Response Styles in Student Evaluation of Teaching
By Edgar Andrés Valencia Acuña

Student Evaluation of Teaching (SET) typically refers to the use of summated rating scales to measure teaching quality based on students’ report. SET is widely used in post-secondary education institutions for informing teacher professional development, curriculum revision, personnel decisions, and for institutional accountability. The literature on SET validity is abundant but often a theoretical, the evidence inconclusive, and provides scarce attention to content and response process. Specifically, research examining whether students respond independently of content relying on response styles is rare. Types of response styles are acquiescence/disacquiescence (tendency to agree/disagree across items), extreme (tendency to endorse extreme response options across items), and midpoint response styles (tendency to use the midpoint option across items). Evidence of a substantial degree of response style would reduce the validity of SET scores as a measure of teaching quality and their utility for informing formative and summative decisions due to overestimation or underestimation of the actual level of teaching quality and artificial changes in the relationship to other variables. Three topics examined in the study are the degree to which SET scores are affected by response styles, differences in the extent to which SET scores are affected by response styles across measurement conditions, and the degree to which response styles moderate differences in SET scores between female and male teachers. Responses to a SET summated rating scale from N=5,921 education graduate students were analyzed. Student-level indexes of response styles suggest a high degree of acquiescence in the direction of teaching quality overestimation, and no disacquiescence, extreme, or midpoint response styles. A 2 (academic department) x 2 (program type) x 6 (academic session) ANOVAs on response styles indexes suggests no statistically significant differences across measurement conditions. Finally, multiple linear regression analysis indicates a statistically significant moderator effect of acquiescence on the difference in SET scores between female and male teachers. The discussion addresses implications of the findings for developers and users of SET summated rating scales, alternative interpretations of the observed pattern of responses, limitations, and suggestions for future research.