The Introductory Class in Higher Education: Some Old Problems and New Considerations

David R. Schmitt
University of Washington

The traditional text and lecture format has remained the norm in introductory college courses despite evidence, such as the personalized system of instruction, that other practices could improve what is learned and retained. The growth of distance learning provides new opportunities to implement effective teaching practices. Unfortunately, an adequate comparison of various teaching practices is not possible (even with regard to the teaching of behavior analysis). Few practices have been assessed with respect to course-end effects (final exam performance), and longer term effects remain almost wholly unexplored. Studies of the retention of academic materials, the practice required for mastery or fluency, and the relation between verbal repertoires and correspondent everyday behaviors suggest course outcomes more modest than those hoped for earlier. Suggestions for changing current practices have little credibility until the size of the gap between present and possible learning outcomes is much better known, and the personal, social, and economic consequences of bridging that gap (or not) are assessed.

Key words: teaching, higher education, evaluation, retention, generalization, personalized system of instruction

Historically, the goal of most introductory liberal arts courses in higher education has been the development of specific verbal repertoires by means of the time-honored combination of texts and lectures. Essays, papers, and objective tests measure the repertoire produced. Although some classes incorporate small-group learning activities or laboratory exercises, the basic structure by which information is presented has changed little over the past century. Many of the current efforts in distance learning replicate this basic structure, with computer screens replacing the classroom as the lecture component. Although most lecture material could be made available to students in written form (and often is, indirectly, through lecture notes), having students personally record information for future testing is widely believed to be important (and is considered one of

the “alternative” learning styles). The importance of interaction between an instructor and students in the classroom is also thought to be important, even though many students are in classes too large for any interaction if more than a small fraction choose to participate.

The importance of the traditional lecture format was challenged three decades ago by Keller’s (1968) personalized system of instruction (PSI), in which all the important materials were presented in writing. Its various components, including frequent tests, demonstration of mastery before test credit is given (i.e., being able to take multiple versions of tests), self-pacing of course progress, and the use of student proctors, were investigated in a number of studies. PSI is usually more effective than traditional lecture methods in promoting superior final examination performance (see Kulik, Kulik, & Bangert-Drowns, 1990; Sherman, Ruskin, & Semb, 1982). In the same year that Keller’s article appeared, Dubin and Taveggia (1968) published a survey of the extensive literature on traditional teaching methods in higher education.

Thanks to Keith Miller for comments on an earlier draft.

Address correspondence and reprint requests to David R. Schmitt, Department of Sociology, Box 353340, University of Washington, Seattle, Washington 98195 (E-mail: schmitt@u.washington.edu).

In what they termed the *teaching-learning paradox*, courses using varieties of face-to-face instruction and independent study did not differ significantly in students' final exam performance. A suggested cause was the most pervasive commonality: the dependence on textbooks and other reading materials for most of the students' exposure to the topics covered. In the 1970s PSI and related behaviorally based instructional methods were used in an array of courses across the country, and were widely seen as having a foothold in higher education (Johnston, 1974). As Fraley and Vargas (1974) put it, "With the national commitment to mass education and the increasing demand that it be met with higher standards, there is an increasing concern with how to produce new behavior in learners rather than simply 'teach at them,' hoping that desirable effects may result" (p. 482). Malott (1984) drew together studies and his first-hand experience in offering 55 ways in which behavioral principles could be used to improve the quality of courses in higher education. These suggestions concerned testing arrangements, motivation, attendance, study objectives, cumulative mastery, labs and practice, writing assignments, curricula, and prevention of drop-outs. Thus, almost two decades ago, the future of the traditional lecture format seemed in doubt as the dominant form of instruction. More effective course arrangements were at hand.

But the traditional format was not supplanted by a behaviorally inspired one. The increasing use of PSI in the 1970s was followed by a decline in the 1980s. PSI has a fraction of its base of 20 years ago, and most behavior analysts do not use it. Observers of this decline have pointed to the failure to teach behavioral approaches to faculty, the greater effort and time required to prepare PSI-related materials, professors' preference for the role of lecturer, and the hostility of administrators and other faculty to PSI (Axelrod, 1992; Buskist, Cush, & Grandpre, 1991).

