

Item Banking in Computer-Based Instructional Systems

Frank B. Baker
University of Wisconsin

This paper examines item banking within computer-based instructional systems from both a systems and a measurement perspective. Traditionally, computer-aided instruction involves little testing, although there is a trend to incorporate posttests in the sessions. However, computer-managed instruction has incorporated testing since its inception. The tests employed are similar in most respects to teacher-made classroom tests. The test results are used as the basis for diagnosis, prescription, and management procedures for individual or small groups of students. At the classroom level, test banking may be more appropriate than item banking. Because of the tight linkage of the tests to instructional procedures, the basic measurement issue appears to be the degree to which the approaches evolved from standardized achievement testing can be applied to the large number of short tests employed in computer-based instructional systems.

The advent of the microcomputer in the late 1970s and its widespread incorporation into the schools in the 1980s has had the effect of dramatically altering the context for computer-aided instruction (CAI) and computer-managed instruction (CMI) at all educational levels. It is now possible to incorporate computers into the ongoing educational process in a manner that is both technologically and economically feasible. Such computer-based instructional systems (CBIS) use the computer both to manage and to deliver instruction. Since testing

is an integral part of instruction, it is natural that computers be employed in test design, construction, and analysis. In order to implement these functions, an item banking capability is essential. While item banking has been a major component in achievement testing programs, its role within the context of CBIS has not been firmly established. Consequently, the present paper examines the characteristics of testing within CBIS and how they affect the conceptualization, implementation, and application of item banking in this context. Particular attention will be given to the measurement issues arising from the use of item banking within CBIS.

Testing Within CBIS

Although CAI originally was implemented through centralized software on mainframe computers, the advent of the microcomputer has resulted in a decentralized (one student, one computer) approach. The microcomputer has resulted in a widespread introduction of CAI; however, the bulk of existing software consists of isolated computer sessions that have little or no relation to each other or to a given curriculum. In addition, these computer sessions exhibit wide variation in their use of testing to determine whether the instructional goals have been achieved.

When testing occurs in a CAI program, there appear to be two types. In the first type, the com-

APPLIED PSYCHOLOGICAL MEASUREMENT
Vol. 10, No. 4, December 1986, pp. 405-414
© Copyright 1986 Applied Psychological Measurement Inc.
0146-6216/86/040405-10\$1.75

puter presents drill and practice exercises, and the student performs tasks that are often equivalent to workbook exercises or test items; the number of correct responses is reported to the student at the end of the session (see, e.g., Chomsky & Schwartz, 1983). The second type involves a game or simulation in which the student participates. Upon completion of the session, a multiple-choice test is administered (see, e.g., Minnesota Educational Computing Consortium, 1981 or Brown, Cabell, Kechula, & Shear, 1984). Such tests typically are short (8–12 items); they are used to determine if the student has achieved the instructional goals, and also provide the student with some feedback on how well he or she performed.

Both types of CAI programs are inconsistent with respect to whether the test results are saved for subsequent use by the teacher or test developer. The bulk of the older CAI sessions made no provision for retaining the test results. Recent microcomputer-based CAI systems, such as Private Tutor (Brown et al., 1984), save the test results on disk and provide access to them. It is interesting to note that in their survey of the history of CAI from 1959 to 1982, Kearsley, Hunter, and Seidel (1983a, 1983b) made no mention of testing as a component of CAI. Also, Roid (1986), in his evaluation of present and future uses of computer technology in testing, made no mention of testing within the context of CAI. Apparently, neither CAI specialists nor measurement specialists perceive testing as an important component of CAI.

