Teaching Experiments for a Course in Introductory Statistics

Josefina V. Almeda
University of the Philippines Diliman
jcvalmeda@yahoo.com

College students' appreciation of the statistical science relies, to a large extent, on how the introductory course is managed. Two groups of students (undergraduate statistics majors and nonstatistics majors) were exposed to teaching an introductory course. Within each group, half is exposed to fun games intended as enrichment activities, while the other half served as the control. Grades after one semester were analyzed and treatment effect is computed through Heckman's Selection Model. While the treatment (games) is beneficial for nonstatistics majors, it is disadvantageous for statistics majors. For students with inherent interest in statistics, the introductory course will only require a clear presentation of concepts that will help them appreciate the discipline. However, the nonstatistics majors or those with negative perception on statistics, enrichment activities can help conceal their dislike for statistical science and help improve the eventual outcomes in the course.

Keywords: fun games, Heckmans' Selection Model, treatment effect

1. Introduction

Appreciation of people in the workplace or society in general on the statistical science can be traced back on how they were introduced to the discipline. Those who were not been convinced that statistics is important will remain to have the same perceptions even after they finished university work. It is a big challenge for teachers of introductory statistics course on how to introduce the course to attract the students' attention that will hopefully be translated into cognition and subsequently appreciation of the discipline.

Students may develop interest in any academic discipline depending on how pleasant their experience was the first time they were exposed to that area. It is postulated that Statistics can be better introduced to college students in a more casual

and fun setting like in an 'ordinary game.' Thus, an experiment was done where one group recieves—aside from the usual instruction—intervention (enrichment activities), while the other group has only plain classroom instruction. This paper aims to compare the performance of students with and without interventions, and to identify possible factors contributing to the outcomes in introductory statistics course.

2. Methods

An experiment is conducted to enhance student learning. Two sections of introductory statistics course for statistics majors and another two sections of introductory statistics course for nonstatistics majors from the University of the Philippines Diliman were identified to participate in the study. Each group were subgrouped further into control where instructions for one semester was undertaken using the usual classroom instructions. The other subgroup received the intervention where in addition to classroom lectures, enrichment activities (e.g., enrichment activities – quotations, games, jokes and quiz bees) are done before, during, and after the presentation of a concept or a method. The purpose of such activities is to make the topics fun and more interesting and for students' easier internalization and appreciation of the subject matter.

Activities given before the new lesson are intended to elicit the students' interest and to have them desire to crave for more on the topic to be discussed. On the other hand, the activities given after learning the lessons are intended to wrap up, to emphasize important concepts and to unwind the class session. All these activities are ways to make the classroom experience more interesting and pleasurable for the students. According to Leblanc (1998), good teaching is also about style. It may be in forms that are entertaining but not necessarily lacking substance.

A quotation may serve as stimulants of students' interest prior to the discussion of a new lesson. A quotation relevant to the new lesson/topic initially sets the mood of the class. The teacher does not only give the quotation but explains its implications as well.

- As an introduction to the definition and uses of statistics, the teacher can use the quotation by Florence Nightingale: "Statistics is the most important science in the whole world; for upon it depends the practical application of every science and of every art; the one science essential to all political and social administration, all education, all organization based on experience, for it only gives results of our experience." After explaining the quotation, the teacher defines what statistics is and follows it up with its varied applications and uses. This allows the students to realize the magnitude of its usability not only in their field but also in other fields.
- On data presentation using statistical charts, a good quotation is from an ancient Chinese proverb that states "One picture is worth ten thousand words." This

explains that once we put data in a chart, we obtain many significant findings. Hence, it is comparable to a painting wherein we can make different interpretations. On the other hand, if the teacher does not want to sound very philosophical, he/she can cite the first line of the song *If* by Bread that says "If a picture paints a thousand words…," the meaning of which is analogous to what was discussed.

- For the discussion of the different summary statistics like measures of central tendency and measures of dispersion, another quotation worth mentioning is by H.G. Wells (1929) that says "The time may not be very remote when it will be understood that for complete initiation as an efficient citizen of one of the new great complex worldwide states that are now developing. It is as necessary to be able to compute, to think in averages, maxima and minima, as it is now to be able to read and to write." This may sound profound however, we let the students realize that being statistics literate is analogous to knowing how to read and write. Every individual equipped with statistical tools has an edge in life. Many data are available but not enough researchers to do the analysis. Only a few has the statistical capability. The teacher emphasizes the fact that there is power in understanding figures.

