Ten elementary special education teachers in a rural educational cooperative implemented specific curriculum based measurement and data utilization procedures with at least two students each over one school year. Three data utilization strategies (no data utilization, therapeutic analysis, and experimental analysis) were compared in terms of their effects on the number of modifications teachers made in the students' programs, and student performance. Teacher preferences for therapeutic and experimental strategies, as well as for two measurement procedures (mastery and performance measurement) were examined also. Results indicated that teachers made more instructional changes and student performance increased more when specific data utilization strategies (therapeutic or experimental) were used. Further, teachers preferred therapeutic analysis over experimental analysis and performance measurement over mastery measurement. (Author)
INSTRUCTIONAL CHANGES, STUDENT PERFORMANCE, AND
TEACHER PREFERENCES: THE EFFECTS OF SPECIFIC
MEASUREMENT AND EVALUATION PROCEDURES

Lynn Fuchs, Caren Wesson, Gerald Tindal,
Phyllis Mirkin, and Stanley Deno

Institute for Research on Learning Disabilities
The Institute for Research on Learning Disabilities is supported by a contract (300-80-0622) with the Office of Special Education, Department of Education, through Title VI-G of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled students.

During 1980-1983, Institute research focuses on four major areas:

- Referral
- Identification/Classification
- Intervention Planning and Progress Evaluation
- Outcome Evaluation

Additional information on the Institute's research objectives and activities may be obtained by writing to the Editor at the Institute (see Publications list for address).

The research reported herein was conducted under government sponsorship. Contractors 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 the official position of the Office of Special Education.
Research Report No. 64

INSTRUCTIONAL CHANGES, STUDENT PERFORMANCE, AND TEACHER PREFERENCES: THE EFFECTS OF SPECIFIC MEASUREMENT AND EVALUATION PROCEDURES

Lynn Fuchs, Caren Wesson, Gerald Tindal, Phyllis Mirkin, and Stanley Deno

January, 1982
Abstract

Ten special education teachers in a rural educational cooperative implemented specific curriculum-based measurement and data-utilization procedures with at least two students each over one school year. Three data-utilization strategies (no data utilization, therapeutic analysis, and experimental analysis) were compared in terms of their effects on (a) the number of modifications teachers made in the students' programs, and (b) student performance. Teacher preferences for therapeutic and experimental strategies, as well as for two measurement procedures (mastery and performance measurement) were examined also. Results indicated that teachers made more instructional changes and student performance increased more when specific data-utilization strategies (therapeutic or experimental) were used. Further, teachers preferred therapeutic analysis over experimental analysis and performance measurement over mastery measurement. The implications of these findings for further development of measurement and evaluation procedures are discussed.
Instructional Changes, Student Performance, and Teacher Preferences: The Effects of Specific Measurement and Evaluation Procedures

The learning principles of educational psychology (Gagne, 1965) and behaviorism (Sidman, 1960; Skinner, 1938) provide a theoretical rationale for incorporating measurement and evaluation into instruction. Additionally, a merger between monitoring and instruction is mandated by federal law (PL 94-142) and supported by research investigating the impact of direct, continuous evaluation on student academic achievement (Bohanon, 1975; Crutcher & Hofmeister, 1975; Frumess, 1973; Lovitt, Shaff, & Sayre, 1970). Therefore, it appears necessary to develop specific measurement and evaluation systems that (a) satisfy the technical requirements of psychometric theory, (b) result in improved student achievement, and (c) can be incorporated efficiently into instructional methods. The Institute for Research on Learning Disabilities (IRLD) has developed procedures (Mirkin, Deno, Fuchs, Wesson, Tindal, Marston, & Kuehnle, 1981) that satisfy certain technical (Deno, Mirkin, Chiang, & Lowry, 1980; Deno, Mirkin, Lowry, & Kuehnle, 1980; Deno, Mirkin, & Marston, 1980; Deno, Mirkin, Robinson, & Evans, 1980; Fuchs & Deno, 1981; Fuchs, Tindal, & Deno, 1981; Marston & Deno, 1981; Marston, Lowry, Deno, & Mirkin, 1981), educational (Mirkin & Deno, 1979; Mirkin, Deno, Tindal, & Kuehnle, 1979), and efficiency requirements (Fuchs, Wesson, Tindal, Mirkin, & Deno, 1981).

