(GRE) score would prompt a closer examination. With a low GPA, 87% take a closer look at the transcript; with a low GRE score, 77% take a closer look. Thus, graduate directors highly value transcript data, and these data are especially important when a student's GPA or GRE scores are low. Here is how one respondent described the process:

> Withdrawals do not affect a review of the transcripts unless there is a clear pattern of incompletes with poor grades or avoidance of certain kinds of courses (e.g., math). We examine student's [sic] last 60 hours of attempted coursework. I will review transcripts when grade point average and/or GRE scores are low.

What is the Effect of One or More Withdrawals on the Transcript?

Three items addressed the effect of withdrawals on the graduate admissions process. Less than 4% of respondents agreed that one withdrawal hurts an applicant's chances of admission into their graduate program, and 20.3% of respondents agreed that two or more withdrawals hurt a student's chance of entry into their graduate program. Thus, a faculty member's advice to a student contemplating a course withdrawal might be that one withdrawal is probably not going to hurt chances at graduate admissions.

One question asked whether withdrawals in certain classes (e.g., Research Methods, Statistics) are more detrimental than withdrawals in other classes (e.g., Abnormal, Cognition). Over 44% of respondents agreed that withdrawals from certain classes are more detrimental than withdrawals from other classes. Two illustrative comments follow.

> Ws in the last 2 years, particularly in major courses, are negatively viewed. Ws in the first 2 years have no effect. Ws in courses outside the major are not detrimental. We examine transcripts closely. One W would not concern us. More than one would and the student would need to address that concern.

Another respondent wrote:

> One withdrawal would not (ordinarily) raise any concerns. A pattern of consecutive withdrawals in a course, especially a "required" course such as statistics would raise concerns. This would be especially true if the withdrawals were followed by a mediocre grade in the course. A set of withdrawals in the same semester would not cause the same concern as the same number of withdrawals over different semesters because one would assume some cause for problems for that semester. Withdrawals would probably have more (negative) impact on a "typical" student than on [an] outstanding student.

Conclusions

Based on this research, what should I tell students who ask about the potential effect of withdrawals on graduate school aspirations? The following suggestions emerge from this study:

- Graduate admissions committees carefully examine transcripts; typically by a minimum of two faculty members.
- Graduate admissions committees place a high value on transcripts, and either a low GPA or low GRE score may prompt a closer examination of a transcript.
- One withdrawal does not appear to be a problem. Two withdrawals is probably not a problem, except for a minority of schools. For some institutions, withdrawals in particular courses are more detrimental than withdrawals in other courses.

After reading the respondents' open-ended responses, however, it became evident that transcript evaluation is a complicated issue. Faculty examine the patterns of Ws over time, and it might make a difference if there are four Ws in one semester or one W in four consecutive semesters in the same course. Perhaps the type of class also interacts with the effect of a withdrawal—a general education course, psychology requirement, or a upper division elective. Stellar GRE scores or an exceeding high GPA may help to ameliorate the effects of Ws on transcripts. Future researchers interested in this topic might want to capture this complexity in an effort to help explain the impact of transcripts and withdrawals on the graduate admissions process.

References

Daehnert, C., & Carter, J. D. (1987). The prediction of success in a clinical psychology graduate program. *Educational and Psychological Measurement, 47,* 1113–1125.

Daubman, K. A., Williams, V. G., Johnson, D. H., & Crump, D. (1985). Time of withdrawal and academic performance: Implications for withdrawal policies. *Journal of College Student Personnel, 26,* 518–524.

Dunwoody, P. T., & Frank, M. L. (1995). Why students withdraw from classes. *Journal of Psychology, 129,* 553–558.

Habler, B. L., Vodanovich, S. J., & Lowe, R. H. (1990). Using a weighted application blank procedure to predict success in graduate school. *College Student Journal, 24,* 255–260.

Hirschberg, N., & Itkin, S. (1978). Graduate student success in psychology. *American Psychologist, 33,* 1083–1093.

