

## Make a quick research proposal

### The Problem

How to reduce cheating in a course?

Assume that students take a multiple-choice exam every 2 weeks as part of their assessment for a course that runs for 4 months duration. All students are lectured in one room but they take exams separately in two sections. Each exam contains 10 questions and it is worth 20 points, 5 exams will make up 100 marks, so only 5 tests are considered. The teacher selects 5 exams to consider for all students.

Recently university is concerned about cheating students who score high marks and get high grades. The university is following a relative grading policy in which a cheating student competes with high scoring students and medium students who don't cheat get buried down to the lower grades. There are many forms of cheating like copying from a mobile device or small piece of paper, or from an inconspicuous resource. Students prefer to get answers from colleagues than ploughing through notes or pdf documents because the later is easier and it is more likely to get correct answer.

To circumvent cheating and make the assessment ground fair, the teacher decided to makes 3 versions of an exam so that students cannot share answers easily. In each version she made three different exams and stacked the exams together first all version 1 exams then version 2 and then version 3. There are two sections who are taking the course. On the first assessment the students who took version2 scored way above the other students on average on both sections. How to make the 3 versions fair?

### The Potential Solution

Using the same set of questions and handing the exam papers randomly will make the average score of students in all versions similar. I shall try to make the two changes and see if scores will even up

### The Method of Testing the Solution

Since there are 3 exams that will eventually be discarded, the teacher has a chance to experiment on the students. The two sections may contain different kinds of students like one of them may constitute many high achievers. To prevent such a selection bias, she performs AA testing based on the first 2 assessments. Here we generally assume that students perform in a similar way from one assessment to the other on average. The average scores on version1, version2 and version3 shall be similar for both sections. It is not clear how large a difference should be in order to be acceptable, the teacher may use her own experience to compare the distributions.

On the third assessment, applying the proposed solution, make the same questions, shuffle them and stack them randomly, then hand the exam papers. Section one is treatment group and students in section two will be given exams arranged in the old way (control group). I can repeat the experiment in the fourth and fifth assessment.

In the control group one of the versions average score is high by a considerable amount. For example, if version 2 is the highest, did the average score of version 2 decreased to match average scores of version 1 and 2? Repeating the experiment on the fourth and fifth assessment and checking the result is helpful to make up for the inconsistencies in students' performance on different weeks.