FACULTY FORUM

Factors Influencing Teaching Evaluations in Higher Education

Ginette C. Blackhart, B. Michelle Peruche,
C. Nathan DeWall, and Thomas E. Joiner, Jr.
Florida State University

Past research indicates several factors influencing teaching evaluation ratings instructors receive. We analyzed teaching evaluations from psychology courses during fall and spring semesters of 2003–2004 to determine if class size, class level, instructor gender, number of publications (faculty instructors), average grade given by the instructor, and instructor rank predicted teaching evaluation ratings. Entering predictor variables into a multiple regression analysis concurrently, results indicated that only average grade given and instructor rank significantly predicted instructor ratings. Specifically, higher average grades given by the instructor predicted higher ratings, and graduate teaching assistants received higher overall ratings than faculty instructors.

Teaching evaluations are widely used in assessing courses and the quality of teaching in higher education. In addition, university administrations often use ratings an instructor receives on these evaluations in personnel and merit decisions. Many educators have raised concern about this method of evaluation, however, as previous research has indicated there are several factors that affect the ratings students give instructors on teaching evaluations. Variables such as physical attractiveness of the instructor (e.g., O'Reilly, 1987), research productivity (Allen, 1996), class size (e.g., Hamilton, 1980; Haslett, 1976), gender of the instructor and student (e.g., Lueck, Endres, & Caplan, 1993), class level (e.g., Haslett, 1976), instructor rank (e.g., Harper, 1991; Schuckman, 1990) as well as expected grades (Goldberg & Callahan, 1991; Hamilton, 1980) or grades given by the instructor (Howard & Maxwell, 1980, 1982) have all significantly predicted teaching evaluation ratings.

Given the importance of teaching evaluations for instructors and their respective educational institutions, the fact that these variables substantially affect the ratings students give instructors on teaching evaluations should be of concern. The majority of these studies, however, have only examined the impact these factors have on teaching evaluations separately. Our study therefore evaluated how many of these variables, such as class size, class level (graduate vs. undergraduate course), instructor gender, average grade given by the instructor, instructor rank (graduate teaching assistant [TA], assistant professor, associate professor, and full professor), and number of publications (for faculty members only) concurrently affected ratings on teaching evaluations, specifically in psychology courses. The purpose of using this method was to allow these factors to control for one another to evaluate how each variable uniquely affected instructor rank, thus detangling unique predictors.

Method

Participants

Students taking courses within the psychology department at Florida State University (FSU) completed the teaching evaluations ($N = 9,240$). We did not gather information on student characteristics, such as gender or age, to maintain anonymity of the ratings.

Materials

The evaluation instrument consisted of a questionnaire (the State University System Student Assessment of Instruction) used to evaluate all courses and instructors at FSU. We used ratings on the item "overall assessment of instructor" as the dependent variable; students could rate instructors on this item as *excellent, very good, good, fair,* or *poor.*

Procedures

We collected data from teaching evaluations obtained from all courses (137 undergraduate courses, 30 graduate courses) offered by the psychology department during two consecutive semesters (Fall 2003 and Spring 2004) at FSU. Students completed teaching evaluations at the end of the semester while the instructor was out of the classroom. Faculty members taught 103 of the courses (19 by assistant professors, 40 by associate professors, and 44 by full professors), and graduate TAs taught 64 of the courses offered in the department.

Statistical Analyses

We first conducted correlational analyses between each variable. Next, we entered class size, class level, instructor gender, average grade given by the instructor, number of publications, and instructor rank into a multiple regression analysis (predictor variables were entered simultaneously to control for one another). We then calculated and entered the average rating on "overall assessment of instructor" as the dependent variable on a scale ranging from 5 (*excellent*) to 1 (*poor*).

Results

Results indicated that there were significant correlations between average grade given and instructor rating, $r(143) = .30, p < .01$, and between instructor rank and instructor rat-

ing, $r(166) = -.22, p < .01$. Although the overall regression analysis was significant, $F(6, 134) = 3.51, p < .01$, class size, class level, instructor gender, and number of publications (for faculty only) did not significantly predict the overall assessment of the instructor. The average grade given by the instructor ($\beta = .30, p < .01$) and instructor rank ($\beta = -.22, p < .05$) did, however, significantly predict ratings of instructors (together accounting for 12% of the variance in instructor ratings). Higher average grades given predicted significantly higher ratings, and this effect was similar for both graduate TAs and faculty instructors. For instructor rank, lower rank predicted significantly higher ratings; graduate TAs consistently received better ratings on overall assessment of the instructor than faculty instructors. Among faculty instructors, however, instructor rank had almost no effect on evaluation ratings (see Table 1).

Discussion

This study investigated how class level, class size, instructor gender, number of publications, average grade given to students, and instructor rank affected teaching evaluation ratings. Results indicated that average grade given by the instructor and instructor rank both significantly predicted instructor ratings on teaching evaluations given to students taking psychology courses at FSU. When evaluating average grade given by the instructor, the higher the average grade for the class, the better the ratings of the instructor. In addition, instructor rank significantly predicted instructor ratings, such that graduate TAs received higher ratings than faculty instructors. Faculty rank (full, associate, or assistant professor), however, did not have a significant effect on instructor ratings.

Although past research has shown that average grade given by the instructor affects teaching evaluations (Howard & Maxwell, 1980, 1982), it remains unclear why. Many educators have concluded that the more lenient an instructor is in grading, the better ratings that instructor will receive. In fact, a number of educators believe student evaluations of their instructors have contributed significantly to grade inflation (Hamilton, 1980). Higher grades given by an instructor, however, may indicate that the instructor was effective in teaching the material to his or her students, thus resulting in higher grades and higher ratings. If students are doing well and getting good grades in the class, they may attribute their performance, in part, to the quality of the instructor. In addition, those students who do well in a class might be more en-

gaged in the course topic and the material being learned, leading to better evaluations of the instructor. Thus, although student evaluations of instructors may well lead to grade inflation, providing instructors assurance of better ratings on their evaluations, there are additional reasons, some directly relating to teaching effectiveness, that could also explain why the grades assigned to students significantly affect teaching evaluations.

This study also found that graduate TAs received better evaluations than faculty instructors. Graduate TAs receiving higher evaluations is not attributable to the level of the courses they taught, as graduate TAs taught introductory lecture courses and labs as well as upper level courses within the major. There are, however, several other possible reasons for these findings. Some educators have suggested that the more productive professors are in research, the more their teaching suffers (see Allen, 1996). Our study found, however, that the number of publications faculty members had did not significantly predict ratings on teaching evaluations, suggesting that productivity does not affect teaching evaluation ratings. In addition, Allen found that research productivity correlated positively with teaching effectiveness.

Another reason graduate TAs may have received better teaching evaluations is due to the enthusiasm, effort, and energy they often bring to teaching. Teaching college-level courses is often a novel experience for graduate TAs, possibly leading to greater enthusiasm about teaching as well as greater effort and energy being expended toward the task. In addition, graduate TAs may receive better ratings from their students due to the fact that they typically are closer in age to the students, with perceived similarity being greater between graduate TAs and students than between professors and students. Harper (1991) also suggested that graduate TAs might better express positive attitudes toward their students than faculty members.

Finally, it should be noted that within the psychology department at FSU, all graduate TAs receive training before teaching. Graduate students learn the fundamentals of teaching and gain hands-on experience by giving guest lectures, preparing course materials in advance, and having these materials evaluated and approved by faculty. They receive additional supervision by faculty when teaching, and the department encourages graduate TAs to attend teaching seminars and conferences offered by the university. Thus, it may be the training graduate TAs received that led to higher teaching evaluation ratings.

In conclusion, our study suggests that, within psychology courses, the average grade given by the instructor and instructor rank are among those variables significantly affecting teaching evaluations when controlling for class size, class level, instructor gender, and number of publications (for faculty instructors). Although several studies have reported class size, class level, and instructor gender as significant predictors of instructor ratings, our study found that these variables failed to significantly predict instructor ratings when we concurrently entered our predictor variables into a regression analysis. Taking each of these variables into account when predicting teaching evaluation ratings, only average grade given and instructor rank emerged as significant predictors of instructor ratings. Although the average grade given by the instructor and instructor rank to-

Table 1. Means and Standard Deviations of Teaching Evaluation Ratings for Faculty Instructors and Graduate TAs on the Question "Overall Assessment of Instructor"

	M	SD
Graduate TA	4.48	0.43
Assistant professor	4.26	0.94
Associate professor	4.13	0.51
Full professor	4.18	0.76

Note. For average rating, ratings were based on a scale ranging from 5 (*excellent*), 4 (*very good*), 3 (*good*), 2 (*fair*), to 1 (*poor*). TA = teaching assistant.

gether accounted for only 12% of the variance in teaching evaluation ratings, these results suggest that teaching evaluation ratings are significantly affected by factors that may be unrelated to teaching effectiveness (e.g., average grade given or instructor rank). University administration should therefore take these factors into consideration when using ratings from teaching evaluations to assess an instructor's teaching effectiveness and overall performance.

References

Allen, M. (1996). Research productivity and positive teaching evaluations: Examining the relationship using meta-analysis. *Journal of the Association for Communication Administration, 2,* 77–96.

Goldberg, G., & Callahan, J. (1991). Objectivity of student evaluations of instructors. *Journal of Education for Business, 66,* 377–378.

Hamilton, L. C. (1980). Grades, class size, and faculty status predict teaching evaluations. *Teaching Sociology, 8,* 47–62.

Harper, B. (1991). Who does the better job teaching the basic course: Teaching assistants or faculty? *ACA Bulletin, 77,* 68–70.

Haslett, B. J. (1976). Student knowledgeability, student sex, class size, and class level: Their interactions and influences on student ratings of instruction. *Research in Higher Education, 5,* 39–65.

Howard, G. S., & Maxwell, S. E. (1980). The correlation between grades and student satisfaction: A case of mistaken causation? *Journal of Educational Psychology, 72,* 810–820.

Howard, G. S., & Maxwell, S. E. (1982). Do grades contaminate student evaluations of instruction? *Research in Higher Education, 16,* 175–188.

Lueck, T. L., Endres, K. L., & Caplan, R. E. (1993). The interaction effects of gender on teaching evaluations. *Journalism Educator, 48,* 46–54.

O'Reilly, M. T. (1987). Relationship of physical attractiveness to students' ratings of teaching effectiveness. *Journal of Dental Education, 51,* 600–602.

Schuckman, H. (1990). Students' perceptions of faculty and graduate students as classroom teachers. *Teaching of Psychology, 17,* 162–165.

Note

Effect of Crib Card Construction and Use on Exam Performance

K. Laurie Dickson and Michelle D. Miller
Northern Arizona University

Both normal classroom use and research typically confound crib card construction with crib card use, making it unclear whether students benefit from the process of creating crib cards. We compared the effect of self-constructed and other-constructed crib cards (written by a student research assistant) on undergraduates' (N = 32) multiple-choice exam performance. Performance was better with other-constructed cards than with self-constructed cards. Crib card construction did not facilitate student learning, nor did the use of self-constructed crib cards enhance exam performance.

College instructors occasionally allow students to use a "cheat sheet" or crib card during an exam. Some educators (Dorsel & Cundiff, 1979; Hindman, 1980; Trigwell, 1987) have argued that the process of constructing a crib card may help students learn the material, whereas others (Dickson & Miller, 2005; Whitley, 1996; Whitley & Keith-Spiegel, 2002) have argued that crib cards are disadvantageous for students.

According to the coding hypothesis, the active processes of selecting, organizing, and representing information during crib card construction benefit students (Dorsel & Cundiff, 1979; Hindman, 1980; Trigwell, 1987). In contrast, Whitley (1996; Whitley & Keith-Spiegel, 2002) argued that crib card construction encourages students to study less, as they assume that they do not really need to know or understand the information. Trigwell (1987) stated that crib cards are not helpful because they merely reflect students' knowledge at the time they construct the card and thus do not enhance student exam performance. Dickson and Miller (2005) found that most students reported that the amount of time they spent studying did not increase when they constructed a crib card compared to exam preparation without a crib card. Thus, the time spent constructing a crib card may have diverted students from more productive study activities. Some students reported that they did not learn the course material as well when they had a crib card, which suggests that the process of constructing a crib card interfered with learning. In contrast, a minority of students thought that the process of constructing a crib card was a useful study strategy. Crib card construction and use are typically confounded because they are not examined as separate processes. Thus, it is unclear whether creating a crib card is beneficial for student learning and exam performance.

The purpose of this study was to separate the effect of crib card construction from the effect of crib card use on multiple-choice exam performance. Students in an upper division psychology course used self-constructed crib cards for two of the four exams. For the other two exams, they unexpectedly used a crib card constructed by another student.

Method

Participants and Procedures

Thirty-two child and adolescent development students (82.4% women; 14.7% sophomores, 35.3% juniors, 50.0% seniors; 73.5% psychology majors) constructed and used crib cards (one side of a 5×8 in. [12.5×20 cm] index card) for Exams I and III (self-constructed crib card). The instructor directed students to draw on the course learning objectives and the textbook when constructing the cards. For Exams II and IV, students were not aware that they would be able to use a crib card. As the instructor handed out these exams, she also distributed premade crib cards to students for use during the exams. A student research assistant developed crib cards based on the learning objectives and textbook (other-constructed crib card). The research assistant had neither taken the course nor seen the exams prior to developing the crib cards. The first author reviewed the other-constructed crib card for accuracy and found no errors.

Students took four multiple-choice question exams (50 questions each) derived from the textbook test bank and based on the learning objectives. Seventy-six test items required lower order skills, and 124 test items involved higher order skills. (For definitions of these terms, see Dickson & Miller, 2005.)

Results

To examine the effect of self-constructed versus other-constructed crib cards on exam performance, we computed each student's average score for self-constructed crib card exams ($M = 70.54$, $SD = 9.03$) and other-constructed crib card exams ($M = 76.00$, $SD = 8.49$). Students performed better with the other-constructed cards than with the self-constructed cards, $t(31) = 4.16$, $p < .001$.

Discussion

Exam performance was significantly better when students unexpectedly used a crib card constructed by someone else. A likely explanation for this finding is that when students do not expect to use a crib card, they make a stronger effort to understand the material prior to the exam. In contrast, when they expect to have information available on a crib card during the exam, they do not study the material as effectively. Thus, the unexpected crib card contributes to their higher scores by serving as a memory aid or as a resource for information that they did not learn. In other words, students simply use the information on their own crib cards to answer the exam questions without learning the material. It is possible that students actively process test-related information during crib card construction, but even if so, this active processing does not compensate for the overall reduction in learning associated with self-constructed crib cards.

Another possible explanation for our results is that the quality of the other-constructed crib cards was significantly better. This possibility seems unlikely given that a student who had not taken the course constructed the crib cards with the textbook and the learning objectives, just as the students in the class did. Even if the other-constructed cards contained more information, the finding that students performed better with the premade crib cards than with their own contradicts the coding hypothesis (Dorsel & Cundiff, 1979; Hindman, 1980; Trigwell, 1987) because it suggests that having information on the crib card helps more than the process of constructing the crib card.

A related issue is that average students may not know how to effectively engage in the coding process while constructing a crib card. In other words, they may not know what information to put on the card or how to study the information on the crib card, deficits that highlight weaknesses in student study habits in general. The fact that our research assistant, an above-average student, had not taken the course or seen the exams, yet produced a crib card that led to higher exam scores, supports this idea. Future research should examine the coding hypothesis with an intervention study by teaching students how to effectively create and use a crib card.