In sharp contrast to CAI, testing has been an integral part of CMI since its inception. Initially, CMI was developed for individualized instruction based upon behavioral objectives (Baker, 1978). Under this approach a pretest was administered prior to an instructional unit in order to determine which non-mastered objectives a student should be assigned. A curriculum-embedded test of each assigned objective and a posttest covering all objectives were used to determine if the student had achieved mastery. Both the pretests and the curriculum-embedded tests were used in a diagnostic fashion as the basis for prescribing remedial instructional procedures (Glaser, 1967). The bulk

of the management procedures supported by the computer were based upon test results. In this view of CMI, the role of testing was to provide the information necessary to manage an individual's mastery of the behavioral objectives, and also to enable the teacher to manage a complex individualized rate of progress instructional paradigm. Achievement testing at the end of a semester or school year could also be encompassed within the approach. The amount of time devoted to testing under this paradigm proved to be excessive.

As a result, the current emphasis of CMI is upon the management of conventional instruction, with its larger curricular units and much lower frequency of testing. A good example of current practice, at the elementary school level, is the Computer Management System (Addison-Wesley, 1982). This is a CMI system for mathematics; it is offered as a complete package including textbooks, instructional materials, tests, and a set of disks containing the microcomputer CMI software. After the class roster is entered, the teacher's only task is to enter item responses using the keyboard or an optical mark reader. The rest of the system is menu-driven and provides access to a variety of reports as well as information on individual students. This CMI system is designed to be used by a limited number of teachers employing a common mathematics curriculum in several classrooms. Hence, it can be considered a decentralized system, as different school buildings would each use their own microcomputer.

At the college level, the Teacher Information Processing System (TIPS; Kelley, 1968) is a good example. TIPS is a centralized CMI system that runs on a mainframe computer, with each teacher having a unique set of data files but timesharing a common set of computer programs. At the end of each curricular unit, students take an ungraded survey quiz of 10 to 15 items on one or more topics. The test results are linked to diagnostic rules based upon item responses and student characteristics. After each test is administered, the results are entered into the computer using an optical mark reader. The computer then generates a prescriptive report for each student as well as summary reports for the

instructor. In practice, midterm and final course examinations typically are not encompassed within the TIPS system, although it is possible. Under the TIPS approach, the instructor is responsible for constructing the tests and for defining the diagnostic rules and prescriptions. It is important to note that the actual tests are external to the TIPS system, while the test results and their eventual uses are stored within the system.

Although both formative and summative testing are employed in CBIS, the bulk of the testing is fundamentally instruction-oriented. Such tests are used to make instruction-related decisions for individual students and/or groups of students. The useful life of the information provided by the tests is very short, on the order of days. As a result, the primary requirement is for a large number of short tests having a very narrow focus within a specific course in a given field of study, whose items are linked to diagnostic, prescriptive, and classroom management procedures (e.g., grouping students for instructional purposes). These requirements are in sharp contrast to those of a standardized achievement testing program such as the Iowa Tests of Basic Skills, SRA Achievement Series, or the Comprehensive Test of Basic Skills. In such standardized achievement testing programs, a test or test battery can deal with an entire single subject matter area, several areas, or even a total educational experience at one or more educational levels. The results of such programs tend to be used for purposes larger than day-to-day instructional management, although they are often advertised as being useful for the latter purposes (Linn, 1983).

The use of microcomputers to implement CAI and CMI at the classroom level results in another difference with standardized achievement testing programs. In the latter, the test(s) are administered on a given date within the school year. The resulting answer sheets are then aggregated over classrooms and schools and sent to a central test processing facility for scoring and reporting. The result is a very efficient method of collecting and processing tests on a large scale (Baker, in press). However, under CAI and CMI, testing takes place on an asynchronous basis as individual students or

small groups of students reach a given point in the curriculum. Given the need for immediate use of the test results, the tests are usually scored by the classroom computer and the results are available quickly. The net result is that when decentralized forms of CAI and CMI are used, it becomes impractical to aggregate test results for item analysis and other technical purposes. This is particularly true of CAI, where students use many different CAI sessions and microcomputers over any given period of time.

