

THE EFFECT OF GRADES ON STUDENT RATINGS OF TEACHING: IMPLICATION FOR UNDERSTANDING STUDENT MOTIVATION

Kevin S. Carlson, Ph.D., Centre for Educational Development, Republic Polytechnic, 9 Woodlands Ave 9, Singapore 738964

ABSTRACT

There is widespread belief that teachers who grade more easily receive better student evaluations than those who are tougher graders. The actual research literature on this issue, however, is equivocal. This study examines this issue in an Asian context and reports similar findings as found in the West - - while grades have a moderate influence, this relationship is often explained by a third variable of student self-perceptions of their own learning outcomes from the module. While most of the previous literature on this issue focuses on validity issues, this paper will examine how such data can inform us about student motivation.

KEYWORDS

Student evaluations; Grade inflation; Motivation.

INTRODUCTION

There is a widespread belief that grades affect student's evaluations of teaching – with lenient teachers rated more highly than strict ones. However, the research literature is somewhat split on this issue. Some contend that there is a robust moderate relationship between grades and ratings of teaching and that this relationship is consistent between instructors and within the same class with the same instructor over time (Rice, 1988).

Others argue that while the modest correlation exists it is often related to actual learning, for example Howard and Maxwell (1980, 1982) found that this correlation was explained via its shared relationship with self-reported student learning. Some have interpreted such relationships as evidence that student ratings are indeed valid as a teacher's effectiveness is then seen as influencing both grades and student ratings. This assumption may be dubious, and others have suggested other interpretations including a “feel-good” factor that questions the validity of self-reported learning (Greenwald & Gillmore, 1997) and that student motivation would further attenuate the relationship between student definitions of learning and student ratings (Cashin, 1995). Furthermore, some feel that student definitions of learning may not match well with institutional goals and hence student evaluations of teaching and learning may be quite independent of institutional concerns and values (Winfred, Travis, Paul, & Edens, 2003).

These debates generally focus around issues of validity - - i.e., to what extent do student reports actually index good teaching practices. While this issue is an important one for both individual teachers and institutions who wish to evaluate teaching and promote better practices, one could ask whether these data can inform us about other issues. For example, do these various relationships potentially inform us about student motivation and the student's "fit" with institutional and curricular goals and philosophy?

This study (a) adds to the literature by examining the relationships between grades, self-reported student learning, and student evaluations of teaching, (b) examines these relationships further especially in relation to institutional and curricular philosophies, and (c) preliminarily investigates how such data inform us about student motivation.

METHOD

Sample. The sample consisted of 1,670 classes from the first three academic years of a newly created Polytechnic in Singapore. Each class contains approximately 25 students. All data presented in this paper are aggregate data – with class being the level of analysis, and not individual students.

Institutionally, this educational setting is quite radical in the cultural context – using a full-scale PBL curriculum with students who for the most part are accustomed to transmission-style instruction. This PBL curriculum and its educational goals necessitate that students re-think their personal definitions of learning.

Student Ratings of Teaching. Student feedback was elicited twice a semester – 4 and 14 weeks into a 16-week term. The survey contained separate sets of questions for teacher and for the module itself and included both quantitative (Likert) and qualitative responses. Only quantitative data are analyzed in the report. Response rates to the teaching evaluation survey ranged from 70-80%, which is quite high considering the purely voluntary nature of the survey.

The teacher aspect of the evaluation survey contained 11 questions (e.g., encourages me to participate in class discussions and activities, communicates clearly, approachable, challenges my thinking). Except for two questions that seemed to assess a grey area between teacher and module, all the questions were aggregated to produce an overall rating score. The extremely high internal consistency ($\alpha = .99$) justifies this aggregation.

Grades. Student grades were indexed via a score assessing the relative strictness-lenience of grading. For each module, each student receives 16 daily grades, assessing performance for that day. This index examines all 16 of these grades in relation to all the other daily grades a student receives in their other modules (ranging from 48 – 64 other daily grades obtained in other modules per semester). Teachers with higher positive scores on this grading index are more lenient than other teachers with the same group of students, while more negative scores indicate

relative strictness. Furthermore, students are very aware of their daily grades, as they see each immediately after it is awarded. In addition, at any time in the semester students can access a summary page for each module displaying all daily grades and predicted grades based upon their performance so far in the term. Unlike some modules in other institutions where students are predicting and guessing their final grade, these students see their module grade unfolding each day in a very explicit manner.

Self-Reports of Learning. Student self-reports of learning was index by one simple item on the student evaluation – “I think I can do well in this module.” Some may argue that this is hardly an index of real learning; however, most of the previous measures of this aspect emphasize achievement and such an item potentially matches such achievement evaluations. Others may argue that this merely asks them to re-state their grade (i.e., doing well means getting a good grade); however, this is not the case – as seen below.

RESULTS

The results of the analyses replicate some of the findings of previous researchers. Namely, there is indeed a moderate correlation between grades and ratings of teachers ($r = .29, p < .0001$); however, controlling for feelings of doing well substantially reduces the effect to practically zero ($pr = .12, p < .0001$).

However, the strength of feelings doing well was quite remarkable. This one item was highly related to teaching ratings ($r = .74, p < .0001$), and the strength of this relationship was extremely robust when controlling for grades ($pr = .71, p < .0001$). Such strong relationships suggest that more analyses should be pursued to further understand the complexity underneath this simple item.

DISCUSSION

These results and other analyses will be discussed in relation to the issue of how these data reflect student adaptation to a new educational setting, forcing them to re-thinking their own definitions of learning.

As a follow-up to the data collected from the previous surveys, other data will be collected and analyzed from the current new semester. These future analyses will include data from all incoming students assessing various motivational components (as measured by MSLQ - Pintrich, Smith, Garcia, & McKeachie, 1993). These data will not only be examined in terms how motivation affects student ratings, but to ask whether student ratings reveal their motivations in different modular and curricular settings.

REFERENCES

- Cashin, W. E. (1995). *Student ratings of teaching: The research revisited*. IDEA Paper No. 32. Kansas State University, Manhattan, KS.
- Greenwald, A. G., & Gillmore, G. M. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology, 89*, 743-751.
- Howard, G.S., & Maxwell, S. E. (1980). The correlation between student satisfaction and grades: A case for mistaken causation. *Journal of Educational Psychology, 72*, 810-820.
- Howard, G.S., & Maxwell, S. E. (1982). Do grades contaminate student evaluation of instruction? *Research in Higher Education, 16*, 175-188.
- Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the motivated strategies for learning questionnaire (MSLQ). *Educational & Psychological Measurement, 53*, 801-813.
- Rice, L. (1988). Student evaluation of teaching: Problems and prospects. *Teaching Philosophy, 11*, 329-344.
- Winfred, A., Travis, T., Paul, D. S., & Edens, P. S. (2003). Teaching effectiveness: The relationship between reaction and learning evaluation criteria. *Educational Psychology, 23*, 275-285.