

Identifying and Removing Individuals with Inappropriate Response Sets
to the GTA Communication Survey

Stephanie L. Sutton, Dr. Ronald G. Downey, Meredith E. Pease, and Laurinda G. Smith

Office of Planning and Analysis

November 2001

Identifying and Removing Individuals with Inappropriate Response Sets to the GTA Communication Survey

Forty percent of the undergraduate courses in research and comprehensive universities in the U.S. are taught by graduate teaching assistants (GTAs), and of the introductory courses for first- and second-year undergraduates classes 60% are taught by GTAs (Wert, 1998, p. xvii). With the trend of increasing tuition and the public outcry about research institutions' commitment to undergraduate education, concerns about GTAs' teaching skills in general, and non-native English speaking GTAs specifically, seem hardly surprising (Bailey, 1984; Chism, 1987; Managan, 1992). Only recently have there been any systematic attempts to evaluate the success of GTAs in the classroom (Signore, Gibbons, & Downey 1996; Smith, Downey and Cox, 1999).

In response to the concerns about non-native English speaking GTAs, a system for assessing non-native English GTAs' language skills has been in place at a Kansas State University for many years. The university and the Kansas Board of Regents addressed the issue of English assessment with a three-tier approach. In this approach, all prospective non-native English speaking GTAs who will have teaching responsibilities must first achieve a score of 550 on the TOEFL (Test of English as a Foreign Language) exam before being admitted into the graduate program. Then, as established by the governing board, they must pass the SPEAK (Speaking Proficiency English Assessment Kit) and have their English skills assessed by a panel of faculty and students. In addition, each department provides some level of training for their GTAs. The training ranges from a very limited form of training, consisting of a syllabus and textbook and a talk with relevant faculty to extensive training programs for GTAs prior to class, including several

weeks of instruction, monitoring by faculty (mentoring program), and continued training through formal courses. In spite of the extensive attempts to select and train GTAs, the complaints continued. Smith et al. (1999) reported on an addition to this process, the GTA Communication Survey.

This new approach and process included the following major elements:

1. The basic skill(s) of concern was the ability of the GTAs to "communicate" effectively in the classroom. The best source of information on GTAs' communication skills was their students.
2. In response to international community concerns, the communication skills of all GTAs (native or non-native English speaking) needed to be assessed.
3. The survey would only be given to GTAs teaching in the classroom for the first time (first-time GTAs).
4. The approach was to be done very early in the semester in case remedial help was needed.
5. The procedures needed to identify GTAs who had communication problems versus those who were judged effective but less skilled. Any GTAs who were identified as having a communication problem needed to have an opportunity to improve.
6. After an opportunity to improve, GTAs identified as having a communication problem were re-evaluated. GTAs who failed to improve to the specified level were reassigned.

At the beginning of each academic semester, GTAs teaching in classrooms for the first time were identified by the departments. Each class section taught by a first-time

GTA was surveyed. The survey used was an adaptation of an existing communication skills survey (Hoyt, 1978).

Using this approach, students complete the survey evaluating the GTA on thirteen items. The first nine statements regarding the instructor's communication skills had alternating negative and positive statements. In utilizing this type of alternating response format students are forced to attend to each question and to read the question carefully before responding. Although the alternating negative and positive item format aided in improving the student's attentiveness to the items, anecdotal information suggested it was not sufficient in preventing a limited number of students from marking all or mostly all items at one extreme or the other. When a student mixes a significant number of extreme negative and positive responses, adding them to the class average would skew the results and not accurately reflect the GTA's communication skills.

The purpose of this research was to develop a means for identifying those individuals with extreme responses without regard to their being either positive or negative in wording. The process for identification needed to be programmable as part of the scoring program.

Methodology

Of the 13 items in the GTA Communication Survey, six items (including the summary item) are stated negatively, four items are stated positively, and two items (10 and 11) reflect the student's effort in the course and were excluded from this project. Responses to the survey used the following rating scale: 1=Definitely False, 2=More False than True, 3=In Between, 4=More True than False, and 5=Definitely True. Negative statements were recoded as follows: 1=5, 2=4, 4=2, and 5=1.

As the first step in this study, two sums were calculated: 1) the sum of all the responses to the four positive items and 2) the sum of all the responses to the six negative items (after they were recoded). The third value calculated was the difference between the sum of the negative items and the sum of the positive items. The logic for calculating the difference was that, since both sets of items were now scored in the same direction, large difference (positive or negative) would indicate individuals who were not responding in a consistent manner. Scores for this difference could range from a low of -14 to a high of 26. Since item responses tended to be greater than 3, it was expected that the mean difference score would tend to be positive.

Results

The data from the Fall 2000 semester was used in this study. The fall 2000 GTA Communication Survey database included data for 123 first-time GTAs teaching 227 sections. A total of 4,919 students provided responses for these sections, but only 4,848 had complete data. The mean difference score for the 4,848 students was 8.4, with a standard deviation of 3.14 (see Chart 1 in Appendix B). The full distribution of these scores was examined to see whether there was a clear and significant drop-off in the distribution at either end.

