Dean Jan Colijn is in the process of reviewing reports on assessment progress from all General Studies coordinators. We have summarized some of the work from the Writing program, and from minors in Women’s Studies and Holocaust and Genocide Studies in this issue. You can get more information about the assessment methods and the findings to date from the coordinators of these minors.

**Writing/ BASK Writing**

Jack Connor, Interim Coordinator and Heather McGovern, Coordinator

The Writing program has been conducting assessment of student portfolios in BASK 1101, College Writing, since 2005. Initially all portfolios were evaluated, but the program has since decided to assess only a random sample. In evaluating portfolios for 2007 – 2008, the program used a rubric to allow for more data collection. A process similar to the early BASK assessment was instituted in the spring.

**Assessment in the Social Work Program**

William Reynolds, Assistant Professor of Social Work

**Assessment in the School of General Studies**

**Putting General Studies Assessment Findings to Use**

March 2009

Assessment of a Lesson on Sampling in a Social Work Research Methods Class

William Reynolds, PhD, LCSW

**Introduction**

"No other part of the social work curriculum has been so consistently received by students with as much groaning, moaning, eye-rolling, hyperventilation and waiver-strategizing as the research courses" (Epstein, 1987, p. 71).

Research Methods in Social Work (SOWK 3102) is a required course for social work majors at Stockton. Teaching this course presents challenges for a number of reasons. For example, research suggests that social work students are less interested in research and have higher levels of research anxiety than students in disciplines like business and psychology (Green, Bretzin, & Leininger, 2001). In addition, undergraduate social work students were found to have greater math anxiety than a cross-section of university students (Royse & Rompf, 1992), which may affect their attitude toward and ability to engage in the research course.

The extent to which students’ feelings and attitudes about research affect their achievement of course objectives is unclear. Locally, assessment data have been collected on students’ perceptions of the extent to which Research Methods in Social Work (SOWK 3102) course objectives were achieved. These end-of-semester evaluations have not been administered consistently over time, so data are limited. However, on their face, recent results suggest that students did not believe they had adequately achieved at least three of twelve course objectives. This finding should be interpreted cautiously; data have not been

(Continued on page 2)
collected over a long enough period of time to allow for the systematic comparison of different sections of the Methods course or comparisons between evaluations of this course and other courses in the social work curriculum. Nonetheless, it is apparent that many social work students experience considerable difficulty with the Research Methods course. To date, a limited amount of empirical work has been done to systematically formulate and investigate propositions about why this might be the case or to assess interventions that could improve the situation.

One of the Research Methods course objectives students have evaluated as incompletely achieved is related to sampling theory and procedures. This objective states that upon completion of the course students will be able to: “Identify the purpose and procedures of probability and non-probability sampling methods and the relationship of such methods to generalizability.” The primary purpose of the present assessment study was to evaluate the extent to which a two period lesson on sampling theory and procedures achieved six specific objectives related to the more general course objective. A secondary goal was to provide a rough validity check on student self-reports of achievement of the sampling lesson objectives by: 1) collecting data on their content knowledge and 2) having an expert observer evaluate instructor effectiveness. Lessons were also digitally video recorded so that outside assessors can evaluate instructor effectiveness at addressing lesson objectives; however, these assessments have not yet taken place.

Methods

Sample

Participants were students in a section of SOWK 3102, Research Methods in Social Work. Of 21 students registered for the course, 16 were present during session one and 16 were present during session two, but different students were absent during each session. Due to these absences, complete data (i.e., pre-test and post-test) were available for only 12 students on one measure and 13 students on the other.¹ No identifiers were attached to study measures, and results will be presented in the aggregate, preserving participant anonymity.

Research Design

This study used a single group, pre-test post-test design intended to evaluate the effectiveness of a two day lesson on sampling theory and methods. This lesson demonstrated key principles of sampling theory and methods by using the class roster as an example of a population and a list of present students as the sampling frame. Students were asked to provide data on four variables: number of siblings, number of social work courses taken, class standing, and gender. Random samples of different sizes were taken using these variables to demonstrate the relationship between sample size and sampling error, and to illustrate concepts such as central tendency and variability (Singleton, 1989).

There were two primary outcome measures designed for this study: 1) the Perception of Objectives Survey, a seven item questionnaire designed to assess students’ understanding of seven important concepts identified as lesson objectives at the beginning of the first of the two classes (one item was not explicitly related to sampling), and 2) the Sampling Content Knowledge Quiz, 10 multiple choice questions taken from a test item bank provided with the course text. Both measures were administered before any classroom instruction had occurred but after students were supposed to have completed reading and homework assignments on the material (Time 1), and after approximately three hours and thirty minutes of classroom instruction and small group activities had taken place (Time 2).

