Student Perceptions of Assessment Strategies in a Multivariate Statistics Course

Ann A. O’Connell
University of Connecticut

Journal of Statistics Education Volume 10, Number 1 (2002)

Copyright © 2002 by Ann A. O’Connell, all rights reserved.
This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the author and advance notification of the editor.

Key Words: Instruction; Multivariate analysis; Student perceptions.

Abstract

The purposes of this paper are to illustrate the use of several assessment strategies in an advanced course in statistics, and to present the results of student ratings for each assessment strategy in terms of difficulty, appropriateness, level of learning achieved, and preference. The assessment strategies used include structured data analysis assignments, open-ended data analysis assignments, reviews of applied research articles, and annotating computer output from multivariate software procedures. Findings indicate that students "prefer" instructor-directed or structured assignments overall, but feel they learn the most when the assessment is unstructured and requires greater self-direction. Suggestions for incorporating these assessment strategies into the multivariate classroom, as well as examples of each strategy, are included in this study.

1. Introduction

A master’s level student came to my office recently to discuss her research project in the area of sports-fitness. She had collected data on several leg and thigh variables and wanted to investigate the relationship between these measures and later occurrence of injury among female athletes. It sounded like an interesting study and was certainly appropriate for a master’s level thesis project. The student’s specific questions related to how to interpret the computer output from her analyses. However, the underlying problem was that the student wasn’t sure which statistical method was needed for her study, so she decided to "click on all the buttons" to make sure she got the "right" answer!

Unfortunately, this is not an uncommon occurrence for many who teach educational statistics courses. Rapid advances in statistical analysis packages may seem to suggest that this "point and click" phenomenon is fairly recent. A decade ago, Searle (1989) and Dallal (1990) discussed similar concerns regarding misguided analyses and misinterpretation of analysis results by students. They debated the merits and cautions of including instruction on statistical computing packages as part of statistics courses. Searle (1989) cautioned against using computing packages in introductory courses, believing that
students would then feel "that easy-to-use computing packages can be a substitute for (rather than a supplement to) a proper knowledge of statistical methodology" (p. 190). Dallal (1990) alternatively argued that "Using a statistical package forces students to come to grips with whether they really understand how to apply the techniques that they have studied. It will be impossible for them to write a clear report, for example, if they cannot make sense of the output" (p. 266).

Many introductory courses are prerequisites for more advanced statistical methods courses; hence computer packages have a necessary place for beginners as well as advanced learners in statistics. But for the student above, interpretation of the output was not the primary problem. There was clearly a missing link hindering her ability to move from the research question to an appropriate analysis and interpretation of the data. With the widespread availability and use of point and click technology, the lack of congruence between her design definitions, research questions, and statistical analysis may simply be more apparent to educators now than in the past.

Too often, students that we would like to think of as "successful" based on their homework, exam performances and class participation seem to rely on a belief that solutions to complex analysis problems outside of the course context can be found by simply clicking on a few buttons. This incongruity suggests a gap between how learning in statistics courses is assessed versus the actual statistical competencies (both procedural and conceptual) we expect our students to have after completing these courses. As a solution, educators need to identify activities and assessments that are likely to transfer to research experiences outside of the classroom.

Relatively little published work is available on the experiences of instructors or students in advanced statistics courses. However, statistics educators can build on the suggestions of colleagues and researchers who have concentrated on introductory statistics courses in order to improve pedagogy in advanced courses. In particular, statistics educators have argued for more active involvement on the part of the student. One way to stimulate students’ active engagement in advanced statistics is to first consider how our own course assessment requirements encourage or inhibit particular kinds of learning, and then revise our strategies according to the goals we want our students to achieve.

2. Purpose

The purposes of this paper are to illustrate several different approaches to assessment for an advanced course in multivariate analysis at the graduate level, and to present the results of student ratings of each strategy within a class of current advanced graduate students. Ratings were obtained on difficulty, appropriateness, and level of learning realized from each strategy. Student preferences regarding these methods are also reported.

A combination of four different assessment strategies were used in this course:

- a. structured data analysis assignments,
- b. open-ended data analysis assignments,
- c. article reviews, and
- d. annotating computer output from specific multivariate software procedures.

Before presenting the results and discussing their implications for instruction and student learning in multivariate analysis, I provide a brief review of some relevant research findings from the literature on statistics education, assessment, and transfer.

3. Review of Related Literature
The literature on pedagogy in introductory statistics courses continues to grow (see, for example, the *International Statistics Review*, special issue, April 1995; *Journal of Educational and Behavioral Statistics*, special issue, Spring 1996; Gal and Garfield 1997), yet comparatively little information is available on how well or how much graduate students are trained beyond the introductory level. For many graduate students in the social sciences, successful learning in advanced statistical methods will enable them to embark on successful academic or professional careers. First, however, students need sufficient preparation in statistical methods for completion of a culminating project (thesis or dissertation) that establishes a personal contribution to their field of study. This initial experience often provides the groundwork for making continued contributions to their field by helping students establish an active and productive research agenda.