Keith-Spiegel, P., & Wiederman, M. W. (2000). *The complete guide to graduate school admission: Psychology, counseling, and related professions* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Rem, R. J., Oren, E. M., & Childrey, G. (1987). Selection of graduate students in clinical psychology: Use of cutoff scores and interviews. *Professional Psychology: Research and Practice, 18,* 485–488.

Notes

1. I thank three anonymous reviewers and the editor for their generous and helpful suggestions concerning this manuscript.

2. Send correspondence and requests for a complete list of the questions including item means and standard deviations to Eric Landrum, Department of Psychology, Boise State University, 1910 University Drive, Boise, ID 83725–1715; e-mail: elandru@boisestate.edu.

Comparing Bayes's Theorem to Frequency-Based Approaches to Teaching Bayesian Reasoning

John Ruscio
Elizabethtown College

Despite the conceptual simplicity of Bayesian reasoning, people often err when calculating or estimating conditional probability. These mistakes can have significant real-world consequences, and

Bayes's Theorem is a notoriously difficult remedy to teach. Experimenters taught 113 students to use either Bayes's Theorem, 2 × 2 tables, frequency grids, or frequency trees to solve a sample mammogram problem. Immediately following written instruction, group demonstration, and a question-and-answer session, performance on new problems was equivalent across groups. However, when retested 4 weeks later, participants in the Bayes's Theorem group solved fewer problems and demonstrated a poorer understanding of Bayesian reasoning than participants in all other groups. Teaching a frequency-based approach to conditional probability appears to promote learning more effectively than teaching Bayes's Theorem.

Performing Bayesian reasoning can be essential to reach sound conclusions, but most people—including many professionals (Abernathy & Hamm, 1995; Dowie & Elstein, 1988; Eddy, 1982)—are not particularly good at it. Within psychology, Bayes's Theorem is often taught in graduate-level statistics courses but seldom at the undergraduate level, especially outside of introductory statistics courses. The underlying principles are not complex, and teaching Bayesian reasoning can allow instructors to show a large number of students how psychological science can improve real-world decision making through the development of mathematical aids to overcome cognitive limitations and biases.

Although performing Bayesian reasoning is not demanding, effectively teaching it remains challenging. Whereas students trained to use Bayes's Theorem fare poorly, people can learn and implement techniques based on alternative formulations. Recently, researchers have argued that humans evolved the capacity to reason better with frequencies, which stem directly from experience with the natural world, than with probabilities, which are a relatively recent and more abstract human invention (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995). Preliminary data appear to support this assertion (Sedlmeier & Gigerenzer, 2001).

For example, to determine the probability that a woman with positive mammogram results has breast cancer, one can construct a "frequency tree" (see Figure 1) that breaks down a hypothetical sample of women first by the presence or absence of breast cancer and then by test results. It is then simple to establish how well mammography predicts cancer (e.g., 8 out of a total of 107 women with positive tests had cancer, for a probability of .075). A "frequency grid" is conceptually similar to the frequency tree (see Figure 1), as is a 2 × 2 table of test results by criterion status (Ruscio, 2002). These frequency-based techniques may be more meaningful and computationally simpler than Bayes' Theorem. The logic of breaking down samples into subgroups using the base rate and test validity is simple to grasp. The process itself draws attention to the number of true and false positive cases, which renders calculations trivial. Because Bayes's Theorem integrates probabilities in a way that is neither transparent nor intuitive, it can be difficult to apply and there is little to be gleaned from an incomplete attempt.

Sedlmeier and Gigerenzer (2001) achieved impressive results with frequency trees and grids, but it is unclear how readily their intensive training can be implemented. This experiment tested the efficacy of four techniques—Bayes's Theorem, 2 × 2 table, frequency tree, and frequency grid—in more ecologically valid ways.