There appear to be fundamental differences between standardized achievement testing programs and the instruction-related testing performed under CBIS with respect to test characteristics, use of test results, and availability of item response data. (For an in-depth discussion of the characteristics of instruction-related testing see Nitko, in press.) In standardized achievement testing programs, the emphasis is upon creating, from an item bank, a few highly refined tests that are administered on a large scale for use in a broad scheme of evaluation. Under CBIS, the emphasis is upon using an item bank to create many short tests that provide information on an as-needed basis for the purpose of instructional decision-making at the classroom level. In addition, there are the aforementioned differences between the two major branches of CBIS with respect to the role of testing.

Some Issues Underlying Item Banking in CBIS

The basic processes of computer-based test construction are essentially independent of the ultimate use of the tests produced. Thus, item banking underlies both formative and summative testing under CBIS as well as standardized achievement testing programs. However, creating and maintaining an item bank for testing within the context of CBIS involves a number of unique issues. One of these is whether a single item bank or a collection of item banks is used.

It is technically possible to use a single large item bank and simply access it from different points of view to construct tests for individual segments

of a number of different courses. Such an approach is feasible only if sufficient computing resources can be devoted to the item-banking tasks. In addition, there are significant managerial problems in maintaining a very large diverse item bank. Seemingly minor problems, such as maintaining a dictionary of the allowable keywords or the relations among descriptors, become considerable managerial tasks. Using multiple specialized item banks within a common system has the advantage that each item bank can be tended by persons knowledgeable in specific curricular areas. However, such an approach leads to considerable parallelism of effort and can be more expensive.

Due to the dramatic decrease in the cost of mass storage, the use of microcomputers in item banking within the context of CBIS is becoming increasingly attractive. Item banks containing several thousand items can be stored easily. Because item banking and test construction are storage-bound rather than computationally bound, most of the recent microcomputers with 10 to 30 megabytes of hard disk storage have enough capability for these tasks. The underlying issue is how to employ these capabilities within the context of CBIS. One option is to use a decentralized approach involving local "stand alone" systems. These have the advantage of providing item banking at the school or classroom level, where the tests are employed, and the teachers can develop a sense of ownership of the item bank. The disadvantage is that many different item banks are created, with varying levels of quality and uncertain commonalities across the same curriculum. Another technically feasible approach employs the microcomputer as a component in a hierarchical computer network having a large computer at its vertex. In such an architecture, each microcomputer can be used to access the central item bank directly and perform test construction locally. In addition, it is possible to download specific item pools to the microcomputer for test construction at the local level.

In the typical item bank, the available number of items is considerably larger than the number of items in a given test constructed from the bank. Because CBIS can involve a large number of tests,

a sufficiently large number of items must be available. For example, the MICA CMI system (Madison Public Schools, 1976), employed in a 4th and 5th grade mathematics curriculum, involved 220 unit pretests and posttests containing a total of 7,000 items. There are both external and internal sources for items. External sources are book publishers, curriculum developers, and other item banks. Internal sources can be found within the educational setting, such as instructors, committees of teachers, and school system curriculum specialists. The advantages of external sources are both economic and practical. Since development of large numbers of items is both time-consuming and expensive, it is often advantageous to use externally developed items. However, externally developed items may not be very well matched to local curricula, classroom conditions, and instructional practices. Internally developed items may provide a better fit, but the costs of item development can be very high, and achieving a consistent level of development for the whole item bank can be difficult.

