ASSESSMENT QUESTION TYPE IN A STATISTICS COURSE FOR NON MAJORS: ANALYSIS OF STUDENTS' PREFERENCE

Dorcas Modupe Okewole and Olusanya Elisa Olubusoye
Department of Mathematical Sciences, Redeemer's University, Ede, Nigeria
Department of Statistics, University of Ibadan, Nigeria
okewoled@run.edu.ng

This study involves the analysis of students' preference on the Assessment Question Type (AQT) in an applied statistics course for non majors. Developing interest in a statistics course may be greatly influenced by performance in statistics courses taken earlier or expectations on the one being considered. It is important that assessment of the students' performance in the course be through a valid evaluation tool which includes appropriate question type. Students' AQT preference information was obtained using an instrument designed for the purpose and was shown to vary mostly between multiple choice and essay questions. Proportion of scores from multiple choice questions in the final examination was modeled with beta regression. Students' level of understanding of the course (as judged by the total score in the examination) was shown to influence proportion of marks from the two AQT employed. Implications of the results were discussed.

INTRODUCTION

Evaluation of the depth of understanding of any concept or phenomenon for which training has been received is done through tests and examinations. Studies on improving students' academic performance and interest in a field such as statistics abound in literature. Research has shown that teacher engagement (Zimmerman et al, 2014) and student motivation (Tuckman, 1996 and 1998; Covington, 2000; Stiggins, 2001, Budé et al, 2007) influence performance in examinations. Furthermore, much interest has been indicated in examination format as it affects scores for example, Myers and Myers (2006); Davidson and Dustova (2017). Myers and Myers (2006) discussed the effects of two exam formats on course achievement and evaluation.

Specifically on Statistics Examination, Dunham et al (2015) suggests that "how students are assessed on a statistics course will be integral to how, and how much, they learn". In their work, Dunham et al made a proposal on how to make use of Bloom's taxonomy to measure the complexity level of an assessment tool in undergraduate statistics course at the introductory level. The examinations considered in Dunham et al were said to have increased in difficulty by percentage of multiple choice questions, a result that requires more consideration. Although it could be said that students can score higher in multiple choice questions than essay or long answer type by simply guessing, in the authors' opinion arising from experience, "students rarely appeared to guess, since they enter the examination with at least partial knowledge, of the topics assessed". The attentions given to assessment tools emphasize their importance in statistical education. For instance, the National Science Foundation (NSF) funded an Assessment Resource Tools for Improving Statistical Thinking (ARTIST) project to address some of the challenges in this regard (Delmas et al, 2007). Moreover, Ziegler and Garfield (2018) also proposed a statistical literacy assessment tool referred to as Basic Literacy In Statistics (BLIS) and provided a list of previously existing assessment tools.

This paper is particularly concerned with an aspect of assessment: the type or format of questions for assessing students on a statistics course taken by Non-majors. Assessment Question Type (AQT) as used here consists of Multiple Choice (MC), Essay type (also known as theory or long answer questions), short answer questions and combinations of different types. The objectives of the research were to analyze the students' preference on AQT for a statistics course and compare their performance in the two types of questions (MC and essay) given in the examination for the course. This is to encourage the development of an assessment question type that is both student and quality focused.

METHOD

The course is statistics for biological and related sciences (with course code: STA 201) being offered by microbiology, biochemistry, anatomy, physiology and physiotherapy students at the second year of their undergraduate studies. It covers descriptive statistics, introduction to probability and

probability distributions, estimation and test of hypothesis, simple regression and correlation analyses, as well as analysis of variance. A short questionnaire was made available to the students online. The major items that captured the objectives of the research were the questions on the students' preference of AQT and whether or not their choice was particular to statistics courses. The AQT options presented in the questionnaire were: multiple choice, short answer, essay (also known as theory or long answer questions) and a combination of at least two types. There were 75 responses out of the 158 students that registered for the course. The low response rate was as a result of the challenges some of the students had in accessing the questionnaire and the fact that it was made optional to them. Nevertheless, the 75 responses were representative enough as they covered all the Departments of students taking the course. Summary statistics as well as other descriptive measures were provided.

The final examination questions for the course comprised of a combination of two AQT: multiple choice questions and essay questions. There were 20 multiple choice questions each with 1.5 marks, to give a total of 30 marks for that section labeled as section A. Section B comprised of 3 essay questions each with a total of 15 marks out of which the students were asked to choose any 2. The total available marks for section B was thus 30 marks also. The relationship between students' scores in each section was analyzed using a beta-regression model because the dependent variable (Y) is a proportion and not raw scores. This is shown as follows;

$$Y_i = \alpha_0 + \alpha_1 X_i + \varepsilon \tag{1}$$

Where Y is the proportion of a student's score in multiple choice questions out of the total score and X is the total examination score.

$$E(Y) = \frac{p}{p+q} = \mu,$$

$p =$ scores in multiple choice questions (section A)

$q =$ scores in the essay questions (section B)

Let $p+q = \phi$ such that $p = \mu\phi$ and $q = (1-\mu)\phi$. As stated in Ferrari and Cribari-Neto (2004), Y follows a beta distribution with density function:

$$f(y;\mu,\phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1} (1-y)^{(1-\mu)\phi-1}, 0 < y < 1, 0 < \mu < 1, \phi > 0 \tag{2}$$

Examining the variation of the multiple choice proportion across departments is another important interest here but we could not include student's department in the regression model because of the strong multicollinearity it had with total score. This objective was therefore taken care of with a test of independence. The beta regression was carried out using R.