An important goal of current pedagogical efforts in statistics education is to improve students’ ability to understand and work competently with statistics and statistical concepts. One of the most significant trends to emerge in statistical education involves an emphasis on activity-based learning, such as cooperative group tasks, computer based problem-solving activities or simulations, or the collection and analysis of "real," as opposed to textbook, data. Research has shown that learning can be enhanced when students are actively engaged in a problem and are given an opportunity to structure their new knowledge in terms of personal experiences and academic interests (Garfield 1995; von Glasersfeld 1987). Using a range of assessment activities designed to model the experiences and concerns of actual researchers and practitioners is one way to actively engage and challenge students in their own learning process.

In the multivariate classroom, immersing students as much as possible in a research-like environment can help connect the often disjointed activities of instruction and assessment, particularly if the assessment activities are carefully constructed to reflect the knowledge and skills that, as instructors, we feel are needed to be competent in such a complex domain. Learning in statistics involves a multiplicity of competencies, and covers understanding, using, and doing statistics. Garfield (2000) describes these processes as an "overlapping hierarchy" (p. 4), and provides corresponding definitions for each category: statistical literacy (understanding statistical language), statistical reasoning (using statistical information to make decisions), and statistical thinking (what statisticians do). Since many graduate students may well be involved in future research that can influence health, social, or educational policy, courses in advanced statistics should guide the students towards statistical development in all three areas simultaneously. In my own courses, I use the phrase "statistical clarity" to describe what I want students to achieve. Statistical clarity emphasizes a synthesis of literacy, reasoning, and thinking, and refers to clarity of purpose in research design and analysis, clarity in decision-making activities regarding data and data collection, and clarity in articulation of results of often complex analyses. Collectively, the assessment tasks discussed in this paper were chosen to support the development of statistical clarity through activities that reflect students’ emerging multifaceted competence for multivariate statistics: reading and interpreting selections from the published research literature which use multivariate analyses, conducting computer analyses on multivariate data and summarizing the results in publishable format, and analyzing authentic research data based on reasonable and appropriate student developed research questions.

Support for this approach can be found in the *Assessment Standards for School Mathematics* (NCTM 1995), which are every bit as applicable to the graduate statistics classroom. For example, NCTM recommends that instructors view assessment as a convergence of information from different kinds of sources or activities; that assessment should motivate greater achievement; and that connections among different kinds of knowledge should be assessed. By assessing the knowledge and skills deemed important for emerging researchers, statistics educators can give students an opportunity to see how their knowledge may be applied in actual research situations outside of the classroom, thus providing a realistic setting for the development of statistical clarity. In other words, the collection of assessment tasks should be as authentic as possible to mirror the activities of applied researchers.
Garfield has often reminded educators to use assessment as a vital part of the learning process and not just as a grading strategy (Garfield 1994; Gal and Garfield 1997). The tasks that are given to students today affect the learning that can and should transfer to new situations and research settings tomorrow. This view of assessment as a learning opportunity is emphasized in the NCTM Assessment Standards (NCTM 1995) as well, since "assessment should be a means of fostering growth toward high expectations" (p. 1). In the multivariate classroom, the expectation typically is for students to consolidate conceptual and procedural knowledge about advanced statistical techniques. How this can best be accomplished remains an issue for consideration by educational researchers and instructors of advanced quantitative courses. However, it is clear that statistics educators need to continue to move toward creative and authentic approaches to assessment of student learning in statistics.

Although sparse, there have been recent calls in the literature for improved graduate training in applied statistics, particularly for those future researchers whose substantive areas of expertise may not be quantitative methods (Keselman, et al. 1998). Keselman and his colleagues found that published research in respected journals in education and psychology often contains no reference to issues such as verification of assumptions, power, or effect size. They argue that improvements in graduate training can influence the quality of research in practice. The need for multiple perspectives on the unit of analysis problem also demands more attention in educational research courses (Pintrich 1994). Clearly, familiarity with both routine and complex statistical methodologies and related issues is critical for understanding the validity and relevance of research results. While it is hoped that instruction related to these issues would encourage transfer to research situations outside of advanced statistics courses, evidence suggests that this is not the case.

There is still much to be learned about transfer of knowledge to new situations (see, for example, Pintrich 1994), yet transfer is clearly an important goal for courses in multivariate analysis. Important outcomes for students in an advanced statistics class are to understand published research literature in their field, choose an appropriate statistical methodology for different research situations, conduct and interpret multivariate statistical analyses, identify and anticipate potential analysis limitations or alternatives, and apply the principles and techniques to new and often rapidly changing research initiatives and directions. The challenges in creating the kind of environment where these outcomes can be achieved may seem overwhelming at first to statistics educators. But, as argued by Shepard (2000), "if instructional goals include developing students’ metacognitive abilities, fostering important dispositions, and socializing students into the discourse and practices of academic disciplines, then it is essential that classroom routines and corresponding assessments reflect these goals as well" (p. 8).