The construction of crib cards did not facilitate student learning, and the use of self-constructed crib cards did not enhance exam performance in this content-driven upper division psychology course. Given the mounting evidence against crib card effectiveness (Dickson & Miller, 2005; Hindman, 1980; Whitley, 1996), we recommend against the use of crib cards during exams, until teaching scholars more fully understand how crib cards affect learning and exam performance after instructors have coached students on the processes of actively selecting, organizing, and representing the information in a condensed form.

References

Dickson, K. L., & Miller, M. D. (2005). Authorized crib cards do not improve exam performance. *Teaching of Psychology, 32,* 230–233.

Dorsel, T. N., & Cundiff, G. W. (1979). The cheat-sheet: Efficient coding device or indispensable crutch? *Journal of Experimental Education, 48,* 39–42.

Hindman, C. D. (1980). Crib notes in the classroom: Cheaters never win. *Teaching of Psychology, 7,* 166–168.

Trigwell, K. (1987). The crib card examination system. *Assessment and Evaluation in Higher Education, 12,* 56–65.

Whitley, B. E., Jr. (1996). Does "cheating" help? The effect of using authorized crib notes during examinations. *College Student Journal, 30,* 489–493.

Whitley, B. E., Jr., & Keith-Spiegel, P. (2002). *Academic dishonesty: An educator's guide.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Notes

Effect of Study Guide Exercises on Multiple-Choice Exam Performance in Introductory Psychology

K. Laurie Dickson, Michael S. Devoley, and Michelle D. Miller
Northern Arizona University

We experimentally investigated whether requiring completion of only multiple-choice study guide questions differentially affected multiple-choice exam performance compared to requiring a variety of study guide exercises (learning goals, vocabulary, fill-in-the-blank concepts, multiple-choice questions, short essay questions, matching, and language enhancement) for introductory psychology students (N = 244). There was not a significant difference in multiple-choice exam performance. We conclude that requiring a variety of study guide exercises produces no significant benefit for multiple-choice exam performance over and above those produced by requiring only multiple-choice exercises.

Published study guides that accompany psychology textbooks potentially enhance student learning and exam perfor-

mance. Completing study guide exercises requires a significant investment of student time and effort, raising the issue of how to maximize returns on this investment. In particular, it is unclear whether all of the exercises typically included in a study guide produce measurable gains in exam performance. It may be the case that only certain exercises, for example the multiple-choice practice questions, actually raise multiple-choice exam performance.

Students perceive study guides that accompany their textbooks as helpful (Dickson, Miller, & Devoley, 2005; Reittinger & Crowley-Long, 1992). However, research regarding the effect of study guides on exam performance is mixed. Reittinger and Crowley-Long (1992) found nonsignificant differences in final grades for an introductory psychology course between students who voluntarily completed a textbook study guide and students who chose not to complete the study guide. Daniel and King (2003) compared the effect of required study guides, Web-based activities (multiple-choice questions, crossword puzzles, flash cards, etc.), and Web-based multiple-choice quizzes on multiple-choice exam performance. Student exam performance was significantly better for the online multiple-choice quiz group compared to the study guide and Web activities groups. Dickson et al. (2005) found that students with a study guide requirement performed significantly better than students without a study guide requirement. Results offered support for the effectiveness of required study guides in courses that use multiple-choice exams; however, the design of the Dickson et al. study did not allow the researchers to address the benefits of specific types of study guide exercises.

Given that Dickson et al. (2005) previously found that requiring a variety of study guide exercises led to higher exam scores, we assessed whether requiring only multiple-choice study guide questions differentially affected multiple-choice exam performance compared to requiring a variety of study guide exercises (learning goals, vocabulary, fill-in-the-blank concepts, multiple-choice questions, short essay questions, matching, and language enhancement). In addition, we examined student perceptions of the study guide exercises.

Method

Participants and Procedures

Two hundred forty-four students in two introductory psychology sections participated (122 students in each section). There were no significant differences between the sections with regard to gender (65.1% women) or class status (73.1% freshmen). Chi-square analyses revealed no significant differences between sections with respect to high school grade point average (GPA), college GPA, Scholastic Aptitude Test (SAT) scores, or American College Test (ACT) scores ($p > .05$).

The second author taught both sections with the same textbook (Nairne, 2003), course policies (with the exception of the study guide requirement), outlines, content, and exams. In one section, the instructor required all of the study guide exercises (Proctor, 2003), including learning goals, vocabulary, fill-in-the-blank concepts, multiple-choice questions, short essay questions, matching, and language enhancement (see Dickson et al., 2005). For the other section, he required only the multiple-choice exercises in the study guide. He graded the study guides for completion after each exam (8% of the final course grade). Students completed four identical multiple-choice question exams (50 questions each) derived from the textbook test bank (see Dickson et al., 2005). Immediately following the last exam, students reported their demographic information (gender, class status, high school GPA, college GPA, SAT scores, and ACT scores) and their perceptions (*negative, somewhat negative, neutral, somewhat positive, positive, not applicable*) regarding the usefulness of the study guide and the specific exercises.

Results

Student Performance

Mean exam performance was 75.05 ($SD = 12.38$) for the multiple-choice-only section and 74.23 ($SD = 10.95$) for the entire study guide section. We conducted an ANCOVA with section as the independent variable, exam performance as the dependent measure, and high school GPA as a covariate (we excluded 8 participants who did not provide high school GPA). High school GPA was a significant covariate, $F(1, 233) = 18.44, p < .001, r = .27$. The difference between the multiple-choice-only and the entire study guide sections was not significant, $F(1, 233) = .28, p = .60, r = .03$.

Most participants (69.6%) reported either positive or somewhat positive feelings about the study guide, 20.4% were neutral, and 7.8% were either negative or somewhat negative toward the study guide (2.2% missing values). When questioned about the usefulness of the various exercises in the study guide, more students (79.6%) thought positively of the multiple-choice questions than the other exercises (45.5% for true–false, 40.5% for vocabulary, 33.3% for learning goals, 21.1% for essay questions, 21.9% for language enhancement). Most participants (76.4%) thought that the multiple-choice exercises improved their exam scores, 16.9% did not think that they helped, and 5.9% reported that they did not complete the exercises (0.8% missing values).

Discussion

We did not find a significant difference in multiple-choice exam scores for students with the entire study guide requirement and students with the multiple-choice requirement. Previous research found that students who were required to complete the entire study guide performed significantly better than students in a control section and that when the study guide exercises were not required most students did not complete them (Dickson et al., 2005). Although study guide exercises led to higher multiple-choice exam scores, our findings suggest that requiring a variety of study guide exercises does not enhance multiple-choice exam performance more than requiring fewer, more relevant exercises (i.e., multiple-choice questions).

Most participants reported positive feelings about the study guide and believed that the multiple-choice exercises

improved their exam scores. Although the true–false, vocabulary, learning goals, essay questions, and language enhancement exercises were not "dangerous detours" (Gurung, 2004) because they did not hinder student performance, they also did not enhance multiple-choice exam performance beyond the benefit of the multiple-choice exercises. The congruence between the type of study guide exercise and exam question may explain the benefit. The transfer-appropriate processing theory (Herrmann, Raybeck, & Gutman, 1993) states that retrieval of information during testing is enhanced if it requires the same kind of mental processing as used during studying (Ellis & Hunt, 1993). However, the possibility exists that the principle of diminishing returns applies. Perhaps any 60 exercises per chapter are sufficient and additional investment results in relatively little payoff. For example, do 60 true–false exercises per chapter show a similar effect on multiple-choice exam performance? A comparison of varying numbers of multiple-choice exercises is warranted to address whether there is a threshold of multiple-choice exercises for enhanced multiple-choice exam performance. In addition, the effectiveness of the multiple-choice exercises could differ depending on how students use them and whether they are mandatory (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Brothen & Wambach, 2001; Grimstad & Grabe, 2004; Thomas & Rohwer, 1986). Requiring study guide exercises of any type may force students to study, which leads to improved exam performance. Is there a difference in the effectiveness of the multiple-choice exercises if students use the exercises to become familiar with the test material or to test their understanding of the material after they have studied?

There was not a significant difference in multiple-choice exam performance between students with the multiple-choice exercises requirement and the entire study guide requirement. Thus, we conclude that students in this introductory psychology course maximized their investment of time and effort by focusing on multiple-choice exercises when preparing for a multiple-choice exam, rather than additionally investing in a variety of study guide exercises.

References

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A, & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61,* 213–238.

Brothen, T., & Wambach, C. (2001). Effective student use of computerized quizzes. *Teaching of Psychology, 28,* 292–294.

Daniel, D. B., & King, J. (2003, January). *What works in general psychology? The differential effects of study guides, Web-based activities and Web-based quizzing on exam scores.* Poster presented at the meeting of the National Institute of the Teaching of Psychology, St. Petersburg Beach, FL.

Dickson, K. L., Miller, M. D., & Devoley, M. S. (2005). Effect of textbook study guides on student performance in introductory psychology. *Teaching of Psychology, 32,* 34–39.

Ellis, H. C., & Hunt, R. R. (1993). *Fundamentals of cognitive psychology* (5th ed.). Madison, WI: Brown & Benchmark.

Grimstad, K., & Grabe, M. (2004). Are online study questions beneficial? *Teaching of Psychology, 31,* 143–146.

Gurung, R. A. R. (2004). Pedagogical aids: Learning enhancers or dangerous detours? *Teaching of Psychology, 31,* 164–166.

Herrmann, D., Raybeck, D., & Gutman, D. (1993). *Improving student memory.* Seattle, WA: Hogrefe & Huber.

Nairne, J. S. (2003). *Psychology: The adaptive mind* (3rd ed.). Belmont, CA: Wadsworth.

Proctor, J. (2003). Study guide for *Psychology: The adaptive mind* (3rd ed). Belmont, CA: Wadsworth.

Reittinger, D. L., & Crowley-Long, K. (1992). The use of student workbooks in introductory psychology courses (Rep. No. JC920516). In J. R. Levine & G. Indenbaum (Eds.), *Teaching of psychology: Ideas and innovations.* Ellenville, NY: Annual Conference on Undergraduate Teaching of Psychology. (ERIC Document Reproduction Service No. ED350060)

Thomas, J. W., & Rohwer, W. D. (1986). Academic studying: The role of learning strategies. *Educational Psychologist, 21,* 19–41.

Notes

1. We thank the anonymous reviewers for their insightful comments.
2. Send correspondence to K. Laurie Dickson, Department of Psychology, Northern Arizona University, P.O. Box 15106, Flagstaff, AZ 86011; e-mail: laurie.dickson@nau.edu.

An Assignment to Help Students Learn to Navigate Primary Sources of Information

Andrew N. Christopher and Mark I. Walter
Albion College

Being able to apply methodological and statistical concepts is a goal that many teachers of psychology may have for their courses. Toward this end, we describe an assignment that involves using primary source journal articles and that leads students to further their understanding of material presented in methodology and statistics courses. Student performance on a small number of exam questions suggested pedagogical value in using this assignment. We present cautions in implementing the assignment and provide resources that teachers can use.

Among the many goals that teachers of statistics and research methods establish is an increased ability by their students to understand information in primary sources. Indeed, 72% of undergraduate research methods courses and 22.7% of undergraduate statistics classes use primary sources (Oldenburg, 2005). Increased understanding of statistics and research methods is more likely when students complete specific assignments using such sources. We developed the assignment described herein to help students understand methodological and statistical information by using psychology journal articles.

Numerous articles have appeared in *Teaching of Psychology* to help teachers of statistics and research methods courses develop specific course activities and assignments. For instance, Dolinsky (2001) provided several practical and easy-to-implement active learning strategies for teaching statistics. Specifically, she discussed various ways to use SPSS to teach different topics (e.g., statistical interactions), how to implement active learning strategies (e.g., student presentations of conceptual information), and the use of writing as-

signments to interpret statistical data. Morgan (2001) proposed that using obituaries can illustrate a variety of methodological and statistical issues such as missing data, outliers, correlations, and the use of SPSS.

Other relevant articles have focused on including primary sources in undergraduate psychology classes. Chamberlain and Burrough (1985) described two techniques for using journal articles to promote critical reading skills. Suter and Frank (1986) described the use of classic articles to spark class discussions of core topics (e.g., types of variables) in their undergraduate methodology courses. Our assignment adds to the resources available to teachers of statistics and methods courses by combining the ideas from these two articles to use journal articles to teach basic statistical and methodological information and to help students apply such information to actual research studies.

The Assignment

On the surface, our assignment is a relatively simple, straightforward way to help students understand and apply information presented in methodology and statistics courses. We first locate primary sources containing methodological and statistical techniques that we cover in our research methodology and statistics courses. We follow Carkenord's (1994) specifications: high correspondence between the article methodology, article analyses, and course material; close match between article technicality and student sophistication; of interest to students; and of relatively short length. These articles can be in the school's library, or teachers can require students to purchase them with the course texts. From each article, we construct questions that require students to interpret information in the article that illustrates methodological and statistical concepts being discussed in class and lab. For example, to help students understand the distinction between within-subject and between-subject designs and the resultant use of the dependent samples or independent samples t tests, we used an article by Burris, Batson, Altstaedten, and Stephens (1994). Table 1 contains example questions from this particular assignment.

We note at least four benefits of using this assignment. First, it forces the integration of information about research methods with statistical information. We teach an integrated two-semester research design and analysis course sequence; thus, it is relatively easy for us to use this assignment. However, when we have taught these topics as separate courses, we found it easier for students to complete these assignments during the second course than during the first, presumably because these assignments force a review of material covered in the first course. Second, when students complete their individual projects in our classes, we cannot always be certain they carefully read and understand the primary sources that appear in their literature reviews. Use of this assignment allows teachers to be somewhat confident that students have read and understood the assigned articles. Third, consistent with Kirk's (2004) recommendation that students think about the meaning of research methodology and statistics, the assignment provides students with an opportunity to grasp the "big picture" rather than just the minutiae of computations. Finally, teachers can use the assignment in any course that stresses methodology or statistical analyses.

Evaluation, Caveats, and Conclusions

Students reported that completing each of these assignments required between 30 min and slightly more than 3 hr of time ($M = 82$ min, $SD = 42$ min). These self-reports of time spent working on the assignments were negatively correlated with final course grades, $r(45) = -.66$, $p < .001$. We examined four exam questions to assess the potential impact of us-

---

Table 1. Example Questions for the Assignment Using Burris et al.'s (1994) Research

1. How did the researchers operationally define each of the following variables: "loneliness," "intrinsic religious orientation," "extrinsic religious orientation," and "quest religious orientation"?
2. What was (were) the independent variable(s), if any, in this study? If you do not think there were any independent variables in this study, explain why.
3. What was (were) the dependent variable(s) in this study?
4. On page 328, at the bottom of the page, the researchers discuss some ancillary measures, including "a question to check the effectiveness of the loneliness-salience manipulation." Then, on page 329, there is an "Effectiveness of the Loneliness-Salience Manipulation" section that provides results to responses on this question.
    (a) First, are these results from an independent samples t test or a dependent samples t test? Justify your choice!
    (b) Second, why is it vital, from the standpoint of psychological realism, that this result be statistically significant? That is, if this statistical test was not significant, what might it tell the researchers about their experimental procedure?
5. Look on page 329, the "Change in Intrinsic Scores" subsection of the Results and Discussion. In particular, examine the last sentence of that subsection.
    (a) What is the probability that the average difference in change scores was due to extraneous variables (i.e., variables other than the loneliness-salience manipulation)?
    (b) How many participants were included in this particular analysis?
    (c) Is this statistical test an independent samples t test or a dependent samples t test? Justify your choice!
6. Look at Table 1 on page 330. In the two "Change" rows are difference scores for the three religious orientations. On all three of these change measures, the high-salience and low-salience groups differed in their mean change scores. Why, then, can we only generalize/infer/conclude that intrinsic orientation changed as a result of the independent variable?
7. On page 331, the second full paragraph, it is reported that "High-salience subjects who wrote of unresolved loneliness experiences had higher Intrinsic-scale change score than did those who wrote of resolved experiences, $t(34) = 1.92$, $p < .06$, two-tailed." Are the researchers justified in generalizing this finding from their sample to the population (as they seem to be doing)? Explain your answer.
8. Was the hypothesis supported or not supported by the data? How do you know, other than the authors saying so?
9. Please comment on both the internal validity and the external validity of this study. Explain why this study did or did not possess each of these two types of validity.

