Common Issues in SATS© Research

Candace Schau
CS Consultants, LLC, 12812 Hugh Graham Rd NE, Albuquerque, NM 87111

Abstract
If we want to improve statistics education, we must improve our research. This paper briefly describes issues found in SATS research. My eventual goal is to create a research and reporting checklist for statistics education researchers. Although I have emphasized SATS research in this paper, many of the issues mentioned are common across statistics education research.

Key Words: Research, Statistics education research, SATS research, Issues in statistics education research, Issues in SATS research

1. Introduction

I've tried to read all articles, presentations, and theses/dissertations that have included the use of either the earlier version of the SATS (the SATS-28©) or the current SATS-36©. Six issues are common in this research. These issues include: design and measurement, participants, scoring, analysis, results, and context reporting. Because of page limits, I only am able to mention these issues here.

1.1 Design and Measurement Issues
Four common design and measurement issues include: revising or dropping items from one or more of the six attitude components, using the SATS only once but drawing attitude change conclusions, using samples that are too small, and not including a code to match pre-test and post-test responses.

Researchers who use the SATS need to use all 36 items that comprise the six component scores in their current form. If they change or omit items, they should retest all psychometric properties of the resulting component scores.

If researchers are interested in attitude change, they must administer the SATS at least twice.

SATS research should include a reasonable sample size. Although this phrase means different things to different researchers, I suggest at least 25 in the smallest group and more is better.

Because student responses should be matched across pre- and post-tests to examine attitude change, researchers need to include a code to allow this matching.

1.2 Participant Issues
Two common issues include: omitting information on participation rate and only examining responses from the matched pre-test post-test group.

We need to know the number and percentage of students who participated.

Most researchers match students who took both the pre- and post-tests and examine their responses. Although this group is important, all three groups should be examined, including the group of students who took only the pre-test and the group who took only the post-test.

1.3 Scoring Issues

Four issues dominate this category: ignoring score distributions, not examining attitude component score quality, calculating component mean scores incorrectly, and calculating a total attitude score.

Researchers need to examine and report information about attitude score distributions, and use this information to determine appropriate analysis techniques. For example, the responses to the Effort component usually are very negatively skewed (i.e., as you might expect, most students report expending a great deal of effort to learn statistics).

Researchers need to examine and report the quality of the six attitude component scores for their sample. I suggest, at a minimum, calculating the internal consistency for each component score at each administration time.

Because the mean scores for individual items within a component differ, a mean attitude component score should be calculated only for students who complete all items in that component.

A total attitude score never should be calculated. Students do not possess an attitude toward statistics but rather a set of interrelated components that comprise attitudes toward statistics. The attitude components consist of different numbers of items, and the components intercorrelate differentially. Obtaining a mean score across all 36 items allows those components with more items (e.g., Value) and those components with higher intercorrelations (e.g., Affect and Cognitive Competence) to contribute more to a total score than the other components. Also, using Confirmatory Factor Analysis, a total attitude score yielded a poor fit to the SATS' data.

1.4 Analysis Issues

Analysis issues fall into two groups: pooling across groups of dissimilar students and using gain scores.

Pool only across similar student groups. For example, it makes no sense to pool across undergraduate students enrolled in an introductory statistics course and graduate students enrolled in a multivariate course.

Instructors and researchers often want to know if students' attitudes have changed. They may calculate gain (or change) scores by subtracting students' pre-test scores from their post-test scores. This approach is not appropriate because SATS' gain scores often share 10% to 25% of their variance with the corresponding pre-test scores. Thus, gain scores are <u>not</u> free from the influence of pre-test scores.

1.5 Results Issues

These issues include the lack of reporting of both central tendency and variability, as well as emphasizing statistical, rather than practical, significance.

Researchers should report both appropriate central tendency and variability measures for each of the six attitude component scores, as well as for any of the other items they use.

Statistical significance depends on sample size. The importance of attitudes is in their effects. Effect sizes should be reported and interpreted, in addition to (or perhaps even in place of) statistical significance.

1.6 Context Reporting Issues

These issues include failure to report: course characteristics, type of institution, instructor characteristics, instructional methods, and student demographics. Without knowing this information, it is impossible to determine whether results likely will generalize. I've listed my initial suggestions here.

Course characteristics (aspects of the course that the instructor usually cannot control) include the department offering the course, level of the course, length and timing, organization, and class size.

Type of institution includes community college, two-year college, four-year college, and graduate institution.

Potentially important instructor characteristics include gender, rank, and statistics teaching experience.

Instructional methods include the use of student response systems, software, the internet, projects, demonstrations, small group activities/discussions, whole class activities/discussions, calculators, and real data sets.

Student demographics include number and percent by gender, age distribution, and number and percent of US citizens and foreign students.

2. Conclusions

This document is a start toward identifying important aspects of statistics education research, specifically research that includes the SATS, that need improvement. These six aspects should be considered prior to doing research that includes the SATS, during the research process, and while interpreting and reporting results. Doing so will yield better and more convincing research.

Please contact me with your comments, additions and subtractions, and other suggestions at cschau@comcast.net.