Early studies by Tversky and Kahneman (1983) and Shaughnessy (1981) showed that formal instruction in statistics did not protect people from their pre-course misconceptions in statistical reasoning or improve their ability to make informed decisions under uncertainty. However, results of other studies by Fong, Krantz, and Nisbett (1986) and their colleagues (see Nisbett 1993) have offered some limited support for transfer of statistical training. More recently, research evidence supports the idea that conceptual knowledge plays an important role in the acquisition, development, or use of procedural knowledge (O’Connell 1999; Rittle-Johnson and Alibali 1999), but how the relationship between these two kinds of knowledge affects transfer to new situations needs further attention. Finally, in terms of transfer of statistical knowledge and skill to actual research activities, Keselman et al. (1998) report that "a substantial gap exists between the inferential methods that are recommended in the statistical research literature and those techniques actually adopted by applied researchers" (p. 351).

Statistics educators can begin to strengthen the bond between conceptual knowledge and procedural knowledge in advanced statistics by reflecting on and adjusting current instructional practices, including the goals for intended student learning and the educator’s corresponding assessment strategies. In particular, assessment that is grounded as much as possible in real-world experiences may offer the best
possibility for transfer of statistical knowledge to novel situations and research questions. One of Garfield’s principles of statistical learning is that students will come to value what they know will be assessed (Garfield 1995). What is it we value in multivariate statistics courses? The ability to accurately apply principles and methods of multivariate analyses in research situations external to the classroom setting is clearly a valuable skill, but it is not the only desirable educational outcome. Students need to appreciate how complex statistical methodologies build from simpler univariate analyses, and how all these methodologies form part of an overall discipline connecting research design, measurement, data analysis, and interpretation and presentation of results. To support such elaborate connections, a variety of assessment strategies may be required.

4. Multivariate Assessment Study

This study provides a summary of how multiple kinds of assessment strategies can be incorporated into a one-semester course on multivariate analysis. As the literature is sparse in this area, it seems reasonable to begin an exploration of the impact that assessment has on student learning of multivariate techniques by studying student reactions to particular forms of assessment. This approach is helpful for initiating reflection on what the goals might be for students, understanding how well these goals are being met, and suggesting methods for improving both teaching and learning of advanced statistics.

The data for this study came from a sample of advanced masters and doctoral-level students enrolled in the first course of a two-semester sequence in multivariate analysis. Typically, the students taking this course are in their second year or later of an applied degree program in psychology, education, or the social sciences. For most, successful completion of the multivariate course is a requirement of their degree program. The educational background of students in this sample is diverse, although none would be called mathematically sophisticated. Students majored in a variety of disciplines, including educational psychology, allied health, educational research, gifted and talented education, business, exercise physiology, family life sciences, and counseling and human development. All students had completed two prerequisite courses in applied statistics. Emphasis in these prerequisite courses is on univariate experimental design, introductory correlational strategies including basic multiple regression, and SPSS®.

The course topics in the multivariate course include:

- Multiple regression (review and additional material)
- Multivariate analysis of variance and covariance
- Factorial analysis of variance
- General linear model
- Repeated measures analysis and profile analysis
- Discriminant function analysis
- Canonical Correlation
- Principal Components Analysis/Factor Analysis

The course objectives stated in the syllabus are:

1. to develop the ability to work competently with a variety of multivariate statistical techniques;
2. to understand research literature/reports involving these techniques;
3. to use the computer (SPSS®) to run and interpret the results of multivariate analyses; and,
4. to write up coherent results from these analyses.

Four different kinds of assessments are used during the course:
a. structured data analysis assignments where specific questions are posed by the instructor and answers found through computer analysis or follow-up hand computation when necessary;
b. open-ended assignments which require students to develop a research question based on a particular data set, analyze the data appropriately for the question and the data, and interpret and write-up a formal report of the results;
c. article analysis which requires students to read a particular research article and answer questions based on the article; and
d. annotating output from computer runs. Detailed feedback is given on every assignment, generally within one week, in an attempt to correct misconceptions as early as possible. Scoring rubrics are discussed in class and a written summary of the point distribution for the open-ended assignments is provided to every student. Examples for each of the four assessment activities are provided in the next section.

Building on the NCTM (1995) recommendations and the work of Garfield and her colleagues (Garfield 1994; 1995; 2000; Gal and Garfield 1997), these specific assessment strategies were selected to encourage students to develop connections between procedural and conceptual knowledge, and to emphasize active rather than passive involvement in activities typical of an applied researcher. Structured questions focus a student’s inquiry about a particular analysis or method, which can help provide scaffolding or guidance for the student when approaching an unstructured research situation. The unstructured assignments simulate much of the actual research process and allow students to apply what they have learned to new research questions and situations. The article analysis is an opportunity to review and critique published articles utilizing many of the multivariate analyses being learned in class, and articles are chosen to cut across a variety of disciplines. In addition to providing models of research and analysis, the article reviews may lessen feelings of intimidation often experienced by students when studying complex statistical material for the first time, since the methods and results used in the articles are becoming more familiar to the student through the classroom discussions and other assignments. Finally, annotating computer output is a concrete way to encourage students to look at actual results, before attempting to put their findings into words.

4.1 Structured Data Analyses

Structured data analysis assignments are those where data (and sometimes the program syntax) are supplied to the student, and specific questions are provided for the students to answer. Generally, all students in the class should have the same or very similar responses to each question. Questions are presented as if the student is the responsible statistician (with some guidance).

