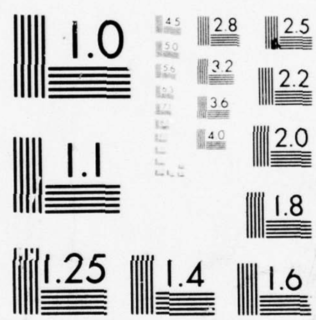
AD-A051 716 ILLINOIS UNIV AT URBANA-CHAMPAIGN COMPUTER-BASED EDUC--ETC F/G 5/9
ANALYSIS OF STUDENT INTERACTION DATA FROM CBE LESSONS,(U)
SEP 77 L FRANCIS, T WEAVER DAHC15-73-C-0077
UNCLASSIFIED MTC-15 NL

| OF |
AD
A051 716

END
DATE
FILMED
5-78
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Computer-based Education
Research Laboratory

University of Illinois Urbana Illinois

ANALYSIS OF STUDENT INTERACTION DATA FROM CBE LESSONS

LARRY FRANCIS
TAMAR WEAVER

MTC REPORT No. 15 SEPTEMBER 1977

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER	MTC #15
2. GOVT ACCESSION NO.	
3. RECIPIENT'S CATALOG NUMBER	715
4. TITLE (and Subtitle)	"Analysis of Student Interaction Data from CBE Lessons"
5. TYPE OF REPORT & PERIOD COVERED	
6. PERFORMING ORG. REPORT NUMBER	
7. AUTHOR(s)	Larry Francis and Tamar Weaver
8. CONTRACT OR GRANT NUMBER(s)	US Army DAHC 15-73-C-0077
9. PERFORMING ORGANIZATION NAME AND ADDRESS	University of Illinois Computer-vased Research Lab. 252 Engineering Reasearch Lab. Urbana, IL 61801
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS	
11. CONTROLLING OFFICE NAME AND ADDRESS	University of Illinois Grants & Contracts Office Davenport Hall, Champaign, IL 61820
12. REPORT DATE	September 77
13. NUMBER OF PAGES	
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	
15. SECURITY CLASS. (of this report)	Unclassified
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE	

16. DISTRIBUTION STATEMENT (of this Report)

~~Unrestricted~~

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

~~Unrestricted~~

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

CAI Computer-aided Instruction
CBE Computer-based Education
PLATO

training

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Provides an update on recent developments in student data analysis Includes an account of how student response data for three versions of a lesson directed revisions and validated changes.

(6)

Analysis of Student Interaction Data from CBE Lessons,

(14) MTC [redacted] -15. 715

(11) September 1977

(12) 35p.

(10) Larry Francis
Tamar Weaver

(15) DAHC15-73-C-0077, NSF-C723

COMPUTER-BASED EDUCATION RESEARCH LABORATORY
UNIVERSITY OF ILLINOIS, Urbana-Champaign

408 130

This research was supported in part by Advanced Research Projects Agency of the Department of Defense under U.S. Army Contract DAHC-15-73-C-0077 and indirectly by the National Science Foundation (USNSF C-723).

Analysis of Student
Interaction Data from CBE Lessons

Acknowledgments

The data handling and analysis routines described in this report have played an important part in the development of the PLATO IV courseware. In addition to the use of these routines by the MTC/PEER group in their evaluation of the ARPA/PLATO sites, many other military and civilian sites have made long and profitable use of the data collection and analysis packages available on the PLATO system. Much of the information in this report has been gleaned from studies conducted as part of the CERL evaluation of the Community College Project. In addition, Educational Testing Service (ETS), the independent evaluators for that project contributed suggestions and findings when they encountered problems with their own data collection programs, and elected to use these routines. Finally, a large but unrecorded number of individual PLATO curriculum developers have employed these well-publicized lessons as part of their development process and have provided feedback to help the creators refine the routines.

The authors would like to acknowledge the contributions of many people toward the advancement of the techniques for collection and interpretation of area summary data. The routines were prepared over several years by R.A. Avner, Kumi Tatsuoka, and Tamar Weaver aided by Elaine Avner, Steven Boggs, and Martin Siegel. The initial users of the routines provided many helpful suggestions and aided interpretation of the data. Those users were: Lou DiBello, Barbara Lederman, Keith Bailey, Mary Manteufel, and Frances Kennedy. Data editor and missing data correlation routines were prepared by Michael Felty, multiple regression programs by Kikumi Tatsuoka, and transformation programs by Tamar Weaver. Recent investigations into the use of time data have been carried out by Kikumi and Maurice Tatsuoka. This report was prepared with the aid of A. Lynn Misselt; the manuscript was edited by Elaine Avner, typed by Julie Garrard and illustrated by Roy Lipschutz.

Abstract

Features of the PLATO IV CBE system allow the gathering of large quantities and varieties of information describing the interactions of students with CBE courseware. Until recently this type and volume of data have not been easily available, and hence few techniques or guidelines for its analysis have been investigated. This report describes some initial attempts and the current status of efforts to collect, condense, and analyze these data for the purpose of diagnosing student problems and improving instruction.

Prologue

Throughout the history of the PLATO system, directions in data collection and analysis have shifted steadily. Shortly after the development of early PLATO systems, a system designed to process detailed individual data, the System for Instructional Response Analysis (SIRA) was devised (Avner, 1969). This package could analyze student responses virtually keystroke-by-keystroke. It was designed to perform many of the highly sophisticated analyses made feasible by the power of the computer. Unfortunately, its extensive flexibility overwhelmed most potential users. These users, under pressure to produce large quantities of instructional material could rarely spare the time needed to master the advanced measurement concepts upon which the SIRA package was based. Hence, they tended to avoid any approach to data analysis that they did not already understand.

Faced with a technically sound package which was not reaching its intended audience, the data collection philosophy of systems designers changed. In order to increase the likelyhood that instructional designers would make full use of the student data that CBE gave access to, it was decided to limit system-supported analysis packages to very basic levels which would be easy to use and "hook" users on the utility of student data. It was expected that once users became accustomed to using student data, they would desire access to more advanced techniques. Rather than attempt to provide the very flexible advanced packages that would be required to support the range of needed applications, authors were to be given tools (commands, databases, etc.) which they would need to develop their own, individual analysis packages. System designers felt that specialized packages for each project, discipline, or group would mean greater efficiency and higher acceptance than a "super-package" for all users. Thus, only basic routines have been system-supported since the advent of the PLATO IV system (Tatsuoka & Siegel, 1974).

