ACT RESEARCH REPORT

No. 60

60

September 1973

IMPLEMENTATION OF A BAYESIAN SYSTEM FOR DECISION ANALYSIS IN A PROGRAM OF INDIVIDUALLY PRESCRIBED INSTRUCTION

R. L. Ferguson
M. R. Novick

PUBLISHED BY THE RESEARCH AND DEVELOPMENT DIVISION

THE AMERICAN COLLEGE TESTING PROGRAM

P. O. BOX 168, IOWA CITY, IOWA 52240

ABSTRACT

The decision process required for Individually Prescribed Instruction (IPI), an adaptive instructional program developed at the University of Pittsburgh, is described. In IPI, short tests are used to determine the level of proficiency of each student in precisely defined learning objectives. The output of these tests is used to guide instructional planning for individual students.

The nature and effect of errors in proficiency decisions are described and a procedure for reducing the probability of such errors is proposed. The plan calls for a Bayesian procedure which would incorporate prior information on the instructional program, for example the distribution of the percentage of items answered correctly by students. Such a procedure would permit inferences about the true level of functioning of each student.

The final section of the paper proposes two methods for implementing these procedures in an ongoing IPI program: one approach calls for the integration of the procedure as a part of a computer-based instructional management system, whereas the second approach describes how the procedure can be made tractable in a typical, non-automated IPI classroom.

IMPLEMENTATION OF A BAYESIAN SYSTEM FOR DECISION ANALYSIS IN A PROGRAM OF INDIVIDUALLY PRESCRIBED INSTRUCTION

The research reported herein was performed pursuant to Grant No. OEG-0-72-0711 with the Office of Education, U.S. Department of Health, Education, and Welfare, Melvin R. Novick, Principal Investigator. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

IMPLEMENTATION OF A BAYESIAN SYSTEM FOR DECISION ANALYSIS IN A PROGRAM OF INDIVIDUALLY PRESCRIBED INSTRUCTION

Richard L. Ferguson
Melvin R. Novick

INTRODUCTION

The feasibility of instructional programs designed to adapt to the individual needs of learners has been adequately demonstrated by educational systems like Individually Prescribed Instruction (Glaser, 1968) and A Program for Learning in Accordance with Needs (Flanagan, 1967). Although these programs accomplish individualization in somewhat different ways, each includes components which can be described by the following sequence of operations:

1. Specification of the learning objectives in terms of observable student behavior.
2. Assessment of the student's entering competencies.
3. Assignment or election of educational materials and/or experiences fitted to the student's individual needs.
4. Continuous assessment and monitoring of the student's performance and progress.

Since programs like IPI and PLAN call for adaptation of the learning environment to meet individual requirements, they necessarily rely heavily on the systematic assessment of student progress. Glaser (1968) has observed that, in IPI, test data serve as the primary source of information enabling teachers to make differential decisions regarding student instruction. Thus, steps (2) and (4) play a prominent role in the successful implementation of IPI. A review of current decision-making procedures for four selected individualized instructional programs has been given by Hambleton (1973).

The fundamental purpose for testing in individualized instructional programs like IPI and PLAN is to ascertain whether or not the student has attained some prescribed level of proficiency in a specified learning objective. Hambleton and Novick (1973) have observed that, "Questions of precise achievement levels and comparisons among individuals on these levels seem to be largely irrelevant." Because test data are used initially to place a student at the appropriate point within an instructional program or sequence, and thus to identify appropriate learning materials or experiences given his needs, the test models which have emerged to serve this function are very different from those used for standard instructional models. Because these tests relate a student's performance on items drawn from a carefully specified domain to a prespecified criterion or standard, these tests have come to be called *domain* or *criterion-referenced tests.*

It is not the purpose of this paper to contrast the differences between norm-referenced tests and criterion-referenced tests. Suffice it to say that criterion-referenced tests are deliberately con-