Example of Structured Data Analysis: Multiple Regression Review

For this assignment, students are provided with a data set consisting of four variables collected on a small sample of junior high school students (adapted from Tatsuoka 1988). The dependent variable of interest is an achievement test score in the physical sciences. In the example, a teacher is interested in studying the extent to which achievement in the physical sciences \(Y\) can be predicted from measures of math ability \(X_1\), mechanical reasoning \(X_2\), or creativity \(X_3\). To address the teachers concerns, students are asked to run a series of three regression analyses and answer the following questions:

1. After controlling for differences in math skill within this sample of students, does mechanical reasoning significantly improve our predictions of achievement? How much more explanation of variability in the physical science measure do we achieve by including mechanical reasoning in a model that already contains math skill?

2. Over and above the explanation obtained from the first two variables, should we include the
creativity test score in our prediction equation? Discuss the evidence as to why or why not. Based on all of your analyses, what suggestions/conclusions would you give the teacher in terms of variables affecting achievement in the physical sciences? What model would you recommend the teacher use for predicting achievement in the physical sciences? What concerns would you share with the teacher about the data and the results?

4.2 Open-ended Assignments

For the open-ended assignments (also called independent projects), students are asked to select one data set from among several available, develop an appropriate research question suitable for a multivariate analysis, analyze the data, and summarize the results. Data sets include High School and Beyond (nces.ed.gov/surveys/hsb); automobile fatality rate data (Judd and McClelland 1989); a survey on attitudes and perceptions regarding distance education (King, Harnar, and Brown 2000); behavioral variables related to condom use among women (Stark, et al., 1996); relationships among different kinds of coded errors during probability problem-solving (O’Connell 1999); and data from the evaluation of an adolescent pregnancy prevention project (O’Connell 1997a). All the data sets and additional information about each study are available on the course Web page at the start of the semester. Two independent projects are required for the course: one for the midterm (generally using multiple regression) and one for the final (on one of the multivariate techniques). Project information (for example, research questions, nature of variables used, and the analyses proposed) is required periodically during the semester, to monitor student progress.

Occasionally, students have asked to use their own data for the final project (see point e below), which is cautiously encouraged if the topic is appropriate. In addition to addressing any concerns about the data (primarily, whether or not the project data is suitable for the fulfillment of the course requirement), proof of adherence to all university institutional review board (IRB) requirements is mandatory before any external data can be used in the course. I emphasize to all students that the purpose of the research/analysis requirement for the course is to gain experience with designing research questions, analyzing the data accordingly, and summarizing the results. As they work through this process and begin to write up their results, new research questions or approaches may present themselves, but they are only required to submit the write-up of one complete analysis. However, self-reflection on their learning process is a priority component of their final paper, and they are asked to summarize any additional analyses or research questions they may wish to investigate in the future.

Example of Open-Ended Assignment: Final Project

The syllabus gives a short summary of goals for the independent projects, and students are also provided with an outline of expected content for the project paper. Briefly, the goals of the open-ended assignments are to establish a research question appropriate to one of the multivariate techniques taught in the course (and corresponding to the chosen data set), and to prepare a report detailing the methods used to answer the research question and an interpretation and critique of the results.

The following summary is provided to the students:

a. The project is a short, independent, research and analysis paper, limited to 5 to 10 pages.

b. Your paper should be written in American Psychological Association (APA) style (or style appropriate for your field).

c. The purpose of the project is to show me that you can understand and write intelligently about the use of a specific multivariate technique selected from those included in this course.
d. Several data sets are available for analysis.

e. If you want to use a data set from a project you are working on, you must clear it with me first. The data must have already been collected, you need permission from the principal investigator, and required human subjects issues must already be addressed with IRB forms filed. I reserve the right to say yes or no to people’s request to use their own data, since I want people to be successful at their particular analysis and not bogged down by extremely messy or complicated data sets.

f. Project Data Information Sheet is to be completed (a downloadable form from the course Web site) - which asks you to describe the variables to be used in the analysis and how they are measured.

g. Be aware that the analysis you might choose for the purposes of this course may not be the only approach for a particular data set/research question. I am interested in how you approach the particular analysis you selected to answer your research question.

h. Critique and self-reflection: include a discussion of what you learned, personally, from this project (in terms of the process of doing a project of this nature). What might you do better or differently, given a chance to do this project over again? What alternative analyses and/or research questions might you consider?

i. For the purpose of this course, you will be graded on the clarity of your analysis and its correspondence to the research question you asked.

j. The appendix to your paper should contain the output you used for analysis, annotated to describe how you used the output to address your research question(s).

4.3 Article Analysis

For the article review assignments, students are instructed to prepare a paper detailing the responses to several questions based on a selected research article. Students in the class submit potential articles for review, and one or two articles are chosen for the entire class to read. Alternative articles are sometimes provided based on the choice of procedures I may wish to emphasize. During the semester in which this study was conducted, time allowed for analysis of only one article (Article 1 below). Some examples of articles are:

- Article 1. Karacostas and Fisher (1993)
- Article 2. Litt, Kalinowski, and Shafer (1999)
- Article 3. Dant, Lumpkin, and Bush (1990) (also from Hair, Anderson, Tatham, and Black (1992), p. 138-152)

Example of Article Analysis: Karacostas and Fisher (1993)

Questions to be addressed by the students include:

1. What is the purpose of the study?

2. Describe the population and the sample. How was the sample obtained? Does the sample provide an appropriate model for the population of interest? How does your answer to this affect your faith in the conclusions presented by the researchers?

