

MEASURING TEACHING EFFECTIVENESS

Sak Onkvisit, College of Business, San Jose State University, San Jose CA 95192;
(408) 924-3535, onkvisit_s@cob.sjsu.edu
John J. Shaw, Department of Business Adm., Providence College, Providence, RI 02918;
(401) 865-2652, jshaw@providence.edu

SIGNIFICANCE OF STUDENT RATINGS

Universities routinely utilize student opinions of teaching effectiveness (SOTE) to assess instructors' classroom performance even though most faculty members perceive them to be "problematic assessment instruments" (Simpson and Siguaaw 2000). SOTE can be both beneficial and harmful since it has inherent limitations that stem from a failure to adhere to measurement theories.

CONCEPTUAL AND OPERATIONAL DEFINITIONS

The SOTE instruments vary from university to university, and there are thousands of versions whose degree of comparability or correlation is unknown. All these versions *operationalize*--in different ways--the construct of teaching effectiveness. In theory, an operational definition should be derived from a conceptual definition. In practice, universities usually do not conceptually define and differentiate "teaching effectiveness" from the other similar but distinct constructs. For example, "teaching effectiveness" and "instructor popularity" are two different but overlapping constructs. Research has shown that personality variables account for 50 to 80 percent of the total variance of the student evaluations (Clayson and Haley 1990; Clayson 1999).

RELIABILITY AND VALIDITY

To enhance reliability, measurement conditions need to be standardized. In practice, **administration variation** is the norm. Instructors differ greatly in terms of the amount of time given to the administration of SOTE. Moreover, some (probably not many) instructors use the beginning of a class for this purpose so that students can base their judgments on the teaching performance for the whole semester. But many professors do the teaching evaluation at the end of the class on the day that they have taught well, thus relying on the recency effect. Some will administer the instrument on the day of a relatively full class, while others will choose to do it when many students (presumably less serious students) are absent.

SOTE, as a means to an end, must ultimately show some reasonable degree of **predictive validity**. Unfortunately, it is difficult to determine the proper **criterion variables**. Theoretically, teaching effectiveness and student success should be correlated. But what should be an indicator of student success: course grade, overall GPA, income after graduation, etc.? One recent report has made a headline by demonstrating clear evidence of grade and evaluation inflation at U.S. universities (Rosovsky and Hartley 2002). Also perceptions of grading fairness have a large impact on ratings of teaching ability (Marks 2000).

OTHER ISSUES

Because SOTE purports to measure the different and independent aspects of teaching performance, it is supposed to contain a number of **distinct dimensions or factors**. However, many SOTE instruments have been found to be unidimensional rather than multidimensional.

SOTE, usually a 5-point scale, is more **stable** but less **sensitive** than, say, a 7-point scale. Should stability be obtained at the expense of sensitivity? Also should the scale be **balanced** or symmetrical? Should there be more "agree" choices than "disagree" categories when people have a tendency to agree? Another question is whether the items should be **randomly** arranged. Furthermore, SOTE items are almost always positive statements by describing some positive teaching experiences. Will it matter if some scales are **reversed**?

CONCLUSION

Courts do not allow the results of polygraph tests to be admitted as evidence because of concerns over their reliability and validity. Arguably, polygraph tests may be more reliable than SOTE instruments. Yet the use of SOTE is almost always mandatory, widespread, and unopposed. Universities need to make better use of measurement theories when developing SOTE. They also need to adopt a more comprehensive perspective. Given the impact of SOTE on both faculty and students, a healthy discussion of the theoretical issues and practices is necessary.