These initial course innovations concerned the application of basic behavioral principles to the implementation of instructional objectives and course management. More recently, attention has focused on the rate at which students make various academic responses, based on evidence that behaviors that can be performed at high rates (i.e., fluently) are retained for longer periods and can be more readily applied elsewhere (for a review, see Binder, 1996). The case for adoption of fluency-based procedures is especially strong in elementary and high school, where repertoires in reading, writing, and math are requisite for future academic and occupational success and are also highly consequential in everyday life. In higher education, however, the situation is often very different. Disciplines differ in the degree to which repertoires taught in introductory classes are indeed basic and likely to be supported later. In fields such as math, physics, chemistry, and biology, there is considerable consensus regarding basic repertoires and the normal sequence of courses from elementary to advanced. In most disciplines in the social sciences and humanities, however, there is much less consensus on content. Many classes have no sequel, course prerequisites are often more nominal than meaningful, and students have infrequent contact with course content outside the classroom.

In PSI the term *mastery* has been used to distinguish testing that gives students multiple opportunities to reach some criterion from the usual procedure that gives them but one opportunity. Although mastery is a relative term, it is given a much more rigorous definition in other areas that consider well-learned behavior. For example, in the study of expert performance (Ericsson, 1996; Ericsson & Charness, 1994; Ericsson, Krampe, & Tesch-Romer, 1993), mastery of complex activities in sports, science, and the arts typically involves guided (i.e., "deliberate") practice extending over at least a decade. In the Morningside instruc-

tional model developed by Johnson and Layng (1992, 1994), four steps on a performance continuum culminate in what they term *true mastery*. The first step is establishing (i.e., training whose aim is an accurate response). The second and third are remembering and enduring (i.e., building fluency so that a response becomes automatic and persists over time). The fourth is applying (i.e., being able to use a response in new environments or as part of more complex skills). Use of the term *mastery* in these contexts bears little resemblance to its use in PSI. For the master in PSI, mastery pushes a bit beyond the rudiments of establishing a response, but certainly does not ensure that it will be remembered, endure, or can be applied.

In most universities, introductory classes in the various disciplines serve two purposes. For the majority of students, they constitute one small part of a liberal education, a brief glimpse into a discipline or specialty. For a minority, they are the first step in a major or minor. The minority are likely to experience at least some direct repetition of material in later courses, but the majority will not. What is the likely fate of weakly established verbal repertoires that are not strengthened and maintained in additional classes? As Malott (1984) has noted, verbal repertoires seem to be easily disrupted by the acquisition of subsequent, competing verbal repertoires. Although some repertoires may be consequential in nonacademic activities (e.g., in jobs, interest groups, or hobbies), most are not. Where they are consequential, the delay in use is often substantial. Semb and Ellis (1994) reviewed studies that bear on long-term retention of what is taught in classrooms. They suggest that retention is more substantial than the popular belief that little is retained more than a week or two after the final exam. Decreases in retention for academic materials were also less severe than those observed in traditional laboratory studies. Retention was affected by level of original learning and by

type of task, with retention greater for recognition than for recall tasks. For example, retention losses after periods ranging between 27 and 52 weeks were approximately 18% for recognition tasks and 45% for recall tasks. Low content coherence and low prior knowledge were associated with retention loss. Considerable evidence from various contexts suggests that a logarithmic function describes the amount remembered plotted over time (Rubin & Wenzel, 1996). The practical consequences of declines in retention depend on the interrelatedness of the material learned. For example, failing to retain a portion of the historical names and dates learned previously may have little consequence for the usefulness of the material retained, whereas failing to recall parts of formulas or complex procedures usually leaves them with little value, unless the source can be consulted again.

These concerns suggest that we should examine much more carefully what impact our courses have on what is retained in both the short and long term. Courses taught through readings and lectures provide a limited number of hours to build or strengthen behavioral repertoires. That time can be used to establish broad but weakly established repertoires or narrower but better established ones.

Some of these issues can be illustrated using the teaching of introductory behavior analysis. Basic texts in behavior analysis (e.g., Grant & Evans, 1994; Malott, Whaley, & Malott, 1997; Martin & Pear, 1996; Miller, 1997) evidence a high degree of consensus in the core material taught. That core includes single-subject methodology and a set of interrelated concepts and procedures that includes reinforcement, punishment, stimulus control, generalization, shaping, conditioning, and schedules. Introductory texts also rely heavily on both everyday and clinical illustrations, and attempt to convey a sense that the principles can be used to understand the causes of various normal and abnormal behaviors, and often

to modify behavior. In fact a common class assignment is to have students select personally relevant behaviors for measurement and modification. Properly defined and illustrated terms and procedures are given high priority in what is learned, a task that is complicated by the frequent disjuncture between technical and everyday meanings. Emphasis is typically placed on recall of terms and principles (via short answers, essays, or flash card drills) in preference to their recognition (via true-false or multiple-choice questions).

