American Educational Research Journal
Summer 1984, Vol. 21, No. 2, Pp. 449–460

The Effects of Frequent Curriculum-Based Measurement and Evaluation on Pedagogy, Student Achievement, and Student Awareness of Learning

LYNN S. FUCHS
Wheelock College

STANLEY L. DENO and PHYLLIS K. MIRKIN
University of Minnesota

This study examined the educational effects of repeated curriculum-based measurement and evaluation. Thirty-nine special educators, each having three to four pupils in the study, were assigned randomly to a repeated curriculum-based measurement/evaluation (experimental) treatment or a conventional special education evaluation (contrast) treatment. Over the 18-week implementation, pedagogical decisions were surveyed twice; instructional structure was observed and measured three times; students' knowledge about their learning was assessed during a final interview; reading achievement was tested before and after treatment. Analyses of covariance revealed that experimental teachers effected greater student achievement. Additional analyses indicated that (a) experimental teachers' decisions reflected greater realism about and responsiveness to student progress, (b) their instructional structure demonstrated greater increases, and (c) their students were more aware of goals and progress.

Principles of educational measurement (Glaser & Nitko, 1971) and psychology (Crow & Crow, 1963; Farnham-Diggory, 1972) provide a theoretical framework for integrating measurement and evaluation with in-

This research was supported by Contract #300-80-0622 between the Office of Special Education, Department of Education, and the University of Minnesota Institute for Research on Learning Disabilities.

struction. Properly conducted assessment generates data that may be useful in enhancing student achievement. With such information, teachers might (a) structure better teaching procedures (Salvia & Ysseldyke, 1981), and (b) provide specific feedback to help students recognize effective learning strategies (Bandura, 1982; Peckham & Roe, 1977; Rosswork, 1977).

In special education a merger between assessment and instruction has been mandated by federal law (PL 94-142, 1975), wherein teachers of handicapped pupils are required to specify Individualized Educational Programs (IEPs) that include procedures for assessing student progress toward goals. Substantive compliance with this law dictates that monitoring progress be ongoing so that instructional programs can be continuously assessed and improved (Deno & Mirkin, 1977).

Despite theoretical rationale and legislative mandate to integrate assessment with instruction, special educators appear to prefer unsystematic observation to objective measurement to assess student progress (Potter & Mirkin, 1982), and they express confidence in the accuracy of such assessments (Fuchs & Fuchs, in press). Unfortunately, evidence suggests that such observation often leads to spuriously optimistic judgments of achievement (Fuchs & Fuchs, in press). This finding is corroborated by research on other types of clinical judgment and decisionmaking (Einhorn & Hogarth, 1978), which indicates that clinicians are prone to experience great confidence in highly fallible, typically self-confirming judgments.

Such findings suggest the importance of integrating objective measurement and evaluation with instruction. To increase the likelihood that teachers will employ objective assessment to influence pedagogy, educational psychologists have developed comprehensive systems for ongoing testing and systematic evaluation (Bloom, Madaus, & Hastings, 1981; Keller, 1968; Lindsley, 1971; Lovitt, 1977; White & Haring, 1980). Research on the effects of such comprehensive models on the reading achievement of handicapped children is inconclusive. Some studies support the effectiveness of such systems (Bohannon, 1975; Lovitt & Hansen, 1976), and others find limited or no effects on student gains (Martin,1980; Tindal, Fuchs, Christenson, Mirkin, & Deno, 1981) or an inconsistent or poor relation between performance on specific skills measured during treatment and more global measures of student achievement (Mirkin & Deno, 1979; Quilling & Otto, 1971).

With respect to the reading skills of handicapped children, the purpose of the present investigation was to examine the educational effects of data-based program modification (DBPM; Mirkin et al., 1981), a repeated, systematic assessment system. DBPM differs from previously investigated models in three important ways. First, it requires practitioners to measure student performance on a long-range goal behavior rather than on a series of short-term objectives. Second, for generating and administering tests, DBPM has teachers adhere to strict guidelines with demonstrated psychometric and edumetric adequacy (Mirkin, Fuchs, & Deno, 1982). Most other models require teachers continuously to create their own testing materials with minimum guidance, a practice that may generate invalid data bases. Third, DBPM requires teachers to adhere to a strict data-utilization rule, with demonstrated edumetric adequacy and interjudge agreement (Mirkin et al., 1982), that dictates when instructional change is to be introduced into a student's program.

The present study, therefore, adds to the previous data base on the effects of frequent testing and continuous evaluation by employing a different and, in some ways, a more prescriptive and technically adequate system. The current investigation also differs from and enhances the previous data base because of two methodological dimensions. First, concurrent with examining effects on student achievement, it examined additional outcome measures of pedagogy and students' knowledge of their own learning. Second, in contrast to most previous studies, treatment implementation in the present investigation was long term.

METHOD

Subjects

Subjects were 39 New York City Public School special education teachers who volunteered to participate in a project in which they would receive inservice training. Teachers were based in seven buildings, with four to seven teachers per school. Within each school, teachers were assigned randomly to experimental and contrast groups, and each teacher then selected three or four pupils for this project.

In the experimental group, teachers (3 men, 15 women) had taught special education for an average 3.79 years ($SD = 2.85$). The students' (51 boys, 13 girls) age-appropriate grade level averaged 5.79 ($SD = 1.66$); 49% were placed in programs for emotionally handicapped students, 32% in programs for brain-injured students, and 19% in resource programs.

Contrast teachers (2 men, 19 women) had taught handicapped children for an average 3.59 years ($SD = 2.72$). The contrast students' (57 boys, 20 girls) age-appropriate grade level averaged 5.45 ($SD = 1.65$); 52% were placed in programs for emotionally handicapped students, 32% in resource programs, and 15% equally distributed across programs for brain-injured, physically handicapped, and educable mentally retarded children.

Statistical tests revealed that experimental and contrast groups were similar with respect to teachers' sex and experience as well as students' sex and grade level. However, there was a relation between treatment group and the distribution of children among program types, $\chi^2(4) = 24.31$, $p < .001$, with a much greater percentage of brain-injured children in the experimental group.

Measures

Passage Reading Test (PRT). The PRT (Fuchs, Deno, & Mirkin, 1982a), comprising three passages from a third grade book of the Ginn 720 series (1976), was employed. The test requires students to read aloud from each passage for 1 minute, and student performance is reported in terms of average numbers of correct words and errors read. Test-retest reliability ranged from .93 to .96 (Fuchs, Deno, & Marston, in press). Concurrent validity with respect to the Woodcock Reading Mastery Tests, Word Identification and Passage Comprehension Tests, ranged between .54 and .92 (Fuchs, 1981). Internal consistency reliability (Cronbach's alpha) obtained in this study for the three-passage test was between .66 and .79.

Stanford Diagnostic Reading Test. Two subtests, Structural Analysis (SA) and Reading Comprehension (RC), of the Stanford Diagnostic Reading Test (Karlsen, Madden, & Gardner, 1976), Green Level, Form A, were employed. The SA subtest measures a child's decoding skills through the analysis of word parts. Internal consistency reliability for the SA was between .93 