

Evaluating the Evaluation: Concerns about Student Evaluation of Teaching

Arlan Parreño

Institute for English Language Education, Assumption University, Bangkok 10240 Thailand
E-mail: arlanparreno@gmail.com

Received 19 March 2019 / Revised 28 April 2019 / Accepted in final form 20 May 2019 / Publish Online 8 June 2019

Abstract

It is believed that students' perception of how they are taught is essential in evaluating teachers for faculty development and personnel decision-making purposes. Thus, student evaluation of teaching, or SET, is a staple in educational institutions, especially in colleges and universities. There are, however, questions about the reliability and fairness of such practice. Many factors are perceived to influence student ratings of their teachers' performance, and grading is a persistent concern. As results of such evaluations are commonly used for administrative decisions, such as for faculty promotions or salary adjustments, teachers are tempted to modify their behavior to obtain favorable ratings. It is therefore suggested that student evaluation of teaching be handled with care in terms of formulation and administration, and be used in conjunction with other methods in order to have a valid and reliable evaluation of teachers' performance.

Keywords: *teacher evaluation, student evaluation of teaching, evaluation of teaching performance, SET*

1. Introduction

Studies about school effectiveness have shown that the most influential factor in the achievement of learners is the teaching quality, well ahead of curriculum, evaluation, and educational management (Hargreaves, 2014). No wonder then that student evaluation of teaching, or SET, is a common feature in tertiary institutions (Biggs & Tang, 2007).

As students are considered primary stakeholders in colleges and universities, their evaluation of how they are taught is seen as an important source of information in directing teacher development and decision-making involving teachers. Ideally, SET should be conducted to improve teaching quality. SET can help teachers understand how students perceive their teaching (Hattie, 2009), and that understanding can help them make adjustments in their teaching accordingly. In reality, however, results of SET are most of the time used for personnel decision-making in terms of salary, tenure, contract renewal, promotions, and awards (Biggs & Tang, 2007).

The evaluation usually comes in the form of a survey using a questionnaire administered either in person or online. Usually, SET instruments are written and formatted in a way that can be used in any department in order to compare teachers using a quantitative scale, for example from 1 to 5, with 1 as highly unsatisfactory and 5 as highly satisfactory. This quantitative approach is common because it offers "administrative convenience" (Biggs & Tang, 2007), especially in institutions with large teacher populations. Individual student responses can be manually keyed in computers or answer sheets can be fed in computerized readers. Quantitative results can be then easily computed using readily available statistical programs.

While SET has been helpful for administrators to make personnel decisions, its formative influence on teaching quality still remains to be seen. In Marsh's study (2007) of 195 teachers in over 13 years, results of SET have shown no significant impact on teacher development. It is rather strange, if not sad, that the opinion of those who are the direct recipients of the act of teaching seems to be considered so lightly, if considered at all. Hattie (2009) noted that SET appears useless for teachers because teachers do not learn anything significant from these supposedly valuable evaluations. This situation begs the question: Why not?

The reason for this seeming lack of teacher development as a result of SET is probably due to contentious issues about the reliability and validity of such an evaluation method. It is fair to say that teachers have the right not to follow what they believe is questionable. What can they learn from an evaluation that they think is not valid and reliable?