Given the early studies of PSI and the numerous interventions in elementary classrooms, it is perhaps surprising that we know very little of what is actually retained from such a course experience (i.e., the lasting effects of the various instructional procedures). Informally, many instructors express confidence that their use of drills, diagrams, everyday examples, discussions, and exercises in which students generate examples and nonexamples of concepts and procedures help them to make at least rudimentary behavioral analyses beyond the classroom and put some procedures into everyday practice. Some generalization is presumed (or at least hoped for), despite the fact that the repertoire is largely verbal and is given very limited practice, but there is very good reason for skepticism even without a loss in retention. Although it is possible to find straightforward examples of behavioral procedures in everyday life, actual situations of interest such as jobs and relationships usually have complicating features, including complex histories, arrays of alternative behaviors, and consequence schedules that are difficult or impossible to observe or even describe. The challenges for understanding are daunting for well-trained behavior analysts. Further, there is little evidence that descriptions of procedures learned in one context (i.e., a classroom) can be accurately enacted (or even recalled) in other, very differ-

ent contexts (Stokes & Baer, 1977; Stokes & Osnes, 1989).

Over time, problems should become more severe as retention losses occur. Successful interventions for changing behavior typically have interrelated procedures, and failure to implement part of the system usually dooms the intervention to failure. In short, no empirical evidence that I know of supports the notion that the procedures common in elementary courses in behavior analysis lead to the adequate implementation of behavioral techniques, and what we know from research on mastery and generalization makes that prospect very remote at best. The most likely product in the months after the class is completed is probably a rapidly dwindling technical vocabulary and procedures recollected, along with the lingering conviction that a behavioral approach is valid in understanding a variety of behaviors. But research has not addressed this important issue. Because behavior analysis often uses dramatic clinical cases as examples (Malott et al., 1997, in particular), recollections may be strongly case related as well. Note that the implicit or explicit expectations of application generated by contemporary courses probably far exceed those from the courses on learning that were offered a generation earlier. There the aim was to illustrate or implement only the most elementary learning principles in laboratory exercises with a single rat or pigeon.

What would be required for greater retention of course material? There are several possibilities. The first is the creation of another environment in which the weakly established verbal responses (e.g., behavioral analyses) are strengthened and maintained. This most commonly occurs when students take additional courses in the area (Malott, 1984). Malott, Vunovich, Boettcher, and Groeger (1995) describe a comprehensive university program addressing this issue, and Ellis and Glenn (1995) discuss such training within the profession. It can also occur

in extracurricular groups such as *Walden Two*—inspired fellowships and communities or classes in parent effectiveness, but few students have such experiences. Another possibility is to make more use of fluency-building exercises, a change that would require far more repetition than current practices provide. For a complex scientific system such as behavior analysis, this would require not just learning definitions but also the rapid recognition of the elements of contingencies in everyday settings. Assuming that this would take place first with illustrations given in class, the problem of generalization to actual interaction would remain. This is a highly ambitious agenda. The work on fluency suggests that it is far too large for a single academic term, and more resembles a graduate training program in behavior analysis (see Malott et al., 1995).

A related possibility is to recognize that although a large number of points cannot be mastered in an academic term, a very small number might be. For example, what half dozen facts or ideas should students be able to accurately recite months or years later? These could be drilled and illustrated much more frequently over the course of a term (in addition to material taught more traditionally), with the likely result that they would be retained. Parenthetically, discussions with colleagues on this issue suggest surprisingly little consensus as to what these points might be or how they might be applied if mastered. A discussion of this issue in the discipline might be useful.

