

Journal of Collective Bargaining in the Academy

Volume 0 *National Center Proceedings 2015*

Article 49

April 2015

Student Evaluations of Instruction: Research Evidence and Their Utility

John Centra
Syracuse University

Follow this and additional works at: <http://thekeep.eiu.edu/jcba>



Part of the [Collective Bargaining Commons](#), and the [Higher Education Commons](#)

Recommended Citation

Centra, John (2015) "Student Evaluations of Instruction: Research Evidence and Their Utility," *Journal of Collective Bargaining in the Academy*: Vol. 0 , Article 49.

Available at: <http://thekeep.eiu.edu/jcba/vol0/iss10/49>

This Proceedings Material is brought to you for free and open access by The Keep. It has been accepted for inclusion in Journal of Collective Bargaining in the Academy by an authorized editor of The Keep. For more information, please contact tabruns@eiu.edu.

Student Evaluations of Instruction: Research Evidence and Their Utility

John A. Centra

Professor Emeritus, Syracuse University

Email: jcentra@twcny.rr.com

As one of my areas of research interest over the past 40 plus years I have been conducting studies on student evaluations of instruction at the college level. I began in the late 1960's when I received a foundation grant to study the effectiveness of the evaluations in improving teaching. I was a research psychologist at Educational Testing Service at the time and continued studies at ETS and at Syracuse University dealing with the validity and usefulness of the evaluations. What follows is a summary of the major studies I conducted.

All of the studies were done using an instrument called the Student Instructional Report (SIR), which I developed while at ETS, and the SIR II, an updated and improved version I developed at Syracuse University in the early 1990's. Over 500 colleges have used these instruments and my research was based on the extensive course data collected at these institutions. SIR II is still provided to interested colleges. Briefly, there are 40 items in the instrument grouped in scales that reflect effective teaching, such as Course Organization, Communication, Faculty-Student Interaction, and Student Effort and Involvement. Large samples of courses were included for each study and all of the studies have been published in referred journals. The issues addressed and the findings are:

What are the dimensions of effective teaching that students can judge?

Students, faculty and administrators generally agree on what effective teaching is, and these dimensions should be the basis of any form used. For example, in addition to those mentioned above, the attainment of course outcomes and the fairness of tests and grading practices have been identified as important by all three groups. These are the basis of items in the SIRs.

Can student evaluations improve teaching?

That depends on how much value teachers place on the results and if they know how to make improvements. The evaluations alone don't guarantee anything. This is why I recently put together a compendium of suggestions for teachers. They are tied into the rating items in SIR II and include suggestions from highly rated teachers as well as those suggestions from the vast literature on good instruction.

Is there gender bias in student evaluations?

In a study that included 741 classes in multiple institutions, I did not find any significant or consistent bias. In other words, male and female students did not generally judge men and women teachers differently. This tended to be true regardless of course discipline.

Do students give higher ratings to teachers who give them higher grades than they expected?

Based on a large sample of SIR II results, in general I did not find that students rewarded teachers who gave them higher grades than expected.

Are electronic evaluations as valid as paper-based in class evaluations?

The problem with electronic student evaluations is that the reply rate is often too low. Paper and pencil evaluations collected at the end of the course in class have much higher rates. One way to increase the rate is to deny students a grade (anonymously) until they submit their evaluations.

Are student evaluations correlated with student learning?

This is an important validity question. Students should give higher ratings to teachers from whom they learn more. In short, their ratings should not simply be a measure of their satisfaction or other extraneous variables. Using multiple section courses taught by different teachers who used a common final exam, I found that there was a correlation, especially with an overall rating of the course. Other studies have replicated my finding.

Should student evaluations be used in tenure and promotion decisions, and if so how?

The critical factor here is using student evaluations from at least four or five classes taught by the teacher. The research evidence is that using multiple classes provides a more valid judgment of the teacher's effectiveness because it minimizes the effect of a particular group of students, time of the class, or other extraneous factors reflected in just a single class. Of course information in addition to student evaluations should be included, such as peer and chair judgments, course materials, examples of student learning, and a teaching portfolio (which gives the teacher an opportunity to make a case for his or her effectiveness). Finally, emphasis should be put on the single item that measures the overall effectiveness of instruction since, as mentioned above, research shows that this item correlates higher with student learning than any single item of group of items.

Numbers are important in using the results of student evaluations. For a single class, I suggest at least two-thirds of the class should respond; moreover, the reliability of the average rating is related to the total number responding, with ten to fifteen as a minimum. In order to provide

accurate feedback to the teacher these are important considerations. In addition, the specific items related to instructional practices will give information on where improvements are needed.

In sum, my research and much of that reported by others have shown that student evaluations at the college level are valid and, given enough replies by students, are reliable as well. They can help provide a basis for instructional improvement and are useful in evaluating teaching for tenure and promotion if a pattern of courses are included, rather than just a single course.