Table 2. Primary Source Articles Used As Assignments

Primary Source Article	Basic Methodological or Statistical Concepts Emphasized
Furnham & Drakeley (2000)	Survey research and zero-order correlations
Kasser & Ahuvia (2002)	
Rhodes & Hammer (2000)	
Roy, Benenson, & Lilly (2000)	Quasi-experimental design and the independent samples *t* test
Page (1997)	Chi-square analysis
Yoder, Hogue, Newman, Metz, & LaVigne (2002)	
Cherulnik & Wilderman (1986)	One-way between-subject experimental designs, one-way ANOVA, and follow-up tests
Burris, Batson, Altstaedten, & Stephens (1994)	Within-subject and between-subject designs and the resultant use of the dependent samples and independent samples *t* test
DeBono, Leavitt, & Backus (2003)	Two-way designs and the use of experimental and nonexperimental data in a single study
Fried (1999)	
Furnham, Trew, & Sneade (1999)	

ing journal articles on students' ability to interpret information in our research design and analysis courses. These questions appeared on the final exam for students in the second semester of our research design and analysis course. One section ($n = 24$) did not complete this assignment, whereas the other section ($n = 23$) completed four of these assignments throughout the two-semester course sequence. The information to answer the four questions was:

Suppose you are reading about a one-way between-subjects analysis of variance (ANOVA) and encounter the following information:
$$F(4, 30) = 8.31, p < .01.$$
a) Why is this analysis called a "one-way" analysis of variance (ANOVA)?
b) How many groups were there in this analysis?
c) How many participants were included in this analysis?
d) What is the probability that the researcher made a Type I error in this analysis?

After combining responses to the four questions (correct or incorrect) into a composite score ($\alpha = 0.78$), we ran an independent samples *t* test. This analysis revealed that the class that completed reading assignments ($M = 73.9\%$) performed better than the class not completing reading assignments ($M = 60.4\%$), $t(45) = 2.17, p < .04$. Thus, based on our small number of identical exam questions for the two groups of students, implementing the assignment described herein may boost students' ability to interpret course material.

We caution teachers using this assignment that, as Carkenord (1994) implied, it is not always an easy task to find articles that closely parallel course material. For instance, much nonexperimental research uses statistical techniques (e.g., path analysis) that are beyond an introductory level. In addition, we try to find articles on topics that would be of interest to students so that they can use such articles to generate ideas for their projects. We have found that locating articles on which to base assignments has the potential to require a large time commitment for teachers. We have constructed assignments using 11 primary sources. Table 2 contains these sources and the basic course emphasis of each one. Furthermore, Ware, Badura, and Davis (2002) described a way to use published undergraduate research to teach statistics and research methods.

The articles that we have located are almost exclusively from various nonundergraduate outlets. Perhaps using articles from undergraduate sources that Ware et al. suggested as initial reading assignments and using articles from nonundergraduate sources as subsequent reading assignments would be the best way to facilitate students' ability to apply statistical and methodological concepts.

In sum, teachers of methodology and statistics courses can help students learn to apply this information by requiring them to answer specific questions pertaining to primary source material. We hope that our assignment sparks other teachers to develop additional ways to accomplish this important pedagogical goal.

References

*Indicates articles used for reading assignments.

*Burris, C. T., Batson, C. D., Altstaedten, M., & Stephens, K. (1994). "What a friend ... " Loneliness as a motivator of intrinsic religion. *Journal for the Scientific Study of Religion, 33,* 326–334.

Carkenord, D. M. (1994). Motivating students to read journal articles. *Teaching of Psychology, 21,* 162–164.

Chamberlain, K., & Burrough, S. (1985). Techniques for teaching critical reading. *Teaching of Psychology, 12,* 213–215.

*Cherulnik, P. D., & Wilderman, S. K. (1986). Symbols of status in urban neighborhoods: Contemporary perceptions of nineteenth-century Boston. *Environment & Behavior, 18,* 604–622.

*DeBono, K. G., Leavitt, A., & Backus, J. (2003). Product packaging and product evaluation: An individual difference approach. *Journal of Applied Social Psychology, 33,* 513–521.

Dolinsky, B. (2001). An active learning approach to teaching statistics. *Teaching of Psychology, 28,* 55–56.

*Fried, C. B. (1999). Who's afraid of rap: Differential reactions to music lyrics. *Journal of Applied Social Psychology, 29,* 705–721.

*Furnham, A., & Drakeley, R. (2000). Predicting occupational personality test scores. *Journal of Psychology: Interdisciplinary and Applied, 134,* 103–111.

*Furnham, A., Trew, S., & Sneade, I. (1999). The distracting effects of vocal and instrumental music on the cognitive test performance of introverts and extraverts. *Personality and Individual Differences, 27,* 381–392.

*Kasser, T., & Ahuvia, A. (2002). Materialistic values and well-being in business students. *European Journal of Social Psychology, 32,* 137–146.

Kirk, R. (2004, October). *Changing topics and trends in introductory statistics.* Keynote address delivered at the meeting of Best Prac-

tices in the Teaching of Research Methods and Statistics in Psychology, Atlanta, GA.

Morgan, B. L. (2001). Statistically lively uses for obituaries. *Teaching of Psychology, 28*, 56–58.

Oldenburg, C. M. (2005). Use of primary source readings in psychology courses at liberal arts colleges. *Teaching of Psychology, 32*, 25–29.

*Page, S. (1997). An unobtrusive measure of racial behavior in a university cafeteria. *Journal of Applied Social Psychology, 27*, 2172–2176.

*Rhodes, L. D., & Hammer, E. Y. (2000). The relation between job satisfaction and personality similarity in supervisors and subordinates. *Psi Chi Journal of Undergraduate Research, 5*, 56–59.

*Roy, R., Benenson, J. F., & Lilly, F. (2000). Beyond intimacy: Conceptualizing sex differences in same-sex friendships. *Journal of Psychology: Interdisciplinary and Applied, 134*, 93–101.

Suter, W. N., & Frank, P. (1986). Using scholarly journals in undergraduate experimental methodology courses. *Teaching of Psychology, 13*, 219–221.

Ware, M. E., Badura, A. S., & Davis, S. F. (2002). Using student scholarship to develop student research and writing skills. *Teaching of Psychology, 29*, 151–154.

*Yoder, J. D., Hogue, M., Newman, R., Metz, L., & LaVigne, T. (2002). Exploring moderators of gender differences: Contextual differences in door-holding behavior. *Journal of Applied Social Psychology, 32*, 1682–1686.

Notes

1. Portions of this article were presented at the meeting of Best Practices in the Teaching of Research Methods and Statistics in Psychology, Atlanta, GA in October 2004.
2. We thank Randolph Smith and three anonymous reviewers for their helpful feedback during the editorial process.
3. Send correspondence, including requests for any of the assignments summarized in Table 2, to Andrew N. Christopher, Department of Psychology, Albion College, 611 E. Porter Street, Albion, MI 49224; e-mail: achristopher@albion.edu.

Supervision of New Instructors: Promoting a Rewarding First Experience in Teaching

Steven Prentice-Dunn
University of Alabama

This article provides suggestions for supervisors of new instructors. Topics include establishing a foundation for a rewarding teaching experience, giving advice to the new instructor on how to shape a course and manage time, and structuring consultation sessions to correct problems while preserving the novice teacher's confidence.

Teaching a college course for the first time can be a tumultuous blend of excitement and uncertainty. That initial experience can heavily influence whether teaching will be part of a person's career aspirations. Increasingly, psychology programs have recognized the importance of training graduate students to teach (Meyers & Prieto, 2000a); approximately two thirds of respondents in a recent survey now offer courses on teaching (Buskist, Tears, Davis, & Rodrigue, 2002).

My purpose in writing this article is to share observations gleaned from supervising a teaching of psychology course over the past 15 years. Although my comments are directed to faculty, I also provide suggestions for supervisors to give graduate students who are teaching their first college course. Beginning instructors may also find those comments helpful to read directly.

Theoretical Context for Supervision

Prieto (2001) provided a useful model for understanding the development of new instructors and the accompanying roles of supervisors. Adapted from Stoltenberg, McNeill, and Delworth's (1998) approach to psychotherapy supervision, Prieto described the orientation and concerns of new instructors who are beginning graduate teaching assistants (GTAs), advanced GTAs, and junior faculty. Each stage is characterized by differences in awareness of self versus others, motivation, and autonomy. For example, beginning GTAs are (a) focused greatly on themselves rather than on their students; (b) highly motivated, although anxious; and (c) highly dependent on supervisors for direction. Advanced GTAs have better awareness of and empathy for their students' concerns, more ambivalence about teaching because of the high demands, and greater autonomy. The primary goal of the supervisor is to facilitate the new instructor's movement beyond the beginner stage.

Foundation for a Positive Teaching Experience

Long before new instructors set foot in their classrooms, you can enhance the likelihood of an affirming experience by shaping how graduate students approach the content and style of their course.

Give Permission to Be Imperfect

Becoming a good teacher is a developmental process, with few graduate students having the skills and experience needed to be an excellent instructor from the outset. Communicate that you are expecting improvement, but not perfection, as the term progresses. Many beginning teachers use a favorite professor as a model for their first attempt, while failing to realize that those respected abilities developed over time and with some failures (Sviniki, 1994). Early, explicit attention to your expectations counters the assumption of beginning instructors that their skills should be fully developed (Prieto, 2001).

Shape Their Expectations Through Readings

Before reading the textbook and crafting a syllabus, beginning instructors should read concise articles on active learning (Bernstein, 1997), depth versus breadth of content coverage (Nelson, 2001), undergraduate students' perspectives on teaching (Grineski, 1999), and myths about learning to teach (Sviniki, 1994). The aim here is to encourage thought about course goals before delving into the details of

content. The readings also provide the structure needed by beginning teachers (Prieto, 2001).

Encourage Early Use of Several Techniques

To counter the frequently held assumption that there is but one correct way to teach (Prieto, 2001), advocate that graduate students try a variety of teaching strategies in the first month of the semester and then settle on those techniques that mesh with the particular class and the new instructor's interpersonal style. For example, a class will more likely embrace interactive teaching that is introduced early rather than later in the term.

Use Active Learning in Training

New instructors benefit from participatory learning just as their own undergraduate students do. Among the valuable techniques suggested by Meyers and Prieto (2000b) are observing other instructors, critiquing syllabi and tests, discussing ethics scenarios, and writing a teaching philosophy. Particularly useful is including an unstructured period during group supervision meetings in which instructors can share their successes and ask advice on problems. Meyers and Prieto noted that this activity normalizes the challenges of teaching, provides peer support, and offers solutions to instructional difficulties.

Advice to the New Instructor

Several excellent sources offer practical suggestions for the beginning college instructor about how to conduct a specific class session (e.g., Buskist, 2000; Davis, 1993). For example, Buskist (2000) noted that new GTAs often rely too heavily on notes, speak while turned away from the class, and fail to reward student participation. My suggestions are more general in nature. Supervisors may provide the following suggestions to novice teachers.

Beware of Multimedia Overload

Today's psychology textbooks come with an array of manuals, slides, DVDs, and Web sites. New instructors can become overwhelmed with the available resources. Because they must prepare each topic from scratch, beginning teachers should quickly identify a few "go-to" sources to check when they begin a new chapter. Limiting oneself can save valuable time and does not preclude a wider search when necessary.

Reduce the Textbook Coverage

New teachers may not realize that most textbooks are exhaustively detailed to reach the broadest possible market. They should not feel compelled to cover every chapter or every fact because it is in the text. Instead, advise covering a few primary points in each chapter in depth in class with accompanying applications and discussion. Remind them that undergraduate students can learn some important concepts on their own.

Carefully Plan Exams

Recommend that instructors use an exam format that reflects the course goals. For example, fact-based multiple-choice tests often do not address the goal of promoting critical thinking. In addition, advise that they provide students with practice in class with the skills needed on the exams to reduce the pervasive anxiety about tests.

Address Diverse Learning Styles

Most graduate students have thrived in a lecture environment and may not realize that other college students require different approaches. Recommend that new teachers change activities every 15 to 20 min to reset attention and accommodate the variety of ways that people learn.

Consultation

Faculty involved in training not only supply materials on how to teach, they also help new instructors work through difficult situations and provide feedback on teaching performance. For the latter activities, a collaborative, rather than expert, approach is better for developing new teachers' skills and confidence (Carroll & Goldberg, 1989; Prentice-Dunn, Payne, & Ledbetter, 2006). A collaborative approach provides sufficient structure to allay the new teacher's anxiety along with the encouragement to explore many options to create a personal teaching style. It guides the instructor away from the beginning stage in Prieto's (2001) typology and toward the advanced stage.

To correct problems while preserving the graduate student's emerging self-efficacy as a college teacher, consider the following pointers.

Be Aware of Self-Criticism

New instructors do not give themselves enough credit. They tend to be overly critical and fail to see what they are doing well (Prieto, 2001). Overcome this tendency by praising all successes, especially early in the term. When possible, have graduate students observe fellow students who are also new to the teaching role. Novice teachers become less anxious and more focused on improvement when they see that their colleagues face similar challenges.

Because of their tendency to self-criticize, novice instructors may take unwarranted blame for a poor first exam performance. They gain a more accurate perception after surveying the class on the time and steps that students took to prepare for the test.

Use Multiple Sources of Feedback

Ratings from the instructor's students can improve teaching if collected during the first half of the term (Cohen, 1980). In addition, new teachers consider narrative comments about strengths and weaknesses to be especially valuable (McElroy & Prentice-Dunn, 2005). However, coach

instructors to look for themes rather than focusing on individual negative or positive comments.

Another source of feedback is video. Prentice-Dunn and Pitts (2001) described several benefits, including seeing the class from the students' perspective. Video enables the viewer to alternate focus between broad themes and the details of teaching. Although novice teachers appreciate being able to watch the tape initially while alone, they need checklists (e.g., Davis, 1993) to focus attention on pedagogical issues rather than excessively on their appearance.

Structure the Consultation Session

The consultation session integrates multiple feedback sources with self-observations. New instructors can review with an experienced instructor how well classroom content fits course objectives and also explore alternative ways to communicate course material. In one study, graduate students reported that such meetings improved their skills more than any other component of a teaching of psychology course (McElroy & Prentice-Dunn, 2005).

Prentice-Dunn et al. (2006) provided several suggestions for conducting consultation sessions. Among these recommendations are allowing the new instructor to take the lead in conversation, focusing on positive behaviors first, framing problems in terms of improvement, and helping the instructor set short-term teaching goals.

Recognize the Rhythms of the Semester

Teaching for the first time can require an enormous amount of energy. Although beginning instructors are enthusiastic, their stamina tends to wane at midsemester due to the ever-present demands of course work, research, and clients. Acknowledge this fact and assist new instructors in recognizing it in themselves and in their students. Provide advice on how to continue to present enthusiastically and engage students.

Although many instructors do not reach Prieto's (2001) second developmental stage (i.e., advanced GTA) for one or two semesters, first-time teachers often display aspects of the advanced stage toward the end of their first teaching experience. Prieto characterized this stage as one of ambivalence. Assist instructors in placing teaching in the context of other demands on their professional and personal time. Help them explore the pros and cons of alternative instructional and grading practices as they begin to reflect on the closing semester and anticipate future courses.