A more promising possibility might be to use a strategy that has been widely employed elsewhere. In many areas of expertise, one does not need to (or cannot) remember all the important points and principles; instead, a reference is consulted. Mathematicians use references for complex equations and formulas. Most households have medical and child-rearing references and “how-to” books for home repairs and improvements. Travelers who can’t

speak the local language well carry phrase books. It is certainly possible, and may be highly desirable, in behavior analysis to pay less attention to memorization and fluency in the limited class time available and to pay more attention to using references easily and effectively. Given a problem, how do you find the material that could provide a way of conceptualizing it behaviorally, and possibly suggesting a solution? Clark’s (1995) book on parent–child relations from a behavioral perspective was written to serve such a function for lay audiences. Although the basic texts noted earlier are organized to teach basic principles, they are also rich in examples and might be used as guides if students keep and use them after the term ended. Child rearing, spousal relations, and worker–supervisor relations are among the everyday areas frequently encountered. The reference approach could lead to a continued exposure to behavioral principles that are accurately portrayed.

There are also possible long-term outcomes from courses that remain unexplored. Most apparent, exposure to a behavioral perspective could lead to reading related materials in later life, urging others to learn more about the area, or seeking behaviorally based treatments. Although untested, existing teaching procedures probably have other more general benefits. Mastery testing requires that students pay greater attention to the written material and follow guidelines. Concept learning, which focuses on examples and nonexamples (Miller & Weaver, 1975), requires more careful attention to distinguishing details. Although these would not be insignificant consequences, long-term effects from any given course would probably be small, as has been the case in attempts to teach writing skills (e.g., Fowles, 1993).

After three decades, we have a wealth of ideas regarding beneficial course procedures, but we are far from having documented accounts of the nature of most of their effects. The most frequently researched ones concern

features of PSI (e.g., Kulik et al., 1990) and testing procedures that facilitate learning (e.g., Bangert-Drowns, Kulik, & Kulik, 1991; Dempster, 1992), but the effects are rarely measured beyond the end of the class. In addition, numerous teaching strategies, techniques, exercises, and demonstrations are provided by elementary texts, instructor's manuals, and sections of *The Behavior Analyst*. Yet, remarkably, few (if any) are supported by objective evidence of their effects on outcomes, not even the most easily measured one: final exam performance. Evidence is almost wholly anecdotal. If behavior analysts were evaluating these recommendations in their roles as scientists and reviewers, they would be treated as proposals awaiting investigation, not as evidence warranting implementation. As Malott stated in his final suggestion in 1984, "Be guided by what hard data we have, and less by debatable possibilities that may reflect more the arguer's biases than the reality of the situation" (p. 245). This has been a hard rule for university instructors to follow, but shouldn't we be increasingly less tolerant of its violation?

The need to understand the immediate and enduring effects of a course's components is more critical now, as distance learning emerges as a powerful competitor to in-class instruction. Although many opponents of distance learning claim beneficial effects of face-to-face contact, there is little supporting evidence beyond the anecdotal. If research finds face-to-face activities to be valuable, it may be possible to capture essential features in computer-mediated exchanges among students. Control and measurement problems are daunting in any course setting, but they should not excuse our failure to study the effects of one of the most ubiquitous, if not intensive, treatments in everyday life.

Although analyses of course features will certainly have a scientific payoff, there is reason to wonder about the practical effect even if certain features are shown to have beneficial effects on

student performance. Putting aside collegial and administrative pressures favoring the status quo, the current practice of infrequent tests over written material (texts and notes) appears to suffice—to produce satisfactory but not necessarily optimal outcomes (Simon, 1956). Internationally, this system is seen as a model, not a failure, and there is little complaint that those who graduate are poorly trained (unlike many high school graduates). And the system is probably the most economical one to administer for classroom instruction or distance learning. Given the accumulating evidence that fluent, applicable repertoires are time consuming to build, the expectation of what can be accomplished in an introductory class that has no sequel or immediate application of content must (or should) surely be lower than it was two decades ago. But, again, enough speculation. How large is the gap between present learning outcomes and possible ones, and what are the personal, social, and economic consequences of bridging that gap or not? Progress on these issues has been made in elementary and high school instruction. We need it now in higher education.