	Level						
	A	B	C	D	E	F	G
Numeration/Place Value	15	9	14	5	6	7	6
Addition/Subtraction	17	12	13	10	4	4	6
Multiplication		4	7	9	7	4	3
Division		3	4	7	9	5	6
Fractions	3	3	6	7	11	8	8
Money	1	1	5	5			
Time		6	6	4	4	2	
Systems of Measurement		3	6	6	5	5	6
Geometry		3	2	4	6	4	2
Applications		3	8	9	5	4	6

Behavioral Objectives

E Level-Systems of Measurement

1. Given a ruler, the student measures a line segment with the indicated degree of precision. LIMIT: smallest unit of precision 1/8 inch; line segments to 10 inches.

2. Given 20 cut-out regions that are each 1-inch squares and an illustration of a rectangular region, the student uses the 1-inch squares to determine the area of the given rectangular region. LIMIT: areas < 20 square inches. Length of sides of rectangles must be multiples of 1 inch.

3. Given the measures of the sides of a rectangular region, the student determines the area of that region. LIMIT: integral measures; one unit of measure per problem; units of measure—inches, feet, yards, miles.

4. Given the measure of the sides of a rectangular region, the student determines the perimeter and the area of that region. LIMIT: At least one of the measures (length, width) must be integral; both measures must be < 100; one measure may be a common fraction < 1 with denominator < 10; 1 unit of measure per problem; units of measure—inches, feet, yards, miles.

5. Given a weight measurement, the student completes a statement to show an equivalent measurement in a different unit of weight measure. Given a word problem that requires conversion of a given weight measurement expressed in standard units to an equivalent weight expressed in another standard unit, the student solves the problem and writes the answer with the appropriate label. LIMIT: units—ounces, pounds, tons.

Fig. 1. Matrix of Units in the IPI Mathematics Curriculum.

proficiency at this point or whether he needs to complete additional steps in the instructional process.

The CET typically consists of from four to six items. Because these short tests serve primarily as self-checks for the student, and because no crucial instructional decision is dependent upon student performance on these tests, they seem to be adequate for the task which they serve.

Unit Posttests

These instruments are equivalent forms of the unit pretests. They are generally administered after the student has concluded learning activities for all skills for which he was identified as being insufficiently proficient on the unit pretest. On the basis of the student's performance on the posttest, he is either advanced to the next unit or required to work with additional instructional materials on those skills for which his test performance did not indicate that he achieved a sufficient level of proficiency. A student generally does not advance to a new unit until he has demonstrated sufficient proficiency for all objectives of the current unit.

As with the pretests, decisions resulting from an analysis of posttest data rely upon tests which generally contain a small number of items. Because incorrect proficiency decisions can be detrimental to the student's progress, a procedure which could add substantially to the accuracy of the decision without increasing the length of the test would be most worthwhile.

THE INSTRUCTIONAL DECISION PROCESS

In this section, the process by which test data are used to make instructional decisions is briefly summarized. In addition, a discussion of the nature and consequences of decision errors resulting from the analysis of test data is presented.

A Summary of the Decision Process

Gross placement tests which sample a broad cross section of the important skills in each unit of the mathematics curriculum are administered upon each student's entry into the IPI program. Score data resulting from these tests are used to determine a profile suggesting the student's level of proficiency in each content area of the curriculum.

At this point, the student completes a pretest for the first unit in the curriculum continuum in which his level of proficiency is insufficient. The profile resulting from the pretest identifies those skills for which learning materials and/or experiences are required if the student is to achieve the specified level of performance. During the instructional process, curriculum embedded tests are available to the student as a means of self-evaluation and an estimate of progress as he works on the skills. After he has completed work on all skills in the unit and is satisfied that he has sufficient competency in all of the unit skills, he is administered a posttest which verifies his progress or identifies those skills for which additional instruction is indicated. Once the unit is successfully completed, the student advances to the next unit on his prescription where he is administered a pretest and the cycle is repeated.