Conclusions

The initial teaching experience establishes strategies and attitudes that can follow the new instructor for years. Faculty members who conduct training workshops and supervise teaching of psychology courses play a vital role in determining what direction that first experience will take. Careful attention to our role can make the first college course not a chore to be persevered, but rather a self-affirming activity that simultaneously benefits a classroom of students.

References

Bernstein, D. A. (1997). Reflections on teaching introductory psychology. In R. J. Sternberg (Ed.), *Teaching introductory psychology: Survival tips from the experts* (pp. 35–47). Washington, DC: American Psychological Association.

Buskist, W. (2000). Common mistakes made by graduate teaching assistants and suggestions for correcting them. *Teaching of Psychology, 27,* 280–282.

Buskist, W., Tears, R. S., Davis, S. F., & Rodrigue, K. M. (2002). The teaching of psychology course: Prevalence and content. *Teaching of Psychology, 29,* 140–142.

Carroll, J. G., & Goldberg, S. R. (1989). Teaching consultants: A collegial approach to better teaching. *College Teaching, 37,* 143–146.

Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education, 13,* 321–341.

Davis, B. G. (1993). *Tools for teaching.* San Francisco: Jossey-Bass.

Grineski, S. (1999). Visiting as a student. *National Teaching and Learning Forum, 8*(2), 7–10.

McElroy, H. K., & Prentice-Dunn, S. (2005). Graduate students' perceptions of a teaching of psychology course. *Teaching of Psychology, 32,* 122–124.

Meyers, S. A., & Prieto, L. R. (2000a). Training in the teaching of psychology: What is done and examining the differences. *Teaching of Psychology, 27,* 258–261.

Meyers, S. A., & Prieto, L. R. (2000b). Using active learning to improve the training of psychology teaching assistants. *Teaching of Psychology, 27,* 283–284.

Nelson, C. (2001). What is the most difficult step we must take to become great teachers? *National Teaching and Learning Forum, 10*(4), 10–11.

Prentice-Dunn, S., Payne, K. L., & Ledbetter, J. M. (2006). Improving teaching through video feedback and consultation. In W. Buskist & S. F. Davis (Eds.), *Handbook of the teaching of psychology* (pp. 295–300). Malden, MA: Blackwell.

Prentice-Dunn, S., & Pitts, G. S. (2001). The use of videotape feedback in the training of instructors. In L. R. Prieto & S. Meyers (Eds.), *The teaching assistant training handbook: How to prepare TAs for their responsibilities* (pp. 89–102). Stillwater, OK: New Forums.

Prieto, L. R. (2001). The supervision of teaching assistants: Theory, evidence and practice. In L. R. Prieto & S. Meyers (Eds.), *The teaching assistant training handbook: How to prepare TAs for their responsibilities* (pp. 103–129). Stillwater, OK: New Forums.

Stoltenberg, C. D., McNeill, B., & Delworth, U. (1998). *IDM supervision: An integrated developmental model for supervising counselors and therapists.* San Francisco: Jossey-Bass.

Sviniki, M. (1994). Seven deadly comments that block learning about teaching. *National Teaching and Learning Forum, 3*(3), 4–6.

Special Topics Courses: Meeting Departmental Needs

Kirsten T. Michels and Joan L. Jackson
University of Georgia

Psychology departments must meet the evolving needs of their undergraduate students. Finding creative solutions to make the best use of available resources is a challenge to which faculty must rise. We discuss ways our university found to offer more upper division special topics courses to undergraduates while mentoring graduate students in developing new courses. We describe this process in an effort to provide an example to other programs of an innovative approach to addressing staffing concerns.

There are ongoing philosophical debates regarding how psychology departments educate their students (Brewer, 1997). Historically, psychology was characterized as a liberal arts discipline with an emphasis on survey over specialization. There was tension regarding the integration of scientific inquiry and application of research findings. McGovern and colleagues discussed the generalist model in detail (McGovern, Furumoto, Halpern, Kimble, & McKeachie, 1991). The authors described the essential components of generalist training in psychology, which included developing a common knowledge base, ethical and critical thinking, interpersonal skills, and research training. Thematic curricula were described as one way to introduce students to a subfield (e.g., developmental or biological psychology) earlier in their training, and then teach research design and statistical analyses as they are applied to a core topic. Thematic models were criticized for logistical problems and the potential to further fragment the discipline (Kierniesky, 1992).

The Psychology Department at the University of Georgia has traditionally followed the generalist model, with a small number of special-interest seminars complementing a larger number of lab courses in foundational areas of psychology. In recent years, several developments have converged to make the offering of special topics courses increasingly desirable. Beginning with a change in the University's core curriculum in 1998, the department requires students to take a larger proportion of their courses at the upper division level to fulfill bachelor's degree requirements than was previously the case. As a result, psychology majors are enrolling in a larger number of psychology courses. In addition, increasing numbers of students from majors other than psychology are also seeking upper division psychology courses as electives.

This curriculum change has the potential advantage of allowing students to add both breadth and depth to their study of psychology. However, the shift toward a larger number of upper division courses has raised questions both about the type of courses that need to be added to the curriculum and how to staff those courses. Consistent with many universities, recent economic conditions have resulted in a tight budget and little opportunity for hiring new faculty to teach the needed courses (Murray, 2000).

The department head, associate head, and chairs of the six doctoral programs are involved in determining the assignments given to the department's graduate teaching assistants (GTAs). GTAs are graduate students with master's degrees who teach undergraduate courses, in contrast to teaching assistants (TAs), who are graduate students who support faculty and other GTAs with responsibilities such as proctoring and grading exams. Doctoral students who enter the psychology program with a master's degree may instruct their own course without previously serving as a TA. Making teaching assignments involves complicated decision making in light of the department's need for additional upper division courses, faculty commitments, and financial limitations. Except under exceptional circumstances, College of Arts and Sciences policy has traditionally restricted teaching by GTAs to lower division, general psychology, and introductory methods courses.

Graduate students can prepare to teach courses within their area of specialization. Special topics courses taught by graduate students have advantages for psychology department administrations and students. Broader courses, such as introductory psychology that require knowledge of multiple areas, may be more adeptly taught by seasoned professors (Freeman, 2002). After careful consideration of the potential benefits and risks, our department decided to offer experienced graduate student instructors the opportunity to develop special topics, upper division undergraduate courses (e.g., anxiety disorders, industrial/organizational psychology) under the guidance of senior faculty in the department. These courses fit into the psychology curriculum by satisfying one of three upper division elective requirements.

The department provided a variety of instructional and mentoring experiences to support these instructors. All psychology graduate students at our university must take a one-semester course on teaching fundamentals (McKeachie & Hofer, 2001). We intend this instruction to begin preparing new teachers and assistants for their roles while in graduate school as well as for careers that potentially include teaching.

GTAs who teach standard courses have faculty mentors with whom they hold regular planning meetings. Mentors can provide instructive feedback following live or videotaped classroom observations. Textbooks for standard GTA-instructed courses are preselected by departmental faculty. For standard courses, syllabi, lecture notes, and tests are often available. Student teachers can also meet with one another and interested faculty to discuss teaching-related issues during a monthly forum known as the Soapbox Session (Verges, Spitalnick, & Michels, 2005).

Enhanced mentoring opportunities were also available to support the instructors of the special topics seminars (Wilson & Kipp, 2004). The GTAs who taught these special topics courses could develop a committee of faculty members from whom they sought professional guidance as needed. There was active student–faculty collaboration in the initial determination of course content area. Experienced in teaching other psychology courses, these GTAs are familiar with developing basic course materials and conducting a classroom. Faculty counseled GTAs on how to work with publishers to obtain examination copies of texts as well as encouragement to contact authors directly to discuss their books. Lectures were videotaped and reviewed by faculty mentors for training and development of GTAs. Faculty provided detailed feedback about teaching style and lecture content. When the first

author confronted a case of academic dishonesty by one of her students, the faculty successfully guided her through the challenging university hearing and appeals process (Keith-Spiegel, Wittig, Perkins, Balogh, & Whitley, 1993).

The students who have taken these GTA pilot courses have provided valuable feedback in the form of midterm and postcourse evaluations. Students consistently report they appreciate the attention they receive in smaller "seminar-style" classes of 35 to 60 students in contrast to the 100- to 300-student lecture courses. They particularly value the opportunity to have more specialized psychology courses taught by knowledgeable and enthusiastic instructors. At the time of publication, seven sections of one particular special topics course enrolled 450 students over five semesters. Themes evident in course evaluations included appreciation for enthusiasm and accessibility of GTAs and for specialized knowledge about their current areas of study. These courses thus appear to provide students with some of the advantages of the smaller classes typically found at smaller, liberal arts colleges. At the same time, students have access to the additional resources of the large research-oriented institutions.

In summary, these courses meet a need of the psychology department to offer additional upper division electives to both majors and nonmajors. Budget restraints can limit the ability to offer these valuable courses without the assistance of willing doctoral students. Graduate students who develop courses also benefit from the opportunity to hone a new skill set. We can now offer our graduate students more steps on a developmental track toward becoming outstanding educators. This type of mutually beneficial relationship is an example of how university departments can thrive in economically difficult times while supporting their junior colleagues in becoming more skilled teachers of psychology.

References

Brewer, C. L. (1997). Undergraduate education in psychology: Will the mermaids sing? *American Psychologist, 52,* 434–441.

Freeman, J. E. (2002). Differences in teaching in a liberal arts college versus research university. In S. F. Davis & W. Buskist (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 247–257). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Keith-Spiegel, P., Wittig, A. F., Perkins, D. V., Balogh, D. W., & Whitley, B. E., Jr. (1993). *The ethics of teaching: A casebook.* Muncie, IN: Ball State University.

Kierniesky, N. C. (1992). Specialization of the undergraduate major. *American Psychologist, 47,* 1146–1147.

McGovern, T. V., Furumoto, L., Halpern, D. F., Kimble, G. A., & McKeachie, W. J. (1991). Liberal education, study in depth, and the arts and sciences major—Psychology. *American Psychologist, 46,* 598–605.

McKeachie, W. J., & Hofer, B. K. (2001). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (11th ed.). Boston: Houghton-Mifflin.

Murray, B. (2000). Can academic values mesh with fiscal responsibility? *Monitor on Psychology, 31*(2), 21–22.

Verges, M., Spitalnick, J., & Michels, K. T. (2005, May). *Soapbox sessions: An informal teaching forum at a Research I institution.* Poster presented at the 16th annual convention of the American Psychological Society, Chicago.

Wilson, S. P., & Kipp, K. (2004). Simple and effective methods for talking about teaching. In B. Perlman, L. I. McCann, & S. H. McFadden (Eds.), *Lessons learned: Practical advice for the teaching of psychology* (Vol. 2, pp. 11–19). Washington, DC: American Psychological Society.

Notes

Avoiding Confusion Surrounding the Phrase "Correlation Does Not Imply Causation"

Julie Hatfield
New South Wales Injury Risk Management Research Centre University of New South Wales

Gavin J. Faunce and R. F. Soames Job
University of Sydney

Psychology textbooks, teachers, and students frequently employ phrases like "correlation does not imply causation." Because of the statistical meaning of the term correlation, students often misunderstand such phrases to mean that limitations on causal inference arise from the type of statistical analysis conducted. Our suggestions for teaching about the causal implications of different research designs aim to avoid confusion that is likely to be detrimental to psychology.

Psychology textbooks and teachers frequently offer students the lesson that "correlation does not imply causation," perhaps because it is simple and easily remembered. For example, in their introductory psychology textbook Gleitman, Fridlund, and Reisberg (1999) stated that "correlation does not imply causation" (p. A13) and that "correlations cannot by themselves signal a cause and effect relationship" (p. A13). In an earlier edition of the same textbook, Gleitman (1991) stated that "correlations can never provide solid evidence of a causal link" (p. A19). In another introductory psychology textbook, Wade and Tavris (1998) claimed that "correlation does not show causation" (p. 57). More recently, Smith, Bem, and Nolen-Hoeksema (2001) stated that "when two sets of scores are correlated, we may suspect that they have some causal factors in common, but we cannot conclude that one of them causes the other" (p. 530).

The Logical Intended Meaning of "Correlation Does Not Imply Causation"

When correlation is taken as a synonym of association, these claims have logical merit. An association between two variables, A and B, can mean any of three things (after dismissing the possibility of a Type I error via available statistical

techniques). First, changes in A may cause changes in B. Second, changes in B may cause changes in A. Third, changes in other, possibly unknown variables, say C, may cause changes in both A and B independently. Mere observation of an association between A and B does not allow discrimination among these three causal possibilities, so manipulation is necessary to allow inferences regarding causal direction. For example, the hypothesis that A causes B is best supported by a finding that unconfounded manipulation of A results in corresponding changes in B.

Thus, claims such as "correlation does not imply causation" identify a weakness of nonexperimental research designs relative to experimental research designs. In experimental research, the investigator deliberately manipulates at least one independent variable and observes the effects on at least one dependent variable. In nonexperimental research (unfortunately often referred to as correlational research), the investigator merely observes selected variables rather than manipulating any directly. It is this lack of manipulation that restricts causal inference.

Confusion Arising From the Statistical Meaning of Correlation and Clarification

Because correlation has a statistical meaning, psychology students are apt to misunderstand statements such as "correlation does not imply causation" as meaning that the limitations on causal inference are the fault of the statistic, rather than of the research design. We have often witnessed this confusion in students, but it is also apparent in textbooks. Although the textbooks we quoted earlier used the term *correlation* ambiguously, some statistics textbooks refer explicitly to the correlation statistic when making claims about the causal implications of correlation. For example, Welkowitz, Ewen, and Cohen (1976) stated that "you cannot determine the cause … from the correlation coefficient" (p. 159). Similarly, Kirk (1978) suggested that a "common error in interpreting a correlation coefficient is to infer that because two variables are correlated, one causes the other" (p. 108).

Limitations of Nonexperimental Research Are Not Avoided by Avoiding the Correlation Coefficient

Kirk (1978) was correct in identifying the common misinterpretation of correlation coefficients (in nonexperimental research). However, the phrasing tends to suggest that it is the fault of the statistic. It would be clearer to replace the word *correlated* with *associated* and to identify that other statistical techniques (e.g., regression, ANOVA, *t* test, chi-square, probit analysis, nonparametric techniques) also do not allow causal inferences from nonexperimental data. For example, the *t* test or one-way ANOVA is appropriate to test the relation between two observed (but not manipulated) variables when one is dichotomous and the other continuous and causal inferences are not possible. A mere observation that people who do crossword puzzles have lower cholesterol than people who do not do crossword puzzles (using a *t* test) does not imply that doing crossword puzzles reduces cholesterol. One plausible interpretation might be that people of higher socioeconomic status have both more time to do crossword puzzles and a better diet.

Furthermore, avoiding the correlation statistic in nonexperimental research does not miraculously remove the limitations on causal inferences that are inherent in such research. We have had students suggest that they dichotomize continuous variables so that they may make causal inferences from nonexperimental data by using a *t* test. Of course, this approach will not work. For example, findings that people with high residential noise exposure have lower sleep quality and psychosocial well-being than people with low residential noise exposure are causally ambiguous regardless of whether researchers detect this association using a *t* test (e.g., Öhrström, 1991) or using the full range of sound level measures and a correlation. People who live in high noise areas may be of low socioeconomic status, which may have an independent effect on their psychosocial well-being. Nonetheless, making a continuous variable dichotomous may be justified by a need to increase statistical power (Zimmerman & Zumbo, 1992).