Data Analysis

Data were analyzed using the Paired Samples T-Test,

¹As this study meets the criterion for exemption under 45 CFR 46.101(b)(1), IRB review and student informed consent were not required. 45 CFR 46.101(b)(1) provides an exemption for: “Research conducted in established or commonly accepted educational settings, involving normal educational practices, such as (i) research on regular and special education instructional strategies, or (ii) research on the effectiveness of or the comparison among instructional techniques, curricula, or classroom management methods.”
which compares the means of two variables. It computes the difference between the two variables for each case, and tests to see if the average difference is significantly different from zero.

**Results**

Table 1 displays data from the *Perception of Objectives Survey* and the *Sampling Content Knowledge Quiz*. Paired Samples T-tests indicate significant increases in agreement with all seven objectives. Students perceived themselves to have a greater understanding of important sampling concepts by the end of the two day lesson than they had before instruction began, when they had (ideally) read the sampling chapter in their text.

Contrary to results on the *Perception* measure, findings presented in the last line of Table 1 indicate that there was no significant difference in the percentage of correct responses on the *Sampling Content Knowl-

**Discussion**

We must display the evidence of teaching and learning (and their embarrassments) through the multiple legitimate narratives we create about our work and our students’ fates (Shulman, 2007, p. 25).

Based on results on the *Sampling Quiz* my lesson failed. My instruction resulted in no improvement in student knowledge of the sampling content evaluated by this measure. Students answered the same number of questions correctly after two periods of instruction as they did when they had only read (or not) the chapter on sampling. On the other hand, they reported a significant increase in their understanding of important concepts presented during the lesson. The questions for this discussion, then, are:

(Continued on page 4)

**Table 1**

**Perception of Objectives Being Met (Mean Response)**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Paired Samples T-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-</td>
<td>Post-</td>
</tr>
<tr>
<td>1. I have an understanding of the purposes of sampling in social work research.</td>
<td>4.17</td>
<td>5.00</td>
</tr>
<tr>
<td>2. I have an understanding of the most common sampling terms.</td>
<td>3.50</td>
<td>4.67</td>
</tr>
<tr>
<td>3. I have an understanding of the criteria for evaluating the quality of a research sample.</td>
<td>3.58</td>
<td>4.75</td>
</tr>
<tr>
<td>4. I have an understanding of the differences between sampling strategies for qualitative and quantitative research.</td>
<td>3.08</td>
<td>4.42</td>
</tr>
<tr>
<td>5. I have an understanding of the difference between probability and non-probability sampling.</td>
<td>3.08</td>
<td>4.58</td>
</tr>
<tr>
<td>6. I have an understanding of the difference between sampling bias and sampling error.</td>
<td>4.08</td>
<td>5.17</td>
</tr>
<tr>
<td>7. I have the ability to interpret multiple graphic displays of descriptive data.</td>
<td>3.58</td>
<td>4.92</td>
</tr>
</tbody>
</table>

**Change in Content Knowledge (Percent Correct)**

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Paired Samples T-test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre-</td>
<td>Post-</td>
</tr>
<tr>
<td>10-item Multiple Choice Sampling Quiz</td>
<td>52.31</td>
<td>49.23</td>
</tr>
</tbody>
</table>
of 2007 as a pilot project for GEN 1120, Rhetoric & Composition. A second pilot was conducted for more sections of GEN 1120 for the 07-08 academic year and for this pilot faculty used a sample of portfolios and a rubric. So far, the data have been used to support changes to the assessment plan, and there has been discussion about extending the portfolio assessment to other courses.

Women’s Studies

Ellen Mutari, Coordinator

The Women’s Studies Program developed and used a rubric to evaluate specific learning outcomes in the Seminar in Feminist Theory capstone projects. Six projects were evaluated by three Women’s Studies faculty in the spring of 2008. Following the evaluation, the program developed suggestions for improving the rubric as well as the capstone project. The rubric was edited to include aspects of the project that were not originally assessed, such as using a literature review to define a core issue and specific writing and presentation skills. For the capstone project, a three phase plan was discussed to better link more of the theoretical course content with the project. These changes are currently being implemented, and there has been discussion about assessing the program’s other core course, Perspectives on Women.

Minor in Holocaust and Genocide Studies

Carol Rittner, Coordinator

For the 2008 - 2009 academic year, the Minor in Holocaust and Genocide studies has identified and is measuring several important student learning objectives at the course level. Faculty will review how well each objective has been met and make changes to better meet the goals for each course. In addition, the coordinator has developed and administered a survey to students who graduated with an HGST minor. The survey utilizes both Likert Scale and open-ended items to assess specific learning outcomes and determine overall program effectiveness.

1) What do these results say about the effectiveness of instruction for the two period sampling lesson? and 2) What can we learn about assessment from the apparently contradictory study findings?