3. Describe the study variables and how they are measured.
4. Report the research questions of interest and the research hypotheses. What specific statistical techniques were used to investigate these research questions?

5. What assumptions are necessary for the statistical validity of the analyses used in this study? Were these assumptions addressed in the paper? Were they met?

6. In Table 1, the second entry is 34/38. Describe what both these values mean.

7. Explain the canonical correlation and how it relates to percent variance explained.

8. Do you agree with the results and conclusions put forth by the authors? Do you feel there are limitations to their conclusions, in addition to any they may have mentioned? Is there any missing information? Can you offer any rival hypotheses?

9. Recommend at least one additional research question and corresponding analysis or an alternative statistical strategy that might be appropriate for this study. What would be the hypotheses for these analyses if the study were replicated?

4.4 Annotated Output.

For these assignments, students are instructed to run a particular analysis after the topic and the printout from an example of that analysis have been reviewed in class. Annotation involves labeling on the output all the elements necessary for understanding the results. Often, the programs for the analysis correspond to data available in the text (we used Stevens (1996) during the Fall 1999 semester). Students are also required to annotate the output for their independent projects.

Example of Annotated Output: Factorial Manova.

The data for this example was modified from Stevens (1996), page 310, for a multivariate 3X2 factorial analysis of variance with two dependent variables. A significant interaction effect is present. Instructions for the students include:

a. Design and run the appropriate analysis including orthogonal contrasts for each factor.

b. Annotate the output (modeled after in-class examples).

c. State the assumptions and how you will verify them. Also write out the null and alternative hypotheses for the homogeneity of covariance test and report the results of this test.

d. Write a brief summary (a few paragraphs). Be sure to explain how your support or non-support of the assumptions required for Manova affect your interpretation of the results.

e. Report the results including interpretation of effect size for any significant multivariate or univariate results that you find.

5. Student Assessment Ratings

At the end of the course, students were asked to rate each different assessment strategy in terms of difficulty, appropriateness to their needs while learning multivariate statistics, and how well they felt they learned using that kind of assessment. Students were also asked to rank the four assessment strategies in
order of preference (1 = most preferred, 4 = least preferred). The survey was anonymous and was administered during the last class of the semester by a student volunteer while the instructor was not present. The purpose was explained as an opportunity for the instructor to understand how to better structure course assessments in the future. Upon agreeing to participate, the students were given the survey. A four-point scale was used with response categories as indicated in the table below. The results presented are based on 14 students who completed the course assessment (three students were absent or chose not to participate). Students were also asked for additional comments regarding their likes and dislikes about each assessment strategy, suggestions for how to improve the use of that strategy, and any other observations or comments they might have. In terms of establishing a course grade, the two open ended assignments (midterm and final) were weighted 25% each; the scores on the other assignments (five altogether) were averaged and worth 50% of the total.

The results for each assessment strategy are provided in the table that follows. The modal response for each technique is indicated in bold print. Other interesting results to be discussed below are presented in bold print as well.

**Table 1. Student Assessment Ratings.**

**Difficulty ratings: frequency (and percent).**

| Response Categories                          | Structured Computer Assignments | Open-Ended Assignments | Article Analysis | Annotating Output |
|---------------------------------------------|---------------------------------|------------------------|-----------------|-------------------|
| 1. Not at all difficult                     | 1 (7.1%)                        | 0 (0.0%)               | 1 (7.1%)        | 0 (0.0%)          |
| 2. Slightly difficult but not too challenging | 5 (35.7%)                      | 2 (14.3%)              | 4 (28.6%)       | 8 (57.1%)         |
| 3. Difficult but challenging                | 7 (50%)                         | 12 (85.7%)             | 7 (50%)         | 6 (42.9%)         |
| 4. Too difficult                            | 1 (7.1%)                        | 0 (0.0%)               | 2 (14.3%)       | 0 (0.0%)          |

**Appropriateness ratings.**

| Response Categories                          | Structured Computer Assignments | Open-Ended Assignments | Article Analysis | Annotating Output |
|---------------------------------------------|---------------------------------|------------------------|-----------------|-------------------|
| 1. Not at all appropriate or useful         | 0 (0.0%)                        | 0 (0.0%)               | 1 (7.1%)        | 0 (0.0%)          |
| 2. Slightly appropriate but not for all my needs | 2 (14.3%)                      | 1 (7.1%)               | 5 (35.7%)       | 5 (35.7%)         |
| 3. Appropriate for many of my needs         | 11 (78.6%)                      | 7 (50.0%)              | 5 (35.7%)       | 9 (64.3%)         |
| 4. Very appropriate for all                 |                                 |                        |                 |                   |
Level of learning ratings.

| Response Categories                      | Structured Computer Assignments | Open-Ended Assignments | Article Analysis* | Annotating Output |
|-----------------------------------------|---------------------------------|------------------------|-------------------|-------------------|
| 1. Didn’t learn anything                 | 0 (0.0%)                        | 0 (0.0%)               | 1 (7.1%)          | 0 (0.0%)          |
| 2. Learned a little bit                  | 3 (21.4%)                       | 1 (7.1%)               | 6 (42.9%)         | 4 (28.6%)         |
| 3. Learned enough to be comfortable with the topic | 8 (57.1%)                     | 7 (50.0%)               | 4 (28.6%)         | 8 (57.1%)         |
| 4. Learned a great deal - more than I would have thought | 3 (21.4%)                     | 6 (42.9%)               | 2 (14.3%)         | 2 (14.3%)         |

* one student did not answer this question for the article analysis.

