FACULTY FORUM

Rapport: Its Relation to Student Attitudes and Behaviors Toward Teachers and Classes

Trisha A. Benson, Andrew L. Cohen, and William Buskist
Auburn University

We surveyed undergraduates to learn the extent to which they experience rapport with their instructors, assess the relation between rapport and student attitudes and behaviors, and determine instructor qualities and behaviors that appear related to establishing rapport. Most students reported experiencing rapport with at least 1 teacher and indicated that rapport is positively related to their enjoyment of the teacher and subject matter. They also reported that rapport motivates them to engage in several proacademic behaviors.

Lowman (1994, 1995) partitioned the qualities of effective teachers into two dimensions: intellectual excitement and interpersonal rapport. He described intellectual excitement as the degree to which students find their instructors' teaching interesting and clear. He noted that students perceive teachers who establish high levels of interpersonal rapport ("extremely warm and open, highly student-centered, and predictable") as knowing "who they [students] are, caring about them and their learning a great deal" (Lowman, 1995, p. 29).

Rapport appears to be linked closely to *immediacy*, the extent to which teachers establish a supportive and caring learning environment through their verbal and nonverbal behavior, which includes establishing eye contact, smiling, and calling students by name (Mehrabian, 1966; Wilson & Taylor, 2001). Engaging in such immediacy behaviors may be an effective step toward establishing rapport with students. For example, Wilson and Taylor found a positive relation between immediacy and positive student attitudes toward teachers. They suggested that immediacy is one means of building a "caring atmosphere in which students can excel" (p. 138).

Instructors tend to favor teaching strategies that focus on the clarity of information transmission within their classroom, but may ignore the importance of establishing rapport with their students (Buskist, Sikorski, Buckley, & Saville, 2002; Schaeffer, Epting, Zinn, & Buskist, 2003). For example, in a survey asking undergraduates to identify characteristics and behaviors inherent in master teaching, Buskist et al. found that 42% of the students ranked rapport within the top 10 qualities of master teaching, but only 7% of faculty members listed this same quality in their top 10 list. Both Buskist et al. and Schaeffer et al. found that, in contrast to students, teachers tended to underscore instructional technique as key to effective teaching (i.e., techniques aimed at increasing the quality of information dissemination and engendering critical thinking).

Initial investigations in rapport building suggested that students who experienced rapport reported increased course enjoyment, class attendance, and attentiveness (Buskist & Saville, 2004). However, Buskist and Saville did not report actual data; they based their beliefs on qualitative student reports. We attempted to replicate their study in an effort to obtain numeric data on the relation between rapport and student attitudes and behaviors. Specifically, we assessed (a) the frequency with which students experienced rapport with their instructors; (b) the relation between rapport and student attitudes and proacademic behaviors such as attending class, paying attention during class, and studying; and (c) instructor behaviors that students identified as contributing to the establishment of rapport.

Method

Participants, Procedure, and Survey Instrument

Our initial sample included 202 Auburn University undergraduate students enrolled in an introductory psychology course. We excluded 36 students (18%) who indicated that they had never experienced rapport with a college instructor, resulting in a sample of 166 students (97 women, 68 men, and 1 nonreport; 7 freshmen, 107 sophomores, 23 juniors, 27 seniors, and 2 nonreports; *M* age of 20.05 years, $SD = 2.51$).

The survey included two dictionary definitions of *rapport* to provide a common framework from which students could conceptualize rapport: "a relationship; especially one of mutual trust or emotional affinity" (American Heritage Dictionary, 1991, p. 1026) and "a relation; connection; an especially harmonious or sympathetic relation" (Random House Dictionary, 1987, p. 1601). Students reported whether they had been in a college class in which the instructor established rapport with them. If students responded yes to this item, they estimated the percentage of instructors with whom they had experienced rapport. Students also reported on the relation between rapport and their behavior by responding to item pairs (e.g., "In a class where rapport has been established, how likely are you to attend class?" and "In a class where rapport has NOT been established, how likely are you to attend class?") Students responded to each item on a 5-point scale ranging from 1 (*very unlikely*) to 5 (*very likely*). Finally, we asked students to think of the instructor (or instructors) with whom they had experienced the strongest rapport and then list up to 10 specific behaviors the instructor exhibited that contributed to the establishment of rapport.

Results and Discussion

Thirty-three percent of the students reported experiencing rapport with 1% to 5% of their instructors, 30% with 6% to 25% of their instructors, and 22% with 26% to 50% of their

Table 1. Means, Discrepancy Scores, and *t* Scores for Self-Reported Likelihood of Engaging in Academic Behaviors and Experiencing Positive Affect When Rapport Is or Is Not Established

Category	With Rapport		Without Rapport		Discrepancy Score	*t*
	M	*SD*	*M*	*SD*		
Attend class	4.80	.47	3.46	1.05	1.34	17.17
Pay attention	4.71	.53	3.23	1.02	1.48	18.55
Study for class	4.54	.69	3.86	0.96	0.68	11.16
Enjoy the subject	4.60	.60	2.75	0.99	1.85	19.55
Enjoy the professor	4.75	.54	2.44	0.96	2.31	27.05
Attend office hours	4.78	.96	2.07	1.10	1.61	17.57
E-mail professor	3.97	.88	2.66	1.16	1.31	13.89
Take another class from professor	4.78	.48	2.07	1.00	2.71	32.37
Take another class in subject	3.97	.76	2.66	0.96	1.31	14.66
Drop the class	3.31	.74	4.57	1.15	1.26	12.74

Note. All *p*s < .001.

instructors. Only 15% of the students reported experiencing rapport with more than half of their instructors.

We conducted paired-sample *t* tests to determine any difference in students' reported likelihood to engage in particular behaviors or their tendency to experience positive affect toward the teacher or the subject matter, depending on whether the instructor established rapport with the class. We found significant differences for both: Students reported positive feelings toward both the teacher and subject matter and an increased tendency to engage in an academic behavior when rapport was established relative to when it was not, $t(164) = 11.16$ to 32.37, all *p*s < .001 (see Table 1). We found these results to hold true regardless of students' year in school.

We also calculated discrepancy scores for affect and for each behavior (the difference in scores between with- and without-rapport conditions) and conducted a repeated measures ANOVA to determine any differences among them. We found a significant difference, $F(9, 1476) = 58.3$, $p < .001$, between several discrepancy scores—again, students' year in school did not influence these findings. Pairwise comparisons revealed the following significant differences (all *p*s < .001): In order, the presence of rapport increased the likelihood of students reporting that they would take another class from the teacher, enjoy that teacher, enjoy the subject material, and attend office hours more so than all other assessed behaviors.

We also gathered students' reports of specific teacher behaviors that they believed contributed to the establishment of rapport. The first two authors independently analyzed these reports based on the classifications used by Buskist et al. (2002) and Schaeffer et al. (2003) to conceptualize teacher behaviors within an established framework for categorizing teacher behaviors. The first two authors agreed on 95% of the categorizations and categorized the remaining reports following mutual discussion. If the behavior description from our list matched a behavioral description of a master teaching quality reported in these studies, we mapped the specific instructor behavior onto this category. If the behavior descriptions did not match, we devised our own quality–behavioral description.

We organized the students' reports into 22 different categories. The 10 most frequently reported rapport-inducing teacher qualities, in order, involved the following: encouraging, open-mindedness, creative, interesting, accessible, happy, having a "good" personality, promoting class discussion, approachability, concern for students, and fairness (for a complete description of the behaviors that comprise these categories, see Buskist et al., 2002). All of these qualities, with the exception of 2—having a good personality and concern for students—overlapped considerably with the qualities described in Buskist et al. and Schaeffer et al. (2003).

Our data suggest that teachers who establish rapport with their classes are likely to have students who attend class, pay attention during class, and enjoy the subject matter. Just as important, rapport may also lay the groundwork for interactions with teachers during office hours or through e-mail. Teachers might develop rapport with their students by engaging in behaviors reflective of the top 10 qualities listed previously. Similarly, the immediacy literature outlines useful behaviors (e.g., Christophel, 1990; Gorham, 1988; Teven & Hanson, 2004; Titsworth, 2001; Wilson & Taylor, 2001). Contrary to immediacy research, our study allowed students to generate freely the perceived professor behaviors that contribute to rapport building. In addition to adding a wider range of useful instructor behaviors than is found in immediacy research, student-generated items might offer a more valid measure of behaviors relevant to students. This approach may be particularly important given prior research (Buskist et al., 2002) that student and instructor perceptions of some important behaviors often differ substantially.

Lowman (1994, 1995) argued that rapport is a key to teaching effectiveness. Our study underscores the likely link between rapport and effective teaching by providing strong evidence for the positive relation between rapport-inducing teacher behaviors and students engaging in proacademic student behaviors as well as positive affect toward teachers and course content. Nonetheless, important questions regarding the relation of rapport to student learning await further research. For example, to what extent is rapport related to student learning in terms of course grades? Empirical answers to such questions would provide more direct evidence for the role of rapport as a causal factor influencing student learning.

References

American heritage dictionary (2nd ed.). (1991). Boston: Houghton Mifflin.

Buskist, W., & Saville, B. K. (2004). Rapport building: Creating positive emotional contexts for enhancing teaching and learn-

ing. In B. Perlman, L. I. McCann, & S. H. McFadden (Eds.), *Lessons learned: Practical advice for the teaching of psychology* (Vol. 2, pp. 149–155). Washington, DC: American Psychological Society.
Buskist, W., Sikorski, J., Buckley, T., & Saville, B. K. (2002). Elements of master teaching. In S. F. Davis & W. Buskist (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 27–39). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Christophel, D. M. (1990). The relationships among teacher immediacy behaviors, student motivation, and learning. *Communication Education, 39,* 323–340.
Gorham, J. (1988). The relationship between verbal teacher immediacy behaviors and student learning. *Communication Education, 37,* 40–53.
Lowman, J. (1994). Professors as performers and motivators. *College Teaching, 42,* 137–141.
Lowman, J. (1995). *Mastering the techniques of teaching.* San Francisco: Jossey-Bass.
Mehrabian, A. (1966). Immediacy: An indicator of attitudes in linguistic communication. *Journal of Personality, 34,* 26–34.
Random House dictionary of the English language (2nd ed.). (1987). New York: Random House.
Schaeffer, G., Epting, K., Zinn, T., & Buskist, W. (2003). Student and faculty perceptions of effective teaching: A successful replication. *Teaching of Psychology, 30,* 133–136.
Teven, J. J., & Hanson, T. L. (2004). The impact of teacher immediacy and perceived caring on teacher competence and trustworthiness. *Communication Quarterly, 52,* 39–54.
Titsworth, B. S. (2001). The effects of teacher immediacy, use of organizational lecture cues, and students' notetaking on cognitive learning. *Communication Education, 50,* 283–297.
Wilson, J. H., & Taylor, K. W. (2001). Professor immediacy as behaviors associated with liking students. *Teaching of Psychology, 28,* 136–138.

Notes

1. We thank Alejandro Lazarte and Jared Keeley for their assistance with statistical analysis.
2. Send correspondence to William Buskist, Psychology Department, Auburn University, Auburn, AL 36849–5214; e-mail: buskiwf@auburn.edu.

How Do Students Really Study (and Does It Matter)?

Regan A. R. Gurung
Human Development and Psychology
University of Wisconsin, Green Bay

Are specific study techniques better than others? I provide a method to answer this question that is easy to adapt for any course. I assessed 229 introductory psychology students' use of 11 different study techniques and correlated their responses with their exam scores. Many, but not all, techniques related to better exam scores. Hours studied were positively related to exam scores but starting studying early and reading material prior to and after class were not. I also found detriments to studying (e.g., listening to music). Results provide a detailed picture of what students do when they study.

There are many different ways to study but not all methods may enhance learning. Although there is a sizeable literature on how students *should* study (Al-Hilawani & Sartawi, 1997; Fleming, 2002; Hattie, Biggs, & Purdie, 1996), not as much is known regarding how students actually *do* study. I assessed how students actually study and tested whether certain study habits were more conducive to learning than others.

Study skills can be divided into four main categories: repetition-based (e.g., flashcards and mnemonic devices such as "CANOE" for the Big 5 personality traits), cognitive-based (e.g., studying with a friend, group work), procedural (e.g., time management, organization, scheduling study routines), and metacognitive (e.g., taking quizzes to test self-knowledge; for more details, see Gettinger & Seibert, 2002). Empirical tests comparing these different methods are equivocal.

Some research suggests that the types of study techniques that a student uses affect exam performance (Bol, Warkentin, Nunnery, & O'Connell, 1999). Other research suggests that there is no one style that is useful for everyone and that a repertoire of techniques is best (Hadwin & Winne, 1996; Nist, Simpson, Olejnik, & Mealey, 1991). For example, repetition and rehearsal, which require minimal amount of processing, may be useful only in remembering small amounts of information (Gettinger & Seibert, 2002). Memorizing facts and definitions does not correlate with students' exam scores, but procedural and organizational-based skills, metacognitive-based skills, and skills that increase elaboration show positive correlations with test scores (Carney & Levin, 1998; Chen & Daelhler, 2000; Elliot, McGregor, & Gable, 1999; Motes & Wiegmann, 1999). Dickinson and O'Connell (1990) also showed that time spent organizing course material (e.g., taking notes on the textbook) related to test scores, whereas actual hours spent studying did not.

The existing literature does not include a comprehensive assessment of a wide variety of studying techniques, and it does not provide studies that both assess techniques and measure learning outcomes. Furthermore, students are often unaware that some of their habits, such as having music on while studying, may hurt their learning. This study provides a rich view of what students do by collectively assessing different behaviors. Consistent with the disparate literatures, I hypothesized that techniques aiding elaboration (e.g., using examples, mnemonics) and metacognition (e.g., self-testing) would predict higher exam scores, whereas those behaviors reducing elaboration (e.g., listening to music) would predict lower exam scores.

Method

Participants

Two hundred and twenty-nine students (169 women and 60 men) from a midsized midwestern university in two sections of my introductory psychology class participated in this study (participation was voluntary). The mean age was 19.26 (SD = 3.91). The majority of the students were freshmen (82%); the remainder were sophomores (7%), juniors (4%), and seniors (7%). The mean ACT score was 22 (range 10 to 31). I combined the data from both sections as exam grades were similar.

Materials

A questionnaire assessed study methods, distractions, and confidence with the material. I based items on previous research (Wade, Trathen, & Schraw, 1990; Winne & Jamieson-Noel, 2002) and feedback from small student focus groups (questionnaire available on request). I asked students which of 11 study methods they used (i.e., memorizing definitions, reading the text, reviewing figures, reviewing highlighted material in the text, testing self-knowledge, rewriting notes, taking notes on the text, mnemonics, studying with friends, reading the notes, rewriting notes) and the extent to which they used them on a 5-point scale ranging from 1 (*never*) to 5 (*all the time*). I also measured distractions ("Do you have music or the television on when studying? Do you have roommates, family, or friends around when studying?" and "Do you respond to instant messaging or e-mail while studying?"); the total hours students studied for the test; the number of days in advance that students started studying; how often they reviewed material before and after a class; and how well they believed they knew the material, understood the material, and how confident they were of their understanding of the material.

Procedure

I added the survey to the end of the last of four exams. After answering 65 multiple-choice questions, participants read instructions stating that the remaining questions on the exam sheet would assess their study habits. I told students that participation was voluntary and that the answers to the questions would not affect their class grades or exam scores.

Results

The majority of students reported studying between 4 to 6 hr for the final (45%). The rest studied between 1 to 3 hr (31%) and 7 to 9 hr (19%). A small number of students reported studying over 10 hr (5%). The frequency and duration for use of the 11 study techniques assessed in this study appear in Table 1.

The frequency of technique use and the duration of technique use were correlated with scores on students' final exam. Partial correlations controlled for student ability (using ACT scores; zero-order correlations available on request). The more students memorized notes, $r(227) = .28$, $p < .001$; made up examples $r(227) = .20$, $p < .001$; read the book $r(227) = .21$, $p < .01$; read their notes, $r(227) = .18$, $p < .05$; used mnemonics, $r(227) = .15$, $p < .05$; and tested their knowledge, $r(227) = .28$, $p < .001$; the higher were their exam scores. No other techniques (i.e., frequency of use) significantly correlated with exam score.

In contrast to the significant correlations with frequency of use described previously, only the amount of time spent memorizing was significantly related to exam scores, $r(227) = .15$, $p < .05$. The global number of hours studied did relate to exam scores, $r(228) = .16$, $p < .05$.

All the distracters and not attending class negatively correlated with exam grades. Students who had music on, $r(226) = -.18$, $p < .01$; the television on, $r(226) = -.21$, $p < .01$; responded to e-mail, $r(226) = -.16$, $p < .05$; or who had friends around, $r(226) = -.13$, $p < .05$; when studying performed worse on the exam. Students who missed class also scored lower on the exam, $r(226) = -.27$, $p < .001$.

Table 1. Frequency and Duration of Use of the Main Study Techniques

		Hours Spent (%)					
	M	0	1	1 to 2	2 to 3	> 3	How Often
Study technique[a]							
Read your notes		1	18	32	26	16	4.01
Read the text		4	23	34	20	11	3.37
Think of mnemonic devices (e.g., "CANOE" for personality traits)		13	41	23	10	3	3.33
Rewrite notes and/or skim notes		10	28	32	14	8	3.25
Review highlighted information from text		8	34	35	11	2	3.15
Memorize definitions through repetition (e.g., flashcards)		9	36	31	13	4	3.11
Review figures and tables in text		8	51	24	7	1	2.96
Make up examples to understand material/incorporate into everyday life		16	43	23	9	2	2.89
Use concept checks, chapter-end questions to test knowledge		23	42	19	6	2	2.62
Take notes from the book		39	38	14	8	1	2.18
Study with a friend		43	29	18	8	2	2.07
Distracters							
Have the television on							4.00
Have music on							3.78
Have roommates/friends/family around							3.02
Respond to instant messaging/e-mail on the Internet							3.75
Self-reports level							
Knowledge	3.43						
Understanding	3.66						
Confidence	3.44						

Note. Hours spent represents the percentage of the class who reported using each option. How often each technique was used/or each distraction was present was measured on a 5-point scale ranging from 1 (*never*) to 5 (*all the time*). Knowledge, understanding, and confidence was measured on a 5-point scale ranging from 1 (*not at all*) to 5 (*extremely*).
[a]Listed in order of use.

Discussion

The results of this study provide a detailed picture of what students spent time on and how effective the different methods were. Not all techniques were effective—the most effective techniques were often not the ones used the most. For example, although the three most frequently used techniques (reading notes and the text, using mnemonics) correlated significantly with exam scores, one of the strongest predictors of exam scores, testing knowledge, was one of the least used techniques. Other techniques commonly used by students (rewriting notes, reviewing highlighted material and figures and tables in the text) did not relate to exam scores. Perhaps most important, the number of hours studied was only weakly associated with exam score. This finding suggests that how students study may be even more important than how long they study and provides a strong rationale for the use of this measure.

The effectiveness of many common study suggestions did not receive empirical support. Results such as these compel a closer look at the recommendations instructors make to their students. Instructors often provide study tips and urge students to use specific techniques, but the correlation between student grades and technique usage is not always significant and is often low. For example, Balch (2001) provided students with six tips (e.g., specific ways to take lecture notes and self-help quizzes) but, except for elaborative encoding, reported use and course grades were not significantly correlated.

The results provide strong empirical evidence of what students should not do. Students who skip class, listen to music, watch television, or use the Internet while studying performed worse on the exam. Although the data do not test the causality of the association between distractions and exam scores, making such data available to students on the first day of class may help them better design their study habits.

This study provides an easy method for individual instructors to assess how their students are preparing for exams. These findings may vary for instructors who use essay exams or different textbooks, but an instructor can modify this method to assess different levels of classes and for different types of exams. The results of this study and self-collected data with this tool will prepare instructors to advise students on how best to study to do well.

