Instruction on statistics writing and data visualization conventions in the GCC Practicum: Is it enough?

N. Saidikowski

Abstract:

To offer effective and well-founded advice, tutors in the GCC learn the conventions for writing across multiple genres, including the genre of science writing which sometimes requires the use of tables and figures. Because they will sometimes need to advise writers about such data visualizations, tutors in the GCC are given some principles and conventions on which to base their advice. Several studies have researched effective methods of teaching statistical conventions to students who might try to resist such teaching. However, the studies tend to focus on students in statistics courses and measure how well they perform in the course. There have been few or no studies into whether a short introduction to statistics can adequately prepare tutors to advise students on writing about statistics or building appropriate data visualizations. The proposed study will examine whether the GCC's current teaching strategy adequately prepares GCC tutors to advise writers on scientific papers which incorporate figures, charts, or other types of data visualization.

Introduction:

Over the course of a semester, tutors at Carnegie Mellon University's Global Communication Center take what is effectively a crash course in genre conventions. The GCC Practicum is designed to offer tutors some exposure to the most prevalent writing conventions in various disciplines, and to allow tutors to make informed and reasonable suggestions to students seeking advice on their writing. The reasoning behind this kind of crash course has its roots in a study that suggested tutors who lack expertise in the conventions of a particular writing discipline will tend to offer less effective advice to students or even prove detrimental by offering bad advice with an unwarranted level of certainty (Mackiewicz, 2004). The overall goal of the practicum, therefore, is to give the tutors an introduction to the writing conventions for various disciplines and impress upon them that there may be some conventions about which they will not know everything.

A fair number of the submissions to the GCC are scientific papers that include data analyses and various types of data visualization, like charts or graphs. Unfortunately, tutors are not always entirely prepared to handle the quantitative arguments within these types of papers. Many textbooks on writing or argument will either implicitly suggest or explicitly state that numbers and data are arhetorical and cannot be manipulated by the writer (Wolfe, 2010). Likely because of this general assumption, many tutors working at the GCC (many of whom come from English or Humanities backgrounds) might tend to view the data analyses or the associated data visualizations as beyond their jurisdiction when, in fact, this is not the case. Often, graphs and charts, the visualization of the scientific writer's results, are the foundation of the entire scientific paper. The belief that numbers and statistics are divorced from argument will inhibit the tutors' abilities to advise scientific writers. If a tutor cannot advise a writer on the graph or chart presenting the central argument of the paper, the tutor will be severely limited in what they can help the writer with.

To this end, the GCC practicum implements a short sequence on data visualization and statistical conventions, but the tutors in the practicum are sometimes resistant to learning them. This kind of resistance is actually all too common, especially for students who tend to feel high levels of statistics anxiety when faced with a statistics course. Several researchers have sought to define "statistics anxiety" as it relates to academic statistics, but the most widely accepted definition is the one given by Onwuegbuzie et al. (2000):

> "*Statistics anxiety* refers to the apprehension that occurs as a result of encountering statistics in any form and at any level."

This anxiety is correlated with students' beliefs of their capability to learn statistics and their performance in statistics courses (Perepiczka et al., 2011), so reducing this anxiety should be part of any campaign to increase tutors' awareness and knowledge of statistical writing conventions. Bartz et al. (1981) showed the beneficial effects of reducing statistics anxiety for a group of psychology students. Students who participated in an intervention designed to lower their anxiety performed better in the course overall and also demonstrated an increased knowledge of statistics (Bartz et al., 1981). If tutors' uptake of statistical knowledge can be increased then two class periods may be sufficient to teach tutors what they need to know about statistical conventions.

