Implementing schema-based assessment in engineering statistics courses

Walter M. Stroup, Christopher Golubski
Department of STEM Education
The University of Texas at Austin
Austin, Texas, USA
wstroup@austin.utexas.edu, golubski@utexas.edu

***Abstract*—In classes like statistics, a staple for many beginning engineering students, deep interdisciplinary understanding as well as critical thinking is required to successfully learn the material. In this sense, traditional assessment, which may consist of multiple choice questions and free response problems, can be drastically inadequate to fully assess student comprehension of complex concepts. We propose a different method of assessment: schema-based items. These items are multiple choice items; however, there may be one or more correct answers based on the proposed question (hence, the items are non-dichotomous). The answer choices are nuanced and designed to measure not simply the student's understanding of a single concept but their progress on an entire learning progression, which may include a single concept as well as several adjacent ones. After some prior promising implementations of these types of assessment items, we are eager to transition these types of items to the post-secondary engineering classroom, where students are required to understand and assimilate content knowledge that is not focused on a single subject area and has a vast interconnection between subjects. Due to the interdisciplinary nature of statistics, these topics can be quite difficult to assess accurately and thoroughly.**

Keywords—assessment; statistics education; engineering education; schema theory

I. Introduction

With the advent of the No Child Left Behind Act and its successor, the Every Child Succeeds Act, standardized testing has become woven into the fabric of our educational system. High-stakes testing has become the norm from elementary school all the way through secondary school, all in the name of college preparedness. Recently, there have been a number of educators who believe that "common sense" should prevail and that we are testing children far too often. [1]–[3]. Not only do these educators believe that we are testing students with too great of frequency, the argument has been made that our educational system has transformed itself into being about only the tests, and that this fervent focus on test taking is on the verge of ruining education as we know it.

Common sense should be employed in this situation. If educators are only teaching to the test and the test can only examine a finite amount of material, there is much being left out that students need to know and should learn in school. Critics argue that students should not be required to take proficiency exams every year and that too much focus is being placed on these exams because school funding is often tied to the students' results on these types of tests.

Additionally, for ease of mass scoring, nearly all of these exams are multiple choice in nature. Of course, this means that students are required to select one answer as the correct answer. Depending on the list of choices, students usually have a 20% or 25% chance of guessing the correct response to a question by randomly selecting an answer. This is not ideal. And when the results are returned to schools, who have implemented initiatives to become more "data driven," how can a teacher interpret these results? In other words, how can a teacher use the results of her class's standardized test to improve her teaching technique for next year's cohort of students? This is a complicated question, and unfortunately, there is no easy answer in sight.

Another problem that has been raised with standardized testing is that standard multiple choice items show a lack of sensitivity to classroom instruction [4]. That is, the results of standardized tests are largely based on test taking ability than the effect a teacher has on a student when teaching her in the classroom. If this theory holds true, in many cases, teachers are being held directly accountable for standardized test results that are not a product of their teaching ability.

If students are going to be regularly tested on material, improvements to assessment are going to be necessary. Not only do we need more sensitive and effective items to be presented to students, but we also need to ensure that teachers can understand (and correct) underlying misconceptions that students have by looking at their responses and knowing where within the students' understanding the disconnect lies. To this end, we are proposing using new types of items that are structurally similar, however, we expect them to behave quite differently than the existing single-answer multiple choice questions. Our items, coined "schema-based items," will be designed not only to effectively assess student knowledge, but also provide educators with means to understand where students' misconceptions are and address them handily.

II. Theoretical Framework

The term "schema-based item" originates in schema theory. A schema is an object that represents a template or blueprint of conceptual knowledge that a person has [5]. An instantiation of a schema might be a person's schema for a house. For example, any given person probably knows when a house is used for, that it usually has at least one door, some windows, and rooms inside. Depending on the person's individual experience, they may be familiar with different styles of houses, like ranch houses or Victorian style houses. All the knowledge a person has about houses is contained within the instantiation of the schema.

In the realm of mathematics, researchers have posited that mathematical objects themselves (e.g. algorithms, processes, and functions) can be considered schemata themselves [6]. With this idea in mind, we also consider these mathematical objects schemata, and by choosing particular objects, we believe that we can conceive of important ideas to communicate to students about these mathematical objects in order to effectively convey (and assess) understanding of these objects. More importantly, by using schema-based items, we can assist students in "adding" or "subtracting" concepts in their own schemata on a particular topic or strand.

III. Research Design

For the purpose of this study, we have chosen the subject area of statistics. This is a poignant choice, because there is a significant body of research that indicates that students struggle significantly with the interpretation of statistics as well as statistical inference [7]–[9]. Considering that statistics education is a burgeoning field and more and more students are taking statistics courses, we feel that using schema-based items in statistics courses could provide us with a rich set of data to see if these items perform differently than traditional assessment items. Not only do we desire to see if the schema-based items perform differently, we also want to see if we can correct common student misconceptions quickly by looking at the results of these items in an attempt to prevent further misunderstanding. Theoretically speaking, we wish to "alter" the schema of the student to include correct information.