5.1. Preference Scores

Students were asked to indicate their order of preference among the four assessment strategies, with 1 = most preferred form of assessment, and 4 = least preferred. Average preference ratings indicated that the most preferred forms of assessment were the structured computer assignments (mean = 2.00), followed by annotating the output (mean = 2.31), use of open-ended assignments (mean = 2.38), and article analysis (mean = 3.15). Yet in terms of difficulty level, appropriateness, and learning ratings, the open-ended assignments actually received the better ratings overall. Approximately 86% of the students found the open-ended assignments challenging (response category 3); 93% found the open-ended assignments appropriate for many or all of their needs (response categories 3 and 4); and 93% reported that through the open-ended assignments they learned either enough to be comfortable with the topic or more than they would have thought (response categories 3 and 4).

5.2. Selected Student Comments

The following collections of student comments are illustrative of those received in response to the open-ended questions on the survey.

Structured Computer Assignments:

- "These smaller assignments helped me to better get into the topic, realize where I have difficulties and learn from your comments."

- "Doing this type of assignment first helped me with the less structured assignments later."

- "Written feedback was very useful."
Open Ended Assignments:

- "This was difficult, but very necessary for us as competent researchers."
- "The open-ended nature of the assignments 'forced' me to think about the bigger conceptual issues. It helped to apply the info to real-life examples, like I will be doing with my dissertation."
- "Excellent, the openness addressed my own misconceptions."

Article Analysis:

- "This assignment was time consuming b/c there was a lot of English vocabulary I have not been exposed to."
- "Would have liked to instead answer all of the questions and then discuss the article in class - but we didn’t have enough time."
- "The topic was not of personal interest."

Annotating Output

- "This assignment was not clear to me so I did not like it or see its usefulness."
- "Excellent strategy; helped to really connect concepts to application."
- "We should be asked to annotate outputs more frequently."

6. Discussion and Implications

Reviewing the results brings to mind an important question: why is there an apparent reversal between preferences and ratings for the assessment strategies? Overall, students preferred the structured computer assignments, but they generally gave better ratings to the open-ended assignments. One explanation for the switch may lie in the time required for the open-ended assignment, relative to the structured assignments where the instructor directs the inquiry. However, basing data analysis within the context of an actual research problem, as in the open-ended assignments, can be a strong motivator for students, and the experience requires the student to become an active participant in their own learning process. The structured assignments, with questions already posed by the instructor, provide focused direction and may have less cognitive load for the student, so in a sense these assignments might be considered easier - or, in fact, safer in terms of achieving a high grade. Structured assignments allow less flexibility and creativity in working towards a solution. Their purpose is very different from that of the open-ended assignments - although structured activities such as annotating output and answering a series of guided questions pertaining to an analysis are important steps toward building competence and the convergence of procedural and conceptual knowledge.

Judging from ratings and comments from the open-ended responses, students are aware that the data analysis process, from start to finish with a written summary of results, is a difficult task, particularly as our research problems become increasingly complex. Educators and researchers know that the research process is challenging, but that it also brings great rewards; this is the piece that needs attention in our course assessments. When the assessments that are utilized slowly guide the student to become decision makers, research directors, data analyzers, and report writers, students can begin to see personal rewards.
through their own development as competent researchers. Certainly, the sports-fitness student mentioned earlier in this paper would have benefited from a course environment better focused on developing these skills.

Reflecting on student perceptions of the assessment strategies that are used in class is a good first step towards improving instruction and making assessments more meaningful. Since students are likely to use the multivariate techniques they are learning as they begin their own academic and professional careers, it is important that the statistics educator ensure that accurate learning has occurred and not lose sight of student needs for applicability outside the classroom. In addition, the process of evaluating student perceptions of classroom activities models good strategies for potential future teachers of a complex domain such as multivariate analysis. Shepard (2000) supports this strategy as well: "If we want to develop a community of learners - where students naturally seek feedback and critique their own work - then it is reasonable that teachers would model this same commitment to using data systematically as it applies to their own role in the teaching and learning process" (p. 12).