At first, as anticipated, various groups began developing their own analysis routines (Smith & Sherwood, 1976; Yeager, 1976; Weaver, 1975). However, the next step of the process did not occur exactly as expected. Rather than each group developing unique packages tailored to their needs, some groups began adopting or adapting existing packages written by others. The resulting packages were upgraded and consolidated until today there again exists a set of multi-user student response analysis routines. (Of course, many groups still maintain their own routines.) The current general-purpose product is somewhat different than the SIRA product: for example, it does not not include many of the detailed functions possible with the SIRA package. Happily, the current package does receive broader usage than the SIRA package.

Thus the evolutionary process has come not quite full circle. Though the authors of this report can venture no opinion concerning the trends for future analysis packages in terms of user-produced or system-produced programs, new system-level data collection commands and continuing improvements to the various user-level analysis packages suggest that further changes in this area are probable. The following report details the creation, development, and use of a set of student data response routines collectively known as "the -area- package." The phase of development which is described herein coincides with the implementation of the PLATO IV system.

Introduction

One of the virtues of automated instruction expressed by CBE advocates is that it allows easy and unobtrusive gathering of vast amounts of data about a student's performance in a lesson. All this information can be used to immediately analyze and improve the instructional product. Nevertheless, it became obvious in the later years of the PLATO III system and the early years of the PLATO IV system efficient student data collection package alone was not sufficient to fully exploit the information available. The full potential of CBE could be realized only by means of computer-aided analyses of the student data.

As the PLATO IV data collection system was being created, the gathering of two kinds of information was thus provided for. The first type of data deals only with a single student interaction: the response he made, a key he pressed, or an error he caused. The second type of data summarizes the activities of a student after many interactions: for example, rather than recording responses, it categorizes them and tallies the number correct and incorrect. The first kind of data must be analyzed manually by the author; the second can be manipulated by the PLATO system, then interpreted by the author.

As part of the software for collecting the summarizing-type student data, the PLATO IV system provides the author with the ability to divide a lesson into contiguous "areas" and to automatically record a variety of measures about the student's use of the lesson within each area. Because this provision is so completely automated, many authors have implemented this feature, and there is a great deal of so-called "area data" being collected on the PLATO IV system.

An 'area' is an author-designated segment of a lesson which fairly consistently uses a single strategy like drill, tutorial, inquiry, etc. to teach a related set of facts or concepts. It is typically about 5 to 20 minutes in length. Whenever a student completes an area or stops working on a lesson, the computer records the following information for that student: the time spent in the area, the number of questions attempted, the number of questions eventually answered correctly, the number of incorrect student responses which were anticipated and unanticipated by the author, the number of questions answered correctly on the first try, counts of instances in which the student requested and received on-line help (and cases when on-line help was requested but unavailable), and the name of the student, the lesson, and the area. This 'area' summary, as well as the individual interaction data (e.g., the student responses) is stored on-line for later sorting, transfer, or printing.

It should be apparent that it is not only possible, but quite easy, to gather appropriate student data to be used to revise lessons and monitor student progress. The remaining problem is that although the once-difficult task of data collection has been solved, the analysis of the resultant volumes of information becomes time-consuming and frustrating. Recording all student responses (typically one per minute) for a class of 30 is clearly out of the question. Individual student interaction data must be recorded very selectively. Even the area summary data collects rather quickly. For example, a class of 210 students taking a two-hour lesson which provides area data about every 10 minutes will generate 2500 area summaries, each containing more than 10 items of data! There are so many values to examine and so many comparisons that might be made that many lesson developers, unable to analyze all their data, analyze none of it.

The analysis of this massive amount of data demands the use of a computer. Before the use of large CBE systems such as PLATO IV, the few CBE users in existence had comparatively little courseware for which to gather data and little software to make data collection feasible. Thus there was a limited need for ways to sort, condense, and analyze the data. For these reasons, PLATO lesson developers who gather area summary data have had few guidelines for its interpretation. The rest of this report describes recent attempts and the current status of efforts to condense, analyze, and interpret area summary data for the purpose of improving instruction and diagnosing student problems.

Interpretation of Area Summary Data

Two measures of student interaction have been found to be especially valuable for evaluating the performance of students and lessons: (a) the percentage of questions correctly answered on the first attempt (abbreviated %okfirst), and (b) the number of incorrect responses per question (abbreviated errors/question or errors/ques). Instructors comparing classroom observations to values for the above parameters find consistent, intuitive conformance. For example, an area which is very difficult will require the classroom proctor to aid the students and will result in a low %okfirst and a high value for errors/question.

For a wide variety of subjects and disciplines a class average of 80 %okfirst has been found to indicate that the difficulty of the material is well matched to the class. Instruction with %okfirst lower than 75 is generally too hard, and material with %okfirst above 85 is typically too easy.

The %okfirst alone is not sufficient. When only a few of the questions are very hard and the rest are easy,

%okfirst has its usual value--only errors/question is sensitive. Conversely, errors/question is not as sensitive as %okfirst in cases where the difficulty is consistently slightly too hard, but is divided evenly between questions. In some cases one finds that reasonable explanations can be made for "unusual" ratios, but in most instances, legitimate problems are found.

Despite large student-to-student variability, averages of %okfirst and errors/question (and completion time) tend to be fairly stable. It appears that if lessons were accompanied by a set of normal ranges for various student populations, an evaluator or instructor could estimate the aptitude and/or preparedness of a class. However, it is not generally feasible to use %okfirst or errors/question to "grade" individual students. This is because extreme values for certain students may point out cases which need further investigation, but variations in the style of student usage often explain what initially appear to be "problems." For example, some students deliberately make errors to see why alternative plausible answers are incorrect. Currently, the greatest generalizable use of area summary data is for guidance in the revision of lessons.

An Example of How Area Data Aided Revision of a Lesson

Initial Version. A math lesson teaching the addition and subtraction of signed numbers was divided into seven areas. Each of the first six areas contained instructions, examples, and exercises dealing with one type of problem (e.g., adding two numbers of opposite sign or subtracting two numbers both of which are negative). The seventh area provided mixed practice with the previous six problem types. Data for several classes is found in Table 1. The lesson

	Area	Min/use	Min/stu	%Okfirst	Errors/ques
Class A	1	2.0	2.3	87	0.40
	2	1.4	1.6	88	0.21
	3	1.7	1.9	88	0.16
	4	1.1	1.2	88	0.22
	5	2.6	2.9	80	0.45
	6	1.6	1.7	78	0.32
	mixed	2.0	2.1	82	0.31
Class B	1	1.4	1.7	85	0.10
	2	1.5	1.8	77	0.35
	3	1.4	1.8	90	0.15
	4	0.9	1.0	88	0.14
	5	2.5	2.7	78	0.35
	6	1.7	1.8	81	0.30
	mixed	2.1	3.2	70	0.48
Class C	1	3.8	5.9	76	0.74
	2	2.3	3.1	78	0.32
	3	2.5	3.2	86	0.16
	4	1.8	2.1	86	0.29
	5	4.7	5.7	70	0.84
	6	4.0	4.6	62	0.56
	mixed	5.1	6.2	56	0.51

Table 1

Class Performance Data for the Original Version of a Math Lesson was first used by large numbers of students in Spring 1975. As anticipated by the authors, the difficulty of the fifth and sixth areas (subtracting with negatives) was greater than that of other areas. However, what was not expected was the very great difficulty of the final mixed exercise. In five of the six classes the %okfirst dropped off dramatically for the for the mixed exercise (the seventh area).