Schema-based items are multiple choice questions; however, while they look like ordinary questions of this nature, they have a different goal in mind. For example, the following question is based on the schema of a confidence interval:

After surveying students at the local university, the Department of Nutritional Services calculated that a 95% confidence interval for the mean cost of food for one week is ($20, $55). Now the department is writing its report and needs to provide an appropriate interpretation. Which of the following is a valid interpretation? (Choose all that apply.)

A. 95% of all students pay between $20 and $55 for food each week.

B. 95% of the sampled students pay between $20 and $55 for food each week.

C. We're 95% sure that the students in this sample averaged between $20 and $55 each week for food.

D. 95% of all samples of students will have average food costs between $20 and $55 per week.

E. We're 95% sure that the average amount all students pay for food is between $20 and $55 per week.

This particular item asks that students select all correct answers that apply. That is, they can choose one answer or all five answers. The structure of the question alleviates one issue immediately: students will have a far more difficult time guessing the correct answer due to the potential combination of possible answers to the question. In additional, thoughtful distractors have been chosen that correspond with common misconceptions that have been noted in the author's freshman statistics course for STEM majors. Since students tend to struggle with the interpretation of confidence intervals in the author's course, we are hoping to have immediate feedback to provide students with an explanation appropriate to the error that the student made.

This work is in its beginning stages now. Items are being written and reviewed by multiple content experts to ensure correct wording and intent. The participants for this study will be recruited from several concurrent sections of an introductory statistics course for science and engineering students at a large competitive school in the southern United States. The sections of this course are usually large, with approximately 100 students in each section. Sections are taught by three or four instructors, and participants will be recruited from all instructors' sections. The course is a service-level course and the curriculum, assignments, labs, and examinations are standard for all instructors of the course.

Participants will indicate their interest in participating in this study by signing up with their instructor. An attempt will be made to keep the number of participants relatively balanced from each instructor. Each participant will be given schema-based questions as an addendum to each in-class examination, which will also contain traditional multiple-choice questions and free response items. Throughout the semester, we will record each participant's responses and scores earned on the schema-based items, as well as the scores earned on the traditional items and the overall exam score.

At the end of the semester, we will analyze several items. First and foremost, we would like to know if there is a difference in student performance between the schema-based items and the traditional items. To determine this, we will use a generalized linear model to see if there is a relationship between the score a student earned on the traditional items versus the score they received on the schema-based items. Secondly, we are interested in assessing patterns of student responses on the schema-based items to see if there are common misconceptions between classes taught by different instructors or differences within classes taught by the same instructor. This allows us to determine if the variance in student performance can be explained by instructor or by the schema-based items themselves.

Finally, we are interested in whether the schema-based items give students the opportunity to improve on topics in statistics where they may have initially faltered. We will analyze the results of students who answered a schema-based item incorrectly and see if there is improvement later in the course when presented an item on the same topic.

IV. Final Remarks

Even though our study of schema-based items is in its beginning stages in the postsecondary arena, a pilot study has been done in a secondary school with some promising results in both mathematics and sciences classes. We are hopeful that we will be able to find the same type of differences between traditional and schema-based items in the postsecondary statistics classroom. If we are successful in showing that these items have greater sensitivity to instruction than traditional multiple choice questions, we believe that we may be able to better assess student understanding of material as well as be able to appropriately intervene in a meaningful manner when they are struggling with concepts. It is also our hope that we are able to devise assessment methods that are sensitive to classroom instruction and fairly measure student performance accurately for the betterment of our students as well as our educators.

References

[1] A. Kohn, *The case against standardized testing: Raising the scores, ruining the schools*. Heinemann Portsmouth, NH, 2000.

[2] M. Neill, “The Dangers of Testing.,” *Educ. Leadersh.*, vol. 60, no. 5, pp. 43–46, 2003.

[3] T. Pandina Scot, C. M. Callahan, and J. Urquhart, “Paint-by-Number Teachers and Cookie-Cutter Students: The Unintended Effects of High-Stakes Testing on the Education of Gifted Students,” *Roeper Rev.*, vol. 31, no. 1, pp. 40–52, 2008.

[4] V. P. Huy, “Computer modeling of the instructionally insensitive nature of the Texas assessment of knowledge and skills (TAKS) exam,” 2009.

[5] W. F. Brewer and G. V. Nakamura, “The nature and functions of schemas,” *Cent. Study Read. Tech. Rep. No 325*, 1984.

[6] A. Sfard, “On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin,” *Educ. Stud. Math.*, vol. 22, no. 1, pp. 1–36, 1991.

[7] A. E. Castro Sotos, S. Vanhoof, W. Van den Noortgate, and P. Onghena, “Students’ misconceptions of statistical inference: A review of the empirical evidence from research on statistics education,” *Educ. Res. Rev.*, vol. 2, no. 2, pp. 98–113, Jan. 2007.

[8] G. Mathilde, E. Verdam, F. J. Oort, A. Mirjam, and G. Sprangers, “Significance, truth and proof of p values: reminders about common misconceptions regarding null hypothesis significance testing,” *Qual. Life Res.*, vol. 23, no. 1, p. 5, 2014.

[9] S. Krishnan and N. Idris, “Students’ Misconceptions About Hypothesis Test,” *J. Res. Math. Educ.*, vol. 3, no. 3, pp. 276–293, 2014.