Evaluating student achievement with a relational database

KENNETH B. LESURE and JAMES A. FROST
Cuyahoga Community College, Cleveland, Ohio

and

ROGER L. WILLIAMS
Quinsigamond Community College, Worcester, Massachusetts

Commercial database programs such as dBase and Paradox, although developed originally for business applications, are versatile and powerful tools that can be used for an academic purpose such as evaluating student performance. They can be used to write and store test questions, assemble and print classroom or on-line laboratory tests, and calculate grades, test statistics, and so forth. Databases are flexible, unlike textbook "ancillary" test bank programs that are inextricably bound to the strictly linear format and brief shelf life of specific textbook editions. A prototypical relational database program is described, with which an instructor can produce tests based on generic terms adapted from Boneau's (1990) study of psychological literacy, as well as on behavioral learning objectives adapted from Bloom's (1956) taxonomy of educational objectives. As a relational database, the program integrates terms, objectives, questions, tests, and test scores, and avoids unnecessary data duplication and waste of computer storage space.

Commercial database programs such as dBase, Paradox, and others were developed primarily for business applications such as recording a company's inventory, customer accounts, sales, employee data, and so forth. These database programs are time-tested and extremely powerful tools for managing large amounts of data, especially in dynamic sets in which information is frequently added and/or updated. Similarly, spreadsheet programs such as Lotus 1-2-3 and Excel were originally developed for business purposes such as financial forecasting. Unlike databases, however, spreadsheets were readily adopted by psychologists and other academics for applications such as electronic gradebooks, statistics, and simulations. What has been overlooked by academics is the enormous potential of versatile database programs for student evaluation. Any one of a number of popular database programs, such as the two named above, can be used to write and/or store test questions of any type, print tests made up of any of the stored questions, and record and analyze data from tests that have been administered in the classroom or even "on-line" in a computer laboratory or on a computer in the instructor's office.

Many instructors of introductory and other undergraduate courses may not be aware that, if they use "ancillary" testing software "bundled" with standard undergraduate textbooks, they are using a database. If they are, however, they may be frustrated by what is likely an extremely limited, "tunnel vision" and "parking meter" type of database application. The tunnel vision metaphor refers

to the textbook-specific limitation of typical textbook-testing software. The typical computerized test bank database is locked to a specific edition of a specific textbook. Questions are written to evaluate a student's memory of facts, concepts, and applications as presented in a linear sequence, chapter-by-chapter and page-by-page. Many questions reflect the textbook's particular emphases, biases, or idiosyncratic ways of presenting material. Worse, questions are usually literal iterations of material presented in the text, primarily because a graduate student or young instructor—desperately in need of the meager supplemental income—was hired on a piecework basis by the publisher at minimum wage for the nuisance "mop-up" chore of writing the questions. Worst of all, the user is forced to read and evaluate each and every question in the linear sequence. Therein, the user must arbitrarily decide which questions are appropriate for inclusion in a test of the student's memory of what was discussed in the classroom and/or memory of what was assigned reading in the textbook. The result is that using a publisher's testing software can be nearly as tedious, time-consuming, and even irrational as writing a test from scratch.

The parking meter metaphor refers to the short shelf life of textbook publishers' testing software. The turnover in textbook editions is occurring at a faster rate than ever before; it is common for editions of standard textbooks to change in 2-year cycles. Although it is laudable that the publishers of undergraduate textbooks encourage their authors to continually update their material by incorporating the latest research findings, theories, and trends, this practice creates a considerable inconvenience for the instructor who has adopted the textbook. If the instructor has assembled a test containing questions on material

Requests for reprints should be sent to Kenneth B. LeSure, FLW Associates, P.O. Box 91292, Cleveland, OH 44101-3292.

Table 1
Taxonomic Levels of Behavioral Learning Objectives

Level	Verbs
D Definition	Define, identify, locate, pinpoint, determine, delineate, name, distinguish
T Translation	Translate, paraphrase, summarize, decipher, discuss, comment
I Interpretation	Interpret, elaborate, justify, review, inter, criticize
C Classification	Classify, apply, show, demonstrate, provide, illustrate, produce
A Abstraction	Abstract, recognize rare forms, identify uncommon examples

unique to the current edition, the rapid turnover rate may force her or him to discard the test only a few months after it was created.