These data suggested a need to add an additional "help" area before mixing together all six types of problems. Because individual problem types were apparently mastered, discrimination between problem types was seen as the stumbling block.

An examination of the histograms of %okfirst for each area provided additional insight about what was happening to the students. Figure 1 presents the results of a representative class for the last three areas of the lesson. These and other data suggest that many students apparently did well throughout the lesson whereas part of the class plummeted rather suddenly on the last mixed exercise. This information, combined with analyses of individual students, led the author to decide that only students who scored lower than 80% on the previous areas would be routed through the additional discrimination exercise (i.e., the "help" area).

A second, simultaneous modification to this lesson was based on instructor requests rather than student data or comments. In order to reduce the time needed for the lesson, the number of examples in each area was reduced from two to one.

Second version. The results of the Spring and Fall classes for the first and second versions of the lesson are shown in Figure 2. The classes using the revised version had a uniformly lower performance (as measured by %okfirst).

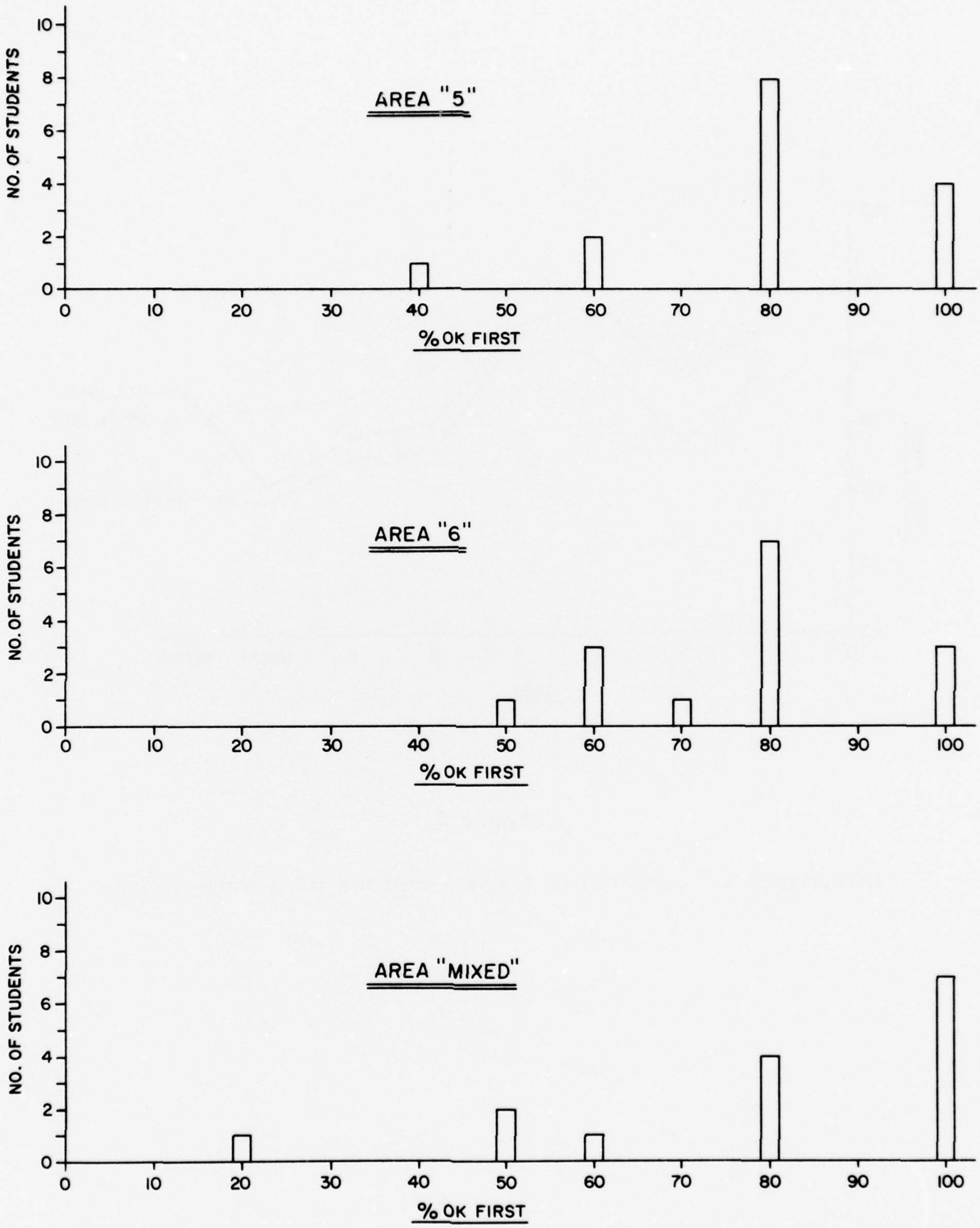


Figure 1

Class Performance on Final Three Areas of a Math Lesson

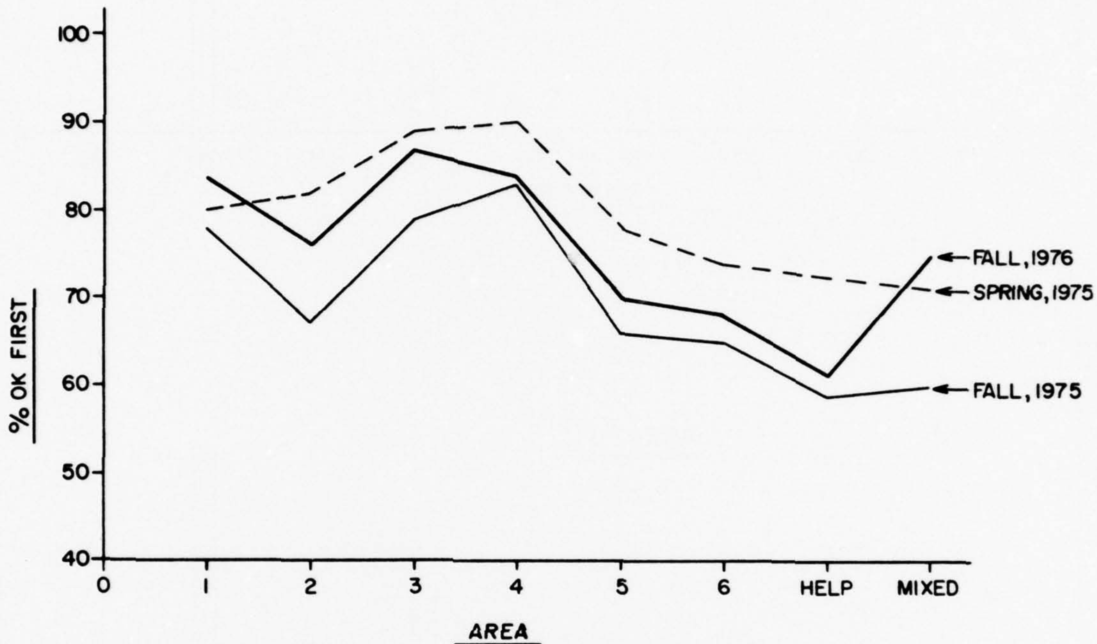


Figure 2

Performance of Classes on Three Versions of a Math Lesson

Although the consistency of the shift in performance might possibly have been caused by Fall 1975 students who were more intelligent than the Spring 1975 students, the reduction in the number of examples was strongly implicated as the culprit. In any case, the lesson's undesirably low performance suggested that it be modified, no matter what the cause might be. Therefore, one or more examples were added back into all teaching areas except area 4, in which students had scored higher than 80% okf for the revised lesson. Student-by-student analyses indicated the "help" section seemed to be aiding poor students as was intended, but the change in overall performance caused by the reduction in number of examples was so large that further interpretation was obscured.

Third version. Data for the third version was collected in Fall 1976. The data from the first two versions was collected from the same community college classes, whereas the data from the third version was gathered from junior high school students. Nevertheless, the results of the analysis suggest the lesson behaved similarly for all three groups of students (Figure 2).

The examples added to the third version seem to have raised the %okfirst for all areas except area 4. The fact that the number of examples in area 4 was not changed confirms the validity of the changes to other areas. It is interesting to note that although the %okfirst for this last version is generally lower than that of the first version, the performance on the final mixed test is higher. This is probably due to the "help" area, which was not available to students taking the original form of the lesson. The %okfirst for the help area is lower than all other areas because only students doing poorly in previous sections are routed through the help area. This extra help and practice in discrimination of problem types tends to keep the poorer students from "falling apart" on the final mixed exercise. The result is that the average score for the whole class, rather than slowly declining through the later areas, makes a sharp upward turn in the final test area. The class average of %okfirst on the final test area is near the average %okfirst of the previous six areas. This is where one would expect it to be if the difficulty with discrimination of problem types had been eliminated: the test samples items from each of the previous areas.