Limitations

Finals are a particularly stressful time of the semester, and studying during the last week of class may not be representative of how students study in general. The fact that the exam was not cumulative (similar to midterm exams) somewhat lessens the problem with this limitation. Having the students complete the assessment as part of the exam (and hence be identifiable) raises the potential for impression management and could contaminate responding. Finally, I made the assumption that exam scores equate to learning. It is possible that even the study techniques that did not significantly relate to exam scores did enhance learning, but this learning was not captured by my exam.

Although students use a variety of study techniques, they are not all effective. Furthermore, students are not using some useful techniques enough. How students prepare for tests can be a crucial element in their achievement. Because certain study techniques are more beneficial than others, instructors should help students more effectively prepare for exams by informing students about the techniques and modifying ways to best help students use the techniques. How students study does actually seem to matter.

References

Al-Hilawani, Y. A., & Sartawi, A. A. (1997). Study skills and habits of female university students. *College Student Journal, 31,* 537–544.

Balch, W. R. (2001). Study tips: How helpful do introductory psychology students find them? *Teaching of Psychology, 28,* 272–274.

Bol, L., Warkentin, R. W., Nunnery, J. A., & O'Connell, A. A. (1999). College students' study activities and their relationship to study context, reference course, and achievement. *College Student Journal, 33,* 608–622.

Carney, R. N., & Levin, J. R. (1998). Coming to terms with the keyword method in introductory psychology: A "neuromnemonic" example. *Teaching of Psychology, 25,* 132–134.

Chen, Z., & Daehler, M. W. (2000). External and internal instantiation of abstract information facilitates transfer in insight problem solving. *Contemporary Educational Psychology, 25,* 423–449.

Dickinson, D. J., & O'Connell, D. Q. (1990). Effect of quality and quantity of study on student grades. *Journal of Educational Research, 83,* 227–231.

Elliot, A. J., McGregor, H. A., & Gable, S. (1999). Achievement goals, study strategies, and exam performance: A mediational analysis. *Journal of Educational Psychology, 91,* 549–563.

Fleming, V. M. (2002). Improving students' exam performance by introducing study strategies and goal setting. *Teaching of Psychology, 29,* 115–119.

Gettinger, M., & Seibert, J. K. (2002). Contributions of study skills to academic competence. *School Psychology Review, 31,* 350–365.

Hadwin, A. F., & Winne, P. H. (1996). Study strategies have meager support: A review with recommendations for implementation. *The Journal of Higher Education, 67,* 692–715.

Hattie, J., Biggs, J., & Purdie, N. (1996). Effect of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66,* 99–136.

Motes, M. A., & Wiegmann, D. A. (1999). Computerized cognition laboratory. *Teaching of Psychology, 26,* 62–65.

Nist, S. L., Simpson, M. L., Olejnik, S., & Mealey, D. L. (1991). The relation between self-selected study processes and test performance. *American Educational Research Journal, 28,* 849–874.

Wade, S. E., Trathen, W., & Schraw, G. (1990). An analysis of spontaneous study strategies. *Reading Research Quarterly, 25,* 147–166.

Winne, P. H., & Jamieson-Noel, D. L. (2002). Exploring students' calibration of self-reports about study tactics and achievement. *Contemporary Educational Psychology, 27,* 551–572.

Notes

1. The University of Wisconsin, Green Bay, Department of Human Development funded this research.
2. A portion of this study was presented at the American Psychological Society Meeting in Chicago in 2004.
3. I thank Heidi Rose and Jessica Peterson for their help collecting, analyzing, and discussing the data and Heather Bloch, Sarah Brill, and Illene Noppe for their helpful suggestions.

4. Send correspondence to Regan A. R. Gurung, Department of Human Development and Psychology, University of Wisconsin, Green Bay, 2420 Nicolet Drive, MAC C318, Green Bay, WI 54311; e-mail: gurungr@uwgb.edu.

Predicting Course Grades: Accurate for Others But Biased for Self

Arnie Cann
University of North Carolina Charlotte

People often are overly optimistic about their outcomes. To test such biases, students made judgments about their classmates' and their own likely performances in a course. Students estimated the grade distribution for the entire class by assigning a percentage to each possible grade and then estimated their own grades. I compared these predicted grade distributions with the actual grade distributions of past students. The grade distributions estimated for the class matched the actual grade distribution, but the distributions for the self were more optimistic. I have used the results to facilitate a class discussion of judgment biases and positive illusions.

For students, the beginning of a new semester means that, irrespective of their previous semester grades, they can believe they are going to be successful this semester. Several years ago, I began collecting data to see if my students were optimistic about everyone's grades or only their grades. Although Prohaska (1994) found clear evidence that students do overestimate their likely grades, he did not compare these predictions to the expectations for other students. The question here is whether students have an accurate or a biased understanding of the likely class grade distribution when they make their projections about their own grades.

In another classroom demonstration of optimistic bias (Snyder, 1997, 1999), students were optimistically biased in estimating their ages of death when given base rate information about cohorts. In this regard, research consistently demonstrates that people do not consider externally provided base rates especially useful in making judgments (Kahneman & Tversky, 1973). On the other hand, base rates generated by the students themselves may provide a more meaningful comparison for determining the degree to which self estimates are biased. This demonstration provides this comparison.

Biases in social judgments are widespread in social cognitive processes (see Gilovich, Griffin, & Kahneman, 2002). Initially, my goal was to use the grade estimation data when discussing optimistic thinking (Armor & Taylor, 2002; Prohaska, 1994; Weinstein, 1980) and self-serving biases (Dunn, 1989; Dunning, Meyerowitz, & Holzberg, 2002), but I also have found that the data are appropriate in discussions of positive illusions (Taylor, 1989), the planning fallacy (Buehler, Griffin, & Ross, 2002), actor–observer biases (Gordon & Kaplar, 2002; Jones & Nisbett, 1972), self-enhancement (Alicke, 1985; Brown, 1986), and self-presentation (Baumeister & Twenge, 2003; Schlenker, 2003). I now have a database spanning several semesters showing the reliable biases of students when making predictions about personal outcomes in my courses, despite the fact that these same students have a realistic set of expectations when predicting outcomes for other students.

Method

Data Collection

Students ($N = 744$) enrolled in sections of an undergraduate social psychology class at a state-supported university in the Southeastern United States participated in the activity. There have been 21 sections of this course (enrollments of 13 to 75 students) from the Fall 1991 semester through the Fall 2003 semester. A subsample ($n = 99$) of these students also completed an evaluation of the grade estimation activity.

On the first day of the semester, after discussing the course syllabus, the testing procedures, and the grading scale (90%+ = A, 80% to 89% = B, etc.), students responded to the following prompt: "What do you think the distribution of grades will be in this class? Indicate the percentage of students you think will get each of the following grades. Be sure that your percentages add to 100." Under the prompt, the grades A, B, C, D, and F appeared on separate lines. After completing this task, I also asked the students to indicate the grades they would receive in the course. Thus, students were not aware that they would predict their own grades when they estimated the grade distribution for the class.

Evaluation of the Activity

The students ($n = 99$) enrolled in the three most recent classes completed evaluations of the activity during the next class meeting after the discussion of the results. The sample included only those students who made predictions and were present on the discussion days. Using 9-point scales (with higher scores indicating greater enjoyment or understanding) students indicated how much they enjoyed the exercise and how much the exercise helped them to understand the relevant social judgment processes.

Results

Predictions Versus the Actual Grade Distribution

Table 1 contains the accumulated data for the entire sample. A comparison of the grade distribution predicted for the class as a whole with the actual grade distribution based on past students' performance revealed no significant difference, $\chi^2(4, N = 744) = 1.17$, $p > .05$. The distribution generated by the accumulated predictions for self, however, differed reliably from both of these distributions: predictions for others, $\chi^2(4, N = 744) = 60.13$, $p < .05$, and actual performance, $\chi^2(4, N = 744) = 66.73$, $p < .05$. Predictions for self were clearly more optimistic than those for others or for actual performance. For example, although students predicted that about 18% of other students would get Ds or Fs, no students ever have predicted Ds or Fs for themselves.

Evaluation of the Activity

On both measures used to evaluate the effectiveness of the class activity, the students rated the exercise positively (enjoyment $M = 7.10$, $SD = 1.67$; understanding $M = 7.00$, SD

Table 1. Predicted Grade Distributions and Actual Grade Distribution

	Actual Grades[a]	Prediction for Self[b]	Prediction for Others[c]
Grade			
A	14.1	45.1	17.9
B	31.7	51.8	31.4
C	32.1	3.1	32.9
D	13.2	0.0	11.9
F	8.9	0.0	5.9
GPA	2.3	3.4	2.4

Note. Grades are reported as percentages and GPA represents the summary for each set of grades. GPA = grade point average.
[a]n = 1,653. [b]n = 743. [c]n = 744.

= 1.63). In fact, for both measures the modal rating was a 9, the highest rating possible.

Discussion

Students are accurate when estimating how others will perform, but overly optimistic when predicting grades for themselves. For the classroom discussion, I present the data generated by that particular class, so that they can see their predictions, but I also provide the accumulated data for all classes to show them that their responses are quite typical. We discuss the data at the end of the course material on social cognition. I begin by posing the question, "How well can you predict your own behavior?" I remind them of the data they provided on the first day of the semester and suggest we look at their predictions as an example. I present their predictions for the class grade distribution first, followed by the distribution of actual grades from the course. They usually are quite pleased to see that they are reasonably accurate in their predictions. After discussing their success briefly, I return to the issue of predicting one's own behavior as opposed to the behavior of others.

At this point, I present the distribution that would result based on the predictions of their own grades. The dramatic differences between the "self" distribution and the first two presented usually evoke much embarrassed laughter and surprise. As we attempt to understand the differences, I am able to introduce a number of social judgment biases into the discussion. For example, we discuss optimism and consider how unrealistic optimism might sometimes be a good bias that prevents decreased motivation. We discuss self-presentation issues to acknowledge that the predictions might be intentionally biased to make a good impression or they might be the result of less conscious biased thinking. We also discuss the planning fallacy, a tendency most students acknowledge, and we discuss why, due to self-serving biases, this fallacy might tend to recur despite a general pattern of failing to meet deadlines. The belief that past failures have been due to external causes (they can usually provide many personal examples of failures that were not their fault) allows them provide an excuse for their current failure (Snyder, Higgins, & Stucky, 1983) and to believe that their good intentions will not be thwarted the next time.

In my course, the discussion takes place just before the first examination. So, as the discussion winds down, I point out to the students that they clearly appreciate what the grade distribution for their class will probably look like, because they accurately predicted how the class as a whole would perform relative to past students. I suggest to them that they can prove that their self-predictions were not simply biased judgments by outperforming past students and producing their predicted distribution. So far no class has been able to perform at the level of their self predictions, but I hope the discussion encourages them to prepare for the exam with a reasonable amount of (unrealistic) optimism, but a strong dose of realism.

Obviously, the degree to which the demonstration can provide clear evidence of bias in predicted grades for self relative to others and actual grades depends on there being a possibility for self predictions to show optimism. The grades for my course, and for similar courses in my department, average in the C range, so there is opportunity for students to show bias for themselves by predicting As or Bs. In a course where the average grades were much higher than mine, there would be less room to show bias.

References

Alicke, M. D. (1985). Global self-evaluation as determined by the desirability and controllability of trait adjectives. *Journal of Personality and Social Psychology, 49,* 1621–1630.

Armor, D. A., & Taylor, S. E. (2002). When predictions fail: The dilemma of unrealistic optimism. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 334–347). New York: Cambridge University Press.

Baumeister, R. F., & Twenge, J. M. (2003). The social self. In T. Millon & M. J. Lerner (Eds.), *Handbook of psychology: Personality and social psychology* (Vol. 5, pp. 327–352). New York: Wiley.

Brown, J. D. (1986). Evaluations of self and others: Self-enhancement biases in social judgments. *Social Cognition, 4,* 353–376.

Buehler, R., Griffin, D., & Ross, M. (2002). Inside the planning fallacy: The causes and consequences of optimistic time predictions. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 250–270). New York: Cambridge University Press.

Dunn, D. S. (1989). Demonstrating a self-serving bias. *Teaching of Psychology, 16,* 21–22.

Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (2002). Ambiguity and self-evaluation: The role of idiosyncratic trait definition in self-serving assessments of ability. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 324–333). New York: Cambridge University Press.

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment.* New York: Cambridge University Press.

Gordon, A. K., & Kaplar, M. E. (2002). A new technique for demonstrating the actor-observer bias. *Teaching of Psychology, 29,* 301–303.

Jones, E. E., & Nisbett, R. E. (1972). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelley, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). Morristown, NJ: General Learning.

Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review, 80,* 237–251.

Prohaska, V. (1994). "I know I'll get an A": Confident overestimation of final course grades. *Teaching of Psychology, 21,* 141–143.
Schlenker, B. R. (2003). Self-presentation. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 492–518). New York: Guilford.
Snyder, C. R. (1997). Unique invulnerability: A classroom demonstration in estimating personal mortality. *Teaching of Psychology, 24,* 197–199.
Snyder, C. R. (1999). Unique invulnerability as applied to personal mortality: The reports of its demise are exaggerated. *Teaching of Psychology, 26,* 217–219.
Snyder, C. R., Higgins, R. L., & Stucky, R. J. (1983). *Excuses: Masquerades in search of grace*. New York: Wiley.
Taylor, S. E. (1989). *Positive illusions: Creative self-deception and the healthy mind*. New York: Basic.
Weinstein, N. D. (1980). Unrealistic optimism about future life events. *Journal of Personality and Social Psychology, 39,* 806–820.

Note

Send correspondence to Arnie Cann, Department of Psychology, University of North Carolina Charlotte, 9201 University City Boulevard, Charlotte, NC 28223; e-mail: acann@uncc.edu.

Writing Exercises for Introductory Psychology

Pam Marek
Anderson College

Andrew N. Christopher
Albion College

Cynthia S. Koenig
St. Mary's College of Maryland

David F. Reinhart
Anderson College

To facilitate introductory psychology students' ability to solidify knowledge by applying it to themselves, 4 instructors from 3 institutions developed and implemented brief, focused, Internet-based writing assignments. Students reacted favorably to assignments on sleep deprivation and body image. Student performance on multiple-choice examination questions related to their papers was at least comparable to performance on other questions based on textbook reading or classroom coverage. From an instructor's perspective, the structured nature of the assignments facilitated grading. Thus, instructors may consider this type of brief paper as a viable tool for extending assessment beyond standard examinations and quizzes.

Educators have emphasized the importance of incorporating multiple forms of student assessment in the classroom (Sternberg, 2001) and have suggested that written assignments heighten students' involvement and motivation (McKeachie, 2002). Whereas instructors of large classes may hesitate to assign full-scale term papers because of the time required to grade them, Lowman (1995) argued that short, focused writing assignments might more effectively involve students in reading and thinking than the more traditional term paper. Given these considerations, we developed and implemented a brief paper assignment to provide introductory psychology students with an alternative to textbook reading and to allow them to solidify their knowledge by applying it to personal experience. We required students to complete a measure of sleep deprivation, use the Internet to explore topics related to sleep, and write a paper in which they integrated factual knowledge with their personal opinions and experience.

The Assignment

We selected the topic based on its presumed relevance to college students, general importance, availability of Internet information from educational sources and sites not apparently designed to sell products, and the subjective judgment of site "stability." Furthermore, in designing questions for students to answer, we selected topics that could be adapted to material on alternate sites without dramatically changing the overall assignment. In addition, to promote student involvement, we located an online scale that students could use for self-ratings (Dement, n.d.).

The sleep paper assignment required students to complete the Epworth Sleepiness Scale (Dement, n.d.) and retrieve a total of four articles from the Web sites of the National Sleep Foundation, PBS, and Scientific American. All articles were written for a nonprofessional audience. Although we subsequently modified the assignment to incorporate brain activity during sleep and dreaming (Chudler, 2004), the framework of the assignment remained intact.

To complete the assignment, students composed a brief paper using a research paper format. The paper included answers to five factual questions addressing changes in average amount of sleep over the past century, the dangers of lack of sleep, operation of the biological clock, the effect of sleep deprivation on learning, and the function of melatonin. Integrated with the factual information, students included their opinions on these same topics (e.g., societal changes contributing to shift in amount of sleep). To conclude the paper, students related their personal results from the Epworth Sleepiness Scale to what they had learned about sleep.

We designed instructions for the assignment that enabled students to complete it with minimal guidance. The structured instruction format also simplified grading, which we accomplished using a checklist for content and elements of writing style.

Assignment Implementation

Four instructors from three institutions required introductory psychology students to complete this assignment during the Fall 2002 semester. We explained the importance of differentiating between fact and opinion and of citing sources for factual information. We covered plagiarism and the challenge of integrating ideas and presenting them in the writer's own words. To fulfill content requirements, students had to visit all Web sites listed and to produce a paper no longer than five double-spaced, typewritten pages. Most students submitted three- to four-page papers. For all instruc-

tors, the number of possible points allotted to the paper comprised 4% to 5% of a student's final grade.

Instructors integrated the assignment into their courses in various ways. For example, two of the authors did not lecture on, discuss, or require students to read the textbook material on sleep, simply using the assignment as an independent extension of textbook and class coverage. In contrast, two instructors used the assignment to supplement textbook material. Students generally did well on the assignment. Their most common error was inappropriate paraphrasing.

Assessment and Student Evaluations

Overall, the percentage of students who correctly answered a multiple-choice exam question related to one aspect of the brief paper (e.g., identifying melatonin as the hormone involved in sleep) was generally comparable to the class average for the entire exam. When evaluating the sleep paper assignment (see Table 1), students ($N = 150$) generally agreed that the assignment increased their understanding of sleep and recommended it for future use. We recognize that students often react favorably to assignments other than quizzes and exams. However, the examples given in the papers, particularly those relating to personal experiences with lack of sleep, suggested that the assignment stimulated students to consider the importance of proper sleep for both learning and personal safety. In the Spring 2003 term, the first author used the same assignment and asked students ($N = 42$) a few additional questions. Responses indicated that students reacted positively to using the Internet and spent a mean of 3.11 hr ($SD = 2.15$) preparing the assignment.

Extension to Other Topics

As an example of an extension to another topic, in Spring 2003, the first author assigned a second paper that focused on body image, control of hunger, and eating disorders. The structure of this assignment paralleled that of the sleep paper. Students completed a Body Shape questionnaire (Monash University, n.d.); retrieved and read four articles from the Internet; prepared their papers to respond to a series of factual questions, citing sources as required; presented their opinions regarding the influence of genetics and environment on obesity; and commented on their experiences and concerns with body image and weight control.

Three weeks after submitting their papers, students took a multiple-choice exam that included two questions about topics included in the body image paper. The mean percentage correct for these questions (65% and 79%, respectively) was at least as high, or higher than, the overall mean percentage correct for all other multiple-choice exam questions ($M = 65\%$, $SD = 13.62\%$). Thus, students appeared to learn at least as well from preparing the paper as they did from reading material in their textbooks and discussing it in class. They also learned useful Internet skills.

Table 1. Student Evaluations of Sleep Paper Assignment

Evaluative Item	*M*	*SD*
The instructions for the paper were clear.	6.19	1.10
The paper helped me understand more about sleep.	6.07	1.09
The paper stimulated interest in the topic of sleep.	5.32	1.37
Filling out the sleep scale helped make the topic more personally relevant.	5.79	1.23
I would recommend using this paper topic in future sections of this course.	6.05	1.21
I generally reached the required Web sites without a problem.	5.87	1.42
The sleep paper was graded fairly.	6.49	1.02

Note. $N = 150$. Students rated items on a scale ranging from 1 (*strongly disagree)* to 7 (*strongly agree*).

