**Assessing the Ability of Psychology and Social Behavior Undergraduates
to Interpret Statistical Findings in Social Science Reports**

Joanne Frattaroli (jfrattar@uci.edu) & Linda Levine (llevine@uci.edu)
Department of Psychology and Social Behavior

Planning the Assessment

Identified Learning Outcomes: Frattaroli, Levine and Dooley (former Chair) led discussions about WASC goals and assessment techniques at two faculty meetings. All faculty who were present at the meetings participated in choosing and revising goals and in developing the assessment strategy summarized below. Four learning goals were drafted, based in part on the learning goals for psychology majors published by the American Psychological Association (Task Force on Undergraduate Major Competencies, 2002; http://www.apa.org/ed/pcue/taskforcereport2.pdf), and they are as follows:

1. Ability to evaluate social science claims critically in terms of research validity
2. Ability to interpret basic descriptive and inferential statistics found in social science reports
3. Knowledge of basic concepts and theories over the broad field of psychology
4. Ability to apply psychological concepts by framing testable hypotheses, gathering and synthesizing appropriate data, and articulating clear results

Assessment Plan: We decided to measure student performance on Goal 2: "Ability to interpret basic descriptive and inferential statistics found in social science reports," by creating a short assessment of statistics knowledge (designed to take approximately 20 minutes). Skills required for Goal 2 are addressed in SE10 (Research Design), are taught as the main focus of SE13 (Statistical Analysis), and are practiced in SE194W, SE195, P196, and H190.

Collecting the Data

Creating the Assessment Tool: To create the assessment tool, we first identified specific statistics goals relevant to overarching goal to *interpret basic descriptive and inferential statistics found in social science reports:* These specific statistics goals were as follows:

(A) To gain an understanding of basic concepts concerning the organization and display of quantitative data in social science research, including having the ability to identify the scales of measurement; identify independent and dependent variables; and identify the uses and properties of, interpret, and distinguish between various graphical displays (e.g., histograms, scatterplots).

(B) To gain an understanding of basic statistics, including having the ability to identify, interpret, and compute measures of central tendency (mean, median, mode); interpret and distinguish between the various measures of variability (e.g., range, standard deviation, variance); interpret effect sizes; and interpret a correlation coefficient

(C) To gain an understanding of the basic properties and uses of the normal curve, including having the ability to compute (and describe the meaning of) a z-score; to identify the properties of the normal curve; and to identify the shape of a frequency distribution (e.g., normal, positively skewed, etc.); and

(D) To gain an understanding of probability and basic inferential statistics, including the ability to calculate a probability using the theoretical approach, interpret the meaning of a p-value (in the context of inferential statistics), identify a random sample, identify the basic properties of the central limit theorem, identify the principles and purposes of null hypothesis testing/inferential statistics, and interpret the results of and identify when it is appropriate to use various inferential tests (e.g., one-sample t-test, chi-squared).

Based on these specific goals, we selected multiple choice questions from undergraduate statistics and research methods textbook test banks, from previous SE10 (Research Design) and SE13 (Statistical Analysis) midterm and final exams, and from relevant questions from Psychology GRE preparatory guides. The resulting pool of questions was pretested in Fall 2009 at the end of a SE13 course. We then chose the 20 questions, from the pool of pretested questions, that best addressed the specific goals. These questions were chosen (and then revised) by Frattaroli, Levine, and Prause, a faculty member who regularly teaches statistics in our School. In addition to the 20 statistics-related questions, the measure asked students to provide descriptive information, such as whether they had taken SE10 or SE13 at UCI and whether they had taken statistics at another college/university.

Administering the Assessment: All PSB faculty, lecturers, and graduate students who were teaching undergraduate courses in the Spring 2010 quarter were invited to administer the assessment in class during the first week or two of the quarter. Instructors of 17 (out of 24) different courses agreed to participate, resulting in the data from approximately 1800 students.

Preliminary Results

Value added of the SE13 series: To assess the "value added" from taking our SE13 course (for which SE10 is a prerequisite), we compared the average performance of students with no methods or statistics training (including from other universities or other departments; the *no training group*) to students who had completed our SE13 course (the *SE13 group*). The performance of the 175 students who completed our SE13 course was significantly better than the performance of the 590 students in the no training group, p < .05 both with and without GPA as a covariate in the analysis.

