How do Students and Educators Interpret Student Evaluations of Teaching?

Katherine P. Neely & Sarah K. Tauber

Introduction

- Student evaluations of teaching (SETs) are widely used at universities to assess faculty teaching performance.
- Often used to inform important decisions about promotions and course organization.
- Faculty tend to overinterpret small mean differences in ratings when reviewing completed SETs, and students may be careless when completing SETs (Boysen 2015, Uijtdehaage & O’Neal 2015).
- The testing effect shows that taking quizzes is a reliable way to improve learning and long-term memory for material (Schwieren 2017).
- Students and faculty may utilize a professor’s ratings and quizzing habits to draw different conclusions about the professor’s teaching effectiveness.

The present research: We examined whether there are differences in how students and faculty interpret SETs and which factors they use to make their judgments.

Materials & Procedure

- Participants & Design: 118 Texas Christian University undergraduate students and 117 faculty members from American universities participated. We manipulated the professor’s rating on the graph (either higher or lower than departmental averages) and the presence of quizzing between-participants.
- Materials & Procedure:
  1. Participants were shown a graph of a fictional professor’s ratings by students. They also saw five student comments that indicated whether or not the professor gave quizzes.
  2. Participants were asked to judge the professor along several dimensions, including teaching efficacy, helpfulness, clarity, and preparedness.
  3. Participants were asked whether they considered the professor’s rating on the graph, the student comments, and the presence of quizzing when making their judgments of the professor.
  4. A manipulation check asked participants to recall whether the professor’s ratings were higher or lower than average and if quizzes were given.
  5. To conclude, participants were asked questions about their demographics and attitudes on quizzing.

Results

Student responses

- Mean Rating
  - Taught Effectively
  - Mean Rating

Factor influencing participants’ judgments

- Mean Rating
  - Taught Effectively
  - Mean Rating

Factors influencing participants’ judgments

Student comments shown to participants

Conclusions

- Students and faculty groups did not significantly differ in their responses.
- The magnitude of the professor’s ratings had a strong influence on our participants’ judgments.
- Participants more strongly agreed that the highly-rated professor possessed positive qualities related to teaching, including helpfulness and clarity.
- The presence or absence of daily quizzing did not make a significant difference.
- Participants showed an understanding of the testing effect, but did not use quizzing to inform their judgments about the professor’s teaching efficacy.

Future Directions

- Consider qualitative variables in an SET other than quizzing, such as study guides and cumulative finals, to see if they impact participants’ judgments of a professor.
- Test effectiveness of a warning that reminds participants about the importance of carefully reading and considering student comments (Boysen 2015).
- Investigate whether reading about the benefits of quizzing and the testing effect immediately prior to reviewing our SET makes participants more willing to reward the presence of quizzing with higher ratings.

Attitudes about quizzing

- High and Quizzes
- Low and Quizzes
- High and no Quizzes
- Low and no Quizzes

Quizzes Enhance Learning