The Nature and Effect of Decision Errors

The placement tests, pretests, and posttests are used primarily to verify that a student either has sufficient proficiency, i.e. mastery, in a given set of skills or that he has an inadequate level of proficiency in those skills. Clearly, it is desirable that the mastery decisions for a student be as accurate as possible. The importance of accuracy of the mastery decision for a student is perhaps best emphasized by a discussion of the consequences of an incorrect decision.

As previously indicated, the IPI tests are constructed by sampling items from the domain of items for the objectives included on the tests. Since any sampling which does not exhaust the population of items for an objective can lead to an incorrect mastery decision and since exhaustive testing is impossible, it is necessary to tolerate the risk of making wrong decisions. In an IPI context, a Type I (α) error occurs when an examinee has sufficient proficiency in a skill but the outcome of the testing suggests that he does not. As a result, he is prescribed work lessons which may serve no significant function. A Type II (β) error occurs whenever the examinee, in fact, lacks proficiency in a skill but on the basis of test results is said to have sufficient proficiency. The consequence of a Type II error is that needed remedial instruction is not provided. A Type II error is perceived to be potentially more serious than a Type I error since the Type II error could easily result in the student having difficulty proceeding through a unit and might paper. In short, although it is not necessary that teachers and students understand the details of the analysis, they must be provided information which facilitates their instructional decision making. In the following section, procedures for dealing with the preceding concerns are discussed.

The collection and analysis of data. During the past several years, considerable investigation has been underway into the feasibility of using a computer as an integral part of the IPI program. A thorough discussion of the most recent developments is available in a progress report (Block, Carlson, Fitzhugh, et al., 1973) recently released by the Learning Research and Development Center at the University of Pittsburgh. Earlier reports include Cooley and Glaser (1969), Ferguson (1970b, 1971), and Ferguson and Hsu (1971). The activities described in these reports emphasize somewhat visionary ideas for how the computer can best be employed in an individualized program of instruction. Although these studies include the more conventional modes of computer-assisted instruction, they extend far beyond into such areas as computer testing and instructional management.

It is in this latter area, instructional management, that Bayesian procedures for determining proficiency decisions would best seem to reside. Work in this area has been concerned with how the computer can assist in the planning and subsequent monitoring of both short- and long-term instruction for individual students. Thus, it would seem appropriate to incorporate a decision-making procedure concerned with individual proficiency level in some skill, or set of skills, as an element of the instructional management component of the IPI program. Specifically, the computer might be used to receive test data on a student and combine this with previously acquired information on other students in this IPI program, analyze the data using Bayesian analysis techniques, and then print out a report indicating the confidence which one could place in deciding that the student is proficient in a given skill at some prespecified level of performance. A more detailed discussion of how this procedure might work is now provided in the context of IPI posttests. Procedures similar to those described below would apply for placement tests and pretests as well.

Development and use of a posttest profile. The primary purpose for administering a placement test, a pretest, or a posttest is to acquire data which can be used to evaluate a student's instructional needs. When a student is administered a posttest, he is presumed to have had instruction in those skills for which he lacked sufficient proficiency at the time he was administered the unit pretest. The posttest either affirms the student's success in acquiring the skills or calls attention to those skills in which additional work is required before he can proceed to the next unit. Thus, the only information which the teacher and student need is a simple statement regarding the level of proficiency at which the student has performed on each skill in the unit. Figure 2 shows an IPI posttest profile based on a test consisting of five, eight-item subtests, each measuring proficiency level on a particular skill.

Level E-Multiplication/Division

Skill	Percent Correct
1	87.5
2	87.5
3	75.0
4	100.0
5	67.5

Fig. 2. Sample of Posttest Profile Currently in Use in IPI.

Presently, the posttest profile names each skill in the unit and lists the *percentage* of items which the student answered correctly. Given the sample profile in Figure 2 and a criterion (cutoff) score of 85%, it is likely that the student would be called upon to undertake additional work in the 3rd and 5th skills of the unit.