We also examined students' performance on individual questions in the context of the four specific statistics goals listed earlier. For two of the three questions addressing **Statistics Goal A**, the *SE13 group* performed significantly better than the *no training group*; both groups performed extremely well on the third question (concerning the interpretation of graphic displays). For two of the five questions addressing **Statistics Goal B**, the *SE13 group* performed significantly better than the *no training group*; for the remaining questions, two questions (regarding variability and the interpretation of correlations) showed good performance overall, and one question (regarding the interpretation of effect sizes) showed very poor performance overall. For four of the five questions addressing **Statistics Goal C**, the *SE13 group* performed significantly better than the *no training group*; the remaining question (regarding the interpretation of z-scores) showed good performance overall. For five of the seven questions addressing **Statistics Goal D**, the *SE13 group* performed significantly better than the *no training group*; for the remaining two questions, one (regarding understanding of probability) showed good performance overall and one (regarding the interpretation of inferential tests) showed very poor performance overall.

Community college training: To assess the appropriateness of our current policy of allowing statistics courses at community colleges to be taken in place of our SE13 course, we compared the average performance of the students who had taken statistics only at a community college (and who had taken a research methods course either at another institution or at UCI) to students who had completed our SE13 course. The performance of the 126 students who had taken statistics at a community college was significantly poorer than the performance of the 175 students who had completed our SE13 course, p =. 027 when controlling for GPA (p = .073 without GPA as a covariate).

We also compared the performance of the two groups on individual questions in the context of the four statistics subgoals listed earlier. For all three questions addressing **Statistics Goal A**, no significant difference was found between the performance of the two groups. For one of the five questions addressing **Statistics Goal B**, the *SE13 group* performed significantly better than the *Community College group;* this question was concerned with identifying, interpreting, and computing measures of central tendency. For one of the five questions addressing **Statistics Goal C**, the *SE13 group* performed significantly better than the *Community College*

group; this question was concerned with computing (and interpreting the meaning of) a z-score. For two of the seven questions addressing **Statistics Goal D**, the *SE13 group* performed significantly better than the *Community College group*; these questions were concerned with identifying the basic properties of the central limit theorem and interpreting the results (and identifying when it is appropriate to use) various inferential tests. For one of the seven questions addressing Subgoal D, the *Community College group* performed significantly better than the *SE13 group*; this question was concerned with identifying a random sample.

Conclusions

Value Added: Although it is clear that overall, students are gaining statistical knowledge from the SE10-SE13 series, there appears to be some statistics subgoals that students have already met elsewhere (e.g., interpretation of graphic displays) and some subgoals that are not being adequately met in our course (e.g., interpretation of effect sizes). We will discuss the following options for improvement at a future faculty meeting: *Option A:* Give students a pretest at the beginning of each quarter to identify which learning subgoals have already been met elsewhere, and have instructors focus instruction time on other learning subgoals. *Option B:* Inform current and future SE13 instructors of the results of the current assessment, with advice to focus less attention on subgoals showing uniformly good performance and more attention on subgoals showing uniformly poor performance. *Option C:* Increase the instructional time offered in the SE13 course by adding required (instead of optional) discussion sections. A combination of these options could also be considered.

Community College Training: Although it is clear that overall, the statistical training that students from community colleges are receiving is not as strong as the statistical training that we offer in our SE13 course, there appears to be several statistics subgoals in which community college students are receiving sufficient (if not superior) training. We will discuss the following options for improvement at a future faculty meeting: *Option A***:** Increase the selectivity with which statistics courses from community colleges are accepted in place of SE13. Instead of approving any statistics course, students could be required to submit a petition (through the Social Ecology Undergraduate Advising Office) to the instructor teaching the course, in which a syllabus and relevant materials are provided. The instructor would decide on a case-by-case basis whether the other-university course appears to be appropriately similar to SE 13 in content and rigor. *Option B:* Incorporate a "statistics refresher" into upper-division courses that all students (including those who transfer from community colleges) take prior to graduation (e.g., SE194W, SE195). This "statistics refresher" could take the form of one or two lectures (taught by the instructor of record or by a faculty member who regularly teaches statistics), during which time the material from these "poor performance" subgoals would be taught/reviewed. (This material would be new to some students while it would be a review for others). *Option C:* Students who wish to get credit for a statistics course taken a community college could be required (or strongly encouraged) to participate in a mini-workshop (e.g., 1-3 hours) in which the "poor performance" subgoals would be taught.

Faculty Involvement and Sustainability

Our general approach has been to make this process meaningful for faculty by selecting a question that had previously been discussed by our faculty and was already of moderately strong interest. As a result, the outcome of the assessment will be useful for informing possible changes in the curriculum that were under consideration. Faculty have been involved throughout the assessment process. A draft of the assessment measure was distributed and then brought to a faculty meeting for discussion, and faculty were involved in the data collection. Similarly, now that assessment data have been collected and analyzed, a summary of the assessment process and the results will be distributed to faculty. The results will be presented at an upcoming faculty meeting, and faculty will have an open discussion about the implications and required actions suggested by the results.   Because this process has now been completed once successfully, we expect strong faculty interest in conducting future assessments of this type on the additional learning goals.