RESULTS AND DISCUSSION
Analysis of the data from the Questionnaire

Summaries of responses by department and AQT preference in STA 201 are given in tables 1 and 2.

Table 1. Distribution of Respondents by Department (Major Course)

	Frequency	Percent
Anatomy	3	4.0
Biochemistry	4	5.3
Microbiology	14	18.7
Physiology	7	9.3
Physiotherapy	47	62.7
Total	75	100.0

Most (62.7%) of the respondents were from Physiotherapy department, while the remaining 37.3% were from the other four departments. The numbers of respondents from the other departments

were so few that a chi-square test of independence could not be carried out on it. The examination result at the end of the course showed that physiotherapy students were distinctly more outstanding than the others which necessitate a comparison of AQT preference between physiotherapy students and others.

Table 2. AQT Preference in STA 201

	Frequency	Percent
Multiple Choice	27	36.0
Short Answer	4	5.3
Essay (Theory)	26	34.7
A combination of at least 2 types	18	24.0
Total	75	100.0

Both multiple choice and essay questions were preferred by almost the same number of the respondents, while the least preferred was the short answer type. In the questionnaire, the students were given opportunity to state the reasons for their choice of preferred AQT. Various reasons were given for the preferences. Most of those who choose multiple choice questions stated less time consumption as the reason for their choice while some stated that they had no reason for their choice. On the other hand, students that went for the option of essay question type stated that they needed to express themselves clearly or that it is easier to score higher marks with this option. An interesting example is the case of one of the students while solving a question on ANOVA, stated the "between subjects sum of squares" as follows:

$$\text{TrSS} = \frac{\text{Total sum of squares of each category}}{n} - \text{Co. fa}$$

The expression might not be perfect but it actually captured the fact that the student participated actively in class and understood the concept. The student was able to interpret the formula, apply appropriately and got full marks. The total marks from the essay question for this student was 24 while that of multiple choice questions was 21.

Test of independence was carried out between AQT preference in STA 201 and:
- Students major course (physiotherapy versus others)
- AQT preference in other courses
- Previous experience in statistics

The short answer option was removed from the AQT preference for STA 201 to enable a valid chi-square test of independence, since there were only 4 (5.3%) responses in that category.

Table 3. Cross-tabulation Between Students' Major Course and AQT preference in STA 201

		Course Physiotherapy	Others	Total
AQT preference in STA201	Multiple Choice	13	14	27
		28.3%	56.0%	38.0%
	Essay (Theory)	20	6	26
		43.5%	24.0%	36.6%
	A combination of at least 2 types	13	5	18
		28.3%	20.0%	25.4%
Total		46	25	71
		100.0%	100.0%	100.0%

The test of independence indicated a weak significance with p = 0.068. Information given in Table 3 shows that most (43.5%) of the respondents from physiotherapy preferred essay questions while

(56%) of respondents from other departments preferred multiple choice questions. The results suggest that students with good understanding of the course preferred essay questions for better expression of their level of understanding.

Most of the students (Table 4) in the different AQT preference for STA 201 stated that their AQT preference would be different if it were another course outside statistics. This suggests that the students do not expect assessment in statistics courses to be handled just like any other course.

Table 4. Cross-tabulation between AQT preference in STA 201 and that of other non-statistics courses

		Multiple Choice	Essay (Theory)	A combination of at least 2 types	Total
AQT preference in other courses	Yes	16	15	9	44
		59.3%	57.7%	50.0%	58.7%
	No	11	11	9	31
		40.7%	42.3%	50.0%	41.3%
Total		27	26	18	75
		100.0%	100.0%	100.0%	100.0%

The test of association between AQT preferences in STA 201and previous experience in statistics indicated significance at P=0.041 (Table 5). This suggests that students' experiences in a statistics course can influence their attitudes towards other subsequent statistics courses.

Table 5. Test of independence between AQT preference in STA 201 and previous experience in statistics

	Value	Df	Asymptotic Significance (2-sided)
Pearson Chi-Square	6.389[a]	2	.041
Likelihood Ratio	6.655	2	.036
N of Valid Cases	71		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 7.35.

Analysis of the results from the final examination

The result of the beta regression on the proportion of marks from the multiple choice questions indicated that the higher the total scores, the more proportion of marks from the essay questions. In other words, lower marks had more proportion of their scores coming from the multiple choice questions rather than essay questions. It suggests that the more a student understands the statistics course, the higher the willingness and ability for detail expression through essay questions, while those still struggling with the course will probably get some marks mostly from few of the multiple choice questions that might require only little efforts. Students with higher level of understanding of the statistics course will obviously still have higher marks in the multiple choice questions than those still finding the course difficult.