Although the size and composition of the sample present some limitations to this study, it remains important for statistics educators to review and adjust the activities we require of our students as they struggle through such a complex domain. In general, my findings from this course-specific survey suggest the following:

a. On average, students preferred the structured assessment formats (instructor-directed assignments, annotating specific computer output) to those requiring greater self-direction (open-ended assignments, article analyses). However, the open-ended activities were found to be more challenging and more appropriate to the task of learning multivariate statistics.

b. A variety of assessment strategies offers an opportunity for different kinds of learning and provides an opportunity for different kinds of feedback to the student about the material being learned. The qualitative comments suggest that it is the combination of formats - rather than one particular kind of assessment - that students seem to find beneficial.

c. Qualitative comments indicate that students are willing to try different kinds of assessment formats, and are eager to help us learn what works best. This tends to illustrate what Pellegrino, Baxter and Glaser (1999) have described as "the perceived power of assessments to effect teaching and learning in optimal ways" (p.338), offering an opportunity to strengthen the connection between assessment and instruction.

This study has helped to refine and refocus my course assignments for better student learning outcomes. In terms of course structure, the findings suggest a need for additional, and perhaps briefer, structured assignments in order to provide more student practice prior to completing the open-ended assignments. More in-class examples using annotated output and additional assignments involving annotation and interpretation of computer output are planned, since the process of annotation, along with writing up a corresponding results section, can mimic the self-explanation strategies that are so important to good learning (Chi, Bassok, Lewis, Reimann, and Glaser 1989; Chi, de Leeuw, Chiu, and LaVancher 1994). For the open-ended assignments, better guidance should be provided related to the instructor’s expectations. A project time-line, perhaps with a comprehensive checklist for project milestones (the research question, a brief lit-review, the analysis method, as examples) and a variable summary checklist might enhance the learning process. Finally, although students did not care for the article analysis component, I believe that these article critiques are important. In many programs of study, article critiques form a part of comprehensive examinations. However, suitable class time needs to be reserved for review of these articles as a class. An option to consider in the future is to establish group presentations of articles. This may have the added benefit of using several articles that may be of interest.
to smaller groups of students from particular majors.

I am still working on the best "mix" of assessments, but this personal study has provided a wealth of information on where to improve my own instruction and assessment practices. Researchers interested in developing alternative forms of assessment in their own classrooms may want to review the following additional models of assessment strategies: individual or group projects (Dietz 1993; Cobb 1992; Garfield 1993; Holmes 1997; O'Connell 1997b); poster session displays (Denson 1992); oral assessment (Jolliffe 1997); and creative multiple choice (Wild 1997). Statistics educators also may need to consider how course structure as well as student perceptions of assessment practices in advanced courses are impacted by course structure and assessment strategies that students may be accustomed to from their introductory courses. Since it is not always the case that instructors of introductory statistics courses also teach the advanced-level courses, statistics educators should work together to achieve the seamless integration of effective assessments into courses at all levels.

Acknowledgments

The author gratefully acknowledges the participation of her EPSY 346 students in developing this work, as well as the helpful comments and suggestions of several reviewers during the preparation of this manuscript.

References

Chi, M. T. H., Bassok, M., Lewis, M., Reimann, P., and Glaser, R. (1989), "Self-Explanations: How Students Study and Use Examples in Learning to Solve Problems," Cognitive Science, 13, 145-182.

Chi, M. T. H., de Leeuw, N., Chiu, M., and LaVancher, C. (1994), "Eliciting Self-Explanations Improves Understanding," Cognitive Science, 18, 439-477.

Cobb, G. (1992), "Teaching Statistics," in Heeding the Call for Change: Suggestions for Curricular Action, ed. L. Steen, Washington, DC: Mathematical Association of America, MAA Notes Volume 22, 3-43.

Dallal, G. E. (1990), "Statistical Computing Packages: Dare We Abandon Their Teaching to Others?" The American Statistician, 44(4), 265-266.

Dant, R. P., Lumpkin, J. R., and Bush, R. P. (1990), "Private Physicians or Walk-In Clinics: Do the Patients Differ?" Journal of Health Care Marketing, 10(2), 25-35.

Denson, P. (1992), "Preparing Posters Promotes Learning," The Mathematics Teacher, 85(9), 723-724.

Dietz, E. J. (1993), "A Cooperative Learning Activity on Methods of Selecting a Sample," The American Statistician, 47, 104-108.

Fong, G. T., Krantz, D. H., and Nisbett, R. E. (1986), "The Effects of Statistical Training on Thinking About Everyday Problems," Cognitive Psychology, 18, 253-292.
Gal, I., and Garfield, J. B. (eds.) (1997), *The Assessment Challenge in Statistics Education*, Amsterdam, The Netherlands: IOS Press.

Garfield, J. B. (1993), "Teaching Statistics Using Small-Group Cooperative Learning," *Journal of Statistics Education* [Online], 1(1). (ww2.amstat.org/publications/jse/v1n1/Garfield.html)

----- (1994), "Beyond Testing and Grading: Using Assessment to Improve Student Learning," *Journal of Statistics Education* [Online], 2(1). (http://ww2.amstat.org/publications/jse/v2n1/Garfield.html)

----- (1995), "How Students Learn Statistics," *International Statistical Review*, 63(1), 25-34.

----- (2000), "The Role of Statistical Reasoning in Learning Statistics," Proceedings of the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Hair, J. F., Anderson, R. E., Tatham, R. L., and Black, W. C. (1992), *Multivariate Data Analysis with Readings* (3rd ed.), New York: Macmillan Publishing.