Conclusions

Overall, students appreciated the brief, written, Internet-based, assignments. Students' evaluations of these writing assignments on the topical understanding and relevance dimensions, backed by students' recommendation for future use of brief papers, hinted that the writing assignments may have, indeed, bolstered involvement. In addition, students readily understood our detailed instructional handout, including a clear topical statement and task definition as Davis (1993) suggested. Whereas performance on multiple-choice exam questions related to material included in the papers was not consistently better than performance on other questions, students' comments suggested that the "take home message" of the papers (the dangers of sleep deprivation and the prevalence of eating disorders) affected their thinking.

Regarding practical issues, the structured series of questions given to students provides a framework to facilitate grading of content and organization, perhaps via use of a grading checklist. Instructors might then weight the value of sentence construction, grammar, and spelling to suit personal preferences. Regarding the stability of Internet sites, the material we used from organizational-type sites was available for at least 12 months, with one minor change of address. We recommend that instructors test the Web site links no less than a week before papers are due, finding alternative sites if original materials are not available. By using only four or five sites for each assignment, we limit the time needed to check availability each semester. Furthermore, we suggest instructors be sensitive and prepared to address students' concerns related to personal sleep or eating patterns that may arise from completing these assignments.

We offer these Web-based assignments as one example of brief, focused writing, as recommended by Lowman (1995). Instructors can use these assignments, and others with a similar design, to extend or replace textbook coverage, introducing an additional form of assessment that students enjoy. Thus, these assignments may offer a productive way to introduce writing in introductory psychology courses, without overburdening instructors with grading.

References

Chudler, E. (2004). What is sleep … and why do we do it? Retrieved December 15, 2004, from http://faculty.washington.edu/chudler/sleep.html

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

Dement, W. (n.d.). *The Epworth Sleepiness Scale*. Retrieved December 15, 2004, from http://www.stanford.edu/~dement/epworth.html
Lowman, J. (1995). *Mastering the techniques of teaching* (2nd ed.). San Francisco: Jossey-Bass.
McKeachie, W. J. (2002). *McKeachie's teaching tips: Strategies, research, and theory for college and university teachers* (11th ed.). Boston: Houghton Mifflin.
Monash University. (n.d.). *Psy 11 laboratory assignment questionnaire*. Retrieved December 15, 2004, from http://www.med.monash.edu.au/psych/ol/vl/psy11body/body.html
Sternberg, R. J. (2001, June). *Expanding assessments of performance in psychology courses: Why it is necessary and not just nice.* Paper presented at the 13th annual meeting of the American Psychological Society, Toronto, Ontario, Canada.

Notes

1. Pam Marek is now at Kennesaw State University.
2. Portions of this article were presented at the 15th Annual Southeastern Conference on the Teaching of Psychology, Kennesaw, GA, February 2003.
3. We thank Randolph Smith and three anonymous reviewers for their comments on an earlier version of this manuscript.
4. Send correspondence and requests for copies of current assignments or grading checklists to Pam Marek, Department of Psychology, #2402, Kennesaw State University, 1000 Chastain Road, Kennesaw, GA 30144–5591; e-mail: pmarek@kennesaw.edu.

Enhancing Online Instruction With Humor

Frank M. LoSchiavo and Mark A. Shatz
Ohio University–Zanesville

We tested the effectiveness of humor as an instructional strategy in an online general psychology course. We randomly assigned students to a standard section of the online course or to a section enhanced with additional humorous content. Humor significantly influenced student interest and participation but had no effect on overall course performance. We discuss the implications of these results, and we provide suggestions for integrating humor into online course components.

Both teachers and students believe that learning should be fun. Teachers rely on humor to create warm, inviting classroom environments (e.g., Buskist, 2002; Roth, 1997), and students report that humor increases how much they like their courses, their instructors, and how much they learn (e.g., Gorham, 1988; Wanzer & Frymier, 1999; White, 2001). Even though objective measures of learning (e.g., exam scores) have produced mixed results, most researchers have concluded that humor is an effective instructional strategy (e.g., Berk, 2000; Berk & Nanda, 1998; McMorris, Boothroyd, & Pietrangelo, 1997; Oppliger, 2003; Zillman & Bryant, 1989; Ziv, 1988).

However, humor is not a pedagogical panacea; no research has evaluated whether humor is beneficial in online teaching environments, where most communication is text-based and there is little or no real-time interaction between instructors and students. Using an experimental design, we tested the effectiveness of humor in an online general psychology course. We hoped that humor would help establish a comfortable rapport that would ultimately alter the immediacy (i.e., perceived social distance; Kelley & Gorham, 1988) of the instructor to the students. Previous research has shown that humor and immediacy are positively related to course satisfaction (Gorham, 1988) and class participation (Korobkin, 1988). Thus, we predicted that students exposed to additional humor would find the course more interesting and enjoyable, and we suspected they would use the course Web site more frequently. Because objective measures of learning have produced inconsistent results (McMorris et al., 1997), we did not expect to find differences in academic performance, particularly because the additional humor was not specifically intended to strengthen one's conceptual understanding of the material.

Method

Participants

Participants were 44 undergraduates (38 women, 6 men) who registered for an online general psychology course.

Procedure

After registration closed, we randomly assigned students to one of two sections of the course, each delivered over the Internet using the Blackboard Learning System (Version 5.5). Student interaction between sections was unlikely because the online students rarely visited campus and a log-in procedure restricted access to one's assigned section.

Both sections featured learning objectives, chapter summaries, text-based lectures, quizzes, and discussion boards. Although we did not intend for the standard section to be humorous, the lectures had a lighthearted, casual tone, making them somewhat entertaining to read. To create the humor-enhanced section, we systematically incorporated humor into the existing course components by adding two or three content-relevant jokes to each of the 26 lectures, humorous cartoons to each of the 14 quizzes, and witty remarks to the 10 electronic announcements. Thus, the only difference between the two sections was that one featured overtly humorous material.

Students followed a strict schedule that included weekly reading assignments, mandatory class participation (using discussion boards), and brief quizzes. In addition, students completed four proctored exams, including a comprehensive final exam. To assess academic and social behavior, we tracked how often students accessed course contents (e.g., lectures) and features (e.g., e-mail). On completion of the quarter, we sent students a brief survey to evaluate how humor influenced their perceptions of the course.

Results

Based on prior research, we estimated that effect sizes would be relatively small ($d = \sim .30$; Cohen, 1988). Consequently, we chose to compensate for low statistical power by performing one-tailed tests.

Academic Performance

We based final grades on class participation, quiz scores, and exam scores. The results appear in Table 1. As predicted, there were no differences in final grades between the two sections. However, students assigned to the humor-enhanced section earned significantly more class participation points, $t(42) = 2.19$, $p < .05$. This analysis demonstrates that students in the humor-enhanced section were more likely to participate in online discussions.

Course Usage

To better understand trends in course usage, we sorted each segment of the course into one of three categories (lecture, social and interactive, or organizational) and analyzed corresponding hit rates. The lecture category included sections for lectures, topic previews, and learning objectives. The social and interactive category included sections for e-mail and discussion boards. The organizational category included sections for course policies, assignments, and current grades. The results appear in Table 2.

Although there were no significant differences in overall course usage, students assigned to the humor-enhanced section used social and interactive features more frequently, $t(42) = 1.71$, $p < .05$. Informal observations also supported this finding. For example, students in the humor-enhanced section were more likely to post personal messages and to reply to other students' questions.

Students' Perceptions of the Course

Sixty-four percent of the students returned the anonymous survey. The results appear in Table 3. Students in the humor-enhanced section were more likely to appreciate the humorous content, recommend that humor continue to be used, and agree that humorous content made the course more interesting.

Subsequent Course Selection

We predicted that additional humor would spark added interest in online courses and in psychology courses. However, an analysis of registration records yielded no significant differences between the two groups. The results appear in Table 4.

Table 1. Student Performance by Category

	Standard[a]		Humor-Enhanced[a]			
	M	*SD*	*M*	*SD*	*t*	*d*
Class participation	92.21	12.45	98.38	4.36	2.19*	0.68
Quizzes	59.86	7.36	60.50	6.65	0.30	0.09
Exams	72.37	12.27	72.10	10.63	0.09	0.03
Final course grade	76.76	11.13	76.71	9.69	0.02	0.01

Note. Means represent the average percentage of total points earned in each category. *d* = Cohen's standardized effect size.
[a] $n = 22$.
* $p < .05$, one-tailed.

Table 2. Course Usage by Category

	Standard[a]		Humor-Enhanced[a]			
	M	*SD*	*M*	*SD*	*t*	*d*
Lecture	638.82	449.99	589.05	429.31	0.38	0.12
Social/ interactive	413.45	242.09	583.45	401.03	1.71*	0.53
Organizational	175.86	55.98	203.59	82.07	1.31	0.40
Total	1,228.14	609.63	1,376.09	674.04	0.76	0.23

Note. Means represent the average number of mouse clicks in each course category during the 10-week academic term. *d* = Cohen's standardized effect size.
[a] $n = 22$.
* $p < .05$, one-tailed.

Discussion

These findings provide direct support for our general hypothesis that pedagogical humor enhances online instruction. Although online learning environments lack many of the interpersonal dynamics of traditional instruction (e.g., face-to-face communication), this study demonstrates that humor can help create an electronic atmosphere that fosters both academic interest and social relations. Specifically, students in the humor-enhanced section were more likely to agree that humorous content made the course more interesting, and they interacted with each other more often.

Nevertheless, additional humor did not increase overall performance. Although humor can foster a positive online environment, it is not a substitute for traditional scholarship. For example, when discussing dreams, an instructor might joke that Martha Stewart dreams about designing trendy prison jumpsuits. Although the joke is easy to understand, it lacks conceptual relevance, so it is unlikely to help students better understand the psychology of dreams (for a related discussion, see Mayer, 2001). In fact, we worry that humor may pacify confused students, leaving them less likely to seek clarification when necessary. Others have voiced similar concerns. For example, Olson and Clough (2003) noted that a potential consequence of the "education as entertainment" movement is that students may develop the belief that learning is easy. Although additional research is needed in this area, evidence suggests that conceptual humor fosters retention (Kaplan & Pascoe, 1977).

It is also possible that our humor manipulation was too weak to influence academic performance. Although the humor-enhanced section featured overt humor, the data indicated that students enjoyed the lighthearted, casual tone in the standard section as well. In the future, researchers may want to test if stronger manipulations influence academic performance.

Nevertheless, the benefits of pedagogical humor merit its use as an instructional strategy. When incorporating humor into online instruction, instructors should consider the following issues. First, the majority of the humor should address a specific instructional objective. After all, the primary goal is

Table 3. Students' Perceptions of Course-Related Humor

	Standard[a]		Humor-Enhanced[a]			
	M	*SD*	*M*	*SD*	*t*	*d*
I appreciated that humor was used in the course.	4.14	1.10	4.71	0.61	1.71*	0.67
The use of humor made the course more interesting.	4.21	0.89	4.71	0.61	1.73*	0.68
The use of humor made the course more enjoyable.	4.36	0.84	4.79	0.58	1.57	0.62
I was motivated to read the course material because of the humor.	3.71	1.14	4.21	0.89	1.29	0.51
I recommend that humor continue to be used in the course.	4.21	1.05	4.79	0.58	1.78*	0.70
I recommend that more humorous material be added to the course.	4.00	1.11	4.50	0.94	1.29	0.51

Note. Students responded to each item using a 5-point scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). *d* = Cohen's standardized effect size.
[a] $n = 14$.
*$p < .05$, one-tailed.

Table 4. Credit Hours by Category

	Standard[a]		Humor-Enhanced[a]			
	M	*SD*	*M*	*SD*	*t*	*d*
Online credits	1.00	2.62	0.73	1.58	0.42	0.13
Psychology credits	2.45	3.43	2.86	3.85	0.37	0.11
Total credits	35.55	21.47	33.59	17.52	0.33	0.10

Note. Means represent the average number of quarter hours in each category during the academic year following enrollment in the present course. No differences between the two groups were statistically significant. *d* = Cohen's standardized effect size.
[a] $n = 22$.

education rather than entertainment. Second, students do not expect their professors or their online courses to be humorous, and because their comedic expectations are low, instructors can use simple forms of humor successfully. For example, online students are likely to appreciate witty observations that would likely fall flat at a nightclub (where expectations are high). Third, because humorous material often pokes fun at a target (e.g., people, places), instructors must recognize its potentially offensive nature and must select targets carefully. This issue is particularly important because online material is not only easy to misinterpret, it is more difficult than spoken words to clarify or to retract. Finally, it is important to understand that misusing humor may cause students to perceive important topics as trivial and instructors as less-than-serious. Furthermore, when instructors misuse humor, students are likely to respond in kind. In general, the judicious, appropriate, and timely use of humor allows instructors to create interesting and inviting online courses that reinforce a critical educational lesson—learning should be fun.

References

Berk, R. A. (2000). Does humor in course tests reduce anxiety and improve performance? *College Teaching, 48,* 151–158.

Berk, R. A., & Nanda, J. P. (1998). Effects of jocular instructional methods on attitudes, anxiety, and achievement in statistics courses. *Humor: International Journal of Humor Research, 11,* 383–409.

Buskist, W. (2002). Effective teaching: Perspectives and insights from Division Two's 2- and 4-year awardees. *Teaching of Psychology, 29,* 188–193.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Gorham, J. (1988). The relationship between verbal teacher immediacy behaviors and student learning. *Communication Education, 37,* 40–53.

Kaplan, R. M., & Pascoe, G. C. (1977). Humorous lectures and humorous examples: Some effects upon comprehension and retention. *Journal of Educational Psychology, 69,* 61–65.

Kelley, D. H., & Gorham, J. (1988). Effects of immediacy on recall of information. *Communication Education, 37,* 198–207.

Korobkin, D. (1988). Humor in the classroom: Considerations and strategies. *College Teaching, 36,* 154–158.

Mayer, R. E. (2001). *Multimedia learning*. Cambridge, England: Cambridge University Press.

McMorris, R. F., Boothroyd, R. A., & Pietrangelo, D. J. (1997). Humor in educational testing: A review and discussion. *Applied Measurement in Education, 10,* 269–297.

Olson, J. K., & Clough, M. P. (2003). Computer-assisted education can undermine serious study. In J. D. Torr (Ed.), *Computers and education* (pp. 56–65). Farmington Hills, MI: Greenhaven.

Oppliger, P. A. (2003). Humor and learning. In J. Bryant, D. Roskos-Ewoldsen, & J. Cantor (Eds.), *Communication and emotion: Essays in honor of Dolf Zillman* (pp. 255–273). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Roth, J. K. (Ed.). (1997). *Inspiring teaching: Carnegie professors of the year speak*. Bolton: Anker.

Wanzer, M. B., & Frymier, A. B. (1999). The relationship between student perceptions of instructor humor and students' reports of learning. *Communication Education, 48,* 48–62.

White, G. W. (2001). Teachers' reports of how they used humor with students' perceived use of such humor. *Education, 122,* 337–347.

Zillman, D., & Bryant, J. (1989). Guidelines for the effective use of humor in children's educational television programs. In P. E. McGhee (Ed.), *Humor and children's development: A guide to practical applications* (pp. 201–221). New York: Haworth.

Ziv, A. (1988). Teaching and learning with humor: Experiment and replication. *Journal of Experimental Education, 57,* 5–15.

Note

Send correspondence to Frank M. LoSchiavo, Department of Psychology, Ohio University–Zanesville, 1425 Newark Road, Zanesville, OH 43701; e-mail: loschiav@ohiou.edu.

Core Terms in Undergraduate Statistics

R. Eric Landrum
Boise State University

*I analyzed 3 introductory statistics textbooks to generate a listing of key terms and concepts. After removing duplications, 374 unique terms formed the master list. A national sample of introductory statistics instructors (N = 190) evaluated portions of the master list, rating each item on a scale ranging from 1 (*not at all important*) to 3 (*extremely important*). I list the Top 100 items and offer recommendations about the usefulness of the importance listing in facilitating statistics instruction.*

An undergraduate course in statistics continues to be an essential component in the curriculum for the psychology major. After reviewing curricula from hundreds of colleges and universities nationwide, Perlman and McCann (1998) concluded that the core of the undergraduate major comprises introductory psychology, statistics, and a capstone course. Instructors of statistics face a host of challenges (Conners, Mccown, & Roskos-Ewoldsen, 1998; Hastings, 1982). For instance, Conners et al. listed four unique challenges in teaching the statistics course: (a) motivating students, (b) overcoming math anxiety, (c) avoiding performance extremes, and (d) making learning last. Given the prevalence of the statistics course and the unique challenges of teaching it, any resource available to facilitate instructor performance in teaching statistics would be highly valuable.

This study shares similarities with previous studies (e.g., Giesbrecht, Sell, Scialfa, Sandals, & Ehlers, 1997). Those authors had 18 professors from four different disciplines rate the importance of topics in research methods and statistics. Giesbrecht et al. obtained topic listings from previous research, review of statistics and research methods texts, and course descriptions, and "found interdepartmental agreement on the relative importance of 97% of the statistics topics, with 74% of those topics deemed to be important" (p. 245). In this study, I focused on statistics instruction at the undergraduate level only and within psychology only, and I base my conclusions on the responses of a national sample of undergraduate statistics instructors.

My goal was to generate a master list of statistical terms relevant to an undergraduate statistics course and then to ask statistics instructors to rate the importance of those terms in that context. Other researchers have successfully used this approach with introductory psychology terms (Boneau, 1990; Griggs & Mitchell, 2002; Landrum, 1993; Quereshi, 1993; Quereshi & Sackett, 1977; Zechmeister & Zechmeister, 2000). For instance, Zechmeister and Zechmeister performed a content analysis of terms using the glossaries of 10 introductory psychology textbooks and generated a list of 2,505 unique terms. After dividing the list into smaller sections, Zechmeister and Zechmeister then sent the sections to a sample of introductory psychology instructors for importance ratings. These top terms are construed as the core terms in introductory psychology.

There are potential benefits in having a set of core terms for an undergraduate statistics course. New instructors of statistics would have a valuable reference in helping determine course content. All statistics instructors would have the ability to compare their decisions about course content (especially across multiple sections) to the importance ratings of a national sample of statistics instructors. Regarding textbook selection, other approaches have been helpful in choosing a statistics book, such as assessing readability and writing style (Harwell, Herrick, Curtis, Mundfrom, & Gold, 1996). A core terms listing for statistics might also assist instructors in the textbook selection process. Instructors face time constraints regarding the content and topics of the course versus the length of the semester. Importance ratings might help instructors (or even statistics textbook authors) determine those topics that are less important. Making this type of informed choice might lead an instructor to come back to a topic at the end of a semester, time permitting. Thus, there are multiple plausible advantages to having a list of core terms in statistics rated by importance.

Method

Participants

Psychology department chairs received a cover letter explaining the purpose of this study, a survey, and a self-addressed business-reply envelope. I asked the department chair to forward the survey to an undergraduate statistics instructor. I mailed 814 surveys to colleges and universities in the United States that offered an undergraduate degree in psychology—I obtained this contact information from the American Psychological Association Office of Research. A total of 190 instructors responded with usable data, yielding a response rate of 23.3%.

Materials and Procedure

Prior to surveying statistics instructors for importance ratings, I developed a master list of terms; I describe that process here and the materials used to survey statistics instructors.