The Correlation Coefficient Can Appropriately Lead to Causal Inferences in Some Research Designs

Welkowitz et al. (1976) were incorrect in suggesting that a correlation coefficient never allows determination of causal direction. In an experimental design in which a researcher manipulates A to investigate corresponding changes in B, a correlation coefficient supports an inference of causation as effectively as other statistical techniques. For example, in a repeated measures design, Berglund, Berglund, and Lindvall (1976) exposed participants to noise stimuli of several different sound levels and found a strong correlation between sound level and self-reported annoyance (r around .9). This finding provides compelling evidence that noise exposure influences annoyance. Of course, other statistical techniques (e.g., *t* tests or ANOVA) are more common than correlational analysis in experimental research because A is generally manipulated to only a few levels.

Furthermore, there are instances in which correlations (and other statistical techniques) may allow causal inferences from nonexperimental data. A researcher can appropriately infer that A causes B if she observes an association between A and B and can eliminate the alternative explanations (i.e., B causes A; C causes A and B) on other grounds. For example, the explanation that B causes A is less likely if A precedes B in time or if there is no logical mechanism to support this explanation. For instance, if A is gender and B is some attitudinal variable, it is illogical to argue that attitude causes gender (unless the attitude influences the gender composition of the sample).

Eliminating the explanation that C causes A and B independently is typically more difficult. This explanation may be undermined by demonstrating that all reasonable probabilities for C (which may not be simple to identify) are not associated with A and B. Even if C is associated with A and B, informed commentary is possible. For example, associations between negative attitudes toward aviation and annoyance with aircraft noise (see Job, 1988) may not indicate that such attitudes worsen annoyance but rather that noise exposure worsens both variables. Indeed, noise exposure is associated with both attitude toward aviation and annoyance with aircraft (Job, 1988). Nonetheless, Job et al. (1998) found that

negative attitudes toward aviation had an equivalent impact on annoyance regardless of noise exposure. These findings strongly suggest that the association between attitudes and annoyance is partially independent of noise exposure (although the causal direction between them remains unclear).

Suggestions for Avoiding Confusion

The intended meaning of the phrase "correlation does not imply causation" is important, and the confusion that is apt to result from it may be detrimental by contributing to inappropriate conclusions from research results. Thus, all teachers of psychology should take care to avoid the confusion surrounding the phrase by adopting the following recommendations:

1. Employ the term *correlation* only in a statistical context, and then refer specifically to correlational analysis or the correlation coefficient.
2. Employ the term *association* to describe the relation between two variables without identifying the particular statistic employed to detect the relation.
3. Refer to research designs in which researchers merely observed variables (rather than manipulating them) as nonexperimental designs rather than correlational designs.
4. Replace phrases like "correlation does not imply causation" with the phrase "without manipulation, association does not imply causation." Given people's predilection for catchy phrases, it appears wise to offer a suitable alternative to the flawed but commonly used existing phrase. Our alternative phrase correctly stresses the importance of the type of research design, rather than the type of statistical analysis performed. Nonetheless, a more substantial phrase such as "association without manipulation does not normally imply causation" would better account for circumstances in which causal inferences from nonexperimental data are defensible. A blander phrase, such as "nonexperimental designs do not imply causation" suffers from not being sufficiently catchy.

Conclusions

In our experience, many students mistakenly form the impression during their undergraduate years that limitations on causal inference are imposed by the correlation statistic rather than by nonexperimental research designs. Although objective data are not yet available, casual observation suggests that through application of the recommendations offered in this article, teachers of psychological statistics and research design can help to eliminate such errors and facilitate more appropriate inferences from nonexperimental and experimental research.

References

Berglund, B., Berglund, U., & Lindvall, T. (1976). Scaling loudness, noisiness, and annoyance of community noises. *Journal of the Acoustical Society of America, 60,* 1119–1125.

Gleitman, H. (1991). *Psychology* (3rd ed.). New York: Norton.

Gleitman, H., Fridlund, A. J., & Reisberg, D. (1999). *Psychology* (5th ed.). New York: Norton.

Job, R. F. S. (1988). Community response to noise: A review of factors influencing the relationship between noise exposure and reaction. *Journal of the Acoustical Society of America, 83,* 991–1001.

Job, R. F. S., Hatfield, J., Peploe, P., Carter, N. L., Taylor, R., & Morrell, S. (1998). Negative attitudes to noise exposure have a pure modifying effect on noise reaction. In N. L. Carter & R. F. S. Job (Eds.), *Noise effects '98: Proceedings of the 7th International Congress on Noise as a Public Health Problem* (pp. 631–634). Sydney, Australia: Noise Effects '98 Pty.

Kirk, R. E. (1978). *Introductory statistics.* Monterey, CA: Brooks/Cole.

Öhrström, E. (1991). Psycho-social effects of traffic noise exposure. *Journal of Sound and Vibration, 151,* 513–517.

Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (2001). *Fundamentals of psychology.* Fort Worth, TX: Harcourt.

Wade, C., & Tavris, C. (1998). *Psychology* (5th ed.). New York: Longman.

Welkowitz, J., Ewen, R. B., & Cohen, J. (1976). *Introductory statistics for the behavioral sciences* (2nd ed.). New York: Academic.

Zimmerman, D. W., & Zumbo, B. D. (1992). Increasing the power of tests of location and correlation by transforming scores to ranks and indicators. *Journal of General Psychology, 119,* 409–418.

Notes

1. R. F. S. Job is currently General Manager of Road Safety Strategy Branch, Road and Traffic Authority of New South Wales, Australia.
2. Send correspondence to Julie Hatfield, NSW Injury Risk Management Research Centre, University of New South Wales, New South Wales, 2052, Australia; e-mail: J.Hatfield@unsw.edu.au.

An Active Learning Classroom Activity for the "Cocktail Party Phenomenon"

Michael A. Clump
Marymount University

This article presents an active learning demonstration of the "cocktail party phenomenon." It involves dividing the students in the class into groups of 3, with 2 individuals acting as speakers and 1 person as the participant. By simultaneously involving all of the students, more students experience the effect in an environment that replicates the cocktail party phenomenon and the students' experiences outside of class, such as in a coffee house. The 3 within-subjects conditions illustrate how certain information slips through an attentional block. The students' evaluations indicated the technique was enjoyable, useful, and a good way to learn about the topic.

Everyone experiences a time when they "tune out" someone else while talking on the phone or watching television. In addition, everyone listens at a party to one person tell a story, but occasionally eavesdrops on another conversation. Cherry (1953) and Cherry and Taylor (1954) began the study of selective attention by using a laboratory task in which participants immediately repeated information presented to one ear, known as *shadowing,* and ignored information presented

to the other ear. The task, known as the dichotic listening task, involved presenting auditory information to participants through a pair of earphones. As a result, researchers could alter the information to be shadowed by the shadowed ear and the information to be ignored by the nonshadowed ear. Cherry found that participants recognized when information presented to the nonshawdowed ear was speech or when a male or female voice presented the information. Moray (1959) found that participants remembered little of the information presented to the nonshadowed ear, unless their names prefaced the information, such as instructions. Consequently, participants in a dichotic listening task remember little of the nonshadowed information because it receives less attention.

A participant's experience in the laboratory dichotic listening task compares to an individual's experience at a cocktail party. At a cocktail party, a person will be actively involved in a conversation while other conversations occur. The individual remembers the information from the engaged conversation, but cannot remember information from the other conversations. The caveat to this scenario occurs when the person hears his or her name, or other information related to the individual, spoken in one of these surrounding conversations. Like the participants in Moray's (1959) study, the individual diverts attention to the surrounding conversation. A cocktail party is a real-world example of the laboratory findings from the dichotic listening task, and thus, a second name given to the dichotic listening task is the *cocktail party phenomenon.*

Goodwin (1988) described a dichotic listening classroom activity that used students to present auditory information instead of laboratory earphones. Goodwin's activity involved two students who presented the shadowed and nonshadowed information to 4 or 5 separate participants, with each participant replicating one of Cherry's (1953) or Moray's (1959) conditions. For instance, 1 participant heard one of the presenters alter his or her pitch, another participant heard a distracting word repeated periodically, and a third condition included the participant's name in the nonshadowed passage.

I modified the laboratory dichotic listening task and Goodwin's (1988) classroom demonstration of this laboratory task to create an easy-to-replicate in-class demonstration of the cocktail party phenomenon. I developed this classroom activity to help students connect what they experience outside of class, such as at parties, at coffee houses, and at other social events, where they are in a conversation and overhear other conversations, with the laboratory findings discussed in class.

This demonstration involves all of the students working simultaneously, so all of the students actively learn about the topic through their experiences, instead of only a few students as in Goodwin's (1988) demonstration. Although having all of the students concurrently involved in the demonstration may appear chaotic, these multiple distracting inputs help them learn through an active, direct replication of their real-world experiences in their dorm rooms, in the cafeteria, in restaurants, or in coffee shops, which may not occur when involved in a replication of the laboratory task, such as in Goodwin's demonstration. Consequently, this activity fits with the design principles Mathie et al. (1993) described for an active learning activity.

My technique allows for the information presented to be altered for each participant, by incorporating first names, middle names, last names, nicknames, friends' names, common classes, and common experiences, instead of using a prerecorded, nonpersonally relevant, message as a direct replication of the laboratory task, such as Goodwin's (1988) technique would use. Thus, as students actively develop and execute this demonstration, they connect the demonstration and the discussion about the findings to their specific lives via self-reference, which Schmeck and Meier (1984) found to be a highly effective way for students to learn and remember material, and the creation of a meaningful context as Mathie et al. (1993) described. This active development of the material, which is different from using predeveloped, general stimuli from a direct replication of the laboratory task and Goodwin's demonstration, also provides students with an opportunity to actively design and conduct research in cognitive psychology.

Description of the Activity

The activity involves dividing the entire class of students into groups of 3 students independent of gender, such that my typical class with 30 students would work in 10 groups. The activity uses a within-subjects design with three conditions. Each group of students decides who will act as the Participant the two Speakers. The Participant in the activity always shadows Speaker 1 and tries to ignore Speaker 2; Speaker 1 tells an interesting, but different, story to the Participant during the three conditions; Speaker 2 also says information to the Participant, but what Speaker 2 says is different in the three conditions. The groups of students arrange themselves so that the Participant sits between the two speakers. The Participants leave the room for a few minutes as I describe the demonstration to the students who will serve as the two speakers for their groups.

Although all of the groups simultaneously conduct the activity, for ease of presentation I describe the activity of one group. Once the Participant exits the room, I assign one student to be Speaker 1 and the other student to be Speaker 2. After explaining Speaker 1's role, which involves continuously telling a story the Participant repeats during each condition, I describe the role of Speaker 2 (i.e., the nonshadowed speaker), who alters the dictated information, based on the following within-subjects conditions:

Condition 1: Speaker 2 says random words and numbers at 1 word/sec, without including any interesting information.
Condition 2: Speaker 2 again says random words and numbers at 1 word/sec, occasionally including the name of the Participant.
Condition 3: Speaker 2 tells an interesting story that includes the Participant's name and information from school, classes, friends, jobs, relatives, or other information specifically relevant to the Participant. (Because students work together on other

group projects and assignments in the course, this type of information typically becomes easily available.)

The two Speakers develop the stories and word lists they will use during the three conditions as the Participant reenters class. The Participant sits between the other two group members, both of whom face the Participant and talk directly into the Participant's ears from less than 1 ft away once the demonstration starts. I then explain how the Participant will immediately repeat exactly what Speaker 1 dictates, while ignoring what Speaker 2 says. Finally, I explain that the Participants will be involved in the task three separate times for 1 min each.

After 1 min of the activity, the group stops and the Participant writes down for 1 min exactly what he or she recalls from the shadowed and nonshadowed ears, during which Speaker 1 and Speaker 2 refresh their stories and word lists. The group repeats the procedure for all three conditions.

Following the demonstration, I ask all Participants to describe the information that they recalled from both the shadowed and nonshadowed ears for the three conditions. Invariably, the Participants describe how their recall was excellent for the shadowed material for Conditions 1 and 2, but the shadowed information recalled for Condition 3 was substantially reduced. In contrast, the Participants' recall of the nonshadowed information significantly increased from Condition 1 (e.g., very little), to Condition 2 (e.g., their names multiple times), to Condition 3 (e.g., most of the material presented by Speaker 2). For each Participant, the entire activity required between 15 and 20 min; consequently, we repeat the activity so the students can rotate through the three roles during a class period and experience the activity from each role.

Evaluation of the Activity

Students (*N* = 22) evaluated the activity on six questions adapted from Madson (2001). Over 90% of the students found the demonstration to be either enjoyable or very enjoyable, and all of the students agreed or strongly agreed with using the demonstration in the future. All the students found the demonstration to be useful or very useful for prompting them to think about the cocktail party phenomenon and dichotic listening. Nearly 91% of the students believed they learned more about the topic having participated in the demonstration. The students did not feel that the demonstration was a waste of their time, as evidenced by the fact 95.5% selected strongly disagree or disagree. Finally, over 80% of the students selected agree or strongly agree for the statement that other instructors should use the demonstration when teaching about the topic in other related courses.

Discussion

The comments made by the students indicated that they enjoyed the activity, especially with the personally relevant information. The Participants indicated they wanted to keep listening to the nonshadowed information, especially when it contained gossip, information related to a course, or a funny story. The students stated that the activity helped them understand Cherry's (1953), Cherry and Taylor's (1954), and Moray's (1959) laboratory findings on selective attention. An instructor could also use the activity to discuss how to design and conduct research in cognitive psychology. The students could redesign the demonstration to effectively measure their inclinations to listen to the nonshadowed information, especially the personally relevant information, which would further involve the students in actively learning about these research studies.

The activity described by Goodwin (1988) and my activity provide faculty with separate demonstrations of selective attention. Whereas Goodwin's technique provides an excellent method to specifically replicate the laboratory nature of the dichotic listening task, my technique focuses on tying dichotic listening to the students' real-world experiences, like the cocktail party phenomenon, by attempting to connect what students experience outside of class with research studies discussed in class. In addition, all of the students benefit from being involved in the active learning focus of the demonstration. The focuses of the activity, involving all of the students in the demonstration and connecting it with their everyday life by having them apply, evaluate, and synthesize the topic, match the main principles of active learning activities espoused by Mathie et al. (1993).

References

Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25,* 975–979.

Cherry, E. C., & Taylor, W. K. (1954). Some further experiments upon the recognition of speech with one and two ears. *Journal of the Acoustical Society of America, 26,* 554–559.

Goodwin, C. J. (1988). Selective attention with human earphones. *Teaching of Psychology, 15,* 104–105.

Madson, L. (2001). A classroom activity exploring the complexity of sexual orientation. *Teaching of Psychology, 28,* 32–35.

Mathie, V. A., Beins, B., Benjamin, L. T., Jr., Ewing, M. M., Hall, C. C. I., Henderson, B., McAdam, D. W., & Smith, R. A. (1993). Promoting active learning in psychology courses. In T. V. McGovern (Ed.), *Handbook for enhancing undergraduate education in psychology* (pp. 183–214). Washington, DC: American Psychological Association.

Moray, N. (1959). Attention in dichotic listening: Affective cues and the influence of instructions. *Quarterly Journal of Experimental Psychology, 11,* 56–60.

Schmeck, R. R., & Meier, S. T. (1984). Self-reference as a learning strategy and a learning style. *Human Learning: Journal of Practical Research & Applications, 3,* 9–17.