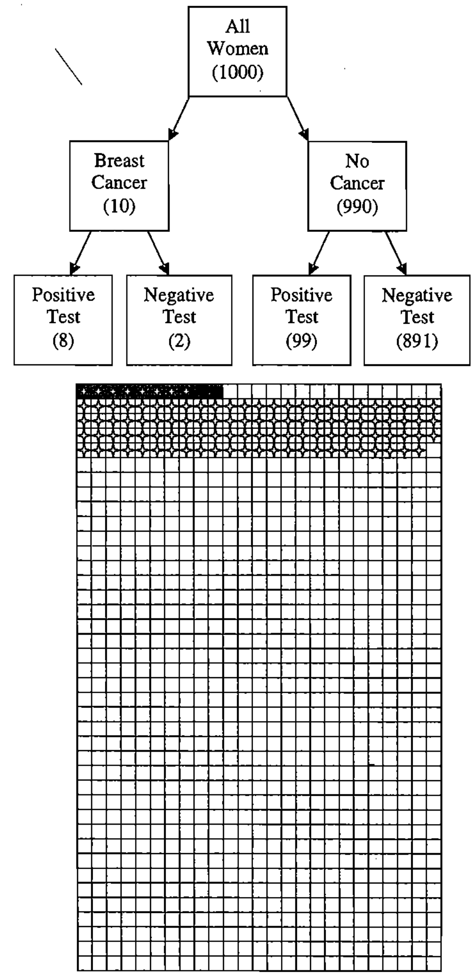

Figure 1. Sample frequency tree and frequency grid solutions to a mammography problem. In the frequency tree, the hypothetical sample of 1,000 women are subdivided first by the 1% base rate of breast cancer, then according to the 80% true positive and 10% false positive rates of mammogram test results. In the frequency grid, the same sample is represented by a grid of 1,000 boxes in which 1% are shaded to represent breast cancer and positive test results are indicated by plus signs.

Method

In exchange for course credit, 113 General Psychology students took part in this experiment.[1] Two trained experimenters conducted 14 sessions that ranged from $n = 4$ to $n = 12$. Both experimenters conducted one or two sessions within each experimental condition.

Prior to instruction, participants completed a pretest problem that contained a base rate (1% of women undergoing mammography have breast cancer), a true positive rate (80% of women with breast cancer test positive on the mammogram), and a false positive rate (10% of women without breast cancer test positive on the mammogram) and posed a question requiring the calculation of positive predictive power (the probability that a woman who tests positive on the mammogram has breast cancer).

[1]Demographic data were not collected. Given the nature of the student body at Elizabethtown College and typical enrollment patt

After completing the pretest, participants read a single-page explanation of one method for solving Bayesian problems that illustrated the technique by correctly solving the pretest mammography problem. In step-by-step form, participants read how to locate the relevant information and apply either Bayes's Theorem or a frequency-based method. Once all participants had read the instructions, the experimenter demonstrated the technique by working through the same steps outlined in the instructions on a large dry-erase board and answered any questions. When all participants felt prepared to work on new problems of this type, they took the first test, which contained three new Bayesian problems of the same form as the pretest problem. Most participants completed the first session in a total of 30 to 45 min.

Approximately 4 weeks later, 106 participants (94%) returned to take a second test with three new problems. Finally, participants completed a questionnaire that contained one additional problem to be solved without using a calculator or writing down any calculations and three questions that probed for a true understanding of Bayesian reasoning, as opposed to the mere ability to apply a mechanical technique that successfully solves problems. Participants completed the second session in about 15 to 20 min.

Results

In keeping with previous studies, "correct" responses had to lie within 5 percentage points of the true probability. Ten participants correctly solved the pretest problem; all subsequent analyses excluded these participants' data, which were evenly distributed across experimental conditions.

The primary analysis tested for performance differences across experimental conditions via a 4 (experimental conditions: Bayes's Theorem, 2×2 table, frequency tree, and frequency grid) $\times$ 2 (time: first vs. second session) mixed-model ANOVA. Scores on each test ranged from 0 (*no problems solved correctly*) to 1 (*all three problems solved correctly*).[2] There was no difference across conditions, $F(3, 92) = 1.42, p = .24, \eta^2 = .04$, but a large drop in performance over time, $F(1, 92) = 22.50, p < .01, \eta^2 = .20$. More important, whereas performance on the first test was comparable across instructional methods, performance deteriorated more markedly for those taught Bayes's Theorem than for those taught a frequency-based technique, $F(3, 92) = 4.41, p < .01, \eta^2 = .13$ (see Figure 2).

