

# Factors Contributing to Instructor Quality Evaluations

## Objectives & Data Analytic Plan

The purpose of this study is to explore variables contributing to differences in student assessments of overall instructor quality from 2017-2019. Specifically, this study asks, “What instructor, student, and course variables predict overall evaluation of instructors?”

Data consists of 362,238 course evaluations, across four semesters (Fall 2017, Spring 2018, Fall 2018, Spring 2019) analyzing one question about overall instructor quality from the UB student survey: “Overall, this instructor was (Very Poor, Poor, Fair, Good, Excellent, or Not applicable).” We investigated Instructor, Student, and Course variables to determine whether they explained differences of overall instructor quality rating scores using a hierarchical linear regression analysis.

*(...continued)*

### Results

Instructor, Student, and Course variables represented 2.9% of the variability in overall instructor quality scores. The following variables were associated with significantly higher instructor quality scores:

#### Instructor Variables

- Younger, female, White and non-tenure track instructors were scored higher.
- Female students rated female instructors higher than male instructors, whereas male students did not significantly differentiate between male or female instructors.

#### Student

- Older students, female students, students with higher GPAs, and non-White students gave higher overall quality scores.

#### Course

- Instructors got higher scores when they taught lower enrollment courses; the Course Levels of Graduate, Law, and Medical; the Course Types of Labs, Seminars, and Lectures; and the Course Mode of In-Person classes.

### Implications

It is common with large samples for variables to appear statistically significant while explaining relatively little amounts of variance in an outcome. Consistent with this, together, all variables accounted for only 2.9% of the variability in instructor scores.

While students rated faculty differently dependent on age, ethnicity, and gender, this does not sufficiently indicate bias. For example, female students may have rated female faculty higher and male faculty lower because their experiences differed from male students. Further surveys, interviews with students, or observations, are needed to begin to explain *why* there are differences.

Appropriate interpretation of instructor overall quality scores require 2 components:

- **Appropriate Comparisons:** Whether by course size, type, level, delivery mode or past performance to determine whether outcomes are positive, negative, improving, etc.
- **Explanatory Data:** Additional information that explains the *why* of course ratings (i.e., new course, traditionally difficult course, new teaching method introduced).