Generation of the master list. I selected a convenience sample of statistics texts (i.e., those that were on my bookshelf): Gravetter and Wallnau (1999), Pagano (1998), and Spatz (1997). The important details for each text follow: Gravetter and Wallnau: 458 pages, 16 chapters, 210 key terms; Pagano: 548 pages, 19 chapters, 238 key terms; Spatz: 488 pages, 14 chapters, 167 key terms. Because not all three books contained glossaries, I could not use previous content analysis techniques (e.g., Zechmeister & Zechmeister, 2000). Instead, I performed a page-by-page content analysis of all three books and selected terms in bold face type or italics. In some instances I used the heading of a particular section. I previously used a similar technique in my examination of introductory psychology textbooks (Landrum, 1993). At times, I had to make decisions about similar terms. For example, "line of best fit" and "best-fitting line" mean essentially the same thing; I eliminated one from the final list of terms. This decision was subjective; if terms were highly

similar but not identical, I erred on the side of caution and included both terms.

The three statistics texts yielded 615 terms; 374 unique terms emerged to constitute the list of statistics terms. After some consideration, I decided not to compare terms chapter by chapter, as previously done in the content analysis of introductory psychology textbooks (e.g., Landrum, 1993; Zechmeister & Zechmeister, 2000). For instance, one book had a separate chapter for measures of central tendency and variability, whereas another combined the two topics into one chapter. Of the 615 terms coded, 44 terms appeared in all three books.

Importance ratings by statistics instructors. I believed that 374 statistical terms was too many to ask any one person to evaluate, so I randomly divided the listing of terms into three separate, mutually exclusive forms. Thus, each instructor rated a list of 124 or 125 terms, presented alphabetically. On average, 63.3 instructors ($SD = 13.0$) rated each list of terms. Participants rated each concept using a 4-point importance scale from 0 (*not at all important*), 1 (*slightly important*), 2 (*moderately important*), to 3 (*extremely important*). I instructed participants to leave an item blank if it was unclear, and I calculated a mean rating for each term. The term *normal curve* was the only term to receive a unanimous rating of 3.0 (for the means and standard deviations of the top 100 terms, see Table 1).

Table 1. Means and Standard Deviations for the Top 100 Terms

Item	*M*	*SD*
Normal curve	3.00	0.00
Statistically significant	2.98	0.13
Bell-shaped curve	2.96	0.18
Significance level	2.96	0.18
Hypothesis testing	2.96	0.19
Normal distribution	2.96	0.19
Standard deviation	2.96	0.19
Sample	2.94	0.22
Alpha level	2.94	0.23
Mean	2.93	0.25
Null hypothesis	2.93	0.37
Central tendency	2.92	0.26
Inferential statistics	2.92	0.31
Variability	2.92	0.31
Arithmetic mean	2.91	0.28
Correlation	2.90	0.29
Pearson correlation	2.90	0.29
Dependent variable	2.89	0.34
Two-tailed probability	2.89	0.44
Positive correlation	2.88	0.37
Data	2.88	0.38
Hypothesis	2.88	0.45
t test	2.87	0.51
Descriptive statistics	2.86	0.34
Variance	2.86	0.34
Negative correlation	2.86	0.43
Not significant	2.85	0.44
Variable	2.84	0.42
Population	2.84	0.52
Statistic	2.84	0.52
Level of significance	2.83	0.37
Critical values	2.83	0.46
Type I error	2.82	0.46
Degrees of freedom	2.81	0.43

(continued)

Table 1 (Continued)

Item	*M*	*SD*
Median	2.81	0.44
Significant effect	2.81	0.53
Rejection region	2.81	0.57
t test for independent-samples design	2.81	0.57
One-way ANOVA	2.80	0.44
Statistical inference	2.79	0.46
Two-tailed test of significance	2.79	0.51
t test for independent groups	2.79	0.58
t statistic	2.79	0.63
Standard error of the mean	2.77	0.55
Critical region	2.76	0.50
Standard error	2.75	0.58
ANOVA	2.75	0.63
Inferential process	2.74	0.57
Alternative hypothesis	2.74	0.60
F ratio	2.73	0.48
Deviation	2.73	0.58
Distribution of sample means	2.73	0.59
Student's *t* test	2.73	0.66
Linear relationship	2.72	0.55
Independent-samples design	2.72	0.64
z score transformation	2.72	0.67
Random	2.71	0.61
Random assignment	2.71	0.62
Sampling error	2.70	0.51
Correlational method	2.70	0.60
z score	2.70	0.72
Null-hypothesis population	2.68	0.65
Frequency	2.67	0.54
Independent groups design	2.67	0.58
Frequency distribution	2.67	0.70
Independent variable	2.66	0.63
Type II error	2.64	0.64
One-tailed probability	2.62	0.67
Random selection	2.62	0.72
Nondirectional hypothesis	2.61	0.64
Sampling distribution	2.61	0.71
Estimated population standard deviation	2.60	0.63
Overall mean	2.60	0.67
Correct decision	2.60	0.77
Sampling distribution of the mean	2.59	0.72
Sampling distributions of a statistic	2.59	0.72
Regression	2.59	0.81
Causation	2.58	0.56
Scatterplot	2.58	0.61
Sum of squares	2.57	0.63
Positive relationship	2.55	0.67
Sampling distribution of *t*	2.55	0.73
Sum of squared deviations	2.55	0.77
Test statistic	2.55	0.82
Chi-square distribution	2.54	0.65
Between-groups sum of squares	2.54	0.75
Simple random sample	2.54	0.75
Population variance	2.53	0.72
Random sampling	2.52	0.78
t distribution	2.51	0.69
Chi-square statistic	2.51	0.78
One-tailed test of significance	2.51	0.78
Probability	2.50	0.70
Standard score	2.50	0.75
F distribution	2.50	0.83
Distribution of scores	2.49	0.64
ANOVA summary table	2.49	0.74
Treatment	2.49	0.75
Levels/treatments	2.49	0.78
Subjects/participants	2.49	0.81

Discussion

The results of this study, as presented in Table 1, provide an interesting look at what statistic instructors consider important. This Top 100 list may be construed as a first approximation of the core of introductory statistics in psychology. These results may also help instructors decide what the important terms to cover in class are and what terms are secondary. The approach of this study is unique because only undergraduate statistics instructors rated the importance of terms. Departments that offer multiple sections of statistics taught by different instructors using different textbooks might find these results useful in an effort to coordinate consistent coverage of the various content areas of statistics. The results of this study are more specific to psychology instruction than those from a previous study (Giesbrecht et al., 1997).

For some semesters, instructors may not cover certain chapters of information they consider less important; other semesters they may cover those chapters. What topics sometimes get deleted due to time constraints? The listing presented in Table 1 should help instructors and authors to answer these questions by understanding the core terms in the context of relative importance.

This study has limitations with regard to the methodological approach used, including the convenience sample of textbooks. In addition, in the instances where terms were highly similar but not identical, I included all terms. More work in this area is warranted to continue to verify the core terms. The listing provided here can be valuable for instructors in designing and making choices about the coverage of topics in an introductory course in statistics.

References

Boneau, C. A. (1990). Psychological literacy: A first approximation. *American Psychologist, 45,* 891–900.

Conners, F. A., Mccown, S. M., & Roskos-Ewoldsen, B. (1998). Unique challenges in teaching undergraduate statistics. *Teaching of Psychology, 25,* 40–42.

Giesbrecht, N., Sell, Y., Scialfa, C., Sandals, L., & Ehlers, P. (1997). Essential topics in introductory statistics and methodology courses. *Teaching of Psychology, 24,* 242–246.

Gravetter, F. J., & Wallnau, L. B. (1999). *Essentials of statistics for the behavioral sciences* (3rd ed.). Pacific Grove, CA: Brooks/Cole.

Griggs, R. A., & Mitchell, M. C. (2002). In search of introductory psychology's classic core vocabulary. *Teaching of Psychology, 29,* 144–147.

Harwell, M. R., Herrick, M. L., Curtis, D., Mundfrom, D., & Gold, K. (1996). Evaluating statistics texts used in education. *Journal of Education and Behavioral Sciences, 21,* 3–34.

Hastings, M. W. (1982). Statistics: Challenge for students and the professor. *Teaching of Psychology, 9,* 221–222.

Landrum, R. E. (1993). Identifying core concepts in introductory psychology. *Psychological Reports, 72,* 659–666.

Pagano, R. R. (1998). *Understanding statistics in the behavioral sciences* (5th ed.). Pacific Grove, CA: Brooks/Cole.

Perlman, B., & McCann, L. I. (1998). *The undergraduate psychology curriculum 1996–97* [Final report]. Washington, DC: American Psychological Association.

Quereshi, M. Y. (1993). The contents of introductory psychology textbooks: A follow-up. *Teaching of Psychology, 20,* 218–222.

Quereshi, M. Y., & Sackett, P. R. (1977). An updated content analysis of introductory psychology textbooks. *Teaching of Psychology, 4,* 25–30.

Spatz, C. (1997). *Basic statistics: Tales of distributions* (6th ed.). Pacific Grove, CA: Brooks/Cole.

Zechmeister, J. S., & Zechmeister, E. B. (2000). Introductory textbooks and psychology's core concepts. *Teaching of Psychology, 27,* 6–11.

Notes

1. The comprehensive list of means and standard deviations for all 374 terms is available from the author.
2. Send correspondence to Eric Landrum, Department of Psychology, Boise State University, 1910 University Drive, Boise, ID 83725–1715; e-mail: elandru@boisestate.edu.

Promoting Conceptual Understanding of Statistics: Definitional Versus Computational Formulas

Katarina Guttmannova, Alan L. Shields, and John C. Caruso
University of Montana

Computer applications have replaced hand calculations as the relevant procedural skill for most of the statistical techniques in introductory statistics courses. Therefore, definitional formulas should replace computational formulas because only the former contribute to conceptual understanding. A review of 12 introductory statistics textbooks indicated that they emphasized computational formulas, particularly for complex techniques and exercises. We argue that the presentation and use of computational formulas is counterproductive to the learning goals of statistics courses and provide recommendations for instructors to facilitate the use of definitional formulas.

In the precomputer era, when students performed calculations by hand, computational formulas provided heuristics for difficult computations but sacrificed understanding of the underlying statistical concepts. Currently, however, 69% of psychology departments use computers in introductory statistics courses, and 90% use them for data analysis at some point in their curriculum (Bartz & Sabolik, 2001). Computers perform most statistical analyses. Thus we question the continued need for computational formulas and advocate instead for definitional formulas computed using hand calculators.

Although hand calculations are less tedious when using computational formulas, definitional formulas provide students with some perspective as to what the statistic is, when to use it, and how to interpret it. Consider, for example, the definitional,

$$s^2 = \frac{\sum_i (X_i - M)^2}{N},$$

and computational,

$$s^2 = \frac{\sum_i X_i^2 - \frac{(\sum_i X_i)^2}{N}}{N},$$

formulas for the variance, where X_i are the raw scores, N is the sample size, and M is the arithmetic mean of all observations. The variance of a distribution is a measure of how dispersed the scores are around the mean in the distribution. Hence, conceptual understanding is advanced by examining and using the definitional formula, whereas the computational formula is only a computational tool with no interpretative significance.

As Abelson (1995) observed, "From long observation of student struggles with statistics, I conclude that the difficulties lie not so much with computational mechanics as with an overall perspective on what they are doing" (p. xii). If instruction should focus on the tasks students will perform after completing their statistics course, and if computer applications have replaced hand calculations as the required procedural skill, then the advantage of the computational formulas is diminished. However, we hypothesized that a focus on computational in contrast to definitional formulas exists in current statistics textbooks, impeding students' grasp of the overall perspective.

Method and Results

We reviewed a convenience sample of 12 introductory statistics books published between 1999 and 2001 (see the Appendix). For each textbook, we tabulated the definitional and computational formulas in the text; in examples; and in exercises for chapters on variance, correlation, and one-way ANOVA. The results of our review appear in Table 1.

Among the textbooks we examined, there was not a statistically significant difference in preference for the definitional formula over the computational formula for the variance. Only 75% of the textbooks provided even a single exercise requiring the use of the definitional formula for the variance. We question how a student can conceptually understand the variance without ever using its definitional formula. In correlation chapters, authors of only 41.7% of the textbooks suggested that students work at least one exercise using the definitional formula. For one-way ANOVA chapters, only 33.3% of textbooks provided a numeric example, and only 25% provided even a single exercise using the definitional approach. We also performed an analysis of 26 textbooks published from 1990 to 2002 and reached similar conclusions (results are available from the first author).

Statistical tests (two-tailed, $\alpha = .05$) of the difference in frequencies between computational and definitional formulas revealed that more textbooks employed computational formulas for correlation exercises (83.3% vs. 41.7%; $z = 2.11$, $p = .04$) and for ANOVA examples (75% vs. 33.3%; $z = 2.05$, $p = .04$) and exercises (75% vs. 25%; $z = 2.45$, $p = .01$). Furthermore, examining the pattern of frequencies in Table 1 reveals a clear trend: The frequencies for definitional formulas dropped as the content became more complex in text (100% for variance vs. 58.3% for ANOVA), $z = 2.51$, $p = .01$; in examples (100% for variance vs. 33.3% for ANOVA), $z = 3.46$, $p < .001$; and in exercises (75% for variance vs. 25% for ANOVA), $z = 2.45$, $p = .01$, whereas those for the computational formulas were relatively stable (never extending below 66.7% or above 83.3%).

Discussion

Our results show a strong emphasis on the computational approach at the expense of the definitional one as procedures became more complex and conceptual understanding is presumably more elusive. Our results indicate that textbook authors are emphasizing a set of formulas designed for hand calculations while sacrificing conceptual understanding. It is unlikely that current students will ever choose to use the computational formulas unless, of course, their statistics instructors require these procedures. Definitional formulas convey the meaning of the statistic at hand, and we believe that repeatedly observing and performing calculations with definitional formulas is beneficial to the attainment of conceptual understanding. However, we recognize that using hand calculations with definitional formulas can be tedious if the examples are not chosen carefully. Consequently, we recommend using manageable instructional datasets.

Table 1. Frequency of Occurrence of Definitional and Computational Formulas in Text, Examples, and Exercises in Three Selected Textbook Chapters

	Text		Examples		Exercises	
Textbook Chapter	*n*	%	*n*	%	*n*	%
Variance						
Definitional formula	12	100.0	12	100.0	9	75.0
Computational formula	10	83.3	10	83.3	8	66.7
Correlation						
Definitional formula	10	83.3	6	50.0	5	41.7
Computational formula	10	83.3	10	83.3	10	83.3[a]
One-way ANOVA						
Definitional formula	7	58.3	4	33.3	3	25.0
Computational formula	9	75.0	9	75.0[a]	9	75.0[a]

Note. $N = 12$.
[a]Indicates that the proportion (frequency/12) for definitional formulas is statistically different from the proportion for computational formulas (two-tailed test, $\alpha = .05$).

Recommendations

The following recommendations to instructors of introductory statistics can make the use of definitional formulas easier for students:

1. *Use small data sets:* 5 or 10 observations are adequate to demonstrate most statistical techniques.
2. *Present data in computationally manageable units:* Instead of using a number such as 1.5 hr, use 90 min.
3. *Construct data sets that produce integer means and standard deviations:* Dudek (1996), McGown and Spencer (1996), and Read and Riley (1983) provided methods for generating such data sets.
4. *Use a textbook that emphasizes the definitional approach:* Aron and Aron (1994) professed the perspective most similar to what we are proposing. They emphasized definitional formulas throughout and stated that: "The purpose of this book is to help you understand statistical procedures, not to turn you into a computer by having you memorize and use computational formulas you will rarely, if ever, employ again" (p. 47). Abrami, Cholmsky, and Gordon (2001) also stressed the definitional approach. Those who would like to use Web-based instructional methods and materials should consider Stockburger (1998).

Conclusions

Historically, using both definitional and computational formulas for teaching statistical concepts in introductory statistics classes was justifiable. The calculation of statistics with pencil and paper or calculator was required of students after the course; hence learning computational formulas was a relevant objective. However, given the widespread use of computers for instructional and especially noninstructional data analysis, the computational formulas have lost their primary purpose. To help students gain conceptual understanding, we believe that instruction as well as statistics textbooks' text, examples, and exercises should be designed around procedures developed specifically for that purpose: definitional formulas.

References

Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Abrami, P. C., Cholmsky, P., & Gordon, R. (2001). *Statistical analysis for the social sciences: An interactive approach.* Needham Heights, MA: Allyn & Bacon.

Aron, A., & Aron, E. N. (1994). *Statistics for psychology*. Upper Saddle River, NJ: Prentice Hall.

Bartz, A. E., & Sabolik, M. A. (2001). Computer and software use in teaching the beginning statistics course. *Teaching of Psychology, 28,* 147–149.

Dudek, F. J. (1996). Data sets having integer means and standard deviations. In M. E. Ware & D. E. Johnson (Eds.), *Handbook of demonstrations and activities in the teaching of psychology* (Vol. 1, p. 88). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (Originally published in *Teaching of Psychology, 8,* 51)

McGown, W. P., & Spencer, W. B. (1996). For statistics classes: Data sets with integer means and standard deviations. In M. E. Ware & D. E. Johnson (Eds.), *Handbook of demonstrations and activities in the teaching of psychology* (Vol. 1, p. 87). Mahwah, NJ: Lawrence Erlbaum Associates, Inc. (Originally published in *Teaching of Psychology, 7,* 63)

Read, K. L. Q., & Riley, I. S. (1983). Statistics problems with simple numbers. *American Statistician, 37,* 229–231.

Stockburger, D. W. (1998, February 19). *Introductory statistics: Concepts, models, and applications.* Retrieved January 26, 2003, from http://www.psychstat.smsu.edu/introbook/sbk00.htm

Appendix
Introductory Statistics Textbooks Reviewed

Abrami, P. C., Cholmsky, P., & Gordon, R. (2001). *Statistical analysis for the social sciences*. Needham Heights, MA: Allyn & Bacon.

Aron, A., & Aron, E. N. (1999). *Statistics for psychology* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Gravetter, F. J., & Wallnau, L. B. (2000). *Statistics for the behavioral sciences* (5th ed.). Belmont, CA: Wadsworth Thomson Learning.

Heiman, G. W. (2000). *Basic statistics for the behavioral sciences* (3rd ed.). Boston: Houghton Mifflin.

Kurtz, N. R. (1999). *Statistical analysis for the behavioral sciences*. Needham Heights, MA: Allyn & Bacon.

Levin, J., & Fox, J. A. (2000). *Elementary statistics in social research* (8th ed.). Needham Heights, MA: Allyn & Bacon.

Raymondo, J. C. (1999). *Statistical analysis in the behavioral sciences*. New York: McGraw-Hill.

Ritchey, F. (2000). *The statistical imagination: Elementary statistics for the social sciences*. New York: McGraw-Hill.

Runyon, R. P., Coleman, K. A., & Pittenger, D. J. (2000). *Fundamentals of behavioral statistics* (9th ed.). Boston: McGraw-Hill.

Sprinthall, R. C. (2000). *Basic statistical analysis* (6th ed.). Needham Heights, MA: Allyn & Bacon.

Thorne, B. M., & Giesen, J. M. (2000). *Statistics for the behavioral sciences* (3rd ed.). Mountain View, CA: Mayfield.

Witte, R. S., & Witte, J. S. (2001). *Statistics* (6th ed.). Fort Worth, TX: Hartcourt College.

Notes

1. Alan L. Shields is currently an Instructor of Psychology in the Harvard Medical School and works in the Alcohol and Drug Abuse Treatment Program at McLean Hospital. John C. Caruso is currently the Director of Services at Dissertation Statistical Services. Katarina Guttmannova is currently a postdoctoral fellow at the Institute for Policy Research, Northwestern University.
2. Send correspondence to Katarina Guttmannova, Institute for Policy Research, Northwestern University, 2040 Sheridan Road, Evanston, IL 60208–4100; e-mail: k-guttmannova@northwestern.edu.