Games, on the other hand, are done in the form of entertainment. This activity enhances thinking, cooperation, and unity among students. When the teacher incorporates games in class, this creates excitement and enthusiasm on the subject matter. Thru games, students will have a better understanding of the new concepts they need to learn.

- On data organization, the game "Arrange Yourselves" is appropriate. The game can be played in two ways – by small groups, and by whole class participation. In playing with small groups, there should be at least 10 members in each group. Students arrange themselves using quantitative variables like age, weight, height, foot size, and waistline. Instructions should be specific whether they will arrange themselves from lowest to highest or from highest to lowest. For the last call, the whole class participates. The entire class needs to arrange themselves according to the variables asked. After the game, the teacher processes the activity by stating that the objectives are 1) to show different ways of arranging individuals and 2) to know the advantages and disadvantages of arranging individuals. Then, the teacher elicits insights from the students regarding what they learned from the game.

  After processing the activity, the teacher discusses the raw data and the array and relates it to the game. Emphasis is on the convenience of obtaining the lowest and highest observations and where the observations are concentrated in the array. The teacher recalls the last game played where there is no more grouping and all the students arranged themselves according to magnitude. Students cite difficulties encountered in the last game and stress is placed on sorting being cumbersome when the number of observations is large. This is

now a good opportunity to introduce the frequency distribution as a better way of organizing data and its benefits.

- Another activity that can serve as an appetizer is the "Find the Word" game. This is helpful in introducing the different terms that students should be acquainted with. The topic "Inferential Statistics," in particular, involve the use of so many terms students are unaware of. The game will allow them to be conscious of the different terms utilized in inferential statistics. The students need not understand the meaning/s of the words yet but rather to give them exposure.

To end a topic in statistics, the teacher gives integration activities (e.g., jokes, quiz crossword, trivia, and quiz bee) to summarize the lesson, to lighten up the class session, and to highlight essential points. Jokes reduces the seriousness of the atmosphere and creates a relaxed mood in the classroom. Kher et al. (1999) say that humor is often overlooked as a teaching tool and that laughter releases stress and tension for both the instructor and the students. The joke should be related to the topic to be discussed for better understanding and appreciation on the part of the students. Adding humor lightens a rather difficult subject matter.

A quiz bee is given as the final activity before the semester ends. The contest serves as a review of all lessons covered and as a preparation for the final examination. As an incentive, the teacher gives prizes to the top three students who will win the contest. This activity motivates and challenges the students to study harder, do further readings, and be competitive. In addition, this creates an atmosphere of fun and excitement.

3. The Data

There are 58 statistics majors and 55 nonstatistics majors involved in the experiment. Same set of variables for both groups were collected at the same time:

- examination scores (long exams and final exam),
- grade, sex, age, height in inches, weight,
- daily physical activity,
- daily stress level,
- average weekly allowance,
- average weekly expense,
- membership in an organization, location of residence (urban, rural),
- status of current residence (1 – lives with parents, 2 – lives with relatives, 3 – lives with friends, 4 – lives in a dormitory, 5 – lives alone, 6 – lives in a boarding house, 7 – lives with family members)
- family status (1 – both parents living together, 2 – separated, both parents alive, 3 – separated, but one parent deceased, 4 – both parent deceased, 5 – one parent deceased)
- highest educational attainment of father
- highest educational attainment of mother
- number of siblings

- course and number of semesters in the University (for nonstatistics majors),
- number of semesters in the current program (for nonstatistics majors),
- status of current program
- grade in algebra, geometry, trigonometry, college algebra
- with honors in high school
- type of high school (1 – private, 0 – public)
- study habits (1 – studies everyday, 2 – studies 4 to 6 times a week, 3 – studies 2 to 3 times a week, 4 – studies once a week, 5 – studies only when there is an exam, 6 – does not study)
- attitude with numerical courses (1 – very interested, 2 – interested, 3 – neither, 4 – uninterested, 5 – very uninterested)
- number of times taken Stat 101 (for Stat 101 students only),
- and overall impression of Stat 114 or Stat 101 (1 – very important, 2 – important, 3 – neither, 4 – unimportant, 5 – very unimportant).