In the present study, additional technical, educational, and logistical requirements of frequent, direct measurement and evaluation procedures were examined. First, this study sought to describe how
the introduction and use of data-utilization strategies affects (a) the number of modifications teachers make in their students' programs and (b) student performance. This investigation of the contribution of data utilization to teacher behavior and resulting student achievement is warranted because few studies have explored systematically whether measuring students or evaluating data accounts for improved student growth. Additionally, this study is important because other research has indicated that teachers who collect student performance data do not necessarily use those data to make instructional decisions (Baldwin, 1976; White, 1977).

The second purpose of this study was to assess the efficiency with which measurement and evaluation methods might be incorporated into instructional routines. Specifically, teachers' preferences for mastery versus performance measurement and their preferences for therapeutic versus experimental evaluation were examined.

Mastery and Performance Measurement

In performance measurement, the measurement task is a random sample of items from a large pool of material, and the goal is to improve the level of performance on that material. Figure 1 illustrates performance measurement. The abscissa represents school days and the ordinate represents the rate of performance on the measurement task; each data point represents the rate of performance on a given day. The line of best fit through the data depicts the student's rate of improvement in performance on the pool of material.
Figure 2 depicts mastery measurement. Here, the abscissa represents school days and the ordinate represents successive segments or objectives of the curriculum mastered; each data point represents the number of curriculum segments mastered on a given day. The line of best fit through the data points depicts the rate of student progress through the curriculum. The goal of repeated mastery assessment is to increase the student's rate of mastery in the curriculum. The teacher measures the student on a random sample of material from the current instructional curriculum unit until mastery is achieved, at which point (a) the student's graph registers that a curriculum unit has been mastered, (b) the student's level of instruction progresses to the next segment in the hierarchy, and (c) the pool of material on which the teacher measures the student also progresses to the next segment in the hierarchy.

----------------------------------------
Insert Figure 2 about here
----------------------------------------

Therapeutic and Experimental Analysis

In therapeutic data analysis, the objective is to ensure that a student's performance reaches a prespecified goal by a certain date. This goal may represent any reasonable performance level selected by the teacher. Or, in a more systematic fashion and in consonance with the principles of normalization (Wolfensberger, 1972), this goal may be a performance level commensurate with a student's mainstream peers or a level that represents a reduced discrepancy between the student's current performance and his/her age-grade appropriate level. This goal, designated as the static aim, is marked on the graph with an X at the intersection
of the desired performance level and the anticipated attainment date. Then, a line of desired progress, the dynamic aim, that connects the student's baseline median score with the static aim is drawn onto the graph. Throughout the delivery of instruction, data interpretation consists of the application of the following rule: If, on three consecutive days, student data are plotted below the dynamic aimline, then the program is judged ineffective and a change is introduced into the student's program.

In experimental data analysis, no student performance level and attainment date are specified. Instead, there is a general directive to improve continuously upon a student's current performance level. One assumes that only by implementing an unending series of program changes and by comparing the effects of the programs on student performance, can an effective, individual program emerge and be improved continuously over time (Deno & Mirkin, 1977). Therefore, program changes are introduced regularly and are treated as experimental hypotheses concerning their effect on a student's performance. The methods of time-series analysis (Sidman, 1960) are employed to summarize and interpret student performance data.

Method

Subjects

Subjects were 10 special education elementary resource teachers (2 male, 8 female) in a rural educational cooperative. These teachers were required by their special education director to implement continuous evaluation procedures and to participate in a series of studies conducted
by the IRLD. The teachers, whose experience in special education ranged from 0 to 10 years, implemented the procedures with elementary age students who had been placed in resource programs for varying amounts of the school day. These children were functioning dramatically below their peers in academic, language, and/or social areas.