REFERENCES

- Axelrod, S. (1992). Disseminating an effective educational technology. *Journal of Applied Behavior Analysis*, 25, 31–35.
- Bangert-Drowns, R. L., Kulik, J. A., & Kulik, C. C. (1991). Effects of frequent classroom testing. *Journal of Educational Research*, 85, 89–99.
- Binder, C. (1996). Behavioral fluency: Evolution of a new paradigm. *The Behavior Analyst*, 19, 163–197.
- Buskist, W., Cush, D., & Grandpre, R. J. (1991). The life and times of PSI. *Journal of Behavioral Education*, 1, 215–234.
- Clark, L. (1995). *SOS help for parents*. Bowling Green, KY: Parents Press.
- Dempster, F. N. (1992). Using tests to promote learning: A neglected classroom resource. *Journal of Research and Development in Education*, 25, 213–217.
- Dubin, R., & Taveggia, T. C. (1968). *The teaching-learning paradox: A comparative analysis of college teaching methods*. Eugene, OR: Center for the Advanced Study of Educational Administration.
- Ellis, J., & Glenn, S. S. (1995). Behavior-ana-

- lytic repertoires: Where will they come from and how can they be maintained? *The Behavior Analyst*, 18, 285–292.
- Ericsson, K. A. (Ed.). (1996). *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games*. Mahwah, NJ: Erlbaum.
- Ericsson, K. A., & Charness, N. (1994). Expert performance: Its structure and acquisition. *American Psychologist*, 49, 725–747.
- Ericsson, K. A., Krampe, R. T., & Tesch-Romer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review*, 100, 363–406.
- Fowles, M. (1993). Designing a writing assessment to support the goals of the project. In B. C. Bruce, J. K. Peyton, & T. Batson (Eds.), *Network-based classrooms: Promises and realities* (pp. 263–285). New York: Cambridge University Press.
- Fraley, L. E., & Vargas, E. A. (1974). Academic tradition and instructional technology. In J. M. Johnston (Ed.), *Behavioral research and technology in higher education* (pp. 481–506). Springfield, IL: Charles C. Thomas.
- Grant, L., & Evans, A. (1994). *Principles of behavior analysis*. New York: HarperCollins.
- Johnson, K. R., & Layng, T. V. J. (1992). Breaking the structuralist barrier: Literacy and numeracy with fluency. *American Psychologist*, 47, 1475–1490.
- Johnson, K. R., & Layng, T. V. J. (1994). Morningside model of generative instruction. In R. Gardner, D. M. Sainato, J. O. Cooper, T. E. Heron, W. L. Heward, J. Eshleman, & T. A. Grossi (Eds.), *Behavior analysis in education* (pp. 173–197). Pacific Grove, CA: Brooks/Cole.
- Johnston, J. M. (Ed.). (1974). *Behavioral research and technology in higher education*. Springfield, IL: Charles C. Thomas.
- Keller, F. S. (1968). "Good-bye, teacher . . ." *Journal of Applied Behavior Analysis*, 1, 79–89.
- Kulik, C.-L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60, 265–299.
- Malott, R. W. (1984). In search of human perfectibility: A behavioral approach to higher education. In H. L. Heward, T. E. Heron, D. S. Hill, & J. Trip-Porter (Eds.), *Focus on behavior analysis in education* (pp. 218–245). Columbus, OH: Merrill.
- Malott, R. W., Vunovich, P. L., Boettcher, W., & Groeger, C. (1995). Saving the world by teaching behavior analysis: A behavioral systems approach. *The Behavior Analyst*, 18, 341–354.
- Malott, R. W., Whaley, D. L., & Malott, M. E. (1997). *Elementary principles of behavior* (3rd ed.). Upper Saddle River, NJ: Prentice Hall.
- Martin, G., & Pear, J. (1996). *Behavior modification: What it is and how to do it* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
- Miller, L. K. (1997). *Principles of everyday behavior analysis* (3rd ed.). Pacific Grove, CA: Brooks/Cole.
- Miller, L. K., & Weaver, F. H. (1975). A behavioral technology for producing concept formation in university students. *Journal of Applied Behavior Analysis*, 9, 289–300.
- Rubin, D. C., & Wenzel, A. E. (1996). One hundred years of forgetting: A quantitative description of retention. *Psychological Review*, 103, 734–760.
- Semb, G. B., & Ellis, J. A. (1994). Knowledge taught in school: What is remembered? *Review of Educational Research*, 64, 253–286.
- Sherman, J. G., Ruskin, R. S., & Semb, G. B. (Eds.). (1982). *The personalized system of instruction: 48 seminal papers*. Lawrence, KS: TRI Publications.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, 63, 129–138.
- Stokes, T. F., & Baer, D. M. (1977). An implicit technology of generalization. *Journal of Applied Behavior Analysis*, 10, 349–367.
- Stokes, T. F., & Osnes, P. G. (1989). An operant pursuit of generalization. *Behavior Therapy*, 20, 337–355.