluations must be changed to include, at the very least, subject variable questions (i.e., student variable questions). Over my time at OSU, I continually have been amazed that student variable questions such as those suggested below have never been asked. Such questions include:

1. How often do you attend class?
2. How many office hours have you attended?
3. How many hours do you spend studying for this class?
4. How seriously do you take your assignments?
5. Based upon your effort, what grade do you expect in the course?
6. What do you think of course evaluations?
7. How seriously do you take these evaluations?
8. How much effort do you put into the class to succeed?
9. How important do you believe this class is for your future?
10. Are you taking your role as an evaluator seriously?

Without these and similar questions, it is not possible to assess a course evaluation properly and therefore a professor’s teaching ability. Some might say that adding these and other questions will make the survey so long that students will not want to complete the evaluation. My answer – too bad. Do we want a proper evaluation that can be used to improve teaching and student performance or not? I would also add that I am certainly not the only faculty member who understands how shockingly unscientific these evaluations are. There is a real danger that some

constructive comments may be overlooked by faculty who no longer believe in these evaluations and stop reading them.

I would also note that more colleges are asking applicants for their course evaluations. Without student variables questions how can a graduate student apply for university positions successfully if she or he has less than stellar evaluations? Parenthetically, there also is the issue that graduate student instructors often receive higher evaluations than a seasoned professor. How can a student with one or two years of teaching experience “outperform” a seasoned faculty member?

To paraphrase Walter Lippmann, the claim that we can measure teaching ability has no scientific foundation at this time. If we were able to measure teaching ability, there would not be a plethora of various instruments to measure it. We cannot measure teaching ability critically when there is no universally accepted definition of teaching ability; we cannot measure teaching ability properly when the background of the evaluator has been merged with literally hundreds of environmental influences from the time of conception to the moment when a student completes the form. Any faculty member regardless of discipline, must surely realize that combining the score of evaluators with different backgrounds is not scientifically valid and therefore useless – in my view these standard evaluations are not worth the “paper” they are printed on.

I would say further, again paraphrasing Lippmann, that all too often these evaluations are used to hold faculty hostage – as it constitutes a last judgement of the ability of a faculty member to teach. Such evaluations can stamp a permanent sense of inferiority on an otherwise good instructor. Several years ago I had a faculty member tell me: “Charles, let me put it in terms you will understand. I am like a pigeon that pecks a response key a thousand times for a single piece of grain.” Often the result of negative evaluations is that it establishes an adversarial relationship between the faculty member and student(s). I would like to remind faculty that in a memorandum from our Provost dated 11/26/2018 it states: “While student evaluations may be one element a department head takes into consideration in developing an assessment of a faculty member’s teaching skills, it must not be seen by heads or faculty as the single criterion for judging teaching effectiveness.” In my role as an ITLE Faculty Teaching Fellow, I have heard from several faculty where these course evaluations are indeed used as a sole criterion for assessing teaching.

Another useful document from the Provost Office is the Evaluation of Instruction Task Force Report which can be found at <https://academicaffairs.okstate.edu/sites/default/files/jul.pdf>. When I ask faculty members how their course is going, the response I hear most often is some version of “I can only answer that when the evaluations are out.” Based on negative evaluations, good professors can become easily depressed and demoralized – especially when some evaluators seldom study, fail to take advantage of office hours, occasionally read the material, do not ask questions, rarely show up to class, or when in class spend more time on social media than class material. Frankly, I dread these evaluations because they do not reflect my ability to teach and often consist of gossip and innuendo. Perhaps one of the more humorous examples was an evaluator who wrote that I gripped (grabbed) a student. As these evaluations are part of a professor’s permanent record I had to explain that the student meant “griped” not grabbed – presumably this evaluator did not know the difference in spelling. This is but one of many dozens of examples I can cite.

How difficult can it be to add a few questions to the standard form – especially now that it is on-line? Although I am not involved in any process to change these evaluations, it would seem that student variable questions can be added to the standard form by someone in authority simply “waving a hand.” Without data on student variables, the evaluation has no scientifically valid meaning whatsoever.