**Procedure**

Teachers were trained during one week of full-day workshops prior to the school year and during bi-monthly, half-day workshops throughout the year. These workshops were conducted by IRLD staff and, prior to February, their focus was on training the teachers to (a) write curriculum-based IEPs, (b) create a curriculum-based measurement procedure including mastery and performance systems, (c) measure and graph student progress toward IEP goals, and (d) develop strategies to improve the feasibility of implementing the frequent measurement systems.

By February, each teacher had developed curriculum-based IEPs for at least two students and was measuring and graphing those students' reading performance at least three times each week. In February, the two data-utilization systems, experimental and therapeutic analysis, were introduced to the teachers. First, one-half of the teachers implemented experimental teaching; one-half implemented therapeutic teaching. After nine weeks, teachers switched data-utilization systems. Therefore, the study included three treatment phases: (1) no data-utilization system, (2) experimental analysis or therapeutic analysis, and (3) therapeutic or experimental analysis, whichever had not been implemented during the second phase.

Beginning in November and every two weeks thereafter, the following
data collection procedures were implemented, (a) IRLD staff inspected, for each teacher, two students' reading graphs and counted the number of instructional changes that had been introduced into these students' reading programs, and (b) teachers measured the student's oral reading rate correct on lists of kindergarten through third-grade words (K-3 Lists) randomly selected from the Core List of K-3 words in Basic Elementary Reading Vocabulary R Series (Harris & Jacobson, 1972).

At the end of the school year, teachers completed a survey that included items on the teachers' preferences for performance vs. mastery measurement and for experimental vs. therapeutic evaluation as well as the advantages and disadvantages of each strategy (see Survey Questions in Appendix).

Results

Effects of Data-Utilization

Number of changes introduced into students' programs. In the no data-utilization system phase, teachers introduced a total of one change in all of the students' programs. In the therapeutic teaching phase, they introduced a total of seven changes; in the experimental teaching phase, six changes were introduced.

Student performance. For each student, the median number of words correct per minute within each of the three phases and then the percentage of increase across phases were calculated. From the no data-utilization system phase to the first data-utilization phase, where one-half of the teachers implemented therapeutic analyses and one-half implemented experimental analyses, the students' oral reading rates increased an average of 38%. From the first to second data-utilization phases, when
Teachers switched data-utilization strategies, student performance increased an average of 24%. This lower percentage suggests that time or maturation alone does not explain the increase in student performance from the no-data-utilization phase to the first data-utilization phase.

Teacher Preferences

Mastery and performance measurement. Table 1 provides information concerning the frequency with which teachers preferred mastery or performance measurement in each of five subject areas. For reading in context, six of eight responding teachers preferred mastery measurement. In the other four subject areas, the majority of teachers preferred performance measurement.

Table 2 lists advantages and disadvantages of the two strategies cited by the teachers. Inspection of this table reveals that teachers cited a wide range of advantages and disadvantages for each strategy, with different teachers attributing some of the same advantages and disadvantages to both mastery and performance measurement. For mastery measurement, advantages cited more than once were: (a) easy to understand, (b) easy to correlate with instruction, and (c) easy to set realistic long-term goals. Disadvantages included: (a) not enough information provided, (b) difficult to compare performance to peers, and (c) difficult to see progress. For performance measurement, advantages included: (a) indicates actual performance on long-term goal, and (b) indicates effectiveness of program changes. The only disadvantage of performance measurement cited more than once was that
testing materials were not related directly to the student's program.

Experimental and therapeutic analyses. Table 3 presents the teachers' preferences for the two evaluation strategies given different purposes. The majority of teachers preferred the therapeutic approach for (a) monitoring progress toward IEP goals, (b) the ease of its use, (c) its efficiency, (d) a guide for when to change a student's instructional program, (e) the ease with which it could be described to parents and other teachers, (f) its more adequate representation of student performance, and (g) its overall usefulness. The experimental approach was preferred by most teachers as a guide for what to change in a student's instructional program.

Teachers also responded to the question: "If you were able to use the data utilization system of your choice, what would that system be?" One-half of the teachers indicated that they would prefer to use a combination of the experimental and therapeutic approaches; four teachers selected the therapeutic approach, and one teacher chose the experimental approach (see Table 4).