Table 6. Coefficients of Beta regression on Proportion of multiple choice questions (mean model with logit link)

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	1.037778	0.093454	11.105	<2e-16 ***
Exam_total	-0.023461	0.002856	-8.216	<2e-16 ***

As earlier mentioned, the students' AQT preference of most physiotherapy students was the essay type. These were the same set of students that scored high marks in the course, mostly > 60%. Those that preferred multiple choice questions were the ones with more proportion of their total exam scores coming from multiple choice questions. A test of independence between student's major course (categorized as physiotherapy and others) and proportion of multiple choice marks (categorized as 0.10 - 0.50, 0.51 – 0.60, 0.61 -0.99) is given in Table 7, with cross-tabulation in Table 8. There is strong evidence (P < 0.0001) that proportion of marks from the multiple choice questions depends on the student's major course. Combining the two AQT (multiple choice and essay) might therefore encourage the students towards good performance while still ensuring quality assessment of the course.

Table 7. Test of independence between student's major course and proportion of scores in MC questions

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	27.551[a]	2	.000
Likelihood Ratio	29.768	2	.000
N of Valid Cases	134		

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 17.01.

Table 8. Cross-tabulation of student's major course and proportion of scores in MC Questions

			Proportion of scores in MC questions			
			0.10 - 0.50	0.51 - 0.60	0.61 - 0.99	Total
Student's major course	Physiotherapy	Count	27	23	7	57
		Percentage	47.4%	40.4%	12.3%	100.0%
	Others	Count	13	22	42	77
		Percentage	16.9%	28.6%	54.5%	100.0%
	Total	Count	40	45	49	134
		Percentage	29.9%	33.6%	36.6%	100.0%

CONCLUSION

Two major assessment question types (essay questions and multiple choice questions) were studied regarding students' preference as well as scores in an applied statistics course. AQT preference and proportion in exam both varied across the major course of the students. Students' level of understanding of the course (as judged by the total score in the examination) was shown to influence proportion of marks from the two AQT employed and indirectly their AQT preference. This study suggests that assessment of students in a statistics course should be made with considerations for students' level of understanding, using a combination of at least multiple choice and essay questions. The results of the study are important for noting in improving both quality and fair assessment in similar

statistics courses. Further studies might analytically examine the relationship between students' AQT preference and examination score across various statistics courses.

REFERENCES

Covington, M. V. (2000). Goal theory, motivation, and school achievement: An integrative review. *Annual Review of Psychology*, *51*, 171-200.

Davidson, C. B. & Dustova, G. (2017). A quantitative Assessment of Student Performance and Examination Format. *Journal of Instructional Pedagogies*, *18*. http://aabri.com/manuscripts/162516.pdf

Delmas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing Students' Conceptual Understanding After a First Course in Statistics. *Statistics Education Research Journal, 6*(2), 28-58. *https://iase-web.org/Publications.php?p=SERJ_issues*

Dunham, B., Yapa, G. & Eugenia, Y. (2015). Calibrating the Difficulty of an Assessment Tool: The Blooming of a Statistics Examination. *Journal of Statistics Education*, *23 (3)*. *http://jse.amstat.org/v23n3/dunham.pdf*

Ferrari, S., & Cribari-Neto, F. (2004). Beta Regression for Modelling Rates and Proportions. *Journal of Applied Statistics*, *31*(7), 799-815. *doi:10.1080/0266476042000214501* Taylor & Francis Online.

Budé, L., Van De Wiel, M. W. J., Imbos, T., Candel, M. J. J. M., Broers, N. J., & Berger, M. P. F. (2007). Students, Achievements in a Statistics Course in Relation to Motivational Aspects and Study Behaviour. *Statistics Education Research Journal, 6*(1), 5-21. *https://iase-web.org/Publications.php?p=SERJ_issues*

Myers, C. B. & Myers, S. M. (2006). Assessing Assessment: The Effects of Two Exam Formats on Course Achievement and Evaluation. *Innovative Higher Education, 31*(4), 227-236.

Stiggins, R. J. (2001). Student-involved classroom assessment (3rd edition). *Upper Saddle River*, NJ: Prentice Hall.

Tuckman, B. W. (1996). The relative effectiveness of incentive motivation and prescribed learning strategies in improving college students' course performance. *The Journal of Experimental Education, 64,* 197-210.

Tuckman, B. W. (1998). Using tests as an incentive to motivate procrastinators to study. *The Journal of Experimental Education, 66,* 141-147.

Ziegler, L., & Garfield, J., (2018). Developing a Statistical Literacy Assessment For The Modern Introductory Statistics Course. *Statistics Education Research Journal, 17*(2), 161-178. *http://www.stat.auckland.ac.nz/serj*

Zimmerman, T., Becker, L., Peterson, J., Nyland, R., & Surdick, R. (2014). Narrowing the Gap between Students and Instructors: A study of Expectations. *Transformative Dialogues: Teaching and Learning Journal, 7(1),* 1-18.