A very conservative criteria was used where there were few individuals in the distribution below or above a certain points at the bottom and top of the distribution, the cutoffs for the difference scores were set between +16 and +26 or -14 and -2. When a student's survey yielded a difference score in either of these ranges, they were excluded from the database. When the cutoff score was implemented, 80 students were excluded. A new analysis was run on the 4,768 students remaining in the database. The new mean

was again, 8.4 and the standard deviation was 2.80, a drop from 3.14 (see Chart 2 in Appendix B).

Discussion

The use of the above cutoff score yields a 20% reduction in the variance of the difference scores for the refined distribution (from a variance of 9.86 to 7.84). There was no difference in the overall mean (8.4 in both cases). Less than 2% of the respondents were removed from this analysis suggesting that a majority of the students appear to be attending to the survey.

While there are clearly still tails to the distribution at both the negative and positive ends, there remains the possibility that the evaluations for the GTAs did differ over the items and there is no clear way or level for setting more stringent cutoffs. Future scoring of the GTA Communication Surveys will use the above cutoffs to eliminate individuals who clearly appear to be responding to one end or the other of the scale and not reading the items. Eliminating these responses allows for a more accurate representation of each GTA's communication abilities and reduces a major source of potential error.

References

- Bailey, K. M. (1984). The foreign TA problem. In K. Bailey, F. Pialorsi, and J. Zukowski/Faust (eds.), Foreign teaching assistants in U.S. universities (pp. 3-15). Washington DC: National Association for Foreign Student Affairs.
- Chism, N. V. N. (Ed.). (1987). International teaching assistants: Institutional responsibilities and responses in the employment and education of teaching assistants (p. 249). Columbus, OH: The Ohio State University Center for Teaching Excellence.
- Hoyt, D. P. (1978). Communication survey. Unpublished manuscript, Kansas State University: Manhattan, KS.
- Mangan, K. S. (1992, March 4). Colleges expand efforts to help teaching assistants learn to teach. The Chronicle of Higher Education, A17.
- Signore, C. E., Gibbons, J. A., & Downey, R. G. (1996). A study of the first-year communication skills of non-native and native English-speaking graduate teaching assistants (Research Rep. No. 76). Manhattan, KS: Kansas State University, Planning and Analysis.
- Smith, L. G., Downey, R. G., Cox, K. S., Edwards, W. R., & Gaylord, T. W. (1999, May). The international GTA problem: A new approach. Paper presented at the 39th Annual AIR Forum, Seattle, WA.
- Wert, E. (1998). Foreward. In Marincovich, M., Prostko, J., & Stout, F. (Eds.). The professional development of graduate teaching assistants. Bolton, MA: Anker Publishing Company, Inc.

Appendix A – Positive and Negative Items on the GTA Communication Survey

Positive Questions

1. My instructor enunciates clearly
3. My instructor speaks loud enough
5. My instructor expresses ideas and thoughts clearly
7. My instructor shows an understanding of student questions or comments

Negative questions

2. My instructor speaks too rapidly
4. My instructor covers too much material in class sessions
6. My instructor lacks good organization in presentations
8. My instructor lacks skill in explaining difficult concepts
9. My instructor uses a vocabulary too advanced for this class
12. My instructor's communication level is seriously hampering my learning in this course

Appendix B – Charts Displaying the Difference Scores Prior (Chart 1) and After (Chart 2) Implementing the Cutoff Scores

Chart 1

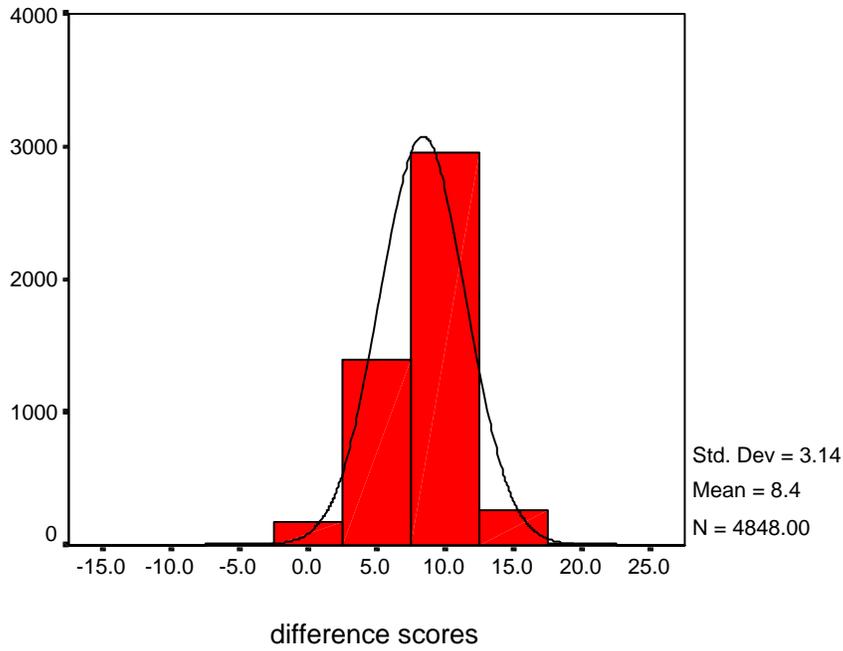


Chart 2