Discussion

The first purpose of this study was to describe how the introduction
and use of data-utilization rules affects the number of modifications teachers make in their students' programs and how the use of those rules affects student performance. The results demonstrated that with data-utilization rules, teachers more often used student performance data to modify students' programs. Additionally, with the use of data-utilization strategies and the concurrent increase in the number of program modifications introduced, student's reading performance improved.

The results of this study corroborate those of earlier investigations. The findings indicated that measuring and graphing students' performance (as in the first phase) does not insure that those data will be used to make instructional decisions or that students' academic growth will be maximized. The results suggested that data evaluation, and perhaps the use of specific data-utilization rules, may be critical in insuring that teachers will use data to adjust instructional programs and to increase the probability that students will realize their educational goals.

The second purpose of this study was to assess teachers' preferences for progress and performance measurement and for therapeutic and experimental analysis. Results indicated that, for spelling, written expression, and reading in isolation, teachers preferred performance measurement; for reading in context, they preferred mastery measurement. Although the teachers considered mastery measurement easier than performance measurement in some respects, they indicated that the information they received through performance measurement was more useful for determining progress and the effectiveness of student programs.
Apparently, then, despite the relative ease with which mastery measurement could be understood and could direct day-to-day instruction, teachers preferred performance over mastery measurement because it was more useful for making program effectiveness decisions. Thus, after a school year's experience with both measurement strategies, teachers preferred performance measurement; given this preference, performance measurement may be the more feasible measurement alternative.

With respect to evaluation strategies, the teachers overwhelmingly preferred therapeutic evaluation. They found the therapeutic approach more useful and efficient for monitoring progress toward IEP goals. They believed that it was a better indicator of when to alter a student's instructional program; it was easier to explain to parents and other school staff; and, it also better represented student performance. The teachers described therapeutic teaching as better for overall use. In fact, teachers chose the experimental approach only for the purpose of determining what aspect of the student's instructional program should be changed. Despite the teachers' overwhelming preference for therapeutic evaluation for specific purposes, five teachers indicated that some combination of the experimental and therapeutic approaches might be best. This finding may be attributed to the fact that therapeutic evaluation addresses the question of when, not what, to change in a student's program, and that teachers preferred experimental evaluation for determining what to change in an educational plan. For handicapped children, the question of what to change may be especially problematic, and this may have led some teachers to conclude that a combination of
the two strategies is optimal.

The manual "Procedures to Develop and Monitor Progress on IEP Goals" (Mirkin et al., 1981) describes an evaluation strategy that is a combination of the experimental and therapeutic techniques. As in the therapeutic approach, teachers are directed to draw the dynamic aimline on the graph. Then, as is common practice in analysis of time-series data, the slope of the student performance data is calculated and drawn on the graph. Decisions regarding the effectiveness of the student's instructional program are made based on the comparison between the slope of the actual data and the dynamic aim. Given teachers' interest in a combination of therapeutic and experimental strategies, future research should investigate the feasibility of this combined evaluation strategy.
References

Baldwin, V. Curriculum concerns. In M. A. Thomas (Ed.), Hey, don't forget about me. Reston, Virg.: Council for Exceptional Children, 1976.