Looking at the individual student performances (not shown here), one sees that students who did very well initially and who thus avoided the "help" area received a <u>lower</u> score on the mixed practice than their average on the six problem types. Conversely, students that did poorly and got the review, did <u>better</u> on the mixed practice than their previous average. This suggests that a second briefer help section to teach "good" students to discriminate problem types would be an efficient way to further increase the average score.

<u>Summary</u>. While lesson revision should continue to be partially based on comments from students and instructors as well as author intuition, the preceding example demonstrates the power of basing revisions on student performance data. For example, the decision to re-introduce examples can be explained and justified to skeptical instructors by means of these data. If a decision were made that the student scores should be increased still further, the absence of data such as that shown here would make it unclear whether more examles or more practice should be added or which types of questions need the most emphasis. In all likelihood, increasing either examples or practice would increase final scores, but only by substantially increasing lesson completion time. As noted above, there is evidence that a

second brief help section for more of the students might substantially boost average scores. If so, that would be a very efficient revision in terms of both the students' and the author's time.

Measures of CBE Interaction

As was previously noted, area summaries provided by the computer for each student contain tallies of various of the student's activities. In the lesson revision example, however, none of the data was used in its raw form--instead a ratio or percentage was calculated. Because the number of questions varies from area to area, the number of questions answered correctly on the first attempt cannot be compared from one area to another. However, the ratio of "questions answered correctly on first try" to "total questions" can be used to construct a parameter which can be usefully compared. For convenience, the ratio is often expressed as a percentage.

The lesson revision example was based on the examination of a single parameter derived from area summary data. In fact, there are many more data available and many more parameters which can be derived. Potentially useful ratios (in addition to those already mentioned) include: questions/minute (a measure of the student's speed and/or the nature of the lesson), interactions/minute (where interactions is the sum of response attempts, on-line help accesses, and student-requested branches), and unanticipated responses/question (a measure of lesson "polish").

The reader can imagine many other measures which can be constructed by algebraically manipulating the parameters previously listed. At this time however, only a few ratios like %okfirst and errors/question have been investigated.

For the other ratios, little is known about the validity, generalizability, or normal range of values.

Features and Options for Data Analysis

The area analysis package automatically calculates many useful ratios and averages as well as allowing the user to construct his own. Other data can also be retrieved, manipulated, and displayed in a vast variety of formats by the data analysis routines. Rather than being exhaustive, this section lists examples of some of the especially useful or interesting options.

To aid the user in finding relationships among the student data, scatterplots and histograms may be formed. After a histogram has been displayed, it is often interesting to investigate the characteristics of data at the ends of the distribution. Therefore, one option allows retrieval of complete information for all records of a type specified by the user (e.g., all time data for any student who completed a specific area with three or fewer errors).