Notes

Understanding the Applicability of Social Psychology: The Benefits of a Semiweekly Journal Assignment

Steven M. Graham
Carnegie Mellon University

This article describes a journal assignment for an introductory social psychology course. The semiweekly assignment was to apply a recent course topic to one's life, to another person's life, or to current world events. The educational goals of this assignment were to (a) help students understand the applicability of social psychology to everyday life, (b) facilitate course discussion, and (c) recognize and correct student misconceptions of course material. Students rated these assignments as useful, recommended including them in future sections of the class, indicated that they became easier over the course of the semester, and performed better on them as the semester progressed.

Social psychology is fundamentally about life as a human being. Social psychological theory and research speak to a number of topics that touch everyday life such as close relationships, social influence, and aggression. In my view, a primary goal of an introductory social psychology course is to facilitate students' understanding of the applicability of social psychology to their everyday lives (cf. Brender, 1982; Lakin & Wichman, 2005; Merrill, 2001). To that end, I used and evaluated a journal assignment with the primary purpose of facilitating such understanding. Two additional goals of this assignment were to encourage course discussion and to recognize and correct student misconceptions of course material.

The Assignment

Instructors have assigned journals in psychology courses for many years (e.g., Connor-Greene, 2000; Hettich, 1990; Miller, 1997; Weber, 1984). Using instructions similar to Weber's (1984), I included the following description of journal assignments in the syllabus:

> The goal of these assignments is to describe in approximately one page how you've seen material described in class at work in the real world. You can discuss how social psychological findings discussed in the last few lectures helped you understand something from your own life, the behavior or attitudes of somebody that you know, or major news events (e.g., violent acts, political campaigns, etc.). Be sure to discuss the social psychological phenomenon covered in class as well as what it has helped you understand. You may also describe an event in which you observed something that was contradictory to social psychological theory or findings.

Students completed these assignments weekly, except in those weeks in which there was an exam, a paper, or a shortened schedule, resulting in eight assignments across the semester. Students submitted assignments via e-mail the night before the last class session of the week. I informed students that I would use some of their journal examples during lectures and class discussions and thus they should indicate if a particular example was confidential.

Most students completed the assignments on time, followed directions, and often provided insightful examples of course concepts. I graded journal assignments promptly and returned them during the next day's class. I assigned grades on a 10-point scale, using accuracy of descriptions of course material and applicability of examples as primary criteria. Although I did not permit students to skip any assignments, I accepted late assignments with a 1-point deduction per day. Combined, journal assignments comprised 15% of students' final grades. Grading, for a class size of 34, took approximately 2 hr to complete each assignment.

Evaluation and Conclusions

Helping Students Understand the Applicability of Social Psychology to Everyday Life

Students (29 out of 34) completed evaluations of course components at the end of the semester. On a scale ranging from 1 (*not at all*) to 7 (*very much*), students rated the journal assignments as useful in terms of understanding the applicability of social psychology to everyday life ($M = 6.45$, $SD = 1.12$), recommended including them in future sections of the course ($M = 6.69$, $SD = 0.92$), and rated them as becoming easier to do over the course of the semester ($M = 5.44$, $SD = 1.17$). Lending support to the reports indicating that the assignments became easier over time was the finding that students earned higher grades on the last four journal assignments ($M = 9.48$, $SD = .55$) than on the first four ($M = 9.29$, $SD = .47$), $t(33) = 2.73$, $p = .01$. Several students spontaneously commented that the assignments helped them understand how they could relate social psychology to their lives. Two examples of such comments were: "Journal assignments were great because they required students to assess their own life/experiences in terms of social psych," and "Journal assignments helped me make a personal connection to the material." Student performance on journal assignments also was correlated with performance on exams, suggesting that putting effort into the assignments led to a greater understanding of course material. To rule out the possibility that students who were more engaged with the class performed better on both the journals and the exams, I examined partial correlations between performance on the journal assignments and each exam, controlling for class attendance and participation grades. This correlation was not significant for the first class exam, $r(31) = .18$, $p = .31$, but was for both the second, $r(31) = .64$, $p < .001$, and third exams, $r(31) = .39$, $p < .05$. This finding suggests that over time (before the second and third exams), journal assignments facilitated understanding of course material.

Facilitating Course Discussion

As a first-time instructor of a social psychology course, I found that the journal assignments were a good source of lecture examples that I will continue to use. One student provided an entertaining (and applicable) media example of the effects of insufficient justification on attitudes (Festinger & Carlsmith, 1959). This example compared the attitudes of women on two different reality television programs, *Joe Mil-*

lionaire (Cowan & Michenaud, 2003) and *The Bachelor* (Fleiss, 2002). The student reported that in the former show, women received valuable jewelry, vacations, and clothing each time they advanced to the next round in the ultimate competition to win the affection of "Joe Millionaire." In the latter, women simply received a flower each time they advanced to the next round in a similar competition. This student reported that the women on *Joe Millionaire* seemed less attached to the object of the competition because they had a great deal of external justification for treating him well, whereas those on *The Bachelor* seemed genuinely to care about the object of their competition. This example was highly entertaining to the class and, more important, provoked discussion from several students.

Recognizing and Correcting Student Misconceptions of Course Material

Recall that in addition to applying an example to everyday life, one part of the journal assignments was to describe a phenomenon discussed in class. Student descriptions of class phenomena provided an excellent opportunity to ensure that they understood course material and to correct misconceptions when they arose. As one example, a student wrote about how the mere exposure effect did not apply to his experience with his first-year college roommate. This student initially disliked his roommate and reported that his dislike only increased with greater exposure to his roommate. This student had misunderstood course material and had not understood or remembered that one must initially find the stimulus (in this case, the roommate) neutral or better for mere exposure to result in increased liking (Perlman & Oskamp, 1971). I was able to recognize and rectify this student's misunderstanding before the next exam by writing detailed feedback on the assignment.

Conclusions

This article describes the successful application of a journal assignment in an introductory social psychology course. The benefits of this assignment were numerous and would likely generalize to introductory courses in other areas of psychology as well as more advanced courses within social psychology. I encourage instructors to use this or a similar assignment in future sections of courses in which they wish to emphasize the applicability of course material to everyday life. Of course, the time required to grade these assignments might be prohibitive in large classes. In such cases, I would encourage instructors to require fewer assignments or to have students compile their assignments into a portfolio to turn in as a final project.

References

Brender, M. (1982). The relevance connection: Relating academic psychology to everyday life. *Teaching of Psychology, 9,* 222–224.

Connor-Greene, P. A. (2000). Making connections: Evaluating the effectiveness of journal writing in enhancing student learning. *Teaching of Psychology, 27,* 44–46.

Cowan, C., & Michenaud, J. (Executive Producers). (2003). *Joe millionaire* [Television series]. Los Angeles: Fox.

Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *Journal of Abnormal and Social Psychology, 58,* 203–210.

Fleiss, M. (Executive Producer). (2002). *The bachelor* [Televeision series]. Los Angeles: ABC.

Hettich, P. (1990). Journal writing: Old fare or nouvelle cuisine? *Teaching of Psychology, 17,* 36–39.

Lakin, J. L., & Wichman, A. L. (2005). Applying social psychological concepts outside the classroom. *Teaching of Psychology, 32,* 110–113.

Merrill, M. D. (2001). First principles of instruction. *Journal of Structural Learning & Intelligent Systems, 14,* 459–466.

Miller, S. (1997). Self-knowledge as an outcome of application journal keeping in social psychology. *Teaching of Psychology, 24,* 124–125.

Perlman, D., & Oskamp, S. (1971). The effects of picture content and exposure frequency on evaluations of Negroes and Whites. *Journal of Experimental Social Psychology, 7,* 503–514.

Weber, A. (1984). Teaching social psychology. *Contemporary Social Psychology, 10*(3), 9–10.

Notes

A Structured Interview Demonstration

Lynn K. Bartels, Cynthia R. Nordstrom, and Jennifer A. Koski
Southern Illinois University Edwardsville

Structured interviews involve asking interviewees the same questions and then scoring their responses on a rating scale. The benefits of structured interviews include better reliability, validity, and legal defensibility. In this classroom activity, students role-played both a structured and an unstructured interview. The class observed and rated the interviewees' responses to both types of interview questions and then contrasted the interview formats and ratings. The activity promoted student learning and generated positive student reactions. Psychology instructors can use the activity to help students understand issues surrounding employee selection, standardization, and scorer reliability.

Employers use different tools to select employees including tests, work samples, and interviews. Of these, the employment interview remains one of the most popular (McDaniel, Whetzel, Schmidt, & Maurer, 1994). Interviews vary in degree of structure. Structured interviews involve a predetermined set of job-related questions and rating procedures. They show clear advantages over unstructured interviews in which interviewers ask different questions of each interviewee. The advantages of structured interviews include better reliability (Conway, Jako, & Goodman, 1995), validity (McDaniel et al., 1994), and le-

gal defensibility (Williamson, Campion, Malos, Roehling, & Campion, 1997). We describe a classroom activity that illustrates structured interviewing.

Interview Demonstration

The instructor asks for two volunteers (interviewer and interviewee). The instructor gives the interviewer sets of unstructured and structured interview questions. Alternatively, the class could develop interview questions. Structured interview questions tend to be more job-relevant and to include more specific rating criteria than questions used in unstructured interviews. For the structured interview, the interviewer asks three questions. A sample structured interview question is "A customer walks into the store. No other salespeople are around to help the person, and you are busy arranging merchandise. What would you do?" (Heneman, Heneman, & Judge, 1997, p. 456). Campion, Palmer, and Campion (1997) provided sample structured interview questions and rating scales.

"What are your career goals?" and "What are your weaknesses?" exemplify common unstructured questions. The questions are generic and not tied to a specific job. The interviewer chooses three unstructured questions to ask from the list provided. Unstructured interview questions are available from career centers, the Internet (e.g., http://www.careercity.com/jobseeker/careerresources/interviewing/traditional.shtml), and interviewing books (e.g., Yate, 1993). We used a sales job, but any familiar job can work (e.g., server, secretary). The interviewee's answers should be average, allowing room for interviewer judgment. Exceptionally good or poor answers can lead to lack of scoring variability.

The class observes the structured and unstructured interviews and rates the interviewee's six answers. The unstructured rating scales have generic rating anchors (e.g., 1 [*low*] to 5 [*high*]) whereas the structured rating scales use more specific behavioral anchors (e.g., typical answers, critical issues). For example, sample behavioral anchors for the preceding question about the busy salesperson might be 1 (*Keep on arranging merchandise*); 3 (*Keep on working, but greet the customer*); and 5 (*Stop working, greet customer and offer to provide assistance*) (Heneman et al., 1997, p. 456). Students rate each response. Then, on the board, tally how many students gave each rating. For example, starting with Question 1, how many students gave a rating of 1, 2, and so forth?

When examining the ratings on the board, there is often a higher level of consistency in the structured interview ratings. Interrater reliability should be higher for the structured ratings. Structured interviews limit rater subjectivity. As predicted, when we used this demonstration in an industrial/organizational psychology class, there was more variability among the unstructured ratings. The unstructured ratings had higher standard deviations (range = .68 to .79) when compared to the structured ratings' standard deviations (range = .37 to .56; both on 5-point rating scales). The unstructured rating with the highest percentage of student endorsement varied from 43.2% to 56.8% across the three questions. Yet, the structured rating with the highest percentage of student endorsement varied from 79.5% to 84.6%.

Clearly, the structured ratings produced a higher level of agreement than the unstructured ratings.

The demonstration could be extended to multiple interviewees to show how unstructured interviews vary across interviewees. In the unstructured interview, interviewers ask different questions of each interviewee, which is similar to giving a different test to each job applicant. The class might discuss fairness issues and the importance of standardization.

Some interviewees perceive structured interviews as fairer assessments of their ability. The structured interview questions are generally more content valid because they have a clearer relation to the job duties. On the other hand, interviewees sometimes believe they are better able to showcase their talents in an unstructured interview because they have more opportunity to manage the interviewer's impressions and present additional information. Interviewers often prefer the unstructured interview because they have more latitude in questioning the interviewee. It is important to keep the interview related to the job and avoid potentially illegal questions (e.g., "How old are you?"), which can be problematic, particularly with the less planned unstructured interviews.

Students often remark about the transparency of the structured interview questions allowing applicants to give socially desirable answers. Campion et al. (1997) suggested adding a conflict (e.g., the CEO requests confidential materials that he or she is not authorized to access) can make the question more difficult. We also talk about the surprising level of honesty that applicants use to answer questions (e.g., a firefighter who admits fear of heights).

Assessment

Thirty-nine students from an undergraduate industrial/organizational psychology course at a Midwestern regional university participated in the demonstration. Afterward, we asked them to complete measures of reactions and learning.

Reaction Measure

We assessed student reactions to the demonstration using a five-item reaction questionnaire. Students indicated their perceptions of the content, presentation clarity, effectiveness, involvement, and learning on a Likert scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). The scale's coefficient alpha was .84. Examination of the means and standard deviations for the reaction measure indicated that students expressed favorable reactions to the demonstration (overall $M = 5.31$, $SD = .84$).

Learning Measure

We assessed learning with a brief quiz consisting of four 1-point multiple-choice items (e.g., "Which type of interview should be used to increase interrater reliability?") administered immediately after the lecture. Students performed well on the quiz ($M = 3.59$, $SD = .64$). We conducted a follow-up pretest–posttest evaluation with 56 social psychology students. There was a statistically significant increase from the

pretest ($M = 2.41$, $SD = 1.06$) to the posttest ($M = 4.36$, $SD = .82$), $t(55) = -12.49$, $p < .0001$. Students performed significantly better on the posttest.

Discussion

Student reactions and quiz scores indicated that the structured interview demonstration is an effective method of illustrating the differences between structured and unstructured interviews. Although we cannot necessarily attribute their learning solely to the interview demonstration, the interview demonstration can be a useful and enjoyable way to teach students about structured interviews, employee selection, interrater reliability, and standardization.

Psychology instructors can use this active learning demonstration to communicate information about the importance of reliability, job-relatedness, and standardization of procedures. Students often struggle to understand these issues, finding them dry, abstract concepts. This activity provides an opportunity to see how these psychometric concepts manifest themselves in a real-world situation. Students can explore a practical and legal decision (i.e., who gets hired and on what basis?). They see a vivid example of how such a decision may be compromised with unstructured interviewing procedures. Another advantage of this exercise is that students learn about different interview approaches used in the job market. Finally, students who may become managers develop a better understanding of how to select employees using fair and legally defensible techniques.

In conclusion, instructors can use this activity in several classes including industrial/organizational psychology, personnel psychology, social psychology, and tests and measurements to discuss structured interviewing. Importantly, students enjoy its practical focus and learn about an important employee selection technique.

References

Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology, 50,* 655–702.

Conway, J. M., Jako, R. A., & Goodman, D. F. (1995). A meta-analysis of interrater and internal consistency reliability of selection interviews. *Journal of Applied Psychology, 80,* 565–579.

Heneman, H. G., III, Heneman, R. L., & Judge, T. A. (1997). *Staffing organizations* (2nd ed.). Middleton, WI: Mendota.

McDaniel, M. A., Whetzel, D. L., Schmidt, F. L., & Maurer, S. D. (1994). The validity of employment interviews: A comprehensive review and meta-analysis. *Journal of Applied Psychology, 79,* 599–616.

Williamson, L. G., Campion, J. E., Malos, S. B., Roehling, M. V., & Campion, M. A. (1997). Employment interview on trial: Linking interview structure with litigation outcomes. *Journal of Applied Psychology, 82,* 900–912.

Yate, M. (1993). *Hiring the best: A manager's guide to effective interviewing.* Avon, MA: Adams Media.