The first step to reducing statistics anxiety is determining whether tutors suffer it, as assuming they have anxiety when they actually do not has consequences. DeCesare (2007) offered a method for empirically determining the existence of statistics anxiety and further showed the drawbacks to assuming anxiety exists when it in fact does not (DeCesare, 2007). He suggested that the assumption that many sociology students feel statistics anxiety is potentially erroneous since much of the research "is based solely on anecdotal and informal evidence" (DeCesare, 2007). After conducting an empirical study, he found that while statistics anxiety does exist, it is perhaps not as widespread as many researchers perceive (DeCesare, 2007). The suggestion he made based on his finding is that instructors should first ascertain if their students do, in fact, suffer statistics anxiety, rather than simply assume that they do (DeCesare, 2007). One danger to assuming anxiety exists when it does not is that the very assumption could encourage a student to feel anxious (DeCesare, 2007). Thus, it would be wise to first determine if the tutors actually feel anxious, or if there is another cause to their resistance to learning conventions for statistics writing and data visualization, before implementing methods to combat anxiety.

Studies into factors associated with statistics anxiety have shown three major trends. Higher levels of statistics anxiety were associated with:

- Bad attitudes towards or misconceptions about statistics
- The attitude of the instructor and whether or not they acknowledged that some students would feel anxious
- The perceived lack of connection between statistics and real world problems

Perepiczka et al. (2011) found a statistically significant positive correlation between students' beliefs in their abilities to learn statistics and their attitudes towards learning statistics. In other words, the better a student's attitude towards statistics, the stronger their belief in their ability to learn the subject matter. Pan and Tang (2004) showed that the instructor's attitude and awareness of students' anxiety were significant factors in students' anxiety. Additionally, Wilson (1999) found that "the instructor's interpersonal style can have a great deal of influence in reducing the stress level of students in his or her class." In another paper, Pan and Tang (2005) showed a connection between student anxiety and the students' perception that the statistics they were learning had few practical applications. These studies suggest that targeting these three specific factors can help to reduce statistics anxiety and increase students' knowledge of statistics.

Several methods have been proposed to combat statistics anxiety. The factors associated with higher levels of anxiety offer three distinct areas to apply treatment. Wilson (1995) showed that students preferred instructors who maintained an "encouraging demeanor" and acknowledged the existence of anxiety about statistics upfront. This finding is supported by Pan and Tang (2004) who then expanded upon the discussion by showing that "innovative instructional methods," which include a positive instructor attitude, can help to reduce statistics anxiety and improve student performance in statistics courses. They recommended the implementation of a new approach to teaching statistics that includes "the application-oriented teaching methods combined with the instructor's attentiveness to students' anxiety" (Pan and Tang, 2004). The results of their study further

showed that including real-world applications with a positive attitude of the instructor could decrease students' anxiety (Pan and Tang, 2004) which would in turn increase their ability to learn and retain statistical concepts (Bartz et al., 1981). It is worthwhile to note that Wilson (1999) additionally found that students liked working in groups to learn statistics, but that this method only worked under certain conditions. The method of "cooperative learning" worked best when the other group members were known and trusted, whereas the method would actually increase anxiety when the other group members were unknown (Wilson, 1999). It may be a worthwhile method to implement if tutors establish enough of a rapport with one another to consider one another known entities.

Even in the event that tutors do not feel statistics anxiety, some of the suggested teaching methods can help to increase tutors' self-efficacy to learn statistics. Perepiczka et al. (2011) defined *statistics self-efficacy* as "an individual's confidence in his or her ability to successfully learn statistical skills necessary in a statistics course." The goal of the GCC Practicum, then, should primarily be to increase tutors' statistics self-efficacy, and to reduce tutors' statistics anxiety if that is the major barrier to their statistics self-efficacy. Fortunately, the approaches to decreasing statistics anxiety and increasing statistics self-efficacy overlap. The study by Perepiczka et al. (2011) suggested that helping students develop more positive attitudes about statistics can increase their self-efficacy to learn it. Thus, a single approach can be taken to reducing tutors' statistics anxiety and increasing their statistics self-efficacy.