Graphing Psychology: An Analysis of Source Material of Graphs in Introductory Psychology Textbooks

Jeffrey S. Nevid and Nicholas Forlenza
St. John's University

We conducted an analysis of source material in data graphs included in introductory psychology texts. Our findings showed an increasing trend toward inclusion of graphs representing more re-

cent source material as well as a greater frequency of data graphs derived from research in "softer" areas of psychology. An examination of the most frequently occurring graphs offers a snapshot of the types of quantitative relationships that textbook authors believe should comprise the student's first encounter with psychology.

Textbooks in psychology and other college-level disciplines have undergone major changes in style and format in recent years. In certain respects, however, textbooks have changed little over the years. They still rely largely on the written word and visual representations of information in the form of graphs and diagrams.

Graphs in college texts can help students understand quantitative data and relations among factors or variables. A recent symposium on the use of spatial text (pictures, illustrations, maps, graphs, etc.) provided support for the educational value of visually presented information in textbooks (Robinson, 2002).

Arguably, a graph leaves a stronger impression with the reader than a narrative discussion of research findings. Graphs may also present data relations more clearly than narrative descriptions. However, in light of the need to limit textbook size and scope of coverage, textbook authors need to be judicious about selecting which graphs to include. In effect, the selective use of particular graphs may be an indirect measure of the degree of importance accorded these data and the studies they represent.

This study examined the source material of data graphs in introductory texts. An earlier analysis of a sample of introductory psychology texts showed an average of 67.7 graphs (Peden & Hausmann, 2000). This study expands on the prior analysis of the frequencies of graphs in introductory texts by examining the source material of the most widely used data graphs, as well as the currency of graphical citations and the subject areas that data graphs represent.

Method

Sample

We sampled all 29 full-sized introductory psychology textbooks with copyright dates of 2000 to 2002 listed on the Web sites of major college publishers (see Appendix). To avoid duplication, we omitted essential and brief versions.

Procedure

One of us (Nicholas Forlenza) reviewed each of the texts in our sample, identifying graphs that displayed quantitative data from cited research findings. We classified figures as data graphs when they met the following criteria: (a) graphical displays of quantitative variables arranged along two axes, (b) numerical values included for specific points along these axes, and (c) findings referenced to cited research. To ensure reliability, a second rater, a doctoral student in clinical psychology, reviewed a random (33%) sample of the textbooks used in the analysis. Interrater agreement for classifying in-text figures as data graphs was 95%.

Results

Table 1 lists the 10 most widely used data graphs in introductory psychology textbooks in our sample. The most widely used data graph represented the forgetting function derived from Ebbinghaus's (1885/1964) study on memory.

To determine currency of graphical displays in introductory texts, we examined frequencies of graphs in relation to the year of publication of the cited works. Figure 1 presents these data in scatterplot form. We observed a positive, accelerating trend showing increasing numbers of data graphs representing more recent studies.

Figure 2 shows the frequency of graphs in relation to subject area, as reflected by chapter headings in introductory psychology texts. To avoid duplication, we counted each graph based on the same source material only once, even if multiple authors used the same graph. We then examined the relative "hardness" of the subject area or subfields in psychology that these graphs represent. We drew on the hardness ratings of psychology journals used in a recent analysis of graphical displays (Smith, Best, Stubbs, Archibald, & Roberson-Nay, 2002). We applied the ordinal rankings of the hardness of the journals included in the Smith et al. analysis to scale the chapter (subfield) topics, extrapolating to chapter topics (health, motivation and emotion, cognitive, learning, memory, sensation and perception) not directly represented in the journal analysis.

We arranged the chapter topics in Figure 2 from top to bottom in approximation of increasing hardness based on this extrapolation. The data revealed relatively greater frequencies of distinct data graphs based on research in "softer" areas of psy-

Table 1. Sources of Graphs Most Frequently Used in Introductory Psychology Textbooks

Graph	Frequency	%
Forgetting As a Function of Time (Ebbinghaus, 1885/1964)	22	76
Familial Relationships and Schizophrenia (Gottesman, 1991)	21	72
Latent Learning (Tolman & Honzik, 1930)	15	52
Sibling Relatedness and IQ Score (Bouchard, Lykken, McGue, Segal, & Tellegen, 1990)	12	41
Lifetime Prevalence of Various Mental Disorders (Kessler et al., 1994)	12	41
Behavioral Study of Obedience (Milgram, 1963)	11	38
Duration of Short-term Memory and Rehearsal (Peterson & Peterson, 1959)	11	38
Bystander Effect (Darley & Latané, 1968)	11	38
Survey on Male and Female Sexual Behaviors (Laumann, Gagnon, Michael, & Michaels, 1994)	10	34
Improved Patients According to Number of Sessions (Howard, Kopta, Krause, & Orlinsky, 1986)	10	34

Note. $N = 29$.

chology (e.g., abnormal, social, developmental) than in harder areas (e.g., memory, sensation and perception, physiological).

Discussion

The listing of the most widely used data graphs revealed some interesting patterns. Textbook authors drew on several classic studies in the field, including Ebbinghaus's (1885/1964) early research on memory, Tolman and Honzik's (1930) latent learning study, and Milgram's (1963) obedience to authority study. On the other hand, our analysis of the currency of data graphs showed an increasing trend toward inclusion of graphs from more recent sources, especially cited works published in the 1990s (see Figure 1).

We found a higher frequency of data graphs from softer areas of psychology (e.g., abnormal, social, developmental) than from harder areas (e.g., experimental, physiological).

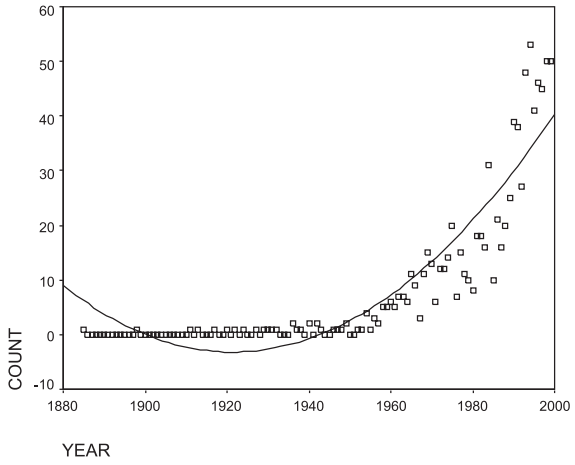

Figure 1. Frequency of data graphs in relation to year of publication of cited work.

These findings contrast with the high correlation that Smith and colleagues (Smith et al., 2002) found between graphical displays in psychology journals and hardness of subject area. The differences between Smith et al.'s findings and ours may reflect differences in the metric used to measure graph usage. Smith et al. analyzed fractional graph area, the proportion of total page area in journal articles devoted to graphical displays. In contrast, we examined the frequency of graphs representing particular studies in psychology, not space accorded graphs. Still, we were surprised by the relative lack of graphical displays from studies representing harder content areas of experimental psychology and physiological psychology. One possible explanation is that authors make greater use of other forms of visual display, especially diagrams, to represent concepts in physiological and experimental psychology. Possibly too, authors may find experimental findings too advanced or complex to be useful at the introductory level. Still, it remains to be determined why textbook authors do not make greater us of graphical displays of findings from experimental areas of psychology.

Recently, Griggs, Proctor, and Cook (2004) found little commonality among introductory psychology texts in citations of books. Griggs Proctor, and Cook argued that their data, together with earlier data showing little commonality in cited journal articles among introductory texts (Gorenflo & McConnell, 1991), point to nonhomogeneity of introductory psychology texts (see also Griggs, Proctor, & Bujak-Johnson, 2002). More recently, Griggs, Bujak-Johnson, and Proctor (2004) found little commonality among glossary terms in introductory psychology texts. Our results provide further support for the nonhomogeneity hypothesis, as none of the data graphs in our sample of introductory psychology texts met the 80% commonality criterion used by Griggs and colleagues (Griggs, Proctor, & Cook, 2004).

Our listing of the most widely used data graphs provides a good starting point for instructors seeking to balance coverage of contemporary research with examples of graphical dis-

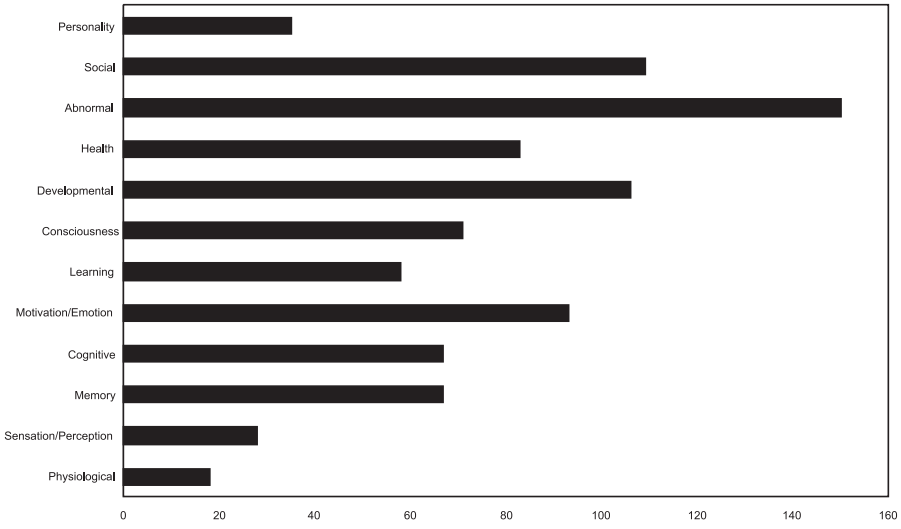

Figure 2. Frequency of data graphs by chapter topics.

plays of classic studies that place present understandings in a historical context. Instructors may also wish to compare the graphical displays in the textbooks they are presently using with the listing in Table 1 of the most widely cited data graphs in contemporary texts and supplement in class for any omissions they believe should be included in the student's first exposure to psychology. We expect that as recent studies become more widely disseminated in the field, they too will take their place in future textbooks alongside the classic works of Ebbinghaus, Tolman, Milgram, and others.

References

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: The Minnesota study of twins reared apart. *Science, 250,* 223–228.

Darley, J. M., & Latané, B. (1968). Bystander intervention in emergencies: Diffusion of responsibility. *Journal of Personality and Social Psychology, 8,* 377–383.

Ebbinghaus, H. (1964). *Memory: A contribution to experimental psychology* (H. A. Ruger & C. E. Bussemius, Trans.). New York: Dover. (Original work published 1885)

Gorenflo, D. W., & McConnell, J. V. (1991). The most frequently cited journal articles and authors in introductory psychology texts. *Teaching of Psychology, 18,* 8–12.

Gottesman, I. I. (1991). *Schizophrenia genesis: The origins of madness.* New York: Freeman.

Griggs, R. A., Bujak-Johnson, A., & Proctor, D. L. (2004). Using common core vocabulary in text selection and teaching the introductory course. *Teaching of Psychology, 31,* 265–269.

Griggs, R. A., Proctor, D. L., & Bujak-Johnson, A. (2002). The nonexistent common core. *American Psychologist, 57,* 452–453.

Griggs, R. A., Proctor, D. L., & Cook, S. M. (2004). The most frequently cited books in introductory texts. *Teaching of Psychology, 31,* 113–116.

Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. *American Psychologist, 41,* 159–164.

Kessler, R. C., McGonagle, K. A., Zhao, S., Nelson, C. B., Hughes, M., Eshleman, S., Wittchen, H., & Kendler, K. S. (1994). Lifetime and 12-month prevalence of *DSM–III–R* psychiatric disorders in the United States: Results from the National comorbidity survey (NCS). *Archives of General Psychiatry, 51,* 8–19.

Laumann, E. O., Gagnon, J. H., Michael, R. T., & Michaels, S. (1994). *The social organization of sexuality: Sexual practices in the United States.* Chicago: University of Chicago Press.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology, 67,* 371–378.

Peden, B. F., & Hausmann, S. E. (2000). Data graphs in introductory and upper level psychology textbooks: A content analysis. *Teaching of Psychology, 27,* 93–97.

Peterson, L., & Peterson, M. J. (1959). Short-term retention of individual verbal items. *Journal of Experimental Psychology, 58,* 193–198.

Robinson, D. H. (2002). Spatial text adjuncts and learning: An introduction to the special issues. *Educational Psychology Review, 14,* 1–3.

Smith, L. D., Best, L. A., Stubbs, D. A., Archibald, A. B., & Roberson-Nay, R. (2002). Constructing knowledge: The role of graphs and tables in hard and soft psychology. *American Psychologist, 57,* 749–761.

Tolman, E. C., & Honzik, C. H. (1930). Introduction and removal of reward, and maze performance in rats. *University of California Publications in Psychology, 4,* 257–275.

Appendix
Textbooks Included in Analysis

Atkinson, R. L., Atkinson, R. C., Smith, E. E., Bem, D. J., & Nolen-Hoeksema, S. (2000). *Hilgard's introduction to psychology* (13th ed.). New York: Harcourt.

Barker, L. (2002). *Psychology.* Upper Saddle River, NJ: Prentice Hall.

Baron, R. A. (2001). *Psychology* (5th ed). Boston: Allyn and Bacon.

Bernstein, D. A., Clarke-Stewart, A., Penner, L. A., Roy, E. J., & Wickens, C. D. (2000). *Psychology* (5th ed.). Boston: Houghton Mifflin.

Coon, D. (2001). *Introduction to psychology: Gateways to mind and behavior* (9th ed.). Belmont, CA: Wadsworth Thomson Learning.

Davis, S. F., & Palladino, J. J. (2000). *Psychology* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Feldman, R. S. (2002). *Understanding psychology* (6th ed.). New York: McGraw-Hill.

Gerrig R. J., & Zimbardo, P. G. (2002). *Psychology and life* (16th ed.). Boston: Allyn and Bacon.

Gleitman, H., Fridlund, A. J., & Reisberg, D. (2000). *Basic psychology* (5th ed.) New York: Norton.

Gray, P. (2000). *Psychology* (4th ed.). New York: Worth.

Hockenbury, D. H., & Hockenbury, S. H. (2000). *Psychology* (2nd ed.). New York: Worth.

Huffman, K. (2002). *Psychology in action* (6th ed.). New York: Wiley.

Kalat, J. W. (2002). *Introduction to psychology* (6th ed.). Belmont, CA: Wadsworth Thomson Learning.

Kassin, S. (2001). *Psychology* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.

Lefton, A. L. (2000). *Psychology* (7th ed.). Boston: Allyn and Bacon.

Lahey, B. B. (2001) *Psychology: An introduction.* (7th ed.). New York: McGraw-Hill.

Morris, C. G., & Maisto, A. A. (2002). *Psychology: An introduction* (11th ed.). Upper Saddle River, NJ: Prentice Hall.

Myers, D. G. (2001). *Psychology: Myers in modules* (6th ed.). New York: Worth.

Nairne, S. J. (2000). *Psychology: The adaptive mind* (2nd ed.). Belmont, CA: Wadsworth Thomson Learning.

Passer, M. W., & Smith, R. E. (2001). *Psychology frontiers and applications.* New York: McGraw-Hill.

Plotnik, R. (2002). *Introduction to psychology* (6th ed.). New York: Wadsworth.

Rathus S. A. (2002). *Psychology in the new millennium* (8th ed.). New York: Harcourt.

Santrock, J. W. (2000). *Psychology* (6th ed.). New York: McGraw-Hill.

Sdorow, L. M. & Rickabaugh, C. A. (2002). *Psychology* (5th ed.). New York: McGraw-Hill.

Sternberg, R. J. (2001). *Psychology in search of the human mind* (3rd ed.). New York: Harcourt.

Wade, C., & Tavris, C. (2000). *Psychology* (6th ed.). Upper Saddle River, NJ: Prentice Hall.

Weiten, W. (2001). *Psychology: Themes and variations* (5th ed.). Belmont, CA: Wadsworth Thomson Learning.

Westen, D. (2001). *Psychology: Brain, behavior, & culture* (3rd ed.). New York: Wiley.

Wood, S. E., & Wood, E. R. G. (2002). *The world of psychology* (4th ed.). Boston: Allyn and Bacon.

Notes

1. A version of this article was presented at the meeting of the American Psychological Association, Honolulu, HI, July 2004.
2. Send correspondence to Jeffrey Nevid, Department of Psychology, St. John's University, Jamaica, NY 11439; e-mail: nevidj@stjohns.edu.

Development of a Departmental Data Archive for Teaching and Research

Thomas M. Brinthaupt and John T. Pennington
Middle Tennessee State University

The analysis of archival data offers many opportunities for students, teachers, and researchers. However, the use of major data archives can be complicated and formidable. We describe the development and use of a departmental data archive that our faculty and students can use. This archive holds demographic as well as self-report behavioral and attitudinal data collected from mass-testing sessions conducted each semester with introductory psychology students. We discuss possible uses and the strengths and weaknesses of such archives for psychology teachers and researchers.

Nearly 30 years ago, Bryant and Wortman (1978) encouraged the development of departmental data archives in psychology. However, since that paper was published, we have found little discussion about these archives and little evidence of their development. This article describes the creation of such an archive and illustrates some of its uses for teachers and researchers.

There are numerous large national and international social science data archives available to psychologists (e.g., Altman et al., 2001; Ryssevik & Musgrave, 2001; Schoenfeldt, 1970). Despite the depth and breadth of existing archives, there are several barriers to their use (Zaitzow & Fields, 1996). As Altman et al. pointed out, creating, sharing, and disseminating archival data is often difficult and expensive for the archivists (see also Sieber, 1991). Also, the size and complexity of some data archives make them difficult to access and therefore less appealing for teachers and students. Unless there is a data set that closely matches one's needs or interests, most students (and probably many faculty) are unable or unwilling to complete the steps necessary to use one of these major archives. In addition, there may be usage costs involved, special programs or technology required, or limitations on who can use the data.

An alternative to the use of national or international data archives is to create and use a local data archive. One advantage such an archive has over larger archives is that the content of the data set is determined by the research-relevant measures and questions provided by local researchers and instructors. There will be a better match with local faculty research interests, and data reflecting the research interests of one's faculty are likely to be of greater interest to students. A second advantage is that control over the archive's operation, content, and usage is local. Local control allows greater flexibility regarding the inclusion of new measures and the development of the archive. Third, a local archive might be a good first experience for students interested in using larger archives. For example, students might learn about some of the advantages of using archival data and, as a result, develop an interest in larger archives. Finally, a local archive's content can be less narrow and restrictive than a larger archive and less dependent on specific funding agency goals.

A Local Data Archive

Our department has used a psychology research pool for more than 15 years. Psychology faculty and students conduct research requiring human participants on a regular basis. The Research Pool contains students from introductory psychology classes who receive descriptions of available research studies. They may choose to participate in several studies in a semester for course credit, or they may submit written critiques of empirical articles.

Like many other colleges and universities, we found that researchers who use the pool often need to recruit a specific sample of students (e.g., married students, those with low self-esteem, smokers) or require baseline information most easily measured before participants arrive for a study. For these and other reasons, our department established a mass-testing procedure and began collecting data from Introductory Psychology students in Spring 1999. At the beginning of each Fall and Spring semester, students may obtain one of their required research credits by participating in one of these 30-min data collection sessions. In these sessions, participants provide demographic information about themselves and complete a variety of personality and attitudinal scales and measures. Whereas we repeat most of the demographic items each semester, many of the other measures change periodically to suit researchers' needs.

At this writing, we have 11 semesters of mass-testing data collected from nearly 5,000 respondents. To organize and facilitate the use of these data, we created a user-friendly mass-testing data archive. This database permits cross-sectional research (e.g., examining differences in alcohol and tobacco use over time), analysis of large personality data subsets, and student use of a "real-world" multivariable data set to learn how to conduct and interpret statistical analyses. Each semester we need approximately 3 weeks to obtain Institutional Review Board (IRB) approval for the data collection, set up and run mass-testing sessions, clean the data, and update the archive. Assistance from graduate students and participating faculty members helps lessen this workload.