To test the hypothesis that participants would learn and retain the frequency-based techniques better than Bayes's Theorem, planned contrasts were conducted on performance at each time. The first contrast compared the Bayes's Theorem group to all other groups; the second compared the 2×2 table group to the frequency grid and frequency tree groups; the third compared the frequency tree group to the frequency grid group. On the first test, none of these contrasts was significant, $t(99) < 1.65, p > .10, \eta^2 < .03$ for each. On the second test, participants in the Bayes's Theorem group ($M = .06$, $SD = .13$) scored significantly worse than did those in the three frequency-based groups ($M = .32$, $SD = .39$), $t(92) = 3.47, p < .01, \eta^2 = .12$. Neither of the other two contrasts was significant, $t(92) < .59, p > .56, \eta^2 < .01$ for both.

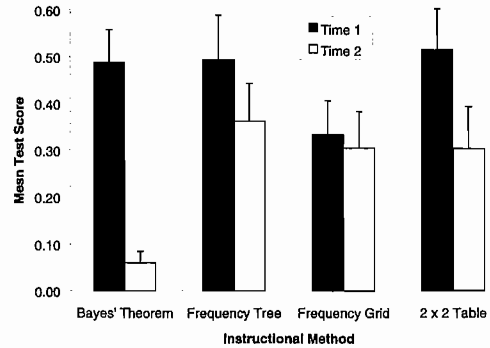

Figure 2. Mean test scores for both sessions and all four experimental conditions. Error bars represent +1 *SE* of the mean.

A final analysis evaluated scores on the final questionnaire using the same three planned contrasts.[3] Again, participants in the Bayes's Theorem group ($M = .46$, $SD = .22$) performed significantly worse than did those in the three frequency-based groups ($M = .58$, $SD = .25$), $t(90) = 2.06, p = .04, \eta^2 = .05$, and neither of the other two contrasts was significant, $t(90) < .99, p > .32, \eta^2 < .02$ for both.

Discussion

The performance of participants taught one of the three frequency-based methods dropped only from 45% immediately following instruction to 32% 4 weeks later, whereas the performance of participants taught Bayes's Theorem plummeted from 49% to 6%. Questions testing participants' deeper understanding of the underlying logic of Bayesian reasoning revealed similar results. Neither performance on Bayesian problems nor an understanding of the logic of Bayesian reasoning differed significantly across the three frequency-based methods.

Whereas this pattern of results mirrors that obtained by Sedlmeier and Gigerenzer (2001), the absolute level of performance was considerably lower here. For example, in Sedlmeier and Gigerenzer's first experiment, participants in the Bayes's Theorem group solved about 60% of problems correctly immediately following training, and this figure dropped to about 20% over 5 weeks; in contrast, participants taught the frequency tree and grid techniques solved between 75% and 90% of their problems correctly, and performance did not drop over time. There are many possible reasons for the achievement gap across studies. Instruction time was briefer in this study (approximately 20 min) than in theirs (up to 2 hr). As compared to the written instructions, group demonstration, and question-and-answer session

erns in General Psychology, it is safe to presume that participants were virtually all between 18 and 22 years of age and that approximately 75% were women.

[2]Scores on the first and second tests were scaled from 0 to 1 by dividing the total number of correct responses by 3 so that group means would correspond to the proportion of problems correctly solved. Scores on the final questionnaire, which ranged from 0 to 4, were divided by 4 so that group means would correspond to the proportion of correct responses.

[3]Two participants' data were missing on the final questionnaire, so *df* for analyses of the final problem and follow-up analyses reflect slightly lower *N*s than in previous analyses.

based on a single sample problem provided in this study, their tutorial consisted of a computer program that provided step-by-step instructions and performance feedback on a number of problems. Participants in this study used only a blank sheet of paper and a calculator to generate a response, whereas the same computer program that had provided instruction prompted Sedlmeier and Gigerenzer's participants for responses and automatically performed calculations.