Increasing tutors' self-efficacy to learn statistics should be a high priority because it is imperative to teach the conventions of data visualization and statistical writing to GCC tutors during the practicum. Equally important is the question of whether the tutors can then effectively advise writers about those conventions. While it would be helpful for tutors to address data visualization or statistical analysis concerns with the writer, the interaction will only be helpful if the tutor has a sufficient understanding of the conventions associated with the discipline. For example, Mackiewicz (2004) documented some of the consequences of a lack of tutor expertise in engineering writing. The tutors that didn't understand the conventions of writing for that discipline tended to ignore larger issues and focus on sentence-level errors (Mackiewicz, 2004). Furthermore, the tutors tended to give bad advice to engineering writers because their advice was grounded in conventions for a different type of writing (Mackiewicz, 2004). Worse still, the tutors would offer the advice in a forceful tone of voice, suggesting authority of the subject matter and leading the writers to bad revisions (Mackiewicz, 2004). The goal of the GCC practicum is to teach tutors enough about data visualization conventions that they can make informed suggestions without leading the writers to bad practices. This instruction will be most possible if methods are implemented to increase tutors' self-efficacy to learn statistics.

While the studies discussed in this paper have extensively researched statistics anxiety and its contributing factors, the negative effects of such anxiety on students' performance in statistics classes, and methods for combating the anxiety to improve students' performance and increase their self-efficacy to learn statistics, the applicability of the findings may be limited. The preceding research used as participants only students in statistics courses and the measure for the efficacy of the proposed methods was either reduction of anxiety or improvement in the course. Of particular interest to the GCC is whether these methods for reducing anxiety and increasing statistics knowledge will not only help tutors learn conventions for statistical writing and data visualization, but help them use those conventions to advise science writers.

Currently, the GCC Practicum incorporates a discussion of statistical writing and data visualization conventions into two class periods. Over the course of these two sessions, tutors taking the practicum are introduced to the concept that the rhetorical practices and argumentation strategies implemented in humanities writing are not necessarily ones that are or should be used in engineering or science writing. Tutors are given some instruction in the importance of figures and graphs for science writing and are taught some of the conventions and practices for creating reasonable

and informative visuals. Then, they are tasked with creating some of their own data visualizations and writing about them. This approach is designed to give tutors enough of a foundation in data visualization and statistical writing conventions that they can then offer informed and useful advice to science writers who seek help at the GCC. Whether or not this assumption is valid has not been tested.

Therefore, the goal of this paper is to propose a study to examine this assumption, and determine whether the teaching strategy currently implemented in the GCC Practicum effectively increases tutors' self-efficacy to learn statistical writing and data visualization conventions and gives them enough of a background in those conventions to allow them to offer informed and useful advice to science writers.

Proposed Methods:

Participants

The participants in this study will be the new tutors taking the GCC Practicum. The number of tutors entering the GCC has historically been small, so they should all be asked to participate. While it may not be reasonable to assume that the set of incoming tutors are representative of the population of all future GCC tutors and although this sample will be rather small, the proposed study should nonetheless offer a reasonable benchmark for future studies into the efficacy of the GCC Practicum's teaching strategy.

Procedures

The tutors will be given a scientific article pulled from the pool of previous GCC submissions. The article should be heavy in data analysis and full of figures and tables. Relevant selections from two example articles can be found in the Appendix. Several of the figures and tables should be problematic; for example, the type of figure used for the given data could be incorrect so that it misleads the reader, a table could be poorly laid out so that it is very difficult to pull out the intended message, a chart could be overdone in terms of colors and emphasis so that it is confusing. Tutors will be asked to analyze the paper and write out their approach to a tutoring session with the scientific writer who submitted it.

The Data

Tutors at the GCC come from various backgrounds, and while many come from an English or Humanities background where statistics and data might not be emphasized, some tutors have come from more technical backgrounds and have more experience in dealing with data and statistics. Thus, ideally, the data collected for the assessment of the knowledge gained by tutors during the practicum will be paired observations in an attempt to account for prior knowledge.