Note

Deviations From APA Style in Textbook Sample Manuscripts

Kim Ernst
Loyola University New Orleans

Lacie Michel
Department of Sociology
Louisiana State University

We examined sample manuscripts in research methods textbooks (N = 29) for deviations from American Psychological Association (APA) style. Formatting errors totaled 225. Because these errors can adversely affect students' learning and their grades, we recommend textbook authors give greater attention to APA style in their sample manuscripts. We also encourage course instructors to resolve formatting errors in textbook sample manuscripts before students begin to learn APA style. Alternatively, course instructors might suggest that students use the Publication Manual *(APA, 2001) exclusively to learn APA format and prepare research papers and lab reports.*

Learning to prepare manuscripts according to American Psychological Association (APA) style is an important course topic (Landrum, Shoemaker, & Davis, 2003), yet it can be a daunting task for students (Austin & Calderon, 1996; Hummel, Kaeck, Whatley, & Monetti, 2004). Indeed, Madigan, Johnson, and Linton (1995) regarded the acquisition and mastery of APA style as challenging for both students and teachers of psychology. Authors of research methods textbooks (RMTs) seem to recognize these sentiments because APA style consumes a substantial proportion of most RMTs (Jackson, Lugo, & Griggs, 2001). Besides providing condensed versions of the *Publication Manual of the American Psychological Association* (PM; APA, 2001), RMTs typically include a sample manuscript to illustrate how an APA-style manuscript should look. Juve, Weiser, Davis, and Rewey (2000) identified the most prevalent APA formatting errors in 69 student manuscripts submitted for publication in the *Psi Chi Journal of Undergraduate Research*. Juve et al. suggested that students are prone to APA formatting errors because the *PM* "overwhelms many students" (p. 2). Although this explanation may account for some students' formatting errors, the explanation also suggests that students are likely to use RMTs as guides rather than the *PM* to prepare APA-style manuscripts. Thus, examining sample manuscripts in RMTs for accuracy in presenting APA format is necessary.

Method

The Sample of Books

The initial sample consisted of 29 RMTs published between 2001 and 2005 (this list is available on request). We chose this time frame because we wanted to include the most recent RMTs available at the time of this study. We estimated the sample represents at least 90% of the published RMTs. We excluded books published outside the United

States, combined research methods and statistics texts, writing manuals, and student study guides.

Procedure

We searched the table of contents and the subject index of each RMT for sections addressing APA style. Then, in random order, we examined only those books providing APA sample manuscripts ($N = 27$). Using the procedure of Juve et al. (2000), we recorded the type of deviation and the number of different deviations for each sample manuscript. Next we assigned each deviation to one of eight categories: manuscript format, title page, abstract, manuscript text, references, table, figure, and author note. We did not include repetitions of a deviation within a single manuscript. To assess coding reliability we each rated identical sample manuscripts, with 50% overlap between us. Average reliability (coding agreement) across all topics exceeded .92.

Results and Discussion

The number of APA formatting errors in sample manuscripts totaled 225 ($M = 8.33$, $SD = 5.55$; a description of these errors is available on request). One RMT sample manuscript contained only a single deviation whereas another sample manuscript displayed as many as 21 deviations. With respect to the eight coding categories, the number of deviations from APA style, ordered by frequency, was manuscript format (63), manuscript text (39), title page (34), table (34), references (24), figure (12), author note (11), and abstract (8).

Examples of manuscript format and manuscript text deviations concerned line spacing (14), page headers (8), page justification (3), spacing following statistical tests (7), and the occurrence of hyphens at the end of text lines (19). More serious deviations involved two sample manuscripts that included information about the participants, procedures, and materials in the method sections but omitted the respective subsection headings. Some deviations appearing on title pages included title positioning (13); wording, case lettering, positioning, and spacing of the running head (13); and title pages without page headers and page numbers (2). Examples of APA style deviations on reference pages were underlined book titles, journal titles, and volumes (9) and references with paragraph indents (6). Table titles with case lettering errors (8) and test statistics not italicized (7) also occurred. Of the 17 RMTs that included a figure, there were four instances of figures appearing on pages with page headers and page numbers and one instance of three figures displayed on a single page. The most prevalent author note deviation concerned the sequence of information presented (10). Examples of abstract errors were paragraph indentations and exceeding the 120-word limit (8).

One explanation that could account for the deviations in these RMT sample manuscripts concerns the recent APA style changes introduced in the fifth edition of the *PM*. However, only about 26% of the 225 deviations in sample manuscripts represented changes set forth in the latest *PM*. Thus, the overwhelming majority of deviations from APA style in sample manuscripts were not solely attributable to the latest

PM but reflected departures from earlier *PM* editions (e.g., APA, 1983, 1994).

Another explanation is that some deviations result from publishers' typesetting decisions and RMT authors' pedagogical preferences. For example, we suspect the occurrence of line spacing errors reflects publishers' concerns about page space economy and design. Similarly, the use of a bold typeface could represent textbook authors' deliberate efforts to highlight sample manuscript headings, sections, and subsections. Yet, such deviations from APA style are likely to mislead students about how an APA-style manuscript is supposed to look.

The general conclusion derived from this investigation is that some sample manuscripts in RMTs are inconsistent with the current *PM*. Our findings underscore the need for RMT authors to correct APA style errors for subsequent editions. If students rely on sample manuscripts in RMTs to prepare APA-style papers, they might easily adopt the incorrect format. Thus, we recommend that instructors caution students about relying on the appearance of RMTs' sample manuscripts. We also encourage instructors to review their course textbooks and resolve deviations from the *PM* before students prepare APA-style papers and lab reports. Alternatively, instructors might ask students to use only the *PM* to learn APA style.

The implications of this study are important because deviations from the *PM* in sample manuscripts can adversely affect students' learning and their grades. Similarly, formatting errors are certain to influence journal editor comments (Brewer, Scherzer, Van Raalte, Petitpas, & Anderson, 2001; Smith, 1998, 2000). That we found several RMTs to contradict the *PM* might also thwart instructors' teaching efforts and lead to students' frustration with learning APA format. Ensuring that students receive accurate formatting instructions will not only help them learn APA style correctly but will also make this seemingly daunting task less frustrating.

References

American Psychological Association. (1983). *Publication manual of the American Psychological Association* (3rd ed.). Washington, DC: Author.

American Psychological Association. (1994). *Publication manual of the American Psychological Association* (4th ed.). Washington, DC: Author.

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: Author.

Austin, J. T., & Calderon, R. F. (1996). Writing in APA style: Why and how. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants* (pp. 265–281). Thousand Oaks, CA: Sage.

Brewer, B. W., Scherzer, C. B., Van Raalte, J. L., Petitpas, A. J., & Anderson, M. B. (2001). The elements of (APA) style: A survey of psychology journal editors. *American Psychologist, 56,* 266–267.

Hummel, J. H., Kaeck, D. J., Whatley, M. A., & Monetti, D. M. (2004). Using the 5th edition of the APA style guide. *Eye on Psi Chi, 9*(1), 18–21.

Jackson, S. L., Lugo, S. M., & Griggs, R. A. (2001). Research methods textbooks: An objective analysis. *Teaching of Psychology, 28,* 282–288.

Juve, J. A., Weiser, A., Davis, S. F., & Rewey, K. L. (2000, April). *"There are HOW MANY spaces after punctuation!": APA format er-*

rors. Poster presented at the annual meeting of the Southwestern Psychological Association, Dallas, TX.

Landrum, R. E., Shoemaker, C. S., & Davis, S. F. (2003). Important topics in an introduction to the psychology major course. *Teaching of Psychology, 30,* 48–51.

Madigan, R., Johnson, S., & Linton, P. (1995). The language of psychology: APA style as epistemology. *American Psychologist, 50,* 428–436.

Smith, R. A. (1998). Another perspective on publishing: Keeping the editor happy. *Psi Chi Journal of Undergraduate Research, 3,* 51–55.

Smith, R. A. (2000). Documenting your scholarship: Citations and references. In R. J. Sternberg (Ed.), *Guide to publishing in psychology journals* (pp. 146–157). Cambridge, England: Cambridge University Press.

Notes

The Nonspecialist and the History of Psychology Course

Bruce B. Henderson
Western Carolina University

Most history of psychology instructors are not expert in general cultural history, historiography, or philosophy of science. Trying to provide students with a comprehensive and intellectually honest view of the development of psychology when the instructor lacks such background presents serious challenges. I argue that the nonspecialist history of psychology teacher needs help from expert history and philosophy of science teachers. Most important, nonspecialists need pedagogical content knowledge in the form of examples and demonstrations to represent to students the connections among the history of psychology, general history, and philosophy of science.

The American Psychological Association's Ethical Principles of Psychologists and Code of Conduct (APA, 2002) state that ethical psychologists teach "within the boundaries of their competence" (p. 1063). I know of no data on how often psychology teachers question their competence to teach particular courses. However, in their study of the ethics of teaching, Tabachnick, Keith-Spiegel, and Pope (1991) found that 38.6% of their 482 academic psychologist respondents admitted to at least sometimes teaching material they had not mastered. Perhaps the courses in which instructors are most likely to feel less than totally competent are those with a strong interdisciplinary component. The history of psychology course is an example.

Many psychology departments use the history of psychology course to provide an integrative capstone experience (Brewer et al., 1993; Raphelson, 1982). Some critics have suggested the course, as typically taught, might not be up to the task. For example, Dagenbach (1999) argued that the course frequently focuses narrowly on the history of experimental psychology and needs to be broadened to help students clearly connect the research of the past 20 or 30 years to its historical roots. Others have argued that the history course provides the last best hope for psychology students to identify some threads of unification in the discipline, allowing students to see how their various courses are connected (Goodwin, 2002; Hilgard, Leary, & McGuire, 1991). In a 1997 survey, the history of psychology was one of the six or seven most common courses in the psychology curriculum (Perlman & McCann, 1999). Its ubiquity suggests the continuing importance of the history course. A more recent survey (Fuchs & Viney, 2002) revealed similar results and also found that few teachers of the history course have specialized training in the history of psychology. Most respondents to the survey taught the course because of personal interest and had educated themselves about history through independent reading.

Like teachers in all courses, the teacher of the history course faces two sets of challenges. One set of challenges is pedagogical: The teacher needs to find ways to engage students in a field in which they believe they have no interest. Assignments in the course need to actively involve the students, ideally as both budding psychologists and as budding historians. Many activities designed to gain student interest are available. They include projects in which students trace the academic genealogies of their department's teachers (Goodwin, Dingus, & Petterson, 2002; Terry, 1980; Weigel & Gottfurcht, 1972), compare introductory textbooks across time (Webb, 1991; Zehr, 2000), present classic papers at mock conventions (Cole, 1983), and create mock-ups of newspapers celebrating historical events (Bryant & Benjamin, 1999). The history of psychology teacher who wants to meet the pedagogical challenges has a lot of help.

The second set of challenges faced by the teacher of the history of psychology is academic. The history of psychology teacher has to have a credible and creditable knowledge base about at least four areas: (a) theories, events, and people in the history of psychology itself; (b) philosophy and philosophy of science, especially with regard to the history of science; (c) the general intellectual and social history within which psychology has developed; and (d) historiography, the approaches and methodologies historians use to do history. Many history of psychology teachers avoid two other important areas of needed expertise for the history of psychology teacher, the history of philosophy (Robinson, 1979) and history of psychology in non-Western countries (Goodwin, 2002), by invoking the limits of a one-semester course.

The first academic area, the history of psychological events and people, probably presents the fewest problems. Most teachers of the history of psychology, at least the volunteers, came to the course with a personal interest in the history of psychology. They are likely to be well-read beyond the

textbook. The philosophy of science issues may be more troublesome. According to anecdotal lore (Goodwin, 2002) and some data (Fuchs & Viney, 2002), teachers of the history of psychology are older than the average psychology teacher. Many were educated in a positivist tradition that does not provide an easy transition to more historically based philosophies of science. Even history of psychology teachers who have learned some philosophy of science from Kuhn (1970) or Laudan (Gholson & Barker, 1985; Laudan, 1977, 1996) usually are novices. No matter how we approach the philosophy of science in our courses, we are likely to feel far less than up-to-date and surely not completely competent.

The other two areas of competence, general historical knowledge and historiography, provide the truly serious challenges. How can the history of psychology teacher with no in-depth training in history or historiography feel comfortable about making connections between events in psychology and general culture? Even more daunting is the prospect of grasping the implications of the debates (Danziger, 1994; Himmelfarb, 2004; Leahey, 2004) about old versus new history (history as the study of the great people and institutions vs. history as the study of the lives of common people) or old versus new versus new-new history (postmodernism) and its even newer versions (including postpostmodernism). The conscientious teacher of history of psychology is left worrying about being too internalist and celebratory (cheering the great accomplishments of the discipline); being too Whiggish (interpreting the past as the prelude to the truth of the present); or being too racist, sexist, classist, or something else-ist that has just been invented. In short, how can the typical teacher of the history of psychology ever hope to be competent in the history part of the history of psychology course?

What would help? The nonspecialist can tap into existing resources, including the best of the history of psychology Web sites (e.g., York University's at http://psychclassics.yorku.ca). Nonspecialists can reach out to colleagues in philosophy and history, audit their courses, invite them to guest lecture, and engage them in less formal discussions of new ideas and accessible writings in their fields (see Appendix for sample readings). Nonspecialists also need new resources. For example, *Teaching of Psychology* could create a section on interdisciplinary updates for teachers in which psychologists who do scholarly research in history or philosophy or historians or philosophers who do scholarly research in psychology could provide basic overviews that inform nonspecialist teachers about recent trends and provide guides to the literature that would not be overwhelming. The authors of textbooks in history of psychology could provide instructor's manuals that present extensive material on trends in philosophy of science, historiography, and general historical trends. Division 26 (Society for the History of Psychology) of APA could offer workshops (perhaps online) to provide updated information for nonspecialists. Perhaps what the nonspecialist needs most is the aid of expert philosophy and history teachers who can provide pedagogical content knowledge (Shulman, 1987). Pedagogical content knowledge from interdisciplinary expert teachers could offer nonspecialists ways to represent for their students the connections among the history of psychology, general history, and the philosophy of science. Included among these representations would be the expert teachers'

most powerful examples, analogies, demonstrations, and other means for making comprehensible history and philosophy of science connections to history of psychology. Measures like these will not allow nonspecialists to reach the elusive high standard of competence but they could make us more competent than we are.

References

American Psychological Association. (2002). Ethical principles of psychologists and code of conduct. *American Psychologist, 57,* 1060–1073.

Brewer, C. L., Hopkins, J. R., Kimble, G. A., Matlin, M. W., McCann, L. I., McNeil, O. V., Nodine, B. F., Quinn, V. N., & Saundra. (1993). Curriculum. In T. V. McGovern (Ed.), *Handbook for enhancing undergraduate education in psychology* (pp. 161–182). Washington, DC: American Psychological Association.

Bryant, W. H. M., & Benjamin, L. T., Jr. (1999). Read all about it! Wundt opens psychology lab: A newspaper assignment for history of psychology. In L. T. Benjamin, Jr., B. F. Nodine, R. M. Ernst, & C. Blair-Broeker (Eds.), *Activities handbook for the teaching of psychology* (Vol. 4, pp. 47–49). Washington, DC: American Psychological Association.

Cole, D. L. (1983). The way we were: Teaching history of psychology through mock APA conventions. *Teaching of Psychology, 10,* 234–236.

Dagenbach, D. (1999). Some thoughts on teaching a pluralistic history in the history and systems of psychology course. *Teaching of Psychology, 26,* 22–28.

Danziger, K. (1994). Does the history of psychology have a future? *Theory & Psychology, 4,* 467–484.

Fuchs, A. H., & Viney, W. (2002). The course in the history of psychology: Present status and future concerns. *History of Psychology, 5,* 3–15.