4. Estimation of Treatment Effect

The experimental units are students from two sections of statistics majors (divided into control and experimental groups) and two sections of nonstatistics majors (also divided into control and experimental groups). The control and treatment groups were given the same examinations (multiple choice type) for all topics covered.

The author taught all four sections included in the study. However, self-selection bias is still inevitable in this experiment since students are free to choose their teacher by enlisting to their preferred section. Heckman (1979) noted that sample selection bias may arise in practice for two reasons. First, there may be self selection by the individuals or data units being investigated. Second, sample selection decisions by analysts or data processors operate in much the same fashion as self selection. When samples are not randomly assigned to their respective groups, biased estimates of treatment effects can lead to misleading information.

We used the Heckman's selection model, a two-step statistical approach, which offers a means of correcting for non-randomly selected samples. The model avoids the sample selection problem by estimating the model parameters by maximum likelihood. The model provides consistent, asymptotically efficient estimates for all parameters in the model.

Heckman's selection model is based on two latent dependent models:

$$Y_1^* = \beta' X + U_1 \qquad (1) \text{ (regression model)}$$

$$Y_2^* = \lambda' Z + U_2 \qquad (2) \text{ (selection model)}$$

where X and Z are vectors of regressors, the errors U_1 and U_2 are conditional on X and Z, jointly bivariate normally distributed with zero mean vector and variance

matrix $\sum$ where the following holds $U_1 \sim N(0, \sigma^2)$, $U_2 \sim N(0,1)$ and $Corr(U_1, U_2) = \rho$.

5. Results and Discussion

The final grade of nonstatistics majors is significantly affected by their grades in college algebra ($p<0.000$), grades in high school trigonometry ($p<0.013$), whether they agree that numerical courses are important ($p<0.093$), and the treatment ($p<0.000$). The partial contribution of the treatment on final exam could increase the grade by 0.856 in the 9-point grading system (1 is highest, 3 is passing, 5 is failing).

Selection to participation in the intervention is significantly affected by the number of semesters they stayed in the university so far ($p<0.028$). Those in their junior or senior years do not need the intervention much as this is needed by the freshman and those in their sophomore years.

Counterfactual simulations show that had all the subjects went through the same intervention, the final grade in Stat 101 is 1.94 compared to 1.98 had they not went through the intervention. This means an estimated treatment effect (due to the intervention) is about 0.0351 (or and increase of 1.78%). This is statistically significant ($p<0.000$).

Final grade of statistics majors is significantly affected by grade in high school algebra ($p<0.034$) and the treatment ($p<0.001$). Grade in high school algebra is a good indicator of performance in introductory statistics for statistics majors taken upon entrance in the program. The partial contribution of the treatment on final grade however, can lower the grade by 0.5449 in the 9-point grading system. While the intervention is an advantageous activity for nonstatistics majors, it has adverse effect on statistics majors.

None of the identified determinants of selection as participants to the intervention can actually dichotomize who needed and who do not need the intervention among statistics majors.

Counterfactual simulations show that had all the subjects went through the same intervention, the final grade in introductory statistics is 1.85 compared to 1.81 had they not went through the intervention. The statistics majors are better off in an introductory statistics course not to participate in the intervention ($p<0.000$). It is recommended that similar interventions can be developed that are more suited to the cognitive level of statistics majors, or simply not to use the current activities as this is detrimental to student performance.

6. Conclusion

Students understand what they are learning if classroom lessons are presented/ discussed in a more fun and interesting ways. Although time consuming, various techniques (e.g., enrichment activities – quotations, games, jokes and quiz bees) may be used to get the students' enthusiam in appreciation of the subject matter.

The objectives of this study are twofold. First, the study compares the performance of students with and without interventions. Second, the study determines factors affecting the grade of students in introductory statistics. The study suggests that statistics majors do not need these enrichment activities to get their interest in studying statistics. However, non-statistics majors may need to have interventions to appreciate statistics, improve cognition, and subsequently better course outcomes.

REFERENCES

Heckman, J. (1979), Sample Selection Bias as a Specification Error, *Econometrica* 47, 153–61.

Kher, N., Molstad S., and Donahue R. (1999), Using Humor in the College Classroom to Enhance Teaching Effectiveness in 'Dread Courses, *College Student Journal.*

Leblanc, R. (1998), Good Teaching: The Top 10 Requirements, *The Teaching Professor.*