To this end, the knowledge tutors already have about statistical writing conventions should be assessed early in the practicum. The methods for collecting the data before and after the sequence of GCC class sessions will be similar. Two options for this part of the data collection process are available:

- The same article can be used for both the before and after assessment. Because the best approach to a tutoring session on the scientific article will not be discussed after students are asked to write it out, they will not know the "right" answers. Hypothetically, their responses post-instruction will be unbiased by their earlier exposure to the paper.
- Two similar articles can be used, one for the before and the other for the after assessment. If there is reason to believe that tutors' exposure to the article during the before assessment will bias their responses to the article post-instruction, then a different article will need to be used to reduce the effects of the bias.

The data collected from the before and after assessments will be mostly qualitative, as tutors are asked to give a written response to the scientific article. The researchers will assess with the written responses whether the tutors gave appropriate and useful advice to the science writer. Did the tutors notice the major issues with the figures and charts? Did tutors approach the figures and charts at all?

Ultimately, the data to be analyzed are the differences between the first assessment and the second assessment. If a tutor ignored the figures and charts

in the first assessment, did they discuss them in the second? Did the advice of the tutors improve? Were they more aware of conventions and data visualization techniques? Did they offer valid advice where they could but hedge where they knew they were unsure?

It may be possible to collect quantitative data for the assessment as well. A survey in the form of a multiple-choice quiz could be used in addition to or in place of the written response. While data from a survey of this kind may be easier to analyze, it may not be as accurate as a written response since categorical answers might constrain tutor answers or even bias them by suggesting particular responses.

Further Research

In the event that the data collected from the proposed study indicates that the GCC practicum is not effectively increasing tutors' self-efficacy to learn statistical writing and data visualization conventions, several remedies may be considered:

- A longer period of time might be spent on the introduction to conventions of statistical writing and data visualization conventions.
- An out-of-class workshop could be offered to tutors (and other students who wish to attend) on the conventions of statistical writing and data visualization.
- More resources could be made available to tutors in the form of optional tutor training activities. For example, tutors could be given data sets and asked to create data visualizations and write about them.
- A resource could be made available which gives some tips on data visualization techniques and conventions. The resource would be designed for writers but also be usable by tutors.

If new techniques are adopted by the GCC practicum to increase tutors' self-efficacy, the methods discussed in this proposal might be used again to test whether tutors' self-efficacy has, in fact, been increased.

References:

1. Bartz, W. R., Amato, P. R., Rasor, R. A., & Rasor, M. O. N. (1981). Effects of reducing student anxiety in a statistics course. Australian Psychologist, 16(3), 347-353.

2. DeCesare, M. (2007). " Statistics Anxiety" among Sociology Majors: A First Diagnosis and Some Treatment Options. Teaching Sociology, 360-367.

3. Mackiewicz, J. (2004). The effects of tutor expertise in engineering writing: a linguistic analysis of writing tutors' comments. Professional Communication, IEEE Transactions on, 47(4), 316-328.

4. Onwuegbuzie, A. J. (2000). Statistics anxiety and the role of self-perceptions. The Journal of Educational Research, 93(5), 323-330.

5. Pan, W., & Tang, M. (2004). Examining the Effectiveness of Innovative Instructional Methods on Reducing Statistics Anxiety for Graduate Students in the Social Sciences. Journal of Instructional Psychology, 31(2), 149-159.

6. Pan, W., & Tang, M. (2005). Students' Perceptions on Factors of Statistics Anxiety and Instructional Strategies. Journal of Instructional Psychology, 32(3), 205-214.

7. Perepiczka, M., Chandler, N., & Becerra, M. (2011). Relationship between graduate students' statistics self-efficacy, statistics anxiety, attitude toward statistics, and social support. The Professional Counselor: Research and Practice, 1(2), 99-108.

8. Wilson, V. A. (1999). Reducing Statistics Anxiety: A Ranking of Sixteen Specific Strategies.

9. Wolfe, J. (2010). Rhetorical numbers: a case for quantitative writing in the composition classroom. College Composition and Communication, 452-475.