Gholson, B., & Barker, P. (1985). Kuhn, Lakatos, and Laudan: Applications in the history of physics and psychology. *American Psychologist, 40,* 755–769.

Goodwin, C. J. (2002, August). *The future of history.* Paper presented at the annual meeting of the American Psychological Association, Chicago.

Goodwin, C. J., Dingus, M., & Petterson, S. (2002). The genealogy project: Tracing academic roots in the history and systems course. *Teaching of Psychology, 29,* 61–63.

Hilgard, E. R., Leary, D. E., & McGuire, G. R. (1991). The history of psychology: A survey and critical assessment. *Annual Review of Psychology, 42,* 79–107.

Himmelfarb, G. (2004). *The new history and the old: Critical essays and reappraisals* (Rev. ed.). Cambridge, MA: Belknap.

Kuhn, T. S. (1970). *The structure of scientific revolutions* (2nd ed.). Chicago: University of Chicago Press.

Laudan, L. (1977). *Progress and its problems: Toward a theory of scientific growth.* Berkeley: University of California Press.

Laudan, L. (1996). *Beyond positivism and relativism: Theory, method, and evidence.* Boulder, CO: Westview.

Leahey, T. H. (2004). *A history of psychology: Main currents in psychological thought* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Perlman, B., & McCann, L. I. (1999). The most frequently listed courses in the undergraduate psychology curriculum. *Teaching of Psychology, 26,* 177–182.

Raphelson, A. C. (1982). The history course as the capstone of the psychology curriculum. *Journal of the History of the Behavioral Sciences, 18,* 279–285.

Robinson, D. N. (1979). The history of psychology and the ends of instruction. *Teaching of Psychology, 6,* 4–6.

Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57,* 1–22.

Tabachnick, B. G., Keith-Spiegel, P., & Pope, K. S. (1991). Ethics of teaching: Beliefs and behaviors of psychologists as educators. *American Psychologist, 46*, 506–515.

Terry, W. S. (1980). Tracing psychologists' "roots": A project for the history and systems course. *Teaching of Psychology, 7*, 176–177.

Webb, W. B. (1991). History from our textbooks: Boring, Langfeld, and Weld's introductory texts (1935–1948+). *Teaching of Psychology, 18*, 33–35.

Weigel, R. G., & Gottfurcht, J. W. (1972). Faculty genealogies: A stimulus for student involvement in history and systems. *American Psychologist, 27*, 981–983.

Zehr, D. (2000). Teaching psychology's history through a comparative analysis of introductory texts. *Teaching of Psychology, 27*, 31–33.

Appendix

Sample Readings for Nonspecialists

I have found many readings in history, historiography, and philosophy useful in teaching the history of modern psychology. The examples provided reflect the bias in my course toward American psychology in the 19th and 20th century.

Tompkins, V., Baughman, J., Bondi, V., Layman, R., & McConnell, T. (Eds.). (1994–2001). American Decades (Vols. 1–10). Detroit, MI: Gale Research. Professional historians probably consider this series to be unsophisticated history, but it is useful for establishing context for the events of the 20th century in popular culture, science and technology, economics, politics, and world events.

Cayton, M. K., & Williams, P. W. (Eds.). (2001). Encyclopedia of American Cultural & Intellectual History (Vols. 1–3). New York: Scribner's. The 221 essays in these volumes are designed to provide more information than is provided in the typical encyclopedia article but to be more accessible than the typical journal article. There is chronological coverage of public figures and institutions and thematic essays on concepts, ideas, and social movements that have had a lasting impact (e.g., the arts, religion, philosophy, nationalism, populism). There also is a section on historical methods and concepts (e.g., hermeneutics, biography, cultural and gender studies, social constructivism).

Flanagan, O. J. (1991). The Science of the Mind (2nd ed.). Cambridge, MA: MIT. Flanagan provides careful analyses of psychological traditions from Descartes to sociobiology as they address fundamental philosophical questions such as mind and body and nature and nurture. His approach is critical yet balanced.

Foner, E. (1998). The Story of American Freedom. New York: Norton. There are scores of good American histories. This one from a prolific American historian is brief and clear. It addresses many of the major social issues important to the context of the early development of American psychology.

Himmelfarb, G. (2004). The New History and the Old: Critical Essays and Reappraisals (Rev. ed.). Cambridge, MA: Belknap. Himmelfarb, a distinguished intellectual historian, defends old history, the study of the great people and institutions, against the new history of everyday events, postmodernism, and cultural and gender studies. Whether you agree with her stance (many historians do not), she neatly portrays the issues. It is especially instructive to compare this edition of the book to the first edition to see just how heated the debates have become.

Laudan, L. (1977). Progress and Its Problems: Toward a Theory of Scientific Growth. Berkeley: University of California Press. Laudan's version of philosophy of science fits the history of psychology well. In his system, progress in science can be assessed rationally. Multiple research traditions coexist and compete based on their ability to solve problems.

Leahey, T. H. (2004). A History of Psychology: Main Currents in Psychological Thought (6th ed.). Upper Saddle River, NJ: Prentice Hall. The introductory chapter of this textbook is an especially succinct overview of the issues at the intersection of history, philosophy of science, and psychology. Leahey also provides a good annotated bibliography.

Novick, P. (1988). That Noble Dream: The "Objectivity Question" and the American Historical Profession. New York: Cambridge University Press. Novick provides a broad perspective on the problems in trying to do the history of any topic. He traces the history of modern historical methods by focusing on debates about the possibility of doing unbiased history. History, historiography, and philosophy of science are involved in his discussion.

Notes

1. I thank Jim Goodwin and the reviewers for helpful suggestions on earlier drafts.
2. Send correspondence to Bruce B. Henderson, Department of Psychology, Western Carolina University, Cullowhee, NC 28723; e-mail: henderson@wcu.edu.

An Innovative International Experience Teaching Community Psychology

Lori Barker-Hackett
Department of Psychology and Sociology
California State Polytechnic University, Pomona

I describe an undergraduate community psychology course taught as part of an international education program. Students participated in international field experiences and applied community psychology values, concepts, and methods to these experiences in class discussions, written assignments, and class presentations.

In the fall of 1999 I had the unique opportunity to teach community psychology within Semester at Sea (SAS). SAS is essentially a "floating university" where students, faculty, and staff live, work, and study aboard a ship. Countries visited that fall included Japan, Hong Kong (People's Republic of China), Vietnam, Malaysia, India, Egypt, Turkey, Croatia, It-

aly, and Morocco. The program focuses on integrating traditional classroom instruction with international fieldwork.

SAS field experiences provided numerous examples of core community psychology values and concepts (see Duffy & Wong, 2000; Rappaport, 1977). Throughout the course I emphasized a number of key concepts and evaluated students on their ability to apply these concepts to real-world phenomena. One key concept is person–environment fit, which stresses the impact of environmental forces on behavior. Both the biopsychosocial model (Engel, 1977; Myers, 1986; Schwartz, 1982) and Kelly's four ecological principles (Kelly, 1966, 1979; Trickett, Kelly, & Vincent, 1985) help to explain the idea of person–environment fit. The biopsychosocial model examines factors contributing to illness on five levels, starting from the most basic and moving outward to the most complex. These levels are biological, cognitive-affective, social-interpersonal, social-institutional, and cultural. Person–environment fit also is illustrated in Kelly's adaptation of four ecological principles: interdependence, cycling of resources, adaptation/niche, and succession. Interdependence states that communities function as systems where each component influences the other. Cycling of resources examines the distribution and use of resources in the community, adaptation/niche examines how community members adjust to changing resources in their environment, and succession examines how communities change over time. Community psychology also believes in prevention, rather than treatment, and in focusing on strengths and competencies, rather than deficits.

Description of the Course

Twenty students (15 women and 5 men) from 18 different colleges and universities enrolled in the class. They represented a variety of majors, the majority ($n = 12$) from psychology. All SAS courses require students to participate in field practica. SAS faculty supervise these practica and either organize their own field experiences or lead trips organized by SAS. Each professor decides how to integrate field practica into his or her course. In the community psychology course, students participated in three field practica, including visits to orphanages, drug rehabilitation centers, and housing projects; they also attended lectures at universities and community and political organizations. Students applied course concepts in journals, a problem analysis paper, a group project, and an oral presentation. For the problem analysis paper, students analyzed a social problem or critiqued an intervention observed in one of the countries. On final group projects, students worked in pairs and proposed an intervention to address a problem observed in one of the host countries. Each pair shared its proposal in a 10-min oral presentation.

Students' field experiences provided rich material for class discussions and for the illustration of community psychology values and concepts. The following section describes some of the topics discussed and how they illustrated the concepts.

Discussion of Field Experiences

Cow Bank in Vietnam

The Cow Bank program in Vietnam is sponsored by the Adventist Development and Relief Agency (ADRA), an international nongovernmental agency. The program loans a poor family a cow that it uses to generate income. ADRA purchases the cows at approximately $150 each and trains program staff, administrators, and participating families. The family repays the loan with the first calf, which the program loans to another family.

This program illustrates Kelly's four ecological principles (Kelly, 1966, 1979; Trickett et al., 1985). (a) *Cycling of resources*: The cow eats the vegetation on the family's land and plows the land to grow crops. Cow dung fertilizes the crops. The families keep what they need for food and take extra produce and milk to market on a cart pulled by the cow to sell for profit. After they return the first calf to the program as payment for the loan, they may sell additional calves for profit or use them as collateral for additional loans. (b) *Interdependence*: Collaboration between local government agencies and ADRA is critical to the program's administration and maintenance. In addition, future families' ability to benefit from the program depends on the skill of the staff and the original family participants. (c) *Adaptation*: Participant families' lives change with increased income generated by the cow (e.g., improved housing, greater access to education, increased entrepreneurial skills). (d) *Succession*: The class speculated about how the entire community might change as more families benefit from the program.

Poverty in India

Students attempted to understand the causes of the widespread poverty observed in India using the biopsychosocial model. The influence of Hinduism, the major religion in India, arises on the cultural level. Hindus believe in reincarnation and that deeds in one lifetime determine one's caste in the next (Cross, 1994). Class members surmised that people in India may not be motivated to relieve poverty because of the underlying belief that those born into poverty deserve it (i.e., a "just world" hypothesis). Thus, a system of beliefs at the cultural level also may affect behavior at the cognitive-affective level (e.g., apathy, lack of motivation) and at the social-interpersonal level (e.g., failure of individuals to help those less fortunate).

Antidrug Policies in Malaysia

Malaysia has very stiff penalties for drug trafficking and drug use. Those convicted of selling drugs receive the death penalty, and drug addicts are sentenced to 2 years of mandatory residential drug treatment (Malaysian Ministry of Health, 2005). Despite such penalties, students found drugs readily available in Malaysia, and rehabilitation programs reported high recidivism rates. Malaysian antidrug policies represent interventions at the social-institutional level of the biopsychosocial model. However, forces at other levels may lessen the deterrent effect of such seemingly harsh penalties. At the cognitive-affective level, the personal economic benefits of selling drugs may outweigh the fear of death. At the social-interpersonal level, an individual might risk death to provide for his or her family. The class also related the apparent ineffectiveness of Malaysian drug rehabilitation programs to person–environment fit.

Taking drug users out of their regular environment for 2 years may not teach them the coping skills necessary to resist subsequent drug use once they return to that environment and encounter the forces that probably led to their drug use in the first place.

Problem-Analysis Papers

The students' problem-analysis papers assessed their ability to apply course concepts to problems observed in the different countries. One student applied the biopsychosocial model to children selling souvenirs on the streets of Vietnam. At the social-interpersonal level she noted the pressure put on children by parents to provide income for the family. On the social-institutional level she related the children's labor to the lack of funding for mandatory education. On the cultural level she explored the seemingly low value placed on education for children in Vietnam.

Another student wrote about her experiences at the Missionaries of Charity Orphanage in India. She discussed the impact of such institutionalization using Kelly's ecological principles (Trickett et al., 1985). She noticed the adaptation of the children in the orphanage to a setting where, due to lack of attention and stimulation, the children became detached, unemotional, and unresponsive. The student predicted the impact such treatment might have on the community when these children become adults (i.e., succession).

Group Projects

Students also applied community psychology concepts in their proposed interventions. One group proposed a program to decrease child labor in Vietnam and increase school attendance. It suggested that a large corporation build a manufacturing plant in Vietnam to provide jobs for the parents and fund a school for the employees' children. The company would guarantee children who graduate from the school a position with the company. They would hire teachers from outside the company at first, but eventually graduates of the program could return to teach. This project incorporated cycling of resources, building skills and competencies, and prevention.

Another group focused on children with disabilities in Vietnamese orphanages. It proposed educating women about prenatal care and mother–infant health, hypothesizing that helping women better care for themselves and their unborn children would reduce the number of children born with disabilities housed in orphanages. In their plan, young women from local villages initially would learn the information. These women then would "adopt" pregnant women and share the knowledge with them. Once these women gave birth, they would adopt other pregnant women, and so on. This proposal also illustrated prevention, cycling of resources, and building skills and competencies.

Conclusions

In many respects, teaching community psychology on SAS was not much different from teaching it on my home campus. I used the same textbook, same assignments, and so forth. Most of the challenges related more to life on SAS (e.g., seasickness, lack of privacy, hectic schedule, adjusting to a different country every few days) than the course itself. I integrate cultural issues throughout my course at home. However, on SAS, students had the opportunity to experience other cultures firsthand and more class discussions focused on field experiences.

Students successfully applied core community psychology values and concepts to their field experiences on SAS; this article includes just a few of the many issues discussed in the class. Although this was a unique teaching experience, I have been able to integrate the examples described here into my community psychology course at home, resulting in interesting and provocative class discussions. The SAS experiences serve as powerful illustrations of many core community psychology values and concepts that others can use in their classes.

References

Cross, S. (1994). *The elements of Hinduism.* Rockport, MA: Element.

Duffy, K. G., & Wong, F. Y. (2000). *Community psychology* (2nd ed.). Boston: Allyn & Bacon.

Engel, G. L. (1977). The need for a new medical model: A challenge for biomedicine. *Science, 196,* 129–136.

Kelly, J. G. (1966). Ecological constraints on mental health services. *American Psychologist, 21,* 535–539.

Kelly, J. G. (Ed.). (1979). *Adolescent boys in high school: A psychological study of coping and adaptation.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Malaysian Ministry of Health, Pharmaceutical Services Division. (2005). Index of /html/Legislations. Retrieved June 22, 2005, from http://www.pharmacy.gov.my/html/legislations

Myers, H. F. (1986, October). [Class lecture]. Lecture presented in the first-year clinical psychology core course, Department of Psychology, University of California Los Angeles, Los Angeles.

Rappaport, J. (1977). *Community psychology: Values, research, and action.* New York: Holt, Rinehart, & Winston.

Schwartz, G. E. (1982). Testing the biopsychosocial model: The ultimate challenge facing behavioral medicine. *Journal of Consulting and Clinical Psychology, 50,* 1040–1053.

Trickett, E. J., Kelly, J. G., & Vincent, T. (1985). The spirit of ecological inquiry in community research. In E. C. Susskind & D. C. Klein (Eds.), *Community research: Methods, paradigms, and applications* (pp. 283–333). New York: Praeger.

Notes

1. For more information on Semester at Sea, visit their Web site at http://www.semesteratsea.com or contact the Institute for Shipboard Education, 811 William Pitt Union, University of Pittsburgh, Pittsburgh, PA 15260.

2. Send correspondence to Lori Barker-Hackett, Department of Psychology and Sociology, California State Polytechnic University, Pomona, 3801 West Temple Avenue, Pomona, CA 91768; e-mail: labarker@csupomona.edu.