Tables of summary data are provided in three categories: all areas within a lesson, all students within an area, and all areas completed by one student. Each summary has a different focus, and inconsistencies in summary data may point to potential problems in areas, lessons, or students. An especially interesting visual display shows the progress of an individual student relative to the rest of his classmates (see Figure 3). An appropriate parameter (okf, time, etc.) is plotted for up to 18 consecutive areas. The mean and standard deviation for the class are visually indicated and the student's score is plotted over this "background." A student may complete an area several times if the results from his first attempt are unsatisfactory according to his

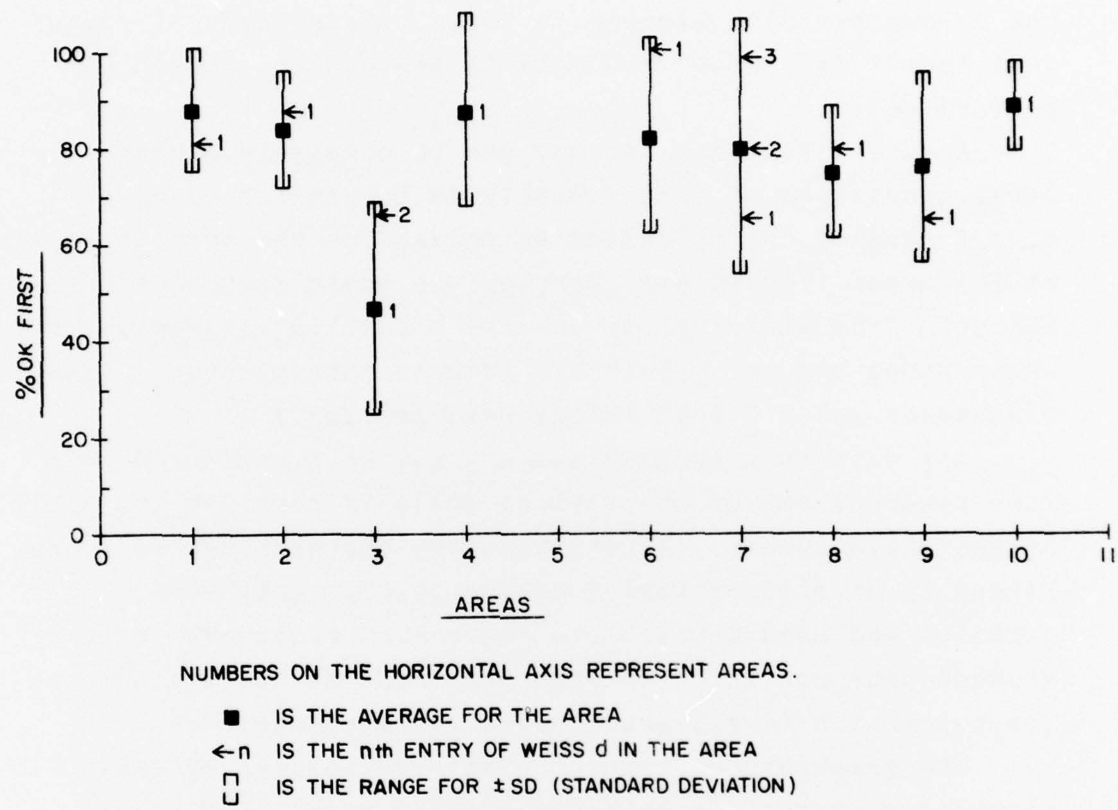


Figure 3

Summary Graph for One Student for One Parameter (%okfirst) own or the lesson's criteria. The results of each successive completion are indicated so that the student's improvement can be verified by the author or by the student's instructor. It was this kind of display that was used in the lesson revision example to follow the success of individual "good" and "weak" students to see how the lesson was performing.

Another technique one may use to concisely portray large quantities of data visually is to generate a cumulative frequency curve (called an ogive) for the completion time of the areas (Figure 4). Varying the scale factors reduces the data from different areas into a familiar S-shaped curve, whose shape and parameters may then be interpretted. (See discussion under "time" in the next section.)

Any data from an area summary may be transferred to a more powerful set of statistical analysis routines available on the PLATO system. After transfer, the data may be transformed by an arithmetical function (e.g., reciprocal, logarithmic) and used in multiple regression routines. Other student data not found in area summaries may be included in the regression (e.g., grade point average, age).

The graphical or tabular displays produced by the analysis packages can be hardcopied by means of a plasma display copying device (e.g., the Varian copier). In addition, most of the tables of data can be automatically formatted for a line printer.

Analysis of Military Data

The data analysis routines were refined and expanded as part of the ARPA/PLATO evaluation project in anticipation of their use during formative and summative evaluation periods. At the inception of the evaluation phase of this project,

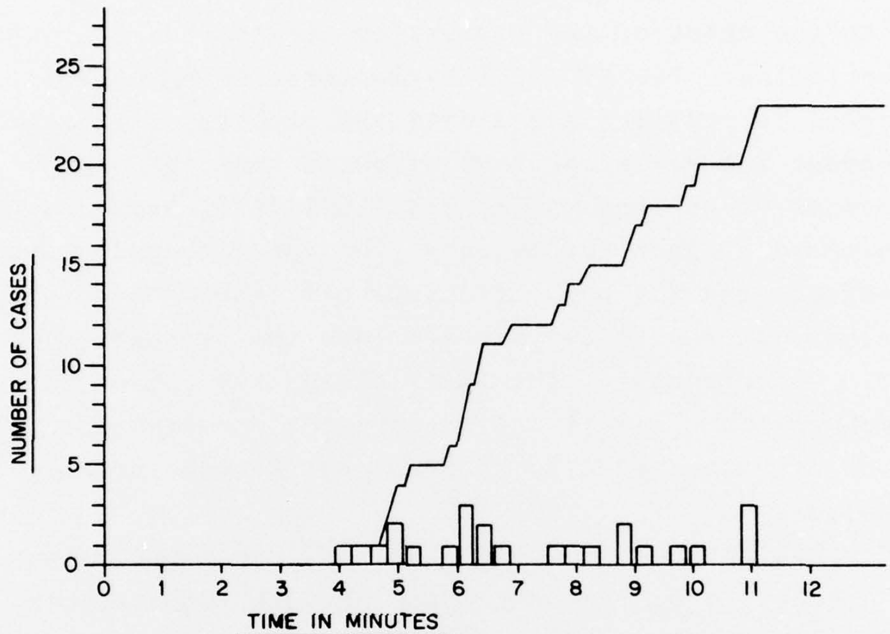


Figure 4

Cumulative Frequency Polygon for a Student's Area Completion Times the lessons at Aberdeen Proving Ground had been completely written and tested; students were no longer using them. At Chanute AFB, all lessons had been written and validated just prior to the start of the evaluation period, but student use was continuing. The MTC/PEER evaluators, being especially interested in studying and aiding the process of lesson development and revision, thus directed their efforts toward the lessons of another major ARPA/PLATO site, Sheppard AFB. The Sheppard project had begun a year later than the other two projects and the bulk of Sheppard's lesson development and validation was slated to fall into the evaluation period. Unfortunately, the small class size (16 or fewer for each of two classes) and the eventual re-direction of that project produced data that was sufficient for only a few analyses.