A salient feature of items in an item bank used under CBIS is the tight linkage of items to the instructional decision-making paradigm in a specific curriculum unit in a given subject matter area. There are two levels of linkage involved. First, there is the linkage of the items to the curriculum and the instructional goals, for purposes of establishing diagnostic procedures based upon the item responses and/or test scores. Second, there is the linkage of the results of the diagnostic procedures to the prescriptions that specify what the student is to do. These linkages in turn severely restrict the "mobility" of an item. For example, items in a posttest for a geometry unit on triangles would rarely be used outside that test, because the diagnostic and prescriptive procedures implemented by the CBIS depend upon the examinee's item responses to the specific items as well as upon the associated test score(s). Thus, the test constructor has little latitude to manipulate the item composition of a test. To do so requires that the corresponding diagnostics and prescriptions also be reestablished. In the case of CAI, it means that computer software may have to be rewritten, while in the case of CMI, the

data files and the reports would have to be modified. It should be noted that both internally- and externally-provided items must be linked in this manner. Thus, acquisition of items is only a part of the overall item banking task under CBIS. In practice, linking the items to the curricular structure, to instructional diagnosis, to prescriptions, and to management procedures is a time-consuming task.

The large number of tests required by curricula employing CBIS and their diversity raises an underlying psychometric issue. As employed in practice, both classical and item response theory assume a single underlying latent trait. Standardized testing programs typically relate an item bank and a single test to a given trait (e.g., mathematical ability). However, under CBIS the same trait may be measured by many short tests, each devoted to some specific facet of the trait (e.g., proofs, equation solving, etc.). The question then becomes whether each of these specific facets should be treated as a unique latent trait, as a component of a single trait, or as one trait within a multidimensional set of traits. The choice made has a direct impact upon the design and use of the item bank and, in particular, upon the computation and interpretation of item statistics and test results. Although some attention has been devoted to the multidimensionality of tests (see Bejar, 1980; McDonald, 1981; Reckase, 1979), this has not been done for the short tests used in the context of CBIS.

The collection of data upon which to base item and test statistics poses a unique problem under CBIS. In the usual standardized testing situation, a norm or calibration group is selected from a population of interest and the test is administered. Under either classical or item response theory, the analysis of this data yields the basal item and test statistics of the item bank. Further expansion of the item bank is accomplished by equating procedures (see Vale, and van der Linden & Eggen, in this issue). Under CBIS, the establishment of the basal items in the item bank can be accomplished in the same manner, but the asynchronous manner in which students take the individual tests makes expansion of the item bank a difficult task. Which

students take a given test at a given time is determined by their path through the instructional decision network. In a sense, the examinees are self-selecting and meaningful aggregation of item response data is not always possible. Because CAI typically is decentralized, incompatibility among the wide variety of software and computers involved, coupled with the lack of provision for storing test results, makes aggregation of test results a tenuous proposition. In the case of CMI, the computer procedures are based upon a database containing item responses and test results. The test data are available but may still be based upon self-selecting groups of examinees. Even at the end of the school year, not all students will have taken all tests. Hence, aggregation of results at that point still has inherent difficulties.

A related practical problem concerns test development procedures. Typically, test items are pilot-tested a number of times before they are included in tests to be equated into the item bank. However, instruction-related tests must be administered at the appropriate point in the instructional process. In addition, the testing cannot involve more time than is normally devoted to testing at a given point in the instructional sequence, as there are severe constraints on how much testing time the classroom can absorb. Thus, the developer of an instruction-related test works in a more constrained environment than does the developer of a standardized testing program.

According to Hiscox (1985), it is a myth that there is widespread interest in computer-administered testing. Perhaps it is also a myth that there are large numbers of teachers interested in constructing tests using computerized item banking for any purpose, let alone for use in the context of CBIS. The appeal of item banking is that by using a computer terminal the items are at the test constructor's fingertips, and tests can be constructed quickly and efficiently. While this is true in theory, there is more to test construction. It is necessary to develop a test definition (either formal or informal), specify the keywords and parameter values used to select the items, extract the items, and determine whether they are actually useful for the purposes at hand. Even if

efficient algorithms for searching the item bank are available, a certain amount of time must be devoted to evaluating each retrieved item and deciding whether it should be retained. Thus, the overall test construction process is time-consuming even though item retrieval may be quick.