Holmes, P. (1997), "Assessing Project Work by External Examiners," in *The Assessment Challenge in Statistics Education*, eds. I. Gal, and J. Garfield, Amsterdam: IOS Press, 153-164.

Jolliffe, F. (1997), "Issues in Constructing Assessment Instruments for the Classroom," in *The Assessment Challenge in Statistics Education*, eds. I. Gal, and J. Garfield, Amsterdam: IOS Press. 191-204.

Judd, C. M., and McClelland, G. H. (1989), *Data Analysis: A Model-Comparison Approach*, San Diego, CA: Harcourt, Brace, Jovanovich.

Karacostas, D. D. , and Fisher, G. L. (1993), "Chemical Dependencies in Students With and Without Learning Disabilities," *Journal of Learning Disabilities*, 26(7), 491-495.

Keselman, H. J., Huberty, C. J., Lix, L. M., Olejnik, S., Cribbie, R. A., Donahue, B., Kowalchuk, R. K., Lowman, L. L., Petoskey, M. D., Keselman, J. C., and Levin, J. R. (1998). "Statistical Practices of Educational Researchers: An Analysis of Their ANOVA, MANOVA, and ANCOVA Analyses," *Review of Educational Research*, 68(3), 350-386.

King, F. B., Harnar, M., and Brown, S. W. (2000), "Self-Regulatory Behavior Influences in Distance Education," *International Journal of Instructional Media*, 27(2), 147-155.

Litt, M.D, Kalinowski, L., and Shafer, D. (1999), "A dental fears typology of oral surgery patients: Matching patients to an anxiety intervention," *Health Psychology*, 18(6), 614-624.

NCTM (1995), *Assessment Standards for School Mathematics*, Reston, VA: National Council of Teachers of Mathematics.

Nisbett, R. E., (ed.) (1993), *Rules for Reasoning*, Hillsdale, NJ: Lawrence Erlbaum Associates.

O’Connell, A. A. (1997a), "Final Report: Evaluating an African-Centered Adolescent Pregnancy Model," (APR 000966-01), *Performance Report Submitted to: National Institutes of Health, Office of Population Affairs*.

----- (1997b), "Not a Spectator Sport: Constructing Knowledge through Practical Statistics," *Paper*
Presented at the Twenty-Sixth Annual Meeting of the Mid-South Educational Research Association, Memphis, TN, November 12-14.

----- (1999), "Understanding the Nature of Errors in Probability Problem-Solving," *Educational Research and Evaluation*, 5(1), 1-21.

Pellegrino, J. W., Baxter, G. P., and Glaser, R. (1999), "Addressing the "Two Disciplines" problem: Theories of Cognition and Learning with Assessment and Instructional Practice," in *Review of Research in Education*, eds. Asghar Iran-Nejad, and P. David Pearson, Washington, DC: American Educational Research Association, 307-353.

Pintrich, P. R. (1994), "Continuities and Discontinuities: Future Directions for Research in Educational Psychology," *Educational Psychologist*, 29(3), 137-148.

Rittle-Johnson, B., and Alibali, M. W. (1999), "Conceptual and Procedural Knowledge of Mathematics: Does One Lead to the Other?" *Journal of Educational Psychology*, 91(1), 175-189.

Searle, S. R. (1989), "Statistical Computing Packages: Some Words of Concern," *The American Statistician*, 43 (4), 198-190.

Shaughnessy, J. M. (1981), "Misconceptions of Probability: From Systematic Errors to Systematic Experiments and Decisions," in *Teaching Statistics and Probability*, eds. A. P. Shulte, and R. J. Smart, *NCTM 1981 Yearbook*, Reston, VA: National Council of Teachers of Mathematics, 90-100.

Shepard, L. A. (2000), "The Role of Assessment in a Learning Culture," *Educational Researcher*, 29(7), 4-14.

Stark, M. J., Tesselaar, H. M., O'Connell, A. A., Person, B., Galavotti, C., Cohen, A., and Walls, C. (1998), "Psychosocial Factors Associated with the Stages of Change for Condom Use Among Women at Risk for HIV/STDs: Implications for Intervention Development," *Journal of Counseling and Clinical Psychology*, 66(6), 967-978.

Stevens, J. (1996), *Applied Multivariate Statistics for the Social Sciences* (3rd ed.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Tversky, A., and Kahneman, D. (1983), "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment," *Psychological Review*, 90(4), 293-315.

Tatsuoka, M. M. (1988), *Techniques for Educational and Psychological Research* (2nd ed.), New York: Macmillan Publishing Company.

Von Glasersfeld, E. (1987), "Learning as a Constructive Activity," in *Problems of Representation in the Teaching and Learning of Mathematics*, ed. C. Janvier, Hillsdale, NJ: Lawrence Erlbaum Associates, 3-17.

Wild, C., Triggs, C., and Pfannkuch, M. (1997), "Assessment on a Budget: Using Traditional Methods Imaginatively," in *The Assessment Challenge in Statistics Education*, eds. I. Gal, and J. Garfield, Amsterdam: IOS Press, 205-220.