More seriously, from a psychometric point of view, the rapid turnover in textbook editions means that it is increasingly improbable that test items are systematically evaluated for their validity and reliability. Basic psychometric data on test questions are rare in a publishing environment in which revision and novelty prevail—not validity, reliability, or the establishment of norms, which are,

of course, the basic elements of test construction and development. Even if the instructor attempts to gather such data on a publisher's test questions, he or she would find it difficult to do so on more than a small proportion of the complete bank. Moreover, such data would be meaningless in regard to questions on idiosyncratic or trendy material with built-in obsolescence.

Our solution to the lack of stability in computerized test banks is to develop a generic body of questions that can be used with any standard textbook in the field. Boneau's (1990) survey of textbook authors suggests that there is a clearly definable body of facts, theories, methods, concepts, and so forth, that is common to all psychology, at least as it is communicated in American introductory textbooks. A general-purpose commercial database program, unlike standard published test banks, is not tied to any specific textbook. Instead, each question can be tied to a specific learning objective. The learning objectives are those which, explicitly or implicitly, are common to all contemporary psychology curricula and introductory textbooks. Each objective, out of a total of several hundred, may have several questions of varying difficulty and type according to a hierarchical taxonomy adapted from

Figure 1. A relational database set for a general psychology course.

Bloom's (1956) well-known taxonomy of educational objectives. Bloom's system contains a hierarchy of seven difficulty levels, with factual knowledge at the base and analysis, translation, interpretation, and so forth, at higher levels. In our multiple-choice question database, Bloom's seven taxonomic levels are collapsed into five (see Table 1). By referring to specific behavioral learning objectives and taxonomic levels, the instructor can easily assemble a test according to rational criteria rather than intuitively or arbitrarily.

The database menus and commands give the instructor the option of assembling tests with questions of one type (multiple choice, essay, fill-in, etc.) and/or level, or of mixing type and/or level. A course could also be designed to accommodate students of several levels of preparedness or motivation by allowing them to contract for different grades or credits according to taxonomic level.

In addition to assembling tests, the instructor can deliver tests to students on-line or directly via personal computer, either on an individual workstation or on a network. Such tests do not necessarily replace traditional classroom paper-and-pencil exams because it is unlikely that, on most campuses, more than a few students can be tested on-line at any given time, but on-line tests can provide an alternative system of assessment and evaluation for nontraditional students such as employed adults, homebound parents, or disabled students who cannot easily attend regularly scheduled classes. On-line tests also enable traditional classroom instructors to use more scheduled classroom time for instruction and discussion and less (or no) time for testing. Furthermore, on-line tests can automate test scoring, test score recording, and grade calculation, thus relieving the instructor of annoying clerical and paperhandling work. They can even assist in the usually neglected task of test item analysis.

A relational database program is so named because the user can "link" (access directly) one or more other databases as if they were part of a single large database (see Figure 1). Rather than simply creating a single database

containing all information, a relational program has the advantage of preventing the unnecessary duplication of information in every record, and, as a result, saving a vast amount of storage space. All that is required for a link between databases is one common "field" (data category). In our prototype database "set," the link among databases is each record's "code," which is analogous to the old Dewey Decimal System used by libraries. The core database contains generic terms and learning objectives for an undergraduate general psychology course. Within the core, each record is relatively small and compact, with a maximum length of about 100 characters. The chapter databases, by contrast, each contain large records made up of the text of questions the user selects for copying to a "test" database, which she or he can give to students in either "report" (hard copy) or "form" (on-line) format. From a test database created by the user, selected data can be copied and linked to a "gradebook" database containing students' test scores and other data categories created by the user. From a gradebook, the user can examine student data by using the intuitive "eyeball" method with a "browse table" that closely resembles a spreadsheet screen layout. In addition, she or he can print out a hard copy "report" of a gradebook and even send individual progress reports to every student by using a user-created "form letter."

Our prototype student evaluation system database will be field-tested in a computerized self-paced introductory psychology course beginning in mid 1993. Evaluative data will be collected regarding its effectiveness, including test question validity and reliability, as well as its overall "affectiveness" (user-friendliness and appeal).

REFERENCES

- BLOOM, B. S. (Ed.) (1956). *Taxonomy of educational objectives: Handbook I. Cognitive domain*. New York: Longman, Green.
- BONEAU, C. A. (1990). Psychological literacy: A first approximation. *American Psychologist*, **45**, 891-900.