A brief review of commonly cited principles of assessment can help answer these questions. First, and perhaps most significantly for the interpretation of present results, asking the right questions is critical (Moore, 2007; Shavelson, 2007). In conceiving the sampling assessment, I used an article that described what appeared to be an effective lesson for teaching sampling principles (see Singleton, 1989 for the lesson format); however, I used the course textbook chapter objectives and quiz questions that came with the text to evaluate that lesson. The problem was that I never explicitly established that the chapter objectives and quiz questions would actually assess the content of the lesson I taught. The lesson was intended to demonstrate the major principles of probability sampling; key elements of non-probability sampling were to be presented via lecture and discussion in the latter portion of the second day’s lesson. The probability sampling demonstration ran long, however, and I didn’t discuss non-probability sampling until the following week. Therefore, the three questions on sampling for qualitative research that were included on the quiz were not credible measures of the content of the lesson. Of course, students did not score substantially better on the questions that were explicitly about probability sampling. I assumed that if I taught a good lesson students would have the content information they needed to answer the questions. However, this assumption holds only when multiple assessment measures carefully reflect lesson content, which was not the case here.

Although I used multiple measures, which is a second consistent recommendation for assessment (Shavelson, 2007; Shulman, 2007), my goals in conducting this assessment were not entirely clear, a limitation that I believe is reflected in the results. Successful assessment has a conceptual framework and goals and objectives that are clearly aligned with the measurable outcomes of the performance activ-
ity (Huba & Freed, 2000; Moskal, 2003). Upon reflection, my primary goal was to evaluate the effectiveness of my instruction. Based on the assessment measures, however, the goals might have appeared to be 1) to assess changes in student content knowledge and 2) to assess changes in student perceptions of their understanding of general sampling concepts. While these measures may reflect some aspects of teaching effectiveness, I evaluate my own effectiveness based on additional factors such as student engagement in classroom activities (small group work was an instructional technique during both classes), verbal participation, and less measurable elements such as facial expressions, tones of voice, and quality of questions asked. None of these elements was captured by the assessment measures; however, a senior faculty member who observed the second day’s lesson noted, “A few students did most of the voluntary feedback, but over time most students participated, including those called upon by Professor Reynolds. He was particularly skilled in translating technical matters into clear and practical language, including using the class itself as a sampling frame.” Perhaps with a better assessment design and measures, qualitative feedback like this would be reflected in assessment results. This result will only occur, however, when assessment goals are clearly articulated and measures carefully address those goals.

Sharing the results of assessment is the final principle relevant to this study (Ekman & Pelletier, 2008). Although methodologically limited, this assessment was an example of a (very) small scale research study. As such, adherence to principles of research ethics requires a transparent process, warts and all. Based on lack of student improvement on the content measure, the findings of the present assessment study could be perceived as negative and, as so many results are, abandoned to the proverbial file drawer never to see the light of day. Doing so, however, would have prevented me from thinking more thoroughly about the function and process of assessment and, perhaps, kept knowledge from others who might improve on the methodological limitations described here. Had I no intention of organizing these results and disseminating them to my peers, it is unlikely that I would have put the time and effort into evaluating their meaning and the role of assessment in my teaching. Thus, I’m fairly certain that the intent to share results led to a more reflective and useful process and product.

**Conclusion**

Administrators may not see that faculty resistance can also indicate that effective teachers perceive assessment activities to be disconnected from the real work of the classroom—advancing student learning—work that brings genuine pleasure to them. Deep divisions can therefore develop on our campuses if those responsible for assessment mistake means for ends (Wiley, 2003, ¶1).

The Social Work Program is currently involved in a self-study for reaffirmation from its accrediting body, which is basically a program-wide exercise in assessment. In addition, in the first faculty assembly meeting I attended in fall, 2008, I learned that the college administration is encouraging the faculty to contribute to a “culture of assessment,” which is intended to involve both individual and institutional efforts. Given these imperatives and the time commitment necessary to do them well, one might reasonably ask at what point the assessment tail begins to wag the teaching dog. As Wiley (2003) notes above (and I have heard frequently from colleagues), assessment efforts that are in service of ends that are either opaque or perceived to be disconnected from the core activities of teaching will be received with indifference, if not outright hostility, by the faculty members who ultimately must carry them out.

My experience suggests that when assessment takes place for the explicit purpose of improving one’s teaching, the value of knowledge gained from scrutiny of both process and outcomes outweighs burdens, perception of which may prevent one from undertaking assessment in the first place. Reflective teachers are involved in a continuous process of assessment whether or not they identify it as such. They establish goals (e.g., course objectives), ask questions of themselves and their students to assess their achievement (e.g., course evaluations), and they
seek multiple sources of feedback to evaluate effectiveness (e.g., peer observations). However, while these processes are valuable, they are not transparent, as they occur on an individual level without the explicit intent to share effective methods and elicit feedback about less successful practices. Although I would have learned something about my teaching from simply performing the assessment described here, its pedagogical value grew exponentially when I made the decision to share it with peers.

REFERENCES