Under the proposed change, rather than evaluating student proficiency solely on the posttest results, additional data would be incorporated within the decision analysis process, and furthermore, the quantity reported would be an index relating the student's estimated proficiency to a stipulated standard. However, it should be emphasized that although the nature of the data reported in the student profile would change, the procedures employed by the teacher and/or student to judge proficiency would remain the same. Specifically, the posttest profile, which presently contains a statement of the percentage of items correctly answered for each skill, would be altered to report the probability that the student has achieved some prespecified level of proficiency in each objective. As far as the teacher or student is concerned, the proficiency decision process is exactly the same—judgments are based on the

The indices reported in the tables would have been generated at some earlier time and would have included consideration of relevant prior data regarding student success on the skills contained in the unit. The tables would be updated on a regular basis as increased numbers of students proceeded through the system, thus making more prior information available. Such an updating might occur once or twice a year.

SUMMARY

Individualized learning programs like IPI generate substantial amounts of data related to student success on skills in the system. Given these data, it seems reasonable to suggest that they should be used to improve the quality of instructional decision making. In particular, prior data should be combined with sample test data to form a more complete information base on which to evaluate student proficiency. By using such data jointly, instructional decisions regarding a student's needs as they relate to a given skill or set of skills will be deserving of more confidence than present decisions which are currently based solely on the student's performance on a short test.

Two procedures for implementing such a plan have been proposed. One calls for the marriage of the Bayesian decision analysis procedures with computer administered tests; whereas, the other would rely on the teacher or student to consult a table to translate student test performance to a "Proficiency Index" which would incorporate both the test data and prior data regarding student success in the system. The ultimate criterion for success of such a plan is the extent to which it leads to improvements in the instructional decision process. To this end, the next step is to implement the procedures and evaluate their impact on students within IPI.

REFERENCES

Block, K., Carlson, M., Fitzhugh, R., et al. *A computer resource for the elementary school: Progress report.* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1973.

Cooley, W. W., & Glaser, R. Computer and individualized instruction. *Science,* 1969, **166.**

Davis, C., Hickman, J., & Novick, M. R. A primer of decision analysis for individually prescribed instruction. *ACT Technical Bulletin No. 17.* Iowa City, Iowa: The American College Testing Program, 1973.

Ferguson, R. *Computer assistance for individualizing measurement.* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1970. (a)

Ferguson, R. Computer assistance for individualizing instruction. *Computers and Automation,* March 1970, **19.** (b)

Ferguson, R. A model for computer-assisted criterion-referenced measurement. *Education,* 1971, **81,** 25-31.

Ferguson, R., & Hsu, T. C. *The application of item generators for individualizing mathematics testing and instruction.* Pittsburgh, Pa.: University of Pittsburgh, Learning Research and Development Center, 1971.

Flanagan, J. Functional education for the seventies. *Phi Delta Kappan,* September 1967.

Glaser, R. Adapting the elementary school curriculum to individual performance. In *Proceedings of the Invitational Conference on Testing Problems.* Princeton, N.J.: Educational Testing Service, 1968.

Glaser, R., & Nitko, A. J. Measurement in learning and instruction. In R. L. Thorndike (Ed.) *Educational Measurement.* Washington: American Council on Education, 1971.

ACT Research Reports

This report is Number 60 in a series published by the Research and Development Division of The American College Testing Program. The first 26 research reports have been deposited with the American Documentation Institute, ADI Auxiliary Publications Project, Photoduplication Service, Library of Congress, Washington, D.C. 20540. Photocopies and 35 mm. microfilms are available at cost from ADI; order by ADI Document number. Advance payment is required. Make checks or money orders payable to: Chief, Photoduplication Service, Library of Congress. Beginning with Research Report No. 27, the reports have been deposited with the National Auxiliary Publications Service of the American Society for Information Science (NAPS), c/o Microfiche Publications, 305 East 46th Street, New York, New York 10017. Photocopies and 35 mm. microfilms are available at cost from NAPS. Order by NAPS Document number. Advance payment is required. Printed copies ($1.00) may be obtained, if available, from the Publications Division, The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240. A check or money order must accompany the request.