For those who believe that I am exaggerating the problem of course evaluations, I would encourage you to do a Google search. When I entered “Course Evaluations” 96,3000,000 hits were obtained; when I entered “Problems with Course Evaluations” 68,5000,000 hits were obtained. In regards to the latter, some of the entries that were returned included:

1. Student Evaluations of Teaching are Not Valid | AAUP
<https://www.aaup.org/article/student-evaluations-teaching-are-not-valid>
2. Student evaluations of teaching are an inadequate assessment tool
<https://www.tandfonline.com/doi/full/10.1080/2331186X.2017.1304016>
3. New analysis offers more evidence against student evaluations
<https://www.insidehighered.com/.../new-analysis-offers-more-evidence-against-student...>
4. The trouble with course evaluations | UC Davis
<https://www.ucdavis.edu/news/trouble-course-evaluations/>

For those readers interested in some empirical research on various aspects of course evaluations, I performed a search in the journal *Psychological Reports* from 1955-2010 (Abramson, C. I., Curb, L., & Barber, K. R. (2011). A bibliography of articles of interest to teachers of psychology appearing in *Psychological Reports* 1955-2010. *Psychological Reports*, 108, 182-212. doi: 10.2466/11.PR0.108.1.182-212.)

Citations to some of the articles I found are below. I cite these articles to illustrate how course evaluations can be manipulated easily by, for example, perceived attractiveness, class size, field of study, interest, shared attitudes, student achievement, and ability to tell jokes. An interesting example is article 14 (Neath, 1996) where tips are provided on how to improve your evaluations without actually improving your teaching!

Brant, W. D. (1979). Attitudes toward female professors scale. *Psychological Reports*, **44**, 1310.

Martasian, P. J., & Goldstein, S. B. (1997). Students' beliefs about animal researchers as a function of researchers' sex. *Psychological Reports*, **81**, 803-811.

Murray, R. E., & Walsh, R. P. (1976). Personality differences between levels of psychology majors. *Psychological Reports*, **39**, 215-220.

Prien, E. P., & Lee, R. J. (1965). Analysis of ten criteria of student performance. *Psychological Reports*, **17**, 273-274.

Adamson, G., O'Kane, D., & Shevlin, M. (2005). Students' ratings of teaching effectiveness: A laughing matter? *Psychological Reports*, **96**, 225-226.

Aleamoni, L. M., Yimer, M., & Mahan, J. M. (1972). Teacher folklore and sensitivity of a course evaluation questionnaire. *Psychological Reports*, **31**, 607-614.

Anastasiow, N. J. (1967). Personality traits of teachers nominated as strong and weak. *Psychological Reports*, **20**, 1343-1346.

Battle, J. (1974). Reliability of college students' evaluation of instructors' competence. *Psychological Reports*, **34**, 1086.

Battle, J., & Fabick, S. (1975). Validity of college students' evaluations of instructors' competence. *Psychological Reports*, **37**, 1112-1114.

Chabassol, D., Docherty, D., & Hora, D. (1976). Evaluation of instructors and similarity of attitudes of student and instructor. *Psychological Reports*, **39**, 462.

Landrum, R. E. (1999). Scaling issues in faculty evaluations. *Psychological Reports*, **84**, 178-180.

Levenson, H., & Leunes, A. (1974). Students' evaluation of an instructor: Effects of similarity of

attitudes. *Psychological Reports*, **34**, 1074.

Meredith, G. M. (1980). Impact of lecture size on student-based ratings of instruction. *Psychological Reports*, **46**, 21-22.

Neath, I. (1996). How to improve your teaching evaluations without improving your teaching. *Psychological Reports*, **78**, 1363-1372.

Sachdeva, D. (1976). Ratings by high- and low-achieving students of teaching qualities of college professors. *Psychological Reports*, **38**, 939-942.

I would like to finish by offering the following suggestions and points for discussion.

1. At the very least we should do what it takes to add student variable questions. If this means pestering faculty council and university administrators to help introduce much needed change, so be it. How long is this problem going to continue – for me it has been 25 years!