Nature of the Archive

Our goal in creating the archive was to allow our faculty and students quick and easy access to the data we have collected in our mass-testing sessions. As with other research situations, researchers who want to use the archive database first must gain IRB approval for their intended project. After researchers receive IRB approval, they submit an archival research request form and receive the requested data files with the condition that they delete the files once they complete their project.

There is an interesting variety of data that researchers can use. Demographic information collected every semester includes respondents' sex, year in school, date of birth, age, race or ethnicity, marital status, and smoking status. Personality, attitudinal, and behavioral measures currently in the archive (many of which appear across multiple semesters) include alcohol use, attachment style, fear of death, homophobia, psychological reactance, racism, religiosity, satisfaction with life, self-esteem, self-talk, and spider phobia. Because students

sometimes repeat Introductory Psychology, they might appear more than once in our archive. Mass-testing participants report whether they have taken Introductory Psychology at our university previously so that researchers may exclude them when conducting cross-sectional analyses or combining data from multiple semesters.

The archive data files are available in SPSS and Excel® formats, with standardized variable labels. We included Excel files because of that program's widespread availability and use among our faculty. We removed all personally identifiable information, created copies of the original survey forms completed by participants, and developed descriptions of each individual scale (or attitude survey) and its scoring instructions. We also cleaned the data by checking for missing or erroneously scanned data points. Because of copyright restrictions, we do not include the scales in the online surveys or descriptions. Researchers can obtain these measures from us. The data files reside on the first author's computer and are not directly accessible by other researchers.

Despite a limited knowledge of Web page design, we developed a Web site for the archive. The *Archive Overview* page provides information on the archive's history, structure, potential uses, and conditions for use. The *Contents* page provides a semester-by-semester listing of data collected, sample sizes, and hyperlinks to the surveys used. In addition, there are "Caution!" notes within the surveys, highlighting items that may be problematic or that respondents might have misinterpreted. On the *Scales Info* page, we list each of the scales and measures used. For each measure, we give a brief (two to three sentence) description, instructions for scoring, and the full citation (including Web links if available). The fourth page is the *Data Request Form*, which researchers must complete to gain access to the data files. On this form, they provide a brief description of the project, the specific data files they are requesting, proof of IRB approval, and the preferred file format.

The archive has the potential to facilitate classroom education and research in several ways. It is clear that the analysis of archival data is an effective way to teach students about research (e.g., Lutsky, 1986). To date, students in graduate-level quantitative methods classes (e.g., factor analysis, multivariate analysis) have used the archive to test the factor structures and validity of measures. In addition, undergraduates in advanced research methods have conducted correlational studies, and graduate students have employed the archive to conduct psychometric and cross-sectional analyses on various attitude scales. Graduate students have also used archive data for their master's theses.

Strengths and Weaknesses of Local Archives

Among the limitations of most data archives are that the data are not precisely what one needs and the original investigators' perspectives on a topic or problem might differ from one's own (Elder, Pavalko, & Clipp, 1993; Zaitzow & Fields, 1996). By their very nature, local archives can offset some of these disadvantages. For example, in response to researchers' and students' needs and interests, we have added or reworded demographic items in our mass-testing data collection. We also have complete knowledge of the data collection materials and methodology. In addition, students might be more interested in data that come from their own school and fellow students.

A disadvantage (which may also apply to larger data archives) is that we include no longitudinal data in our archive, meaning that only cross-sectional analyses are possible. There is also some inconsistency of measures used from semester to semester. Typically, because of time and space limitations in data collection, we include only smaller (10- to 25-item) standardized measures. Whereas our archive includes data only from human participants, there is an interest in and need for animal data archives (e.g., Church, 2002; Kurtzman, Church, & Crystal, 2002).

Why would departmental faculty want to create their own data archive? We believe that an archive would be beneficial if a department already collects mass-testing data (or intends to begin doing so), and faculty can tie the data to their class assignments and projects in subsequent semesters. The effort required to establish a mass-testing data collection procedure would probably be worthwhile only if enough researchers needed selection or baseline data for their projects. If departments have mass-testing, we would certainly recommend the development of a local archive. Once an archive is established, perhaps its best use is by faculty who teach research methods and statistics courses. These instructors might include key items each semester with an eye toward establishing data subsets for undergraduate and graduate student assignments and projects. Few drawbacks exist when using the archive for these purposes.

In summary, we believe a local data archive is easy enough to implement that it can be beneficial for students and teachers. It allows control over and access to data that reflect the interests of local research-active instructors. Because the archive represents the varied interests of our students and faculty, our students can conduct many possible projects, with a wide variety of measures. Researchers who regularly collect mass-testing data might consider the benefits of creating a local archive rather than having those data languish.

References

Altman, M., Andreev, L., Diggory, M., King, G., Sone, A., Verba, S., Kiskis, D., & Krot, M. (2001). A digital library for the dissemination and replication of quantitative social science research: The Virtual Data Center. *Social Science Computer Review, 19,* 458–471.

Bryant, F. B., & Wortman, P. M. (1978). Secondary analysis: The case for data archives. *American Psychologist, 33,* 381–387.

Church, R. M. (2002). The effective use of secondary data. *Learning and Motivation, 33,* 32–45.

Elder, G. H., Jr., Pavalko, E. K., & Clipp, E. C. (1993). *Working with archival data: Studying lives.* Newbury Park, CA: Sage.

Kurtzman, H. S., Church, R. M., & Crystal, J. D. (2002). Data archiving for animal cognition research: Report of an NIMH workshop. *Animal Learning and Behavior, 30,* 405–412.

Lutsky, N. (1986). Undergraduate research experience through the analysis of data sets in psychology courses. *Teaching of Psychology, 13,* 119–122.

Ryssevik, J., & Musgrave, S. (2001). The social science dream machine: Resource discovery, analysis, and delivery on the Web. *Social Science Computer Review, 19,* 163–174.

Schoenfeldt, L. F. (1970). Data archives as resources for research, instruction, and policy planning. *American Psychologist, 25,* 609–616.

Sieber, J. E. (1991). Openness in the social sciences: Sharing data. *Ethics & Behavior, 1,* 69–86.

Zaitzow, B. H., & Fields, C. B. (1996). Using archival data sets. In F. T. L. Leong & J. T. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants* (pp. 251–261). Thousand Oaks, CA: Sage.

Notes

1. Development of the archive and writing of this article were greatly facilitated through the support of a Middle Tennessee State University Faculty Research and Creative Activity Committee Grant awarded to Tom Brinthaupt and John Pennington.
2. We thank Doug Cothern, Dennis Kramer, and Will Langston for their comments on an earlier version of the manuscript.
3. Readers can view the archive site at the following address: http://mtsu32.mtsu.edu:11311/archive/psych_data_archive.htm.
4. Send correspondence to Tom Brinthaupt, P.O. Box X034, Department of Psychology, Middle Tennessee State University, Murfreesboro, TN 37132; e-mail: tbrintha@mtsu.edu.

'Tis Better to Give Than to Receive: An Undergraduate Peer Review Project

Theresa L. White and Brenda J. Kirby
Le Moyne College

Forty-six undergraduates studying research methods had the opportunity to learn about the publication process by mimicking submission to a journal. Students then acted as blind reviewers for their peers' papers and received blind reviews of their own. Based on the reviews, students edited and resubmitted their papers. They indicated that they gained more knowledge through writing the reviews than by receiving them.

Research methods students learn that the scientific method forms the basis for viewing psychology as a science. One component of the scientific method is communication of findings to the professional community, which typically involves publication in a peer-reviewed journal. With few exceptions, the peer review process improves manuscripts submitted for publication. In this manner, peer review both protects and enhances science through constructive criticism and collective knowledge.

This article describes a peer review exercise for use with undergraduate students in an introductory research methods course. The goals of this exercise are to allow students to develop writing abilities, learn about the publication process, and to enhance their learning of research methodology. Students completed a group project and then had a writing opportunity similar to a professional psychologist in the submission of papers for blind peer reviews and revision of their manuscripts. We derived this project from a published peer review activity with very different goals in which graduate students in content courses submitted independently written manuscripts that other students received for blind review (Haaga, 1993).

Giving constructive criticism, the foundation of the peer review process, is a difficult skill to develop. A reviewer can be overly critical and offensive rather than helpful. In an effort to foster reviews that would be professional, impersonal, and offer honest, direct advice, students read about writing fair reviews (e.g., Sternberg, 2002a, 2002b) before beginning the peer review project.

Method

Participants

Forty-six research methods students enrolled in four sections (two sections taught by each author) participated as partial fulfillment of course requirements.

Materials

A seven-item checklist questionnaire assessed student reviewer opinions of a peer article (see Table 1). An additional questionnaire assessed student opinions of the peer review project through Likert ratings (see Table 2).

Procedure

The class completed two studies (one observational, one experimental), with students preparing manuscripts for each using the *Publication Manual of the American Psychological Association* (American Psychological Association [APA], 2001) as though for submission to a journal.

To ensure blind review, the instructors removed the cover pages from two of the three submitted copies of each manuscript and labeled them with preassigned identification numbers that varied between the two class projects to protect anonymity. Each student completed a checklist-style questionnaire (see Table 1) along with a peer review (maximum two pages plus cover sheet) for two randomly assigned manuscripts. Students also wrote comments directly on the reviewed papers. The instructors evaluated each peer review for effectiveness of content, corrected incorrect information, and returned to the reviewers the cover pages with comments and grades. Student authors received three reviews from their initial submission: two from student peer reviewers and one from the course instructor. Students revised and resubmitted their manuscript within a week for an additional grade. At semester's end, students completed anonymous ratings of the peer review process (see questions in Table 2).

Results

Two student reviewers completed the checklist evaluation for each submitted manuscript to allow calculation of a percentage agreement interrater reliability measure. Agreement varied widely by question, ranging from 35.6% to 84.8%, indicating that students differed in their abilities to assess specific aspects of the publication process. The highest agreement level occurred when students assessed whether the manuscript author understood the experiment, and the lowest agreement was in evaluating whether there were errors in APA style (see Table 1). Pearson's correlation coefficients revealed no significant correlation in overall agreements between the two student reviewers on either paper: observation paper, $r(44) = .11$, $p = .47$; experimental paper, $r(43) = .05$, $p = .76$.

Table 1. Percentage of Agreement Between Reviewers on Checklist Questions Evaluating Peer Articles

	Observation Article	Experiment Article
Did the author review the literature well?	65.2	64.4
Did the author seem to understand the question being asked by the experiment?	84.8	75.6
Did the author accurately report the experimental method?	56.5	68.9
Were the results accurately presented?	47.8	57.8
Did the author understand what the results meant and how they might relate to the reviewed literature?	63.0	53.3
Were there errors in APA style?	56.5	35.6

Note. Two class projects were completed, one observational and one experimental. Each percentage agreement appears separately.

Table 2. Evaluation of Peer Review Process

	M	*SD*	*t*(39)
Please rate the draft/receipt of reviews/revision process in terms of how helpful it was for your own paper.	5.55	2.21	1.58
Did you find reviewing someone else's paper useful for improving your own paper?	6.85	2.03	5.76**
Did you find the review process useful for learning APA style?	5.75	2.27	2.09*
The review process was worth 105 points (25 first draft, 15 points for each review, and 50 points for the final draft). I would prefer to just submit one paper worth 105 points.	0.75	1.24	–21.76**
I believe that getting reviews from other students helped me to rewrite my paper.	5.50	2.78	1.13
I believe that I had enough time to write the final draft.	8.05	1.93	9.97**
The peer review process has helped me to understand how publication in psychology works.	7.42	1.92	7.99**

Note. The first three statements were evaluated on a scale ranging from 0 (*worthless—learned nothing from it*) to 10 (*highly educational—learned lots*). The remaining four statements were evaluated on a scale ranging from 0 (*disagree*) to 10 (*agree*). Single-sample *t*s compared actual to neutral responses (5, on a 11-point scale).
*$p < .05$. **$p < .001$.

The evaluations of the accuracy, formatting, and writing quality on a 30-point scale of each student's peer reviews give a rough indication of students' abilities to assess the work of their peers. This ability marginally improved from the first ($M = 26.94$, $SD = 2.20$) to the second ($M = 27.67$, $SD = 1.83$) lab review, $t(45) = 2.00$, $p = .05$.

Forty students evaluated the submission of their papers for peer review. Mean responses to each question appear in Table 2, as do single-sample *t* tests for each question that compare the given responses to neutral responses (5 on a 0- to 10-point scale). It is noteworthy that students were neutral only in their responses to questions dealing with the helpfulness of receiving information from their fellow students. Most students indicated that the peer review project helped them understand the publication process and that they would not rather just submit a single paper instead of participating in the exercise. The ratings showed that students believed that reviewing the papers of their peers was more helpful than the process as a whole, paired $t(39) = -4.22$, $p < .001$. Students also indicated that seeing and reviewing the papers of others improved their papers more than receiving reviews from their peers, paired $t(39) = 3.55$, $p < .001$.

Discussion

Most students indicated that the peer review process helped them understand the publication process and that they preferred it to simply submitting a single paper. Thus, the peer review project provided a writing experience for undergraduate students that taught the advantages of the peer review aspects of publication including the benefit of revising and resubmitting a paper.

The students indicated that reviewing improved their papers more than receiving peer comments, which may reflect the inconsistency of information conveyed by peers. At least initially, students had a poor understanding of a professional APA manuscript, and increasing this understanding was one of the goals of this exercise. The uneven learning across students may account for the low interrater reliability level reported in this study, although it is not vastly inconsistent with published accounts of interrater reliabilities associated with professional reviewers (Marsh & Ball, 1989). A lack of reliance on information contained in peer reviews may also reflect students' keen awareness that the final evaluation would come from the instructor. Just as professional psychologists afford more weight to a journal editor's suggestions than those of individual reviewers, students may attend more to information that they know with certainty will improve their grades.

Although Haaga (1993) described a similar peer review activity for graduate students as a way to promote publication and development of professional identity, the literature rarely contains reports of such exercises with undergraduate students (Dunn, 1996). Nevertheless, peer review activities are possible with students as young as the high school level (Gift & Krasny, 2003), and online applications are available to assist in developing peer review skills (Robinson, 2001).

Despite the increased work load for the instructor of roughly 5 hr per project, this research indicates that undergraduates benefit from peer review. They learn through teaching others about publication and prefer the "revise and resubmit" nature of the process to writing one paper. By mimicking the journal publication process, undergraduates learn about APA style, receive additional exposure to writing prac-

tice and research methodology, and begin to develop an appreciation for the importance of communicating research results as the final stage of the scientific method.

References

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

Dunn, D. S. (1996). Collaborative writing in a statistics and research methods course. *Teaching of Psychology, 23,* 38–40.

Gift, N., & Krasny, M. (2003). The great fossil fiasco: Teaching about peer review. *The American Biology Teacher, 65,* 270–278.

Haaga, D. A. F. (1993). Peer review of term papers in graduate psychology courses. *Teaching of Psychology, 20,* 28–32.

Marsh, H. W., & Ball, S. (1989). The peer review process used to evaluate manuscripts submitted to academic journals: Interjudgmental reliability. *Journal of Experimental Education, 57,* 151–169.

Robinson, R. (2001). Calibrated peer review: An application to increase student reading & writing skills. *The American Biology Teacher, 63,* 474–480.

Sternberg, R. J. (2002a). On civility in reviewing. *APS Observer, 15*(1), 3, 34.

Sternberg, R. J. (2002b). A contradiction in civility? Sternberg's response [Letter to the Editor]. *APS Observer, 15*(2), 6.

Note

Send correspondence to Theresa L. White, Le Moyne College, 1419 Salt Springs Rd., Syracuse, NY 13214; e-mail: whitetl@lemoyne.edu.

A Multicomponent Approach to Teaching Sensitive Topics: Elder Abuse As an Example

Kristi S. Multhaup
Davidson College

Sensitive topics are important, if unpleasant, to discuss in psychology courses. This article describes a multicomponent approach to teaching sensitive topics like elder abuse. This approach also highlights issues of cultural influence on perceptions of (in this case) elder abuse, offers hands-on experience in interpreting data, and encourages students to identify and discuss methodological issues in research.

Even though the National Academy of Sciences (2002) reported that between 1 and 2 million Americans ages 65 or older have been mistreated by a caregiver, the teaching literature rarely addresses elder abuse. A PsycINFO search that included *teach** and *elder abuse* yielded only three entries: One focused on nurses (Galbraith & Zdorkowski, 1984), and two focused on medical school curricula (Alpert, Tonkin, Seeherman, & Holtz, 1998; Jogerst & Hartz, 1999). By contrast, the focus of this article is using a multicomponent approach to incorporating the topic of elder abuse into an *undergraduate* curriculum. Although the motivation for this approach was to explore elder abuse issues, a beneficial by-product is that the components foster discussion about how culture affects perceptions of elder abuse. Thus, this approach adds to a growing literature on teaching diversity-related issues (for a review, see Ocampo et al., 2003).

I teach about elder abuse in Psychology of Aging within the broader topic of relationships for two reasons. First, my texts (e.g., Papalia, Sterns, Feldman, & Camp, 2002; Rybash, Roodin, & Hoyer, 1995) have mentioned abuse of elderly family members. Second, I emphasize the complexity of older adulthood by balancing the negative topic of elder abuse with more positive topics (e.g., friendships, sibling relationships). This latter point fits a course theme that older adulthood, like any other age period, may contain a mix of positive and negative experiences and no one person experiences them all. Our week studying relationships includes two lectures on empirical data and theories related to relationships (e.g., Carstensen, Isaacowitz, & Charles, 1999) and a class period devoted to discussion of elder abuse.

In the first few weeks of class, well before our week on relationships, students complete a modified version of Moon and Williams's (1993) survey about perceptions of elder abuse. The survey contains all 13 scenarios from Moon and Williams's appendix and asks students whether each situation involves abuse; if so, whether the abuse is mild, moderate, or severe and whether they would ask for help if they were in each situation. The scenarios range from a spouse getting very angry and yelling over a small accident to sexual exploitation of a victim of Alzheimer's disease.

When we study relationships, at the end of the first lecture students view two cases from the documentary *A House Divided: Elder Abuse* (Wright, 1989; the Elder Abuse section of American Psychological Association's Adult Development and Aging Division Web pages describes this and other videos, http://apadiv20.phhp.ufl.edu/vidlist.htm). The first case portrays an older Canadian woman whose son physically abused her. The second depicts an older Chinese man in San Francisco whose wife and son locked him in the basement and took his Social Security checks. The video explores community agencies designed to help older adults.

The week's reading includes the text chapter about relationships, the Moon and Williams (1993) article, and a data summary of student responses from each time I have taught Psychology of Aging (e.g., see Table 1). The writing assignment is for students to (a) compare my Davidson student data with Moon and Williams's data and comment on what they see as the implications of their comparison and (b) submit a comment or question for class discussion. The instructions indicate that the *Affirmative Response* and *Yes, Seek Help* columns in the table correspond to the first data columns of Tables 2 and 4, respectively, in the Moon and Williams article and that a .5 in the *mild*, *moderate*, and *severe* columns indicates a response was between two choices (e.g., I coded a response between moderate and severe as .5 in the moderate column and .5 in the severe column). Students submit written responses at the beginning of the class discussion.