The reports since October 1970 in this series are listed below. A complete list of the reports can be obtained by writing to the Publications Division, The American College Testing Program, P.O. Box 168, Iowa City, Iowa 52240.

No. 37 *Practices and Outcomes of Vocational-Technical Education in Technical and Community Colleges,* by T. G. Gartland, & J. F. Carmody (NAPS No. 01441; photo, $6.80; microfilm, $2.00)

No. 38 *Bayesian Considerations in Educational Information Systems,* by M. R. Novick (NAPS No. 01442; photo, $5.00; microfilm, $2.00)

No. 39 *Interactive Effects of Achievement Orientation and Teaching Style on Academic Achievement,* by G. Domino (NAPS No. 01443; photo, $5.00; microfilm, $2.00)

No. 40 *An Analysis of the Structure of Vocational Interests,* by N. S. Cole, & G. R. Hanson (NAPS No. 01444; photo, $5.00; microfilm, $2.00)

No. 41 *How Do Community College Transfer and Occupational Students Differ?* by E. J. Brue, H. B. Engen, & E. J. Maxey (NAPS No. 01445; photo, $5.50; microfilm, $2.00)

No. 42 *Applications of Bayesian Methods to the Prediction of Educational Performance,* by M. R. Novick, P. H. Jackson, D. T. Thayer, & N. S. Cole (NAPS No. 01544; photo, $5.00; microfilm, $2.00)

No. 43 *Toward More Equitable Distribution of College Student Aid Funds: Problems in Assessing Student Financial Need,* by M. D. Orwig (NAPS No. 01543; photo, $5.00; microfilm, $2.00)

No. 44 *Converting Test Data to Counseling Information,* by D. J. Prediger (NAPS No. 01776; photo, $5.00; microfiche, $2.00)

No. 45 *The Accuracy of Self-Report Information Collected on the ACT Test Battery: High School Grades and Items of Nonacademic Achievement,* by E. J. Maxey, & V. J. Ormsby (NAPS No. 01777; photo, $5.00; microfiche, $2.00)

No. 46 *Correlates of Student Interest in Social Issues,* by R. H. Fenske, & J. F. Carmody (NAPS No. 01778; photo, $5.00; microfiche, $2.00)

No. 47 *The Impact of College on Students' Competence to Function in a Learning Society,* by M. H. Walizer, & R. E. Herriott (NAPS No. 01779; photo, $5.00; microfiche, $2.00)

No. 48 *Enrollment Projection Models for Institutional Planning,* by M. D. Orwig, P. K. Jones, & O. T. Lenning (NAPS No. 01780; photo, $5.00; microfiche, $2.00)

No. 49 *On Measuring the Vocational Interests of Women,* by N. S. Cole (NAPS No. 02071; photo, $5.00; microfiche, $1.50)

No. 50 *Stages in the Development of a Black Identity,* by W. S. Hall, R. Freedle, & W. E. Cross, Jr. (NAPS No. 02072; photo, $5.00; microfiche, $1.50)

No. 51 *Bias in Selection,* by N. S. Cole (NAPS No. 02073; photo, $5.00; microfiche, $1.50)

No. 52 *Changes in Goals, Plans, and Background Characteristics of College-Bound High School Students,* by J. F. Carmody, R. H. Fenske, & C. S. Scott (NAPS No. 02074; photo, $5.75; microfiche, $1.50)

No. 53 *Toward an Integration of Theory and Method for Criterion-Referenced Tests,* by R. K. Hambleton, & M. R. Novick (NAPS No. 02075; photo, $5.00; microfiche, $1.50)

No. 54 *College Student Migration,* by R. H. Fenske, C. S. Scott, & J. F. Carmody (NAPS No. 02215; photo, $5.00; microfiche, $1.50)