Foreseeing some of these problems, the MTC/PEER evalu- staff elected to collect data from Chanute AFB students. Because the Chanute staff did not use area summary information during validation, no provision had been made for its collection. Therefore MTC/PEER staff reprogrammed eight lessons we had previously reviewed, dividing each lesson into areas and installing other programming required for collection of area data.

Because we were forced to rely on the forbearance of Chanute staff for a number of data-handling procedures unrelated to their own goals, data collection could be maintained for only a limited period of time.

<u>Chanute data</u>. Because the Chanute AFB lessons had already met their validation criteria, major revisions based on the area summary data appeared unlikely. Therefore the analyses performed with these data were directed toward developing general hypotheses about how to use area data rather than toward diagnosing student or lesson problems.

The existence of large quantities of criterion-referenced test data, course test data, attitude surveys, and area summary data for Chanute students has provided a rich source for CBE investigation. Current research efforts make use of routines which extract relevant parameters from area data and allow them to be compared to a selection of predictors such as those suggested previously.

Sheppard data. We transferred and analyzed all data collected for both classes of Physician Assistant students--summaries for 256 areas from 62 lessons were obtained. In several cases there were sufficient data to serve as a basis for revision of lessons; however, the modified goals of the project plus a delay in the availability of the analyses meant that no revisions were actually made on the basis of this data. A post hoc analysis comparing the student performance of the first and second classes demonstrates that the student population shifted or that revisions to the lessons were largely successful. For example in 12 lessons where comparisons were possible, both the %okfirst and the errors/question improved dramatically in 10 cases and worsened slightly in 2 cases. As explained previously, further interpretations and analysis were not made because of the uncertain future for the lessons. Nevertheless, with the potential for efficient revision and validation demonstrated in this initial effort by MTC/PEER staff, the Sheppard PLATO group has now taken over the responsibility for gathering area data, analyzing it, and using it for formative development.

Preliminary Findings

Time. Initially, users of the area analysis package concluded that time is not generally an indication of student success. As an example they cited two students (Gerny and Lenerts in Table 2) who spent more time, did many more exercises than other student, and performed well. All students had the option to get more practice. These two chose to do so and apparently enjoyed it. However, student Adam took about as much time as Gerny and Lenerts, but did very poorly. Thus the elapsed time in an instructional area is seen to be a less-reliable indicator of success than %okfirst or errors/question. Other data, not shown here, indicate that in "testing" areas, as opposed to teaching or practice areas, fast students tend to perform better than students who take a lot of time to complete the test.

Investigations of cumulative frequency curves for time data (Tatsuoka & Tatsuoka, 1977) promise to supercede the early conclusion about the usefulness of time data cited above. The new research attempts to relate time and performance data by fitting student data to a mathematical model and equation. There is an indication that data which does not fit the model well may point out problems with a lesson segment or student population. For example, Chanute AFB data from a lesson whose data fit the model especially poorly was found to have at 37% failure rate (vs. the 10% or lower failure required for validation). Students taking a statistics lesson (originally written at an ARPA site) were divided into two groups based on their performance. The time data from the group with higher performance fit the theoretical distribution significantly better than that for the poorer group.

An attempt is being made to use time data to provide additional information when choosing whether to pass or fail students whose score is near the cutoff point of a criterion-referenced test. If one uses only the student's score, measurement errors inherent in testing mean that the decision

Student	Time	No. of ques.	No. of Okf	Antici-pated errors	Unanti-cipated errors
adams	27.6	56	22	24	47
baker	7.1	15	12	3	3
carlsen	2.2	12	12	0	0
douglas	2.9	12	12	0	0
eggers	3.1	12	11	0	1
franklin	3.0	12	12	0	0
gerney	26.9	146	137	0	8
harrison	3.5	6	5	0	1
ihold	1.9	2	2	0	0
jackson	2.4	2	2	0	0
king	5.1	8	5	0	1
lenerts	33.2	192	172	0	20
michelson	2.7	3	3	0	0

Table 2

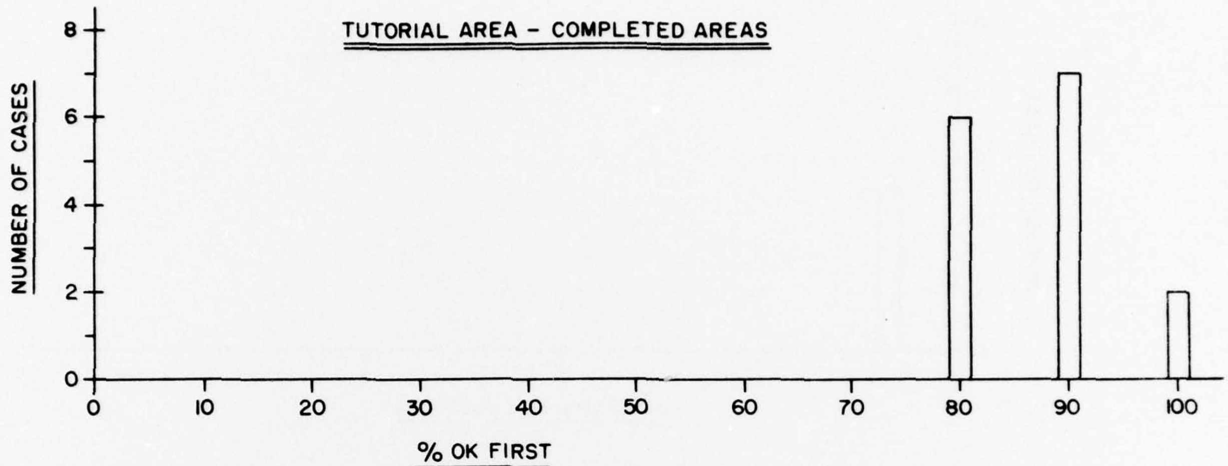
Student Parameters for an Area of a Math Lesson
(Note: student names are fictitious.)

to pass or fail some students near the cutoff point will be made incorrectly. Pass/fail decisions may be made more accurately by including additional data about the time needed by the student to complete the lesson and/or the lesson test. Final reports of the research, being supported by the Basic Skill Group of the National Institute of Education, will be available late in 1977.

Lesson strategies. If a lesson is divided into areas so that some of the areas contain mostly didactic, tutorial material with few opportunities to practice new skills while other areas consist primarily of practice with few help sequences, one will find a sharp contrast in the patterns of %okfirst and errors/question histograms. Figures 5 and 6 exemplify these differences. The distribution of both %okfirst and errors/question is broader for the practice area than for the tutorial area.

