
An Empirical Evaluation of Teaching Effectiveness

This document is a summary of the results of my teaching evaluations during my time at Gardner-Webb University. Rather than simply presenting the raw data from each semester's evaluations, I have here pooled my teaching evaluations by category and compared them to my overall average for each question. The results suggest that my strengths lie in the promotion of learning, conceptual understanding, and maintaining a culture of mutual respect, while my relative weakness is course organization. I also seem to excel in upper-level courses as opposed to those within the "general education" curriculum.

It should be mentioned that course evaluations based on student surveys are a notoriously flawed instrument. Numerous studies have found that student evaluations discriminate against female or minority instructors (a summary of these studies can be found at [this link](#)). While as a white man I tend to come up on the beneficial end of these biases, I still find it difficult to productively compare the results of my teaching evaluations to that of others. Instead, I present here a breakdown of my teaching evaluations as measured against my own baseline.

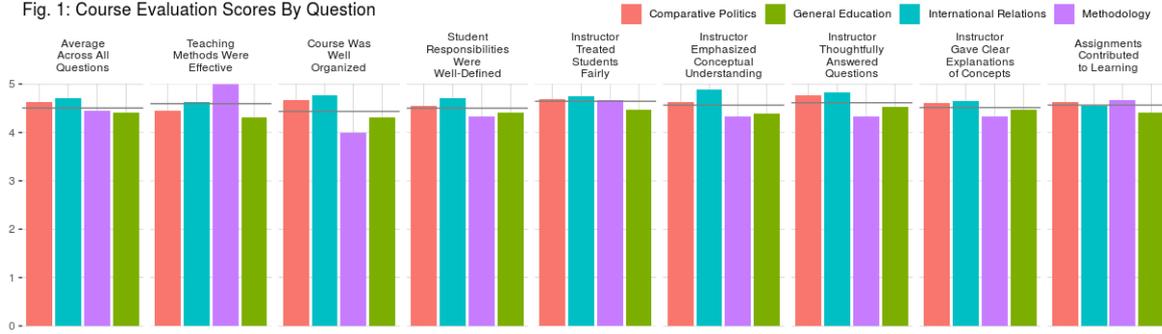
The course evaluation system at Gardner-Webb is conducted online. Students are prompted to submit their evaluations at the end of each semester, but the optional nature of this system presents unavoidable selection bias: As with any online rating system, the surveys are completed by those with the highest motivation to complete them, for good or for ill. Respondents answer two sets of questions on instructor performance. The initial battery of 12 questions is measured on a five-point Likert-type ordinal scale: Respondents are given a prompt in the form of a declarative statement, and students select a number from 1-5 where higher values indicate increased agreement with the prompt. The results from this initial battery are presented in Figure 1.¹ Since I have taught a wide array of courses during my time at Gardner-Webb, I have pooled my results by course type and level: International Relations, Comparative Politics, Methodology, and General Education, and here present my scores averaged across courses of each type.² The horizontal grey lines across each cluster represents the average score for each question across all classes. A full list of courses taught can be found at the bottom of this document. A second smaller battery of tests measures agreement from 1-4, but the process is nearly identical.³ I include the results of this second survey in Figure 2.

1. I exclude four of these questions to avoid clutter and because they are either redundant to existing questions or they do not capture the effectiveness of instruction. Raw teaching evaluation scores are presented in a second file.

2. I realize that using arithmetic means is an inappropriate measure of central tendency when dealing with ordinal data, but the system we are given prevents a more nuanced approach

3. It is unclear to me why this second set of questions is measured separately, and even more unclear why it operates on a different scale.

Fig. 1: Course Evaluation Scores By Question



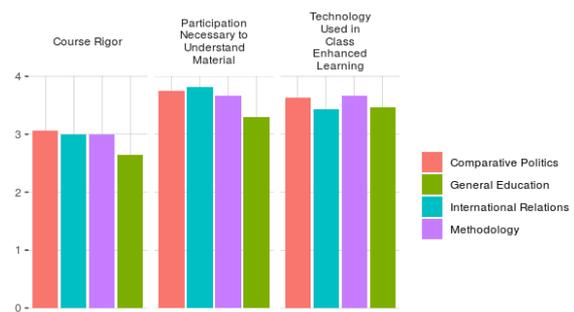
The results of the initial battery of questions are presented above. Across all classes I score highest on the "Instructor Treated Students Fairly" (4.65/5) and "Instructor Thoughtfully Answered Questions" (4.62/5) prompts, which I take as a point of pride. Cultivating an environment of honesty, trust, and mutual respect is a key tenet of my teaching philosophy. In my experience, students tend to live up to the expectations put upon them by instructors. If we expect them to be irresponsible children (and treat them as such), we should not be surprised with what we get in return. Accordingly, I try to always respect my students and the journey they took to land in my classroom. As young professionals, they will meet (and often exceed) that standard when given the opportunity. The success of this approach is borne out in my next-highest average score, for the "Teaching Methods Were Effective" prompt (4.6/5).

My lowest average scores come on the "Course Was Well-Organized" prompt. While my casual and improvisational approach surely contributes to this in some ways, there is evidence to suggest that disorganization is concentrated in courses in which I have the least experience. The "Methodology" section is based on a single course, the first time I had prepped and taught that particular section. Likewise the "General Education" courses score lower than average on this measure, likely because I arrived at Gardner-Webb having never taught general education courses before. Over time I seem to improve greatly on this point, as multiple iterations of any given course help me to tighten up its organization.

The second battery of questions is perhaps somewhat less informative, but nonetheless valuable. My courses score very highly in regards to the importance of class participation and the use of technology in the classroom, two points of emphasis in my teaching. My courses generally score a 3 out of 4 in regards to rigor (where 4 is most rigorous), which I take as a success: Students feel challenged, but not overwhelmed. Course rigor for my introductory "General Education" courses is slightly lower than 3, which I again see as a positive result. General education courses are meant to be a layman's overview, giving students to basics of the discipline while building their interest.

So as to satisfy technical requirements, I include raw course evaluation reports at the end of this document. However, I hope this discussion was a useful and productive look at the quality of my teaching during my time at Gardner-Webb.

Fig 2: Supplemental Course Evaluation Questions



Courses Taught

- **International Relations**
 - American Foreign Policy
 - Human Rights
 - International Conflict and Security
 - International Organizations
 - International Political Economy
 - International Relations
 - Terrorism
- **Comparative Politics**
 - Asian Politics
 - Civil Wars
 - Comparative Political Economy
 - Comparative Politics
 - Latin American Politics
 - Political Violence
 - Politics of the Developing World
- **Methodology**
 - Research Design
- **General Education**
 - Global Understanding