The reality is that even with item banks and computerization, test construction requires some level of expertise and, more importantly, the time to sit at the computer to perform the necessary tasks. It is this latter issue that is really at the root of dispelling the myth. Public school teachers at both the elementary and secondary level must cope with many demands upon their time; they teach many classes and have limited time both in and out of class. Consequently, they will resist the addition of any tasks that add to their workload. Much of the item banking literature is written from the point of view that classroom teachers have an abiding interest in constructing item banks "from the bottom up," that is, in such a way that the item bank results from the classroom teachers working individually or in concert (see Millman & Arter, 1984). This does not appear to match the realities of the typical classroom, particularly in the elementary school.

At the college and university level the situation is somewhat better (see Seely & Willis, 1976). Instructors at this level tend to teach a limited range of courses; testing is less frequent and the subject matter is reasonably stable. Because of this, there has been greater interest in item banking at this level. For example, most of the contributors to Lippey's book (1974) on computer-aided test construction were college instructors. However, teachers at neither level have the resources needed to develop tests to the level of technical excellence of standardized achievement testing programs on the scale required by CBIS.

Despite the widespread availability of microcomputers, there are relatively few item banking programs for them that encompass the total item storage, test construction, and test analysis process. Hsu and Nitko (1983) and Hsu (1986) surveyed microcomputer software useful in testing. Deck and Estes (1984) surveyed the item banking/test con-

struction software. In these surveys, it was found that software exists for doing parts of the overall process, but accomplishing the full set of tasks requires the use of several different packages from several different sources. Unfortunately, the interfacing problems among these are considerable and it is not feasible to use them together. However, Assessment Systems Corporation (1984) has developed an integrated microcomputer-based item banking, test construction, administration, and analysis system, called MICROCAT. Nitko and Hsu (1984) have also developed an integrated system, called PETA, that has a CMI orientation and is intended for use by classroom teachers. Thus, although the necessary software is beginning to appear, it is available on a limited basis and has not found its way into the fabric of the classroom or CBIS.

Test Banking

Given that there is a range of issues unique to testing within CBIS, what is the potential for item banking within this context? This most viable approach seems to be one in which the unit of transaction at the classroom level is the test rather than the item. Under this approach, the instructional context would indicate which test was appropriate. The classroom teacher would then use the computer to retrieve a specific test, rather than construct a test on-line from items in an item bank. The difficulties of item banking at the classroom level have been recognized by a number of textbook publishers, and they have turned to "test banking."

It has been a long-standing practice of textbook writers to issue instructor's manuals that contain tests for the several sections of the textbook. In the past, the teacher typed the test and duplicated copies for local use. The current trend is for these tests to be stored on a floppy disk accompanied by microcomputer software for the selection and printing of the appropriate test (e.g., Psychological Corporation, 1982; Science Research Associates, 1982). Under the test banking approach, the linkage of the test results to instruction and to reporting is established by the publisher as part of the overall

curriculum design and implementation process, and hence would require no time on the part of the user. The teacher's only task would be reproducing the computer-printed test. Test banking is employed in the CAMELOT CMI system (Miami-Dade Community College, 1986). However, in this microcomputer-based system, instructors are responsible for using the computer to construct tests from their own item banks and then storing them for subsequent use within the CMI portion of the system.

While test banking is a natural fit with CMI, CAI poses a somewhat different problem. Under CAI the tests must be included in the computer session. The tests can be appended to the computer lesson as a separate component or they can be integrated into the session. The former case presents the mechanical problem of coordinating the test with the student's use of the CAI lesson. Typically, this would involve bringing in the test from mass storage along with test administration, scoring, and reporting software. This results in a rather clumsy procedure, as it involves two sets of software. Integrating the testing into the CAI software is attractive but more difficult, as it must be done at the software level. However, it may be possible to include standard software subroutines in CAI lessons to perform the test-related activities.