<table>
<thead>
<tr>
<th>Subject Area</th>
<th>Preferred Strategy</th>
<th>Mastery</th>
<th>Performance</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading in Context</td>
<td></td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Reading Isolated Words</td>
<td></td>
<td>1</td>
<td>7</td>
<td>1</td>
</tr>
<tr>
<td>Spelling</td>
<td></td>
<td>1</td>
<td>6</td>
<td>2</td>
</tr>
<tr>
<td>Math</td>
<td></td>
<td>1</td>
<td>8</td>
<td>0</td>
</tr>
<tr>
<td>Written Expression</td>
<td></td>
<td>2</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Measurement Strategy</td>
<td>$N^a$ Advantages</td>
<td>$N^a$ Disadvantages</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>----------------</td>
<td>------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mastery</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>easy to understand</td>
<td>2</td>
<td>does not provide enough information</td>
<td></td>
</tr>
<tr>
<td></td>
<td>easy to correlate with instruction</td>
<td>2</td>
<td>difficult to compare performance with peers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>makes it possible to set realistic long-term goal</td>
<td>2</td>
<td>difficult to see progress</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>easy to implement</td>
<td>1</td>
<td>does not indicate achievement</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>allows one to see progress clearly</td>
<td>1</td>
<td>does not indicate effectiveness of plans</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>allows one to know when to progress</td>
<td>1</td>
<td>does not test over old material</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>tests on current material</td>
<td>1</td>
<td>tests only short-term memory</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>avoids memorization of material</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Performance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>indicates actual performance on long-term goal</td>
<td>3</td>
<td>tests on material unrelated to program</td>
<td></td>
</tr>
<tr>
<td></td>
<td>indicates effectiveness of program changes</td>
<td>1</td>
<td>more time consuming</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>easy to compare with peers</td>
<td>1</td>
<td>difficult to see progress</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>easy to explain to parents</td>
<td>1</td>
<td>easy to misjudge long-term goal</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>indicates when program change is necessary</td>
<td>1</td>
<td>more variability in graph</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>easy to understand</td>
<td>1</td>
<td>can be initially discouraging</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>easy to utilize data</td>
<td>1</td>
<td>difficult to determine how much material to skip</td>
<td></td>
</tr>
</tbody>
</table>

$^a$Frequency with which advantages or disadvantages were cited.
### Table 3

Numbers of Teachers Preferring Data Utilization Strategies for Different Purposes

<table>
<thead>
<tr>
<th>Purpose</th>
<th>Therapeutic Strategy</th>
<th>Experimental Strategy</th>
<th>No Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monitoring progress toward IEP goals</td>
<td>10</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Easy to use</td>
<td>8</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Efficient (takes less time)</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Best guide for when to change the student's instructional program</td>
<td>7</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Best guide for what to change in the student's instructional program</td>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>Easy to describe the procedure to parents and other teachers</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Most adequately represents a student's performance</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Overall use</td>
<td>9</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Table 4
Number of Teachers Preferring Data Utilization Strategies Given Free Choice

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Number of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Therapeutic Teaching</td>
<td>4</td>
</tr>
<tr>
<td>Experimental Teaching</td>
<td>1</td>
</tr>
<tr>
<td>A combination of experimental and therapeutic teaching</td>
<td>5</td>
</tr>
<tr>
<td>An entirely different system</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 1: Illustration of performance measurement.
Figure 2: Illustration of mastery measurement.
APPENDIX

Survey Questions

2) Given the following factors, which data utilization strategy do you prefer? (Check one for each factor.)

Therapeutic (Aimline)  Experimental

a. monitoring progress toward IEP goals

b. easy to use

c. efficient (takes less time)

d. best guide for when to change the student's instructional program

e. best guide for what to change in the student's instructional program

f. easy to describe the procedure to parents and other teachers

g. most adequately represents student performance

h. overall use

3) If you were able to use the data utilization system of your choice, what would that system be? (Check one.)

a. Experimental

b. Therapeutic

c. A combination of experimental and therapeutic

d. An entirely different system

If you checked c or d, how would your system be different?

______________________________

______________________________

______________________________
PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $3.00 per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor, IRLD, 350 Elliott Hall; 75 East River Road, University of Minnesota, Minneapolis, MN 55455.


Note: Monographs No. 1-6 and Research Report No. 2 are not available for distribution. These documents were part of the Institute's 1979-1980 continuation proposal, and/or are out of print.


Thurlow, M. L., & Greener, J. W. Preliminary evidence on information considered useful in instructional planning (Research Report No. 27). March, 1980.


Epps, S., McGue, M., & Ysseldyke, J. E. Inter-judge agreement in classifying students as learning disabled (Research Report No. 51). February, 1981.

Epps, S., Ysseldyke, J. E., & McGue, M. Differentiating LD and non-LD students: "I know one when I see one" (Research Report No. 52). March, 1981.


