UNDERSTANDING COURSE EVALUATION DATA

At the end of each semester, students[1] are emailed links to course evaluations for each of their courses via Qualtrics, the College's enterprise survey application. After course evaluations close, the Office of Institutional Research and Effectiveness (OIRE) analyzes the data and summarizes the results in course- and department-level evaluation reports. To ensure respondent anonymity and minimally representative samples, OIRE only analyzes course evaluation data for classes with three or more respondents.

When your course evaluation report arrives, it can be tempting to judge the success of the course from the numerical data alone. However, care must be taken to determine the significance and validity of the data. What follows are some general principles and guidelines to help you get the most out of the numerical data presented on your evaluation report.

RESPONSE RATE

The following class size and response rate standards help ensure valid interpretation of course evaluation data.

Table 1. Minimum acceptable response rate by class size

Class Size (Students)	Required No. of Respondents	Minimum Acceptable Response Rate
5	4	75%
10	8	75%
15	10	67%
20	12	58%
25	13	54%
30	14	48%
35	15	43%
40	16	40%
45	16	35%
50	18	35%
60	19	31%
70	20	28%
80	20	25%
90	21	23%
100	21	21%
300	24	8%
500	25	5%

Source: Nulty, D.D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education, 33*(3), p.310.

Note. Nulty(2008) used and justified an 80% confidence interval for his calculations, and through a vast number of assumptions and corrections for bias, states that classes with fewer than 20 students need a minimum of a 58% response rate to be considered valid. Courses with greater than 50 enrollees can use 35% as their bar. Since instituting online evaluations, Morehouse averages a 39% response rate across all class sizes.

THE MEAN

Developed by the Faculty Welfare Committee, the evaluation questions use a 5-point Likert Scale. Each of the five responses has a numerical value that is used to measure the respondent's attitude: Strongly agree/Excellent = 5; Agree/Very good = 4; Neither agree nor disagree/Good = 3; Disagree/Fair = 2; Strongly disagree/Poor = 1.

The mean reported is an arithmetic mean, which is the sum of all the individual student ratings for a particular question divided by the number of students answering that question. By itself, this number has very little meaning, which is why evaluation reports present the mean along with the distribution of student ratings (i.e., the extent to which students chose "Excellent" or "Poor" for example).

Note: Research has clearly shown that numerical student evaluation data is positively biased, meaning that responses are skewed toward the positive end. Thus, the average rating on a five-point scales tends to fall around 3.4 rather than the expected 3.0.

DEFINING DEPARTMENTS

Course evaluation reports also include a department- and college-level mean for comparison purposes. It is important to note that, in this case, departments are defined by the relevant course subject code and not formal Morehouse departmental units. For instance, ratings for all courses under subject HEGR are summed to calculate the "department mean" for engineering courses, even though these courses occur within the Physics department.

[1] Course enrollment data for Morehouse students are extracted from the College's student information system, Banner. Course enrollment data for Spelman College and Clark Atlanta University students enrolled in Morehouse courses must be collected from the student records offices at the respective institutions.