Switching the emphasis at the classroom level under CBIS from item banking to test banking has a number of implications. It means that the persons responsible for the test bank must (1) develop the items and use the computer to maintain the item banks from which the tests in the test bank are constructed; (2) establish the connections among the curriculum, the items, and the tests; and (3) create the linkages between test results and instructional procedures (e.g., diagnosis and prescription) as well as between test results and management procedures. These latter instructional linkage tasks are not currently performed by test constructors in the setting of a standardized achievement testing program. Thus, there is a considerably greater workload and cost associated with maintaining an item/test bank under CBIS than that necessary for a traditional item bank.

Discussion

Item banking within the context of computer-based instructional systems shares many characteristics with current practice as described by Millman and Arter (1984) and by Wright and Bell (1984). However, there are unique characteristics that set it apart from other applications. In the case of CAI, testing has not been a traditional part of the instructional procedure. One of the major difficulties in employing item banking in CAI is the diversity of sources of software and the range of microcomputers employed. Very little of the existing CAI software has been developed as an integral part of a curriculum or even as a systematic series of sessions. As a result, there is little coordination of the computer sessions, testing, and classroom instruction. In such a context, it is quite difficult to employ item banking in an effective manner.

Because computer access is still rather limited in most classrooms and because of the historical lack of interest in testing, it is questionable whether efforts should be made to superimpose conventional post-testing upon CAI lessons. The amount of time it takes to administer and score a conventional instrument as part of a CAI lesson does not seem to be a cost-effective use of scarce microcomputer resources.

As Baker (1984) has suggested, CAI appears to be a place for non-intrusive testing. That is, the CAI software should collect data from the student's interaction with the computer during the instructional process. These data can then be used within the CAI computer program for diagnosis and prescription as well as evaluative purposes. Under this approach, the student is not aware of being tested in the conventional sense, yet upon the completion of the computer session the usual testing goals have been achieved. Such an approach places a greater burden upon the lesson designer and software developer, but has the potential for better evaluation.

Non-intrusive testing would appear to be an approach intermediate to traditional post-testing within CAI and CAI programs based upon cognitive science. A good example of the latter is BUGGY (Brown & Burton, 1978), a diagnostic computer program for arithmetic. The BUGGY program consists of

hundreds of diagnostic rules for arithmetic problems and it is capable of determining a student's misunderstanding of the rules of arithmetic. In a real sense, it understands why a student makes errors. The next step, called intelligent CAI (Sleeman & Brown, 1982), is to institute computerized instruction to remedy the student's misunderstandings. The amount of intellectual and fiscal resources needed to develop intelligent CAI programs, even for simple topics, is very large. However, over the long term, cognitive science and intelligent CAI could have a profound impact upon the conceptualization and implementation of diagnosis, prescription, and evaluative procedures. It is also implicit that there is a limited role in such computer programs for conventional measurement techniques and procedures.

One of the recurring themes in recent measurement literature is the separation that has developed between testing and instruction (see Airasian & Madaus, 1983; Glass, 1986; Nitko, in press). Much of the discussion has focused upon obtaining a linkage between instruction and standardized achievement tests (see the Summer 1983 issue of the *Journal of Educational Measurement*, which is devoted to this topic). The underlying goal in such linkage is to have achievement tests measure either what students are being taught or should be taught.

In the case of CMI systems, testing has long been an integral part of the instructional management process and the tests are designed with this linkage in mind. The item responses and test scores are used as the basis for diagnosis, prescription, and evaluation as well as for instructional management purposes. As a result, these tests are much closer in spirit to teacher-made classroom tests than to standardized achievement tests. Linn (1983) described teacher-made tests as having the following characteristics: "(1) The degree of match between test items and instructional objectives, (2) the use of test results to provide feedback to the students and to the teacher, (3) the use of tests to flag facts or concepts that are considered important, and (4) the use of tests to determine grades" (p. 179). Because the characteristics of tests employed under CMI have a nearly one-to-one relationship to those of teacher-made classroom tests, item banking un-

der CMI must be approached more from an instructional design point of view than from a standardized achievement test point of view.