2. Provide a preface to the standard course evaluation. A preface or preamble describing how important these evaluations are should be included. Moreover, this preface should discuss the student's role in the evaluation process. Evaluating professors is a student's responsibility and should be taken seriously.

3. Sign the evaluation. This suggestion may be controversial. I believe the argument against students signing their evaluation is that we will retaliate against the student. Really? In my case, I have better things to do with my time such as developing innovative teaching exercises, publishing and grant seeking. Moreover, there are rules in place for faculty retaliating against students – why would we engage in such behavior? Are faculty members held in such low regard that students have to be protected from faculty members when they express their opinion? What are we teaching students when they can hide behind anonymity with unsupported comments? Is this the Cowgirl/Cowboy way? I hope not.

4. The course evaluations violate basic principles of learning theory. As a psychologist who is a recognized international expert in the comparative psychology of learning, I consistently am amazed how these evaluations violate the most fundamental learning principles. Let us take but two examples:

a. Delay of reinforcement. In my case I ask students regularly (often after each class) how they liked an individual lecture or class activity and what can be improved. With this method, I can make changes before the next session. With course evaluations, we receive them after the course has ended. How are faculty members expected to improve when the evaluation is after the fact? The answer is we cannot improve – especially when the course is offered once per year or every other year. Anyone who knows even a little about the psychology of learning understands that delay of reinforcement is a major training variable. In general, the shorter the delay the better the conditioning (i.e., learning).

b. Focus on rewarding the behavior associated with good teaching. As I mentioned earlier, in my opinion these evaluations are too often used as a club against faculty members. The evaluations are used as a punishing stimulus or negative reinforcer. It has been known for centuries that if you want to eliminate a behavior do not punish it or use coercion (negative reinforcement). Instead, use positive reinforcement to obtain the behavior you want – “you achieve better results with honey than with vinegar.” In the case of punishment, the faculty member being punished does not learn the correct behavior – this is one of the major drawbacks of using punishment.

5. Freshman Orientation. During freshman orientation, take a few minutes to include a discussion on how to properly conduct a course evaluation and why these evaluations are important. The individuals who conduct the orientation session can include models of what constitutes a good evaluation from a poor one.

6. Encourage students to speak with faculty when deciding to take a class. Students should not be encouraged to use such garbage as *rate my professor*. In fact, I would go a step further and advocate that the university should add a “trigger warning” to such sites and dissuade students from looking at this trash to base their decisions. Students should be encouraged to speak with the professor if they have questions about a potential course. Using an example from my own experience, a student was overheard saying “I would rather have my eyes poked out than have a class with Dr. Abramson.” This student never spoke with me and as a result, missed, if I do say so myself, an excellent course in the history of psychology. As a result of episodes such as this, I wonder how many other students missed out on a possible life changing experience interacting with me.

7. Use ITLE Faculty Teaching Fellows to help evaluate an instructor. I recently did one of these evaluations and was shocked to see students walking in 30 minutes late, “playing” on their phones, and other disruptions – not to mention that less than 50% of the students even bothered to attend the session. I felt sorry for the professor who will be evaluated by these students. From my perspective, this borders on the criminal as a professor’s career may be adversely affected and nicely illustrates why student variable questions must be asked on the standard form.

8. Hold a faculty/student summit on course evaluations. I would be interested in knowing what students and faculty think about course evaluations and what they believe constitutes a proper evaluation. Anecdotally, it appears that some students take their responsibility seriously and others do not. Again, this is a nice illustration of why we need student variable questions. It is not scientifically valid to mix, within the same sample, two such opposing views.

9. Create a depository of the non-standard evaluation forms used across the university. Such a database would be useful for departments wishing to eliminate the standard course evaluation form.

In conclusion, if anyone would like to discuss the issues I have raised here I would be happy to do so. These standard evaluations must be changed. I believe that the standard evaluations hurt faculty, can establish a “me versus them” attitude, and do little if anything to improve teaching. The first step is to include student variable questions and set-up a statistical system to incorporate these questions into the ratings.