Because of these differences, the results of this experiment may be more generalizable to classroom settings and therefore provide more realistic estimates of what instructors can achieve with minimal instruction. One might view these results as the floor and Sedlmeier and Gigerenzer's (2001) results as the ceiling for what one can hope to attain in teaching Bayesian reasoning. To the extent that they devote more instructional time (i.e., greater than the 10 to 15 min of instruction in this experiment) and resources (i.e., more than one problem, phrasing, or technique) to teaching this material, instructors have every reason to be optimistic that a greater proportion of their students will truly understand, correctly apply, and successfully retain the principles of Bayesian reasoning. Future research could investigate potentially influential instructional factors (e.g., instruction time, presentation format, feedback).

Regardless of how much energy one commits to teaching Bayesian reasoning, it seems clear that a frequency-based approach is likely to promote learning better than Bayes's Theorem. However, at present there appears to be little reason to prefer one frequency-based method over another. Whether there are practically significant differences between these approaches is another topic for future research.

References

Abernathy, C. M., & Hamm, R. M. (1995). *Surgical intuition: What it is and how to get it.* Philadelphia: Hanley & Belfus.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition, 58,* 1–73.

Dowie, J., & Elstein, A. (Eds.). (1988). *Professional judgment: A reader in clinical decision making.* Cambridge, England: Cambridge University Press.

Eddy, D. M. (1982). Probabilistic reasoning in clinical medicine: Problems and opportunities. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 249–267). New York: Cambridge University Press.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102,* 684–704.

Ruscio, J. (2002). *Clear thinking with psychology: Separating sense from nonsense.* Pacific Grove, CA: Wadsworth.

Sedlmeier, P., & Gigerenzer, G. (2001). Teaching Bayesian reasoning in less than two hours. *Journal of Experimental Psychology: General, 130,* 380–400.

Notes

1. I thank the two experimenters, Jennifer Mills and Donna Ondik, for their invaluable assistance in collecting, entering, and checking these data. I am also grateful to Peter Sedlmeier for providing a large set of Bayesian problems from which the ones in this experiment were selected.

2. Send correspondence and requests for the full text of all experimental materials to John Ruscio, Department of Psychology, Elizabethtown College, One Alpha Drive, Elizabethtown, PA 17022; e-mail: rusciojp@etown.edu.

Teaching Successful Grant Writing to Psychology Graduate Students

Thomas Eissenberg

Department of Psychology and Institute for Drug and Alcohol Studies
Virginia Commonwealth University

Grant writing is a part of many psychology careers, but psychology departments seldom offer grant-writing courses. In 2000, Virginia Commonwealth University's (VCU's) Department of Psychology began offering a course focusing on the National Institutes of Health (NIH) National Research Service Award (F31). In 2 years, 16 students completed the course, and 6 submitted F31 proposals. Mean VCU priority scores were significantly better than the mean for all competing F31s submitted to the same institutes. NIH funded all 6 VCU proposals. This article describes the course and its key aspects. Grant-writing courses offer short-term rewards and potential long-term benefits to students, faculty, and psychology departments.

Grant funding is important for psychological research, researchers, and clinicians. For example, in 2001, total federal grant funding for psychological research was approximately $751 million (National Science Foundation, 2001). Psychologists' success at writing grants can influence hiring, tenure, and promotion decisions, but the online catalogs of the top 24 psychology graduate programs (as ranked by *U.S. News & World Report*, 2002) showed no courses devoted to this topic. Other areas give more attention to training in grant writing. For example, the online catalog of Saint Louis University's Department of Biochemistry and Molecular Biology reveals that PhD students are *required* to take a grant-writing course, and such a course is available to psychiatry postdoctoral fellows at the University of Pittsburgh (Reynolds et al., 1998). Thus, although psychology graduate students receive little grant writing training, students in other areas learn these necessary skills.

A formal course that teaches psychology graduate students how to write grants may be a vital but missing part of current curricula. A formal course might encourage students to apply for funding that will aid their graduate and professional careers, while increasing funding for psychological research in general. It would also allow instructors to receive acknowledgment for their pedagogic work and contribute to a higher and more standardized level of instruction.

This article describes a grant-writing course developed for graduate students in Virginia Commonwealth University's (VCU's) Department of Psychology. The course has two main goals: (a) teaching students how to write effective grant proposals and (b) increasing the amount of external funding for VCU's graduate students.