It is also the case that psychometric specialists either tend to ignore classroom tests or try to superimpose the technical accoutrements of standardized achievement tests upon them. However, it appears that an overemphasis upon the technical characteristics of the items and tests may be inappropriate in the context of CMI. Typically, minor but continuous shifts are made in curricula as a function of recognized deficiencies, new ideas, new instructional materials, and personnel changes that have an impact upon the tests, diagnosis, prescription, and classroom management procedures. Consequently, the life cycle of a given test may be short compared to that of a standardized achievement test.

Occasionally, major shifts in the curriculum occur. For example, new textbooks are adopted, different educational philosophies come into vogue, and new instructional paradigms are implemented. Because of the tight linkage between the curriculum and the item/test bank in CBIS, such major shifts can be significant. These shifts can require a total or near-total redesign of the tests as well as the linkages to the instructional management-related procedures. The net result is that it may not be economically feasible to invest more than some acceptable minimum amount of resources on the technical aspects of any given test within the CMI system. However, procedures for accommodating these shifts must be built into the overall CMI design and specifically into the item bank design. Thus, an item banking capability cannot exist as a separate entity to be interfaced to a CMI system. Rather, it must be designed from its inception as an integral part of the overall curricular approach in which CMI provides management support for the teacher. To do otherwise will result in a piecemeal design and a system that is difficult to employ effectively.

One of the enduring hopes of item bank software designers is that their systems will achieve widespread use at the classroom level. This is not likely to be the case. Test construction, even with computer assistance, requires a certain level of expertise and, more importantly, time at the computer

console. As a result, online test construction is more of a myth than a reality in many situations. In the context of CBIS, test banking seems to be more attractive than item banking. Using existing microcomputer technology, tests can be easily delivered to the classroom on demand in a form that teachers can readily use. In addition, the test content and the scores they yield would already be incorporated into the management procedures of the CBIS system.

Test banking also facilitates the cooperation of curriculum designers, textbook writers, measurement specialists, and management system developers. Using item banking, such teams can develop tests that are matched to the instructional decision-making and management procedures required by the subject matter area and the academic level of the course. The goal is to systematize the development and use of the equivalent of teacher-made classroom tests while retaining the desirable characteristics of such tests. As a result, a crucial aspect of item and test banking within CBIS is the requirement to be able to close the feedback loop within the classroom. Without such a loop, the item/test bank operates in a vacuum, and a gulf quickly develops between the users and maintainers of the system. Such a gulf has been the cause of death of more than one computer-based system.