Lessons that teach by "trial and error" method, or lessons that mainly teach by help sequences have still different patterns of %okfirst. As could be predicted, inquiry-type teaching produces a higher error rate.

Use of incomplete area records. During the analysis of area data one must decide whether or not to include records of area data in which the student failed to complete the whole area. Available information suggests excluding such incomplete area records when searching for lesson problems. This is counter-intuitive: one would seemingly want to find cases where students were unable to finish an area because of problems. However, the data from incomplete areas is likely to be skewed so badly by short-time users that problems may be hidden, rather than elucidated, by including this data. When sufficient numbers of completed areas are not available, data from incomplete areas can be "cleaned" by discarding any students who have answered no questions in



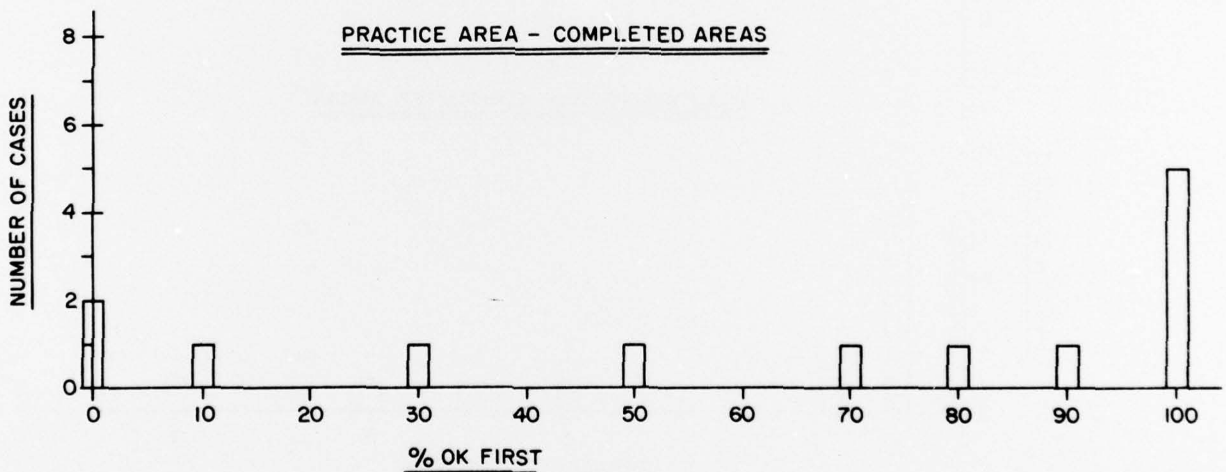
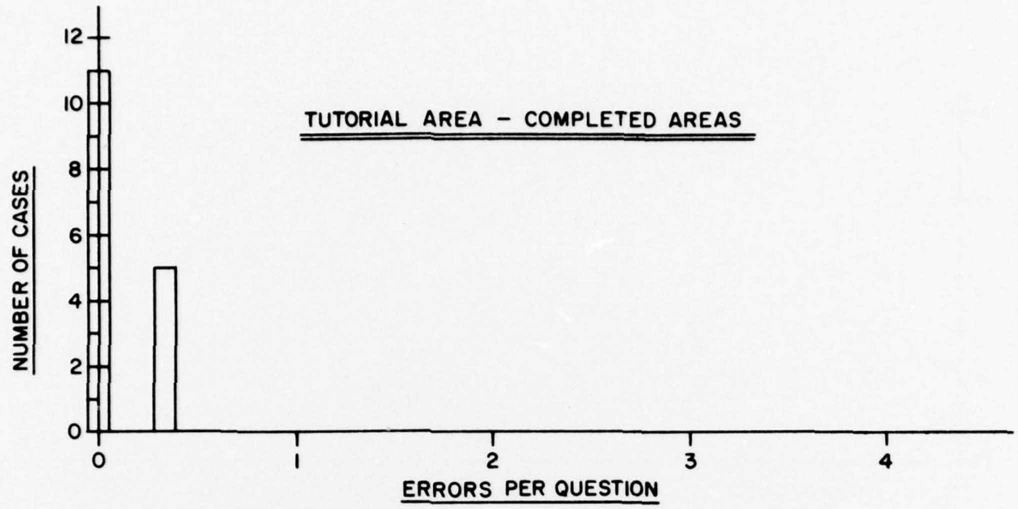


Figure 5

Comparison of %okfirst in Tutorial- and Practice-type Areas



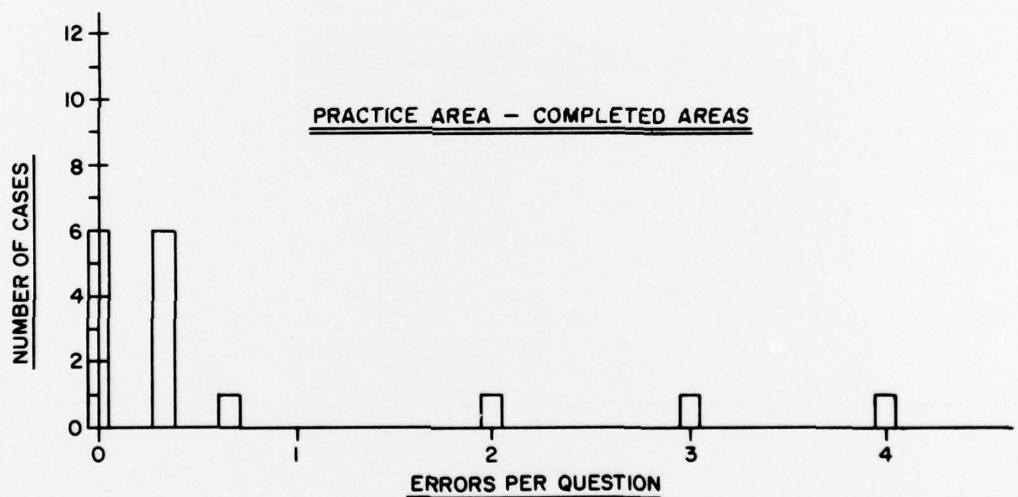


Figure 6

Comparison of Errors-per-Question in Tutorial- and Practice-type Areas an area or who have spent only a short time in the area. The following histogram (Figure 7) contrasts data collected from "complete only" and "all" area summary records. The shift to the right for the "complete only" uses is visually apparent and demonstrates typical differences observed when trying to interpret incomplete area records.

When calculating teaching efficiency or when viewing the progress of an individual student, all area data should be examined so that both successful and unsuccessful endeavors are profiled.