Although discussions vary from semester to semester, several points inevitably arise. First, students discuss the issue of who decides what abuse is, particularly in the case where generational and cultural perspectives may differ—should it be well-meaning social servants or the older adult involved? Second, students note that for some of the scenarios they sympathize with the victim's child because they know what it

Table 1. Frequencies of Davidson College Undergraduate Responses to Moon and Williams's (1993) Scenarios

Scenario	Affirmative Response		If Answer Affirmative			Yes, Seek Help[a]
	n	%	Mild	Moderate	Severe	
1	138	99	4.0	50.5	83.5	89
2	37	26	20.5	12.5	4.0	68
3	118	84	58.0	47.0	13.0	55
4	113	81	49.0[b]	46.5	16.5	80
5	134	96[c]	10.0	38.0	86.0	84
6	88	63	55.0[b]	21.0	11.0	54
7	32	22	28.0	4.0	0.0	2
8	134	96	14.0	67.5	52.5	85
9	105	75	57.0	36.0	12.0	80
10	70	50	53.0[b]	14.0	2.0	80
11	128	91[c]	42.0	61.5	24.5	68
12	73	52[c]	64.0	9.0	0.0	18
13	99	71[c]	67.0	29.0	3.0	63

Note. $N = 140$.
[a]*N*s ranged from 113 to 140, so percentages are given. [b]Total of severity ratings does not equal affirmative responses due to missing severity ratings. [c]$N = 139$.

is like to have busy lives that may lead to neglectful behaviors. Students must then explore this unpleasant identification. Finally, students discuss methodological issues such as (a) Moon and Williams's (1993) participants completed the survey orally whereas students completed a written version; (b) the lack of men in Moon and Williams's data in contrast to the Davidson College data; and (c) whether it is more appropriate to compare Davidson data, in which Caucasians are the majority, to Moon and Williams's Caucasian group's column or to their column that combines African American, Korean American, and Caucasian groups' responses. This point often leads to a broader discussion of generalizability of data sets studied earlier in the semester.

In the spring of 2004 I sent students ($N = 31$) a brief e-mail questionnaire requesting their feedback about my approach to teaching about elder abuse. Students replied directly or, if they had concerns about anonymity, e-mailed the questionnaire to a friend's account and replied from there ($n = 1$). All students ($n = 19$; 61% response rate) rated each component as *not helpful, a little helpful, quite helpful,* or *the most helpful aspect for learning about elder abuse*. The modal responses were as follows: for the video (a tie), *the most helpful …* and *quite helpful* (both 42%); for the article, *quite helpful* (68%); for the data comparison, *quite helpful* (53%); and for the discussion, *a little helpful* (58%).

My evaluation is that the video opens students' eyes—some express surprise that elder abuse occurs—with real cases that students describe as powerful. Reading the article and comparing data help students work toward course goals of developing skills in understanding and evaluating data. I value the discussion component because (a) students' insights model for their classmates the course goal of a critical approach to the evaluation of research (see previous discussion examples), (b) students practice speaking in groups, and (c) it is part of a larger campus-wide effort to take opportunities to discuss how culture makes people different, leading to different perspectives. Admittedly it is only one conversation, but it helps students become more aware of how their beliefs are products of their culture and experiences rather than truths that everyone shares. The discussion component is often the most difficult. I have taught this class seven times, and in recent years I have divided the class into small groups to increase participation. I recommend having groups start by selecting the best written question or comment from their group to pass to another group to spark that group's discussion (teachers should ensure variety in the selected questions). When we come back together as a whole class, this method leads to less redundant group reports and more actual discussion because students from various groups have the chance to comment on topics that their group did not consider.

The multicomponent approach offers something for students who prefer visual stimulation and for those who prefer reading. While students explore the important topic of elder abuse, they must face issues of cultural influence on perceptions of elder abuse, consider methodological issues, and practice data interpretation. Teachers who adopt this approach may have students compare their responses to Davidson College students' responses as well as to Moon and Williams's (1993) older women's responses. Differences across student samples reinforce the idea that members of a seemingly similar group (e.g., undergraduates) sometimes hold different perspectives. Lectures could include additional information about cross-cultural comparisons of elder abuse. In a chronologically based life span development course, I recommend teaching about elder abuse in the older adult section (although middle-aged adult children can be abusers, grandchildren and spouses can be as well). Such timing would allow teachers to discuss elder abuse within the complexity of older adults' relationships. Finally, teachers could use the multicomponent approach to discuss other difficult topics (e.g., child abuse, sexual abuse): Students complete a relevant survey, read a published report that used the survey, watch a topically related video, and then compare among the three during class discussion. There is a better chance of engaging a range of students in sensitive issues with a multicomponent approach compared with a single-component approach, and engagement is a necessary step in moving students toward thinking critically about complex topics.

References

Alpert, E. J., Tonkin, A. E., Seeherman, A. M., & Holtz, H. A. (1998). Family violence curricula in U.S. medical schools. *American Journal of Preventive Medicine, 14,* 273–282.

Carstensen, L. L., Isaacowitz, D. M., & Charles, S. T. (1999). Taking time seriously: A theory of socioemotional selectivity. *American Psychologist, 54,* 165–181.

Galbraith, M. W., & Zdorkowski, R. T. (1984). Teaching the investigation of elder abuse. *Journal of Gerontological Nursing, 10,* 21–25.

Jogerst, G., & Hartz, A. J. (1999). Assessment of geriatric training by medical students. *Gerontology and Geriatrics Education, 19,* 19–34.

Moon, A., & Williams, O. (1993). Perceptions of elder abuse and help-seeking patterns among African-American, Caucasian American, and Korean-American elderly women. *The Gerontologist, 33,* 386–395.

National Academy of Sciences. (2002). *Elder mistreatment: Abuse, neglect, and exploitation in an aging America.* Retrieved August 14, 2004, from http://www.nap.edu/books/0309084342/html/

Ocampo, C., Prieto, L. R., Whittlesey, V., Connor, J., Janco-Gidley, J., Mannix, S., & Sare, K. (2003). Diversity research in *Teaching of Psychology*: Summary and recommendations. *Teaching of Psychology, 30,* 5–18.

Papalia, D. E., Sterns, H. L., Feldman, R. D., & Camp, C. J. (2002). *Adult development and aging* (2nd ed). New York: McGraw Hill.

Rybash, J. M., Roodin, P. A., & Hoyer, W. J. (1995). *Adult development and aging* (3rd ed). Chicago: Brown & Benchmark.

Wright, L. (Director). (1989). *A house divided: Elder abuse* [Motion picture]. Canada: National Film Board of Canada. (Available from Filmakers Library, 124 East 40th St., New York, NY 10016).

Notes

1. Thanks to Ruth L. Ault and to three anonymous reviewers for their helpful comments on earlier drafts of this manuscript.

2. Send correspondence to Kristi S. Multhaup, Box 7000, Davidson College, Davidson, NC 28035–7000; e-mail: krmulthaup@davidson.edu.

From Freud to Erikson to Marcia: Concept Maps in Personality Psychology Courses

Kristine Anthis
Southern Connecticut State University

This study examined concept maps produced by personality psychology students as a function of different types of concept map instruction. The results indicated the number of quantitative items in the students' concept maps significantly increased from the pretest to the posttest, as would be expected over the course of an academic semester. More importantly, the type of instructions the students received at the pretest played a significant role in the number of cross links depicted in the maps at the posttest. I discuss the results in terms of the utility of instruction type when using concept maps in social science courses.

The technique of concept mapping for educational purposes involves providing a representation of the psychological structure of knowledge (Wandersee, 1990) not unlike a flow chart, albeit the organization of a concept map is not sequential, but hierarchical. That is, a general concept is listed at the top of the map, with links to subordinate concepts below the general concept. The subordinate concepts may link to one another or link to an additional level of subordinate concepts.

Classroom Uses of Concept Maps

In 1990, the *Journal of Research in Science Teaching* published a special issue devoted to concept mapping. Researchers (Lloyd, 1990; Okebukola & Jegede, 1989) have amply documented the pedagogical utility of concept mapping in natural science courses, yet few studies are available that investigate concept mapping in social science courses.

Jacobs-Lawson and Hershey (2002) conducted an empirical study of concept maps in an introductory psychology class. The results showed that there was an increase from the pretest to the posttest in the number of concepts students generated. Jacobs-Lawson and Hershey noted that the lack of differences between the pretest and the posttest in the number of other markers (such as cross links between concepts) may have been attributable to the general nature of the central concept that they employed (i.e., psychology) and that a more specific central concept (e.g, personality) might have been more likely to reveal differences.

This Study

If students' concept maps of personality were to significantly improve from the beginning to the end of a semester, this pattern would indicate learning, just like in a standard classroom testing procedure. Yet concept maps provide an alternative means of assessing learning with their ability to tap not only the knowledge of a discipline's content (i.e., propositional knowledge), but also the extent to which students clearly recognize the relation within the content area (i.e., structural knowledge). For example, students in undergraduate personality classes should be able to delineate the hierarchical relation between the theories of Freud, Erikson, and Marcia by the end of the semester.

Jacobs-Lawson and Hershey (2002) suggested that students who have never mapped would benefit from first seeing a sample map before being asked to produce a map. Therefore, the purpose of this study was to determine if there was evidence that the method of concept map instruction plays a role in the concept maps students produce. This study also sought to replicate the findings of the Jacobs-Lawson and Hershey study by using a larger sample and a course other than introductory psychology.

Method

Participants

Personality students ($N = 103$) volunteered to participate in this study in exchange for extra credit. One of the questions in the packet of materials that students received asked if they had ever taken a personality course. Eight participants had, so I eliminated their data from further analyses. Of the 95 participants, 79 (83.2%) were women, and 16 (16.8%) were men. In terms of race and ethnicity, 82 (86.3%) were European American, 7 (7.4%) were African American, 1 was

Latino (1.1%), and 5 (5.3%) did not respond to the question. Participants ranged in age from 18 to 55 years, with a mean age of 21.67 ($SD = 5.76$).

Materials

Participants completed a consent form, a demographics questionnaire, and the concept map. I randomly gave students one of four sets of instructions: a list of suggested terms plus instructions and Jacobs-Lawson and Hershey's (2002) "psychology" map as an example map (LIE), a list of suggested terms with the instructions (LI), instructions plus an example concept map (IE), or instructions only (I). The written instructions were a modified version of Jacobs-Lawson and Hershey's instructions. The list of suggested terms included concepts basic to most personality psychology courses, such as *research, psychodynamic, humanistic,* and *philosophical assumptions.*

Procedure

The study used a pretest–posttest design. I taught four sections of an undergraduate personality theories course during the Fall 2002 semester; the text was *Theories of Personality* (Feist & Feist, 2001). I taught all four sections using the same set of notes and PowerPoint® slides. I administered packets of materials that consisted of the consent form, the demographic sheets, and one of the four sets of instructions, along with a blank page to draw the concept map. I used counterbalancing to determine which set of instructions to include in each packet and then randomly distributed them to students at the pretest. At the posttest, students received a packet designated especially for them that had the same instruction set that they received at the pretest.

Coding of Concept Maps

I used the structural scoring method to analyze the concept maps and used six markers based on previous research (McClure, Sonak, & Suen, 1999). The markers included "Number of Valid Concepts," which included any and all terms found on a master list of terms I created; the "Number of Hierarchical Levels" found under the central concept of personality; the "Number of Horizontal Levels," which consisted of the number of concepts listed across concept map under the heading of Personality; the "Number of Valid Links," which counted all of the links to the central concept of personality; the "Number of Cross Links," which represented the number of concepts with links to other concepts; and the "Number of Sublinks," which consisted of examples listed under concepts.

A graduate student in psychology, who had taken a course in personality, coded the maps after I trained her. I also coded a random selection of the maps and compared these scores to those of the graduate student for reliability purposes. The resulting alphas for the six markers were between .72 and .93.

Results

Before examining the relation between instruction type and the students' concept maps, I attempted to identify any change patterns in the six markers from the pretest to the posttest. Paired sample *t* tests showed a significant increase in each of the six markers used to measure the concept maps (see Table 1). That is, the number of valid concepts, the number hierarchical levels, the number of valid links, the number of cross links, the number of sublinks, and then number of horizontal concepts all increased from the beginning to the end of the semester.

The purpose of this study was to determine if instruction type played a role in the concept maps that students produced, so I then conducted six ANCOVAs (using instruction set as the predictor variable, pretest scores as the covariate, and each of the concept map markers as a outcome variable). Out of these six analyses, one was statistically significant. There was a main effect of condition, $F(3, 88) = 2.90, p < .05$, when I used the number of cross links as the outcome variable. Pairwise comparisons of the means revealed that those students in Condition LIE ($M = .73$, $SD = .28$), in Condition LI ($M = .78$, $SD = .31$), and in Condition I ($M = .93$, $SD = .33$) had significantly lower mean numbers of cross links than those students in Condition IE ($M = 1.86$, $SD = .32$).

Discussion

Because those participants in Condition IE had significantly more cross links than all other participants—including those in Condition LIE—there was some evidence that instruction type played a role in the quality of concept maps that the students produced. This pattern of findings is intriguing, as I expected that the more information students had available to them, the better their maps would be. One possible interpretation of why more (information) was not better is that students may have felt obligated to include most of the suggested terms into their maps, which may have precluded them from producing links between these concepts. This is also perhaps why it was the only one out of the six markers (i.e., cross links) that varied as a function of instruction type. Nonetheless, the pattern found confirms Jacobs-Lawson and Hershey's (2002) suggestion that students who have never mapped before benefit from first seeing a sample map (yet not necessarily from receiving a list of suggested terms in addition to the sample map) before producing a map.

Therefore, this study provided empirical evidence that instruction type plays a role in students' ability to demonstrate structural knowledge in a concept map. As a result, the effec-

Table 1. Pretest and Posttest Scores for Six Different Markers

		Pretest		Posttest			
Variable	*N*	*M*	*SD*	*M*	*SD*	*t*	*p*
NVC	94	5.11	2.59	8.09	3.62	–7.25	< .01
NHL	93	3.30	0.75	3.70	1.01	–3.38	< .01
NVL	93	5.73	3.14	8.60	4.29	–5.93	< .01
NCL	93	0.70	1.28	1.04	1.54	–2.00	< .05
NSL	93	3.06	2.22	5.39	3.24	–6.18	< .01
NHC	93	2.15	0.71	2.42	1.15	–2.13	< .05

Note. NVC = number of valid concepts; NHL = number hierarchical levels; NVL = number of valid links; NCL = number of cross links; NSL = number of sublinks; NHC = number of horizontal concepts.

tiveness of assessing changes in students' comprehension of the conceptual relations between theories of personalities—theories that are seemingly disparate to many beginning personality students as Freud, Erikson, and Marcia—appears to be a function of the method used to assess such relations.

References

Feist, J., & Feist, G. J. (2001). *Theories of personality* (5th ed). New York: McGraw-Hill.

Jacobs-Lawson, J. M., & Hershey, D. A. (2002). Concept maps as an assessment tool in psychology courses. *Teaching of Psychology, 29,* 25–29.

Lloyd, C. V. (1990). The elaboration of concepts in three biology textbooks: Facilitating student learning. *Journal of Research in Science Teaching, 27,* 1019–1032.

McClure, J. R., Sonak, B., & Suen, H. K. (1999). Concept map assessment of classroom learning: Reliability, validity and logistic practicality. *Journal of Research in Science Teaching, 36,* 475–492.

Okebukola, P. A., & Jegede, O. J. (1989). Students' anxiety towards and perception of difficulty of some biological concepts under the concept-mapping heuristic. *Research in Science and Technological Education, 7,* 85–92.

Wandersee, J. H. (1990). Concept mapping and the cartography of cognition. *Journal of Research in Science Teaching, 27,* 923–936.

Notes

1. I thank Regina Carini for her assistance with data coding and Patty Kahlbaugh, Larry Brancazio, and the anonymous reviewers for their feedback on an earlier version of this manuscript.

2. Send correspondence and any requests for visual examples or instructions to Kristine Anthis, Department of Psychology, Southern Connecticut State University, New Haven, CT 06515; e-mail: anthisk1@southernct.edu.

Preparing Graduate Students for Academic Positions in Psychology: Suggestions From Job Advertisements

John Clifton and William Buskist
Auburn University

We examined 2 successive years of academic position advertisements in the American Psychological Association's Monitor *and the American Psychological Society's* Observer *for descriptions of materials that search committees requested applicants include in their application packets. Advertisements stressed the inclusion of both research-related and teaching-related materials, although baccalaureate and master's institutions requested teaching-related materials more frequently than did doctoral institutions. We recommend that the "teaching of psychology" course or workshops on teaching offered in many doctoral and master's programs prepare graduate students not only to become effective teachers, but also prepare them to apply for academic positions.*

One important function of graduate education in psychology is to prepare graduate students to join the professoriate. This process includes providing instruction and mentoring in research and teaching and, to a lesser extent, service. The specific form that instruction and mentoring takes depends on the nature of the graduate program; the opportunities for conducting research, teaching, or assisting in the teaching of undergraduate courses; treating clients; serving on committees; and so on. Unfortunately, many graduate students are not well prepared to enter the academic marketplace (e.g., Buskist, Tears, Davis, & Rodrique, 2002; Meyers, Reid, & Quina, 1998; Wimer, Prieto, & Meyers, 2004).

Graduate students (and faculty) applying for academic positions often submit poorly prepared dossiers. For example, Brems, Lampman, and Johnson (1995) found that dossiers for a tenure-track academic position frequently lacked specific information relative to that requested in the position announcement. Likewise, Perlman, Marxen, McFadden, and McCann (1996) found that candidates applying for a teaching position failed to emphasize their teaching experiences in their application materials.

Looking at the application process from a different perspective, Sheehan, McDevitt, and Ross (1998) examined the procedures that search committees used to hire new faculty. In terms of the top three criteria used to assess applicants' dossiers, search committees emphasized, in order, letters of recommendation, the fit between applicants' research interests and departmental needs, and applicants' experience teaching particular courses relative to departmental needs. Landrum and Clump (2004), in a survey of search committee chairs seeking to fill assistant professor positions, found generally that the three most highly rated selection factors centered on applicants' fit with the departmental needs and interests, research experience, and teaching experience. However, these results varied depending on whether the department was housed in a public or private institution and whether the institution involved undergraduate programs only or a combination of undergraduate and graduate programs.

We sought to broaden the understanding of the marketplace in academic psychology and graduate students' preparation for it by examining the printed advertisements for job openings. Our research differed from previous studies in that we examined a wider range of institutional classifications and requested components of applicants' dossiers.

Method

Procedure

We examined 1,259 advertisements for faculty positions in U.S. psychology departments listed in the October and November 2002 and 2003 issues of the American Psychological Association's *Monitor* and the American Psychological Society's *Observer*. We examined these issues because our casual observation suggested that more institutions place job advertisements in these publications during these months than any other 2 consecutive months during the rest of the year. We reviewed each advertisement to determine the type of information that search committees requested applicants to include in their dossiers: cover letter, vita, letters of reference, article reprints, graduate transcripts, statement of research interests, statement of teaching philosophy, copies of course syllabi,

teaching portfolios, and course evaluations. We counted advertisements that appeared in both publications only once and thus also analyzed the contents of duplicated advertisements only once. We categorized advertisements according to the Carnegie classification (www.carnegiefoundation.org/classification) of the institution listing the position and to the academic rank of the position advertised.

Results and Discussion

Two of the advertised positions (< 1%) were at associate institutions, 173 (14%) at baccalaureate, 254 (20%) at master's, 786 (62%) at doctoral, and 23 (2%) at specialty institutions; 21 (2%) of the institutions did not have Carnegie classifications. In terms of faculty rank, we found 810 positions (64%) at the assistant professor level, 145 (12%) at either assistant or associate level, 14 (1%) at associate, 33 (3%) at associate or full, 6 (< 1%) at the full professor level. Forty-one (3%) advertisements listed rank as open, and 207 (16%) did not specify rank—describing the position as a "faculty position."

We analyzed our findings both by professorial rank (Table 1) and by Carnegie classification (Table 2). As Table 1 indicates, the majority of advertisements, regardless of the rank of the position, requested a cover letter, vita, and copies of article reprints in their dossiers and to have letters of reference sent. Fewer advertisements requested applicants to include a description of research interests, statement of teaching philosophy, or copies of graduate transcripts. Very few advertisements requested copies of course syllabi, teaching evaluations, or teaching portfolios. Overall, these data suggest that psychology departments seeking to fill an assistant professor position tended to request both research and teaching related materials slightly more than they did when seeking to fill positions at a higher rank.