References

- Addison-Wesley Publishing Company. (1982). *Computer management system to accompany mathematics in our world*. Menlo Park CA: Author.
- Airasian, P. W., & Madaus, G. F. (1983). Linking testing and instruction: Policy issues. *Journal of Educational Measurement*, 20, 103–118.
- Assessment Systems Corporation. (1984). *MICROCAT: A computer program for computerized adaptive testing*. St. Paul MN: Author.
- Baker, F. B. (1978). *Computer-managed instruction: Theory and practice*. Englewood Cliffs NJ: Educational Technology Publications.
- Baker, F. B. (1984). Technology and testing: State of art and trends for the future. *Journal of Educational Measurement*, 21, 399–406.
- Baker, F. B. (in press). Computer technology in test construction and processing. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.). Washington DC: National Council on Measurement in Education.
- Bejar, I. I. (1980). A procedure for investigating the unidimensionality of achievement tests based upon item parameter estimates. *Journal of Educational Measurement*, 17, 283–296.
- Brown, J. S., & Burton, R. R. (1978). Diagnostic models for procedural bugs in basic mathematical skills. *Cognitive Science*, 2, 155–192.
- Brown, T. K., Cabell, R. W., Kechula, M. A., & Shear, J. L. (1984). *Private Tutor (Version 2.0)* [Computer Program]. Boca Raton FL: IBM.
- Chomsky, C., & Schwartz, J. L. (1983). *M-ss-ng l-nks*. Pleasantville NY: Sunburst Communications.
- Deck, D., & Estes, G. (1984, April). *Microcomputer software for item banking*. Paper presented at the annual meeting of the American Educational Research Association, New Orleans.
- Glaser, R. (1967). Adapting the elementary school curriculum to individual performance. In *Proceedings of the 1967 Invitational Conference on Testing Problems*. Princeton NJ: Educational Testing Service.
- Glass, G. V. (1986). Testing old, testing new: School-boy psychology and the allocation of intellectual resources. In B. S. Plake & J. C. Witt (Eds.), *The future of testing*. Hillsdale NJ: Lawrence Erlbaum.
- Hiscox, M. D. (1985). Computer-based testing systems: Much ado about nothing. *Educational Measurement: Issues and Practices*, 4, 27–28.
- Hsu, T. (1986, April). *Developments in microcomputer applications to testing*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco.
- Hsu, T., & Nitko, A. J. (1983). Microcomputer testing software teachers can use. *Educational Measurement: Issues and Practices*, 2, 15–31.
- Kearsley, G., Hunter, B., & Seidel, R. J. (1983a). Two decades of computer-based instruction projects: What have we learned? Part I. *T.H.E. Journal*, 10(1), 90–94.
- Kearsley, G., Hunter, B., & Seidel, R. J. (1983b). Two decades of computer-based instruction projects: What have we learned? Part II. *T.H.E. Journal*, 10(2), 88–96.
- Kelley, A. C. (1968). An experiment with TIPS: A computer-aided instructional system for undergraduate education. *American Economic Review*, 58, 446–451.
- Linn, R. L. (1983). Testing and instruction: Links and distinctions. *Journal of Educational Measurement*, 20(2), 179–190.
- Lippey, G. (Ed.). (1974). *Computer-assisted test construction*. Englewood Cliffs NJ: Educational Technology Publications.
- Madison Public Schools. (1976). *Managed Instruction With Computer Assistance (MICA)*. Madison WI: Author.
- McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology*, 34, 110–117.

- Miami-Dade Community College. (1986). *CAMELOT: The individualized information system*. Miami FL: Author.
- Millman, J., & Arter, J. A. (1984). Issues in item banking. *Journal of Educational Measurement*, 21, 315–330.
- Minnesota Educational Computing Consortium. (1981). *GENETICS (APPLE II+)*. St. Paul MN: Author.
- Nitko, A. J. (in press). Designing tests that are integrated with instructions. In R. L. Linn (Ed.), *Educational Measurement* (3rd ed.). Washington DC: National Council on Measurement in Education.
- Nitko, A. J., & Hsu, T. (1984). A comprehensive microcomputer system for classroom testing. *Journal of Educational Measurement*, 21, 377–390.
- Psychological Corporation. (1982). *PRISM* [Computer Program]. New York: Author.
- Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics*, 4, 207–230.
- Roid, G. H. (1986). Computer technology in testing. In B. S. Plake & J. C. Witt (Eds.), *The future of testing*. Hillsdale NJ: Lawrence Erlbaum.
- Science Research Associates. (1982). *Microcomputer Test Administration System (MTAS)*. Chicago: Author.
- Seely, O., & Willis, V. (1976). SOCRATES test retrieval at the California State University and Colleges. *AEDS Journal*, 9, 65–70.
- Sleeman, D., & Brown, J. S. (1982). *Intelligent tutoring systems*. London: Academic Press.
- Wright, B. D., & Bell, S. R. (1984). Item banks: What, why, and how. *Journal of Educational Measurement*, 21, 331–346.

Author's Address

Send requests for reprints or further information to Frank B. Baker, University of Wisconsin-Madison, 1025 W. Johnson Street, Madison WI 53706, U.S.A.