Future

At least three concurrent efforts are needed to advance this field of computer analysis of student interaction data:

1. New, meaningful measures must be found and validated.
2. Ways to use these measures for the formative development of lessons and for project management must be found.
3. Computer software packages to simplify the analysis task for all authors must be built, tested, and documented.

There are dozens of measures and parameters which seem meaningful, and potentially hundreds of correlations that might be examined. The exploration of normal values for these parameters and correlations has just begun. Intuition has served as the only guide until recently--now careful, systematic investigations are underway. A desireable result might be a set of algorithms for reducing the great volume of area data into a handful of consistent, sensitive measures that relate to the major characteristics of a lesson (effectiveness, efficiency, difficulty, etc.) Enough exploratory work has been done to demonstrate both immediate

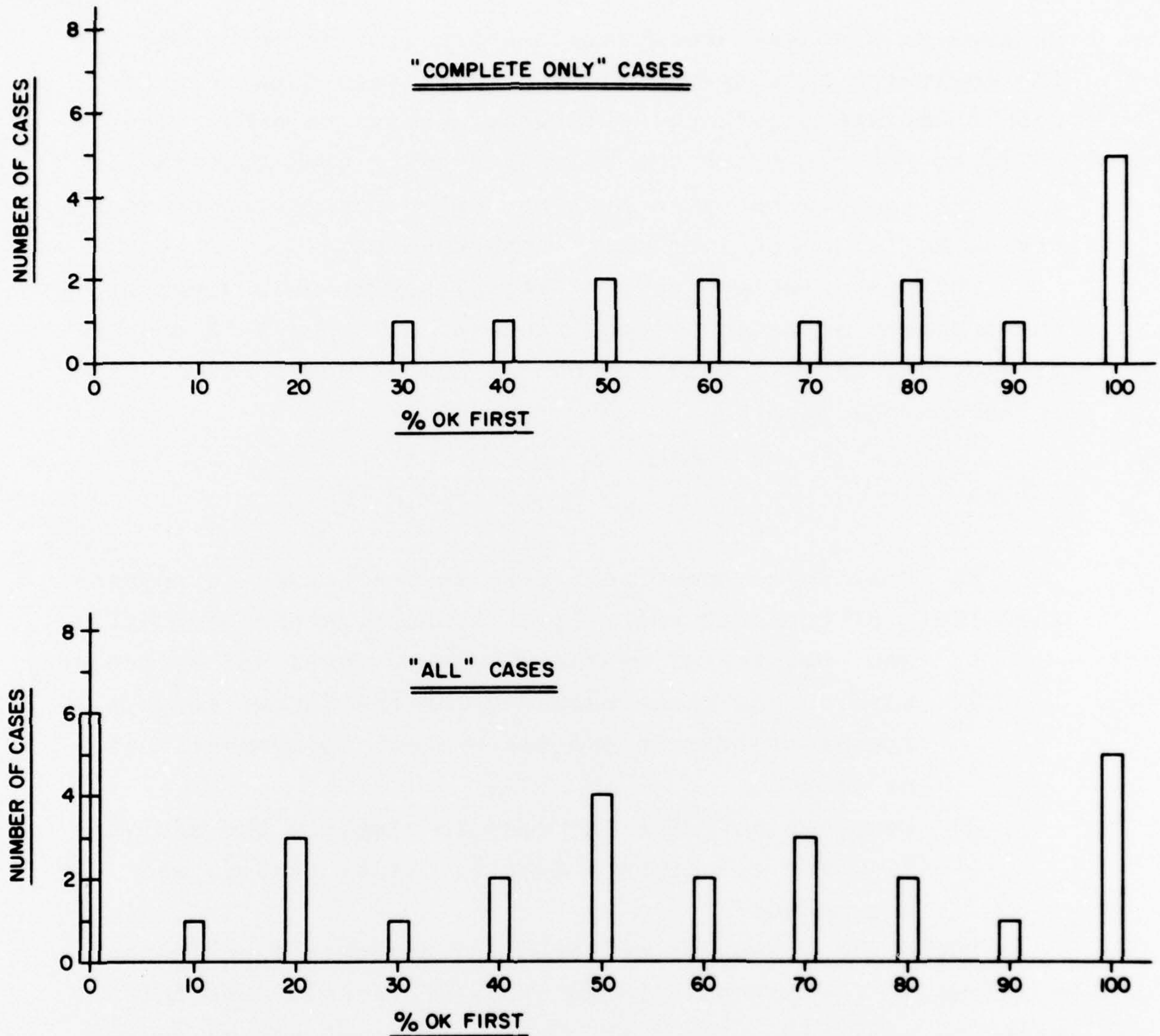


Figure 7

Comparison of %okfirst from "Complete only" vs. "All" Student Area Summaries applications and the potential for the future.

Current research at CERL involves regression analysis of the area data to predict lesson and course scores, and the analysis of elapsed time for teaching and testing to predict student success. Other users of area data are investigating the shape (i.e., mean, standard deviation, skewness) of distributions for parameters such as %okfirst to diagnose lesson weaknesses (Smith, Ghesquiere, & Avner, 1974).

A new manual describing the use of the data transfer and storage routines, and the interpretation of the data is in preparation by Weaver and Tatsuoka for publication in late 1977.

References

Avner, R. A. "SIRA--A System for Instructional Response Analysis." Paper presented to the Amer. Ed. Research Assn. Annual Meeting, Chicago, Ill., 1968.

Smith, S.G., Ghesquiere, J.R. & Avner, R.A. The use of computers in the teaching of chemistry. Chemical Education, 1974, 51, 243.

Smith, S.G., & Sherwood, B.A. Educational uses of the PLATO computer system. Science, 1976, 192, 344.

Tatsuoka K. & Siegel, M. Manual to Accompany "area" summary data and analysis package, Internal report. Computer-based Education Research Laboratory, University of Illinois, Urbana, Ill., 1975.

Tatsuoka, K. & Tatsuoka M. Time-score analysis in criterion-referenced tests. Progress report to Basic Skill group of the National Institute of Education for October 1, 1976-April 30, 1977. University of Illinois, Urbana, Ill., 1977.

Weaver, T. & Lederman B. A supplemental package for -qarea- (Chapter 5) in On-line Data Collection compiled by Dibello, L. V., Community College Math Group Interim Report, Computer-Based Education Research Laboratory, University of Illinois, Urbana, Ill., 1975.

Yeager, R. Student Performance Data and Security. Internal Report. Computer-Based Education Research Laboratory, University of Illinois, Urbana, Ill. 1976.

IED

-78