Examining the data according to Carnegie classification revealed an interesting difference among baccalaureate, master's, and doctoral institutions (see Table 2). Advertisements from doctoral institutions tended to stress research-related components of the dossier whereas baccalaureate and master's institutions tended to stress teaching-related components. For example, three fourths of doctoral advertisements requested article reprints in contrast to slightly less than half of baccalaureate and master's announcements. Baccalaureate and master's institutions requested statements of applicant's teaching philosophy and copies of teaching evaluations about twice as often as doctoral institutions. These data support recent findings showing that both baccalaureate and master's institutions tend to emphasize the importance of teaching during the interview process (Benson & Buskist, 2005). Interestingly, although academic lore holds that research experience is important only for getting a position at doctoral institutions, such experience also is important for jobs at many baccalaureate and master's institutions.

These results have three important implications for graduate students (as well as new faculty) preparing to enter the academic job market. First, because the statement of teaching philosophy is the most requested teaching-related component of the dossier, regardless of institutional type, applicants should know how to write one. Thus, psychology departments that offer courses or workshops on teaching for their graduate students might include a written assignment of this nature as part of the training regimen (Korn, 2002).

Second, because institutions, particularly baccalaureate and master's schools, request copies of teaching evaluations

Table 1. Percentage of Application Materials Requested by Advertised Rank

Rank	*N*	Vita	Cover Letter	References	Reprints	Graduate Transcripts	Research Interests	Teaching Philosophy	Syllabi	Teaching Portfolio	Evaluations
Assistant	810	95.8	69.4	95.3	68.1	22.6	41.4	29.8	2.0	2.6	22.8
Assistant and associate	145	97.9	68.3	91.7	64.1	11.0	45.5	31.0	1.4	1.4	14.5
Assistant, associate, and full	41	97.6	51.2	90.2	65.9	7.3	46.3	31.7	0.0	0.0	7.3
Associate	14	92.9	78.6	100.0	57.1	14.3	35.7	28.6	0.0	0.0	14.3
Associate and full	33	97.0	81.8	81.8	66.7	9.1	36.4	21.2	0.0	0.0	3.0
Full	6	100.0	50.0	83.3	50.0	16.7	33.3	16.7	0.0	0.0	0.0
Faculty	210	93.3	63.3	90.0	51.9	20.5	32.9	16.2	1.4	0.0	11.0
M		96.4	59.6	90.3	60.5	14.5	38.8	25.0	0.7	0.6	10.4

Table 2. Percentage of Application Materials Requested by Institutional Classification

Carnegie Classification	*n*	Vita	Cover Letter	References	Reprints	Graduate Transcripts	Research Interests	Teaching Philosophy	Syllabi	Teaching Portfolio	Evaluations
Associates	2	100.0	50.0	100.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0
Bachelors	173	93.1	72.3	97.7	49.7	25.4	39.3	43.9	5.2	6.4	30.6
Masters	254	96.5	77.2	94.5	46.5	47.2	37.4	35.0	2.8	2.8	28.7
Doctoral	786	96.1	63.1	92.7	76.0	10.6	42.9	21.9	0.6	0.6	13.6
Specialty	23	95.7	82.6	82.6	43.5	8.7	17.4	13.0	0.0	0.0	0.0
Unspecified	21	95.2	90.5	85.7	14.3	4.8	19.0	19.0	0.0	0.0	9.5
M		96.1	72.6	92.2	38.3	24.5	26.0	30.5	1.4	1.6	13.7

in dossiers, graduate students (a) need to have teaching experience, preferably as a teacher of record, and (b) need to have favorable evaluations (see Benson & Buskist, 2005). The fundamental working assumption of graduate teaching assistant programs is that training graduate students how to teach increases their teaching effectiveness when given the opportunity to teach (Wimer et al., 2004). Unfortunately, not all graduate students who pursue academic careers receive opportunities to teach while in graduate school (Buskist et al., 2002).

Third, although relatively few institutions requested copies of teaching portfolios in dossiers, the teaching portfolio nonetheless remains an effective means for organizing the teaching-related components of the dossier. A secondary, but valuable, advantage of training graduate students to assemble a teaching portfolio is that they will be better prepared to do so when they prepare promotion and tenure dossiers (Seldin, 2004).

Many master's and doctoral psychology departments offer workshops or courses for their graduate students on the teaching of psychology (Buskist, Beins, & Hevern, 2004; Buskist et al., 2002). Although these offerings differ widely across departments, they have the singular aim of improving their graduate students' teaching effectiveness. These platforms tend to focus primarily on classroom teaching. Our data suggest an additional goal: to prepare graduate students to enter the academic workforce. Thus, in addition to receiving instruction regarding the nuances of classroom teaching, such offerings might also require students to write a statement of their teaching philosophy and a teaching portfolio (for different models of such courses, see Buskist et al., 2004).

References

Benson, T. A., & Buskist, W. (2005). Understanding "excellence in teaching" as assessed by psychology faculty search committees. *Teaching of Psychology, 32,* 47–49.

Brems, C., Lampman, C., & Johnson, M. E. (1995). Preparation of applications for academic positions in psychology. *American Psychologist, 50,* 533–537.

Buskist, W., Beins, B. C., & Hevern V. W. (Eds.). (2004). *Preparing the new psychology professoriate: Helping graduate students become competent teachers*. Retrieved December 5, 2004, from Society for the Teaching of Psychology Web site http://teachpsych.lemoyne.edu/teachpsych/pnpp/index.html

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

Korn, J. H. (2002). Beyond tenure: The teaching portfolio for reflection and change. In S. F. Davis & W. Buskist, (Eds.), *The teaching of psychology: Essays in honor of Wilbert J. McKeachie and Charles L. Brewer* (pp. 203–213). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Landrum, R. E., & Clump, M. A. (2004). Departmental search committees and the evaluation of faculty applicants. *Teaching of Psychology, 31,* 12–17.

Meyers, S. A., Reid, P. T., & Quina, K. (1998). Ready or not, here we come: Preparing psychology graduate students for academic careers. *Teaching of Psychology, 25,* 124–126.

Perlman, B., Marxen, J. C., McFadden, S., & McCann, L. I. (1996). Applicants for a faculty position do not emphasize teaching. *Teaching of Psychology, 23,* 103–104.

Seldin, P. (2004). *The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions* (3rd ed.). Bolton, MA: Anker.

Sheehan, E. P., McDevitt, T. M., & Ross, H. C. (1998). Looking for a job as a psychology professor? Factors affecting applicant success. *Teaching of Psychology, 25,* 8–11.

Wimer, D. J., Prieto, L. R., & Meyers, S. A. (2004). To train or not to train: That is the question. In W. Buskist, B. C. Beins, & V. W. Hevern (Eds.), *Preparing the new psychology professoriate: Training graduate students to become competent teachers*. Retrieved December 5, 2004, from Society for the Teaching of Psychology Web site http://teachpsych.lemoyne.edu/teachpsych/pnpp/rtf/pnpp01.rtf

Notes

1. We thank Randolph A. Smith and three anonymous reviewers for their helpful comments on earlier versions of this article.

2. Send correspondence to William Buskist, Psychology Department, Auburn University, Auburn, AL 36830–5214; e-mail: buskiwf@auburn.edu.

An Effective Psychophysiology Laboratory on Cardiovascular Reactivity to Cold Pressor Pain

Stephen Ray Flora and Holly Elizabeth Long
Youngstown State University

A psychophysiology class used a cold pressor pain manipulation, hand immersion in water, for laboratory demonstration of the activation of the sympathetic branch of the autonomic nervous system, as measured by blood pressure increases. This inexpensive laboratory exercise provided a clear demonstration and experience of sympathetic activation requiring few participants and little time. Students rated the laboratory exercise as beneficial, reported that it increased their understanding of psychophysiology and stimulated their interest, recommended its use in future courses, and found the equipment easy to use.

Although students can easily read about the divisions of the nervous system and their basic functions, convincing and affordable demonstration of this functioning is more problematic. Rowland, Kaariainen, and Houtsmuller (2000) reported an exception. They obtained significantly higher blood pressures (BP) on average in undergraduates viewing a suspenseful film compared to when the students viewed a nonsuspenseful film. Presumably the cardiovascular reactivity elicited by viewing the suspenseful film was a function of activation of the sympathetic division of the autonomic nervous system.

Although the students reported the exercise to be highly effective, Rowland et al. (2000) reported systolic measurements only. They combined results from several laboratory sections and required a large *N* to obtain statistical significance. Therefore this manipulation may not always be effective because many psychophysiology, biological psychology, or physiological psychology laboratory classes have small enrollments. In addition, some students may remain skeptical that simply viewing a film could directly affect their physiology; the fact that a large *N* was necessary to obtain signifi-

cance suggests many students experienced small responses or did not experience sympathetic activation.

Compared to "psychological" stress manipulations such as viewing of suspenseful films, threat of electric shock, or experimenter insults that may require participants believe the manipulation before it is effective, the cold pressor stressor is a standard experimental pain induction procedure that reliably produces cardiovascular reactivity (Andreassi, 2000; Edens & Gil, 1995) and high subjective ratings of pain (Flora, Wilkerson, & Flora, 2003). In our laboratory exercise, students used cold pressor pain, hand immersion in ice water, to demonstrate and experience autonomic nervous system-mediated cardiovascular reactivity.

Method

Participants

Eight students (7 women, 1 man) enrolled in a psychophysiology course with a concurrent psychophysiology laboratory component and the male course instructor participated in this exercise.

Stimuli

A portable automatic BP monitor (Omron model HEM–705CP, approximately $40; Omron Healthcar, Inc., Vernon Hills, IL) recorded BP and pulse. A plastic container (approximately 3.5 liters) was on the seat next to the participant. In the cold pressor pain condition, the container contained ice and water with the temperature maintained at approximately 5°C. In the control condition the water had no ice with the temperature at approximately 33°C. A pool/hot tub thermometer recorded water temperature.

Procedure

The exercise took place during the course unit on heart activity, BP, and behavior. A previous course unit covered the nervous system, its major divisions, and their functions.

The laboratory began with a demonstration on how to use the BP monitor. The instructor served as the first participant and as a model for the procedure. The sequence of events was identical for all participants and used an ABA single-subject experimental design.

After a participant had been seated for at least 5 min, the experimenter placed the BP monitor cuff on the participant's arm of choice and recorded BP and pulse. Within 1 min of the first recording, the participant immersed his or her noncuffed hand into the container of ice water up to the wrist. Forty sec later the experimenter took another BP and pulse recording. The experimenter asked participants to keep their hand in the water until completion of the BP recording (approximately 60 sec total), but they could remove their hand if the pain became intolerable. Immediately after completion of the second recording, participants removed and dried the immersed hand. One min after removing the hand, the experimenter recorded a final BP and pulse recording. In a nonsystematic manner each student served as an experimenter and participant for the other classmates. To demonstrate that it was the cold pressor and not merely immersing a person's hand in water that was responsible for any cardiovascular reactivity, during the next 2 weeks each participant, two each class day, repeated the exercise in the identical ABA sequence except the water in the container was maintained at 33°C.

Following completion of the laboratory, the students answered a brief survey about the laboratory. On a 10-point Likert-type scale ranging from 1 (*not at all*) to 10 (*very much*) students answered the following questions: How beneficial was the experience, how much did it increase understanding of psychophysiology, did the laboratory stimulate interest in psychophysiology, how much would the students recommend using it in future labs, were instructions clear and helpful, and how difficult was the equipment to use?

Results

Effects of Stimulation

As expected, hand immersion in warm water produced no cardiovascular reactivity. The cold pressor test produced no systematic pulse rate changes.

Cold pressor cardiovascular reactivity occurred in both diastolic and systolic BP measurements. There was no difference between mean systolic BP before the cold pressor and after it (111.67 vs. 109.33) but there were differences between these measurements and systolic BP during the cold pressor. Mean systolic BP was significantly higher during the cold pressor (127.78) than before it (111.67). A repeated-measured ANOVA indicated significant differences between the three systolic BP measurements, $F(2, 16) = 13.00, p < .001$ (Figure 1). The quadratic contrast was also significant, $F(1, 8) = 20.96, p < .002$, which indicated a curvilinear relationship between the measures. Both a graphic analysis (Figure 1) and

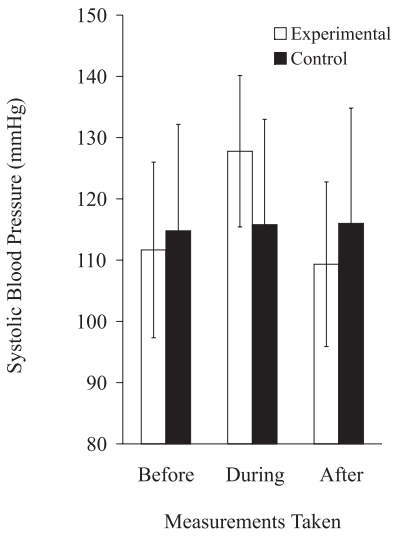

Figure 1. Mean systolic blood pressure before, during, and after hand immersion in either cold (experimental group) or warm water (control condition).

paired *t* tests revealed that there were significant differences before and during the cold pressor, $t(8) = -4.85$, $p < .001$, and during and after the cold pressor, $t(8) = 3.81$, $p < .005$, but not before and after the cold pressor, $t(8) = .67$, $p < .52$.

Likewise, there was no difference between mean diastolic BP before the cold pressor and after it (72.0 vs. 75.0) but there were differences between these measurements and BP during the cold pressor. Mean diastolic BP was lower before the cold pressor (72.00) than during it (84.89), and higher during the cold pressor (84.89) than after the cold pressor (75.00). However, the repeated measures ANOVA quadratic contrast only approached statistical significance, $F(1, 8) = 4.52$, $p < .066$ (Figure 2). The graphic analysis and paired *t* tests show that there was a significant difference before and during the cold pressor, $t(8) = -3.02$, $p < .02$, but not during and after the cold pressor, $t(8) = 1.41$, $p < .20$, or before and after the cold pressor, $t(8) = -0.67$, $p < .52$.

Furthermore, and perhaps more important from both a medical and behavioral perspective, single-participant data revealed a clear functional relation between cold pressor pain and BP. The ABA design revealed a functional relation between the cold pressor and systolic BP for 8 of the 9 participants and unclear results for 1 participant. Likewise, the ABA design revealed a functional relation between the cold pressor and diastolic BP for 7 of the 9 participants and equivocal results for 2 participants, 1 of whom was on medication to control BP.

Students' answers to the survey about the laboratory indicated it was a beneficial experience ($M = 9.00$, $SD = 1.41$), increased understanding of psychophysiology ($M = 8.88$, $SD = 0.99$), and stimulated interest in psychophysiology ($M = 8.38$, $SD = 1.69$). Ratings indicated the instructions were clear and helpful ($M = 8.88$, $SD = 1.46$) and the equipment was easy to use ($M = 1.38$, $SD = 0.52$). The students recommended using the activity in future labs ($M = 9.13$, $SD = 1.73$).

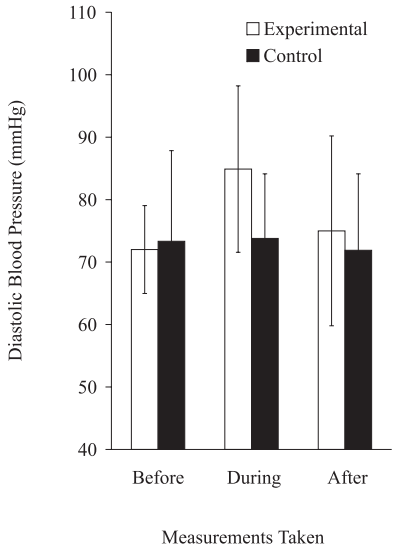

Figure 2. Mean diastolic blood pressure before, during, and after hand immersion in either cold (experimental group) or warm water (control condition).

Discussion

A Direct Demonstration of Sympathetic Division Activation

Coupled with corresponding lecture and readings, this laboratory exercise provides a clear demonstration and experience of the activation of the sympathetic division of the autonomic nervous system. Hand immersion in ice water causes vasoconstriction in that hand to preserve warmth, whereas under normal conditions vasoconstriction results in a cooling of skin temperature (Hassett, 1978).

The sympathetic division of the autonomic NS mediates vascoconstriction. That is, "all of the blood vessels of the body are innervated by the nerve fibers from the SNS [sympathetic nervous system] alone" (Andreassi, 2000, p. 302). Due to linkages of ganglia, once any part of the sympathetic division is activated by either a physical stimulus (e.g., a cold pressor) or psychological stimulus (e.g., a suspenseful film, insults), it acts as an integrated whole and cardiac output and blood pressure increase as part of the sympathetic division's "fight-or-flight" response.

Psychological or Physical Independent Variables?

Although this laboratory demonstrated strong functional effects of cold pressor pain on BP, some students may complain that the demonstration was purely a physiological demonstration having nothing to do with psychological variables. This concern is an erroneous complaint for several reasons. First, the experience of pain is often more a function of psychological variables, including reinforcement contingencies for pain behaviors, than pain is a function of physical damage (e.g., Flora, 2004, pp. 205–207).

In addition, psychological pain, such as social rejection, is registered in the same brain areas as is physical pain, such as a kick in the stomach (Eisenberger, Lieberman, & Williams, 2003). If these findings hold up, then a student interested in the effects of psychological pain as an independent variable (IV) and some measure of behavior or physiological reactivity as a dependent variable (DV) could use physical pain as the experimental IV and reasonably expect that psychological pain would produce similar effects on the DV, and vice versa. That is, physical and psychological pain induction may function as equivalent, or at least as highly similar, eliciting stimuli on the behavior and physiology of organisms. Even if psychological pain is the phenomenon of interest in some experiments, it may be preferable to use physical pain as the IV because it is more reliably and precisely controlled. Cold pressor pain may be preferable because the pain is induced and terminated within minutes and participants may terminate the pain at any time by removing their emerged limb. Psychologically induced pain has the danger of persisting well beyond the experimental session.

Pedagogical Value

Student answers to the postlaboratory exercise indicated that it was pedagogically valuable and practical. Because the necessary equipment is widely available (e.g., at most drug

stores) and relatively inexpensive, and because students recommended using it in the future labs and indicated that the equipment is easy to use, other instructors should be able to replicate this exercise without difficulty. Furthermore, students reacted favorably to the exercise, rating the laboratory as a beneficial experience that increased their understanding of psychophysiology and stimulated their interest in psychophysiology. Based on these results, instructors can be reasonably confident that the exercise will be valuable for their students.

References

Andreassi, J. L. (2000). *Psychophysiology: Human behavior and physiological response* (4th ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Edens, J. L., & Gil, K. M. (1995). Experimental induction of pain: Utility in the study of clinical pain. *Behavior Therapy, 26,* 197–216.

Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003). Does rejection hurt? An *f*MRI study of social exclusion. *Science, 302,* 290–292.

Flora, S. R. (2004). *The power of reinforcement.* Albany: State University of New York Press.

Flora, S. R., Wilkerson, L. R., & Flora, D. B. (2003). Effects of cold pressor pain on human self-control for positive reinforcement. *Psychological Record, 53,* 243–252.

Hassett, J. (1978). *A primer of psychophysiology.* San Francisco: Freeman.

Rowland, D. L., Kaariainen, A., & Houtsmuller, E. J. (2000). Interactions between physiological and affective arousal: A laboratory exercise for psychology. *Teaching of Psychology, 27,* 34–37.

Note

Send correspondence to Stephen Ray Flora, Department of Psychology, Youngstown State University, One University Plaza, Youngstown, OH 44555–0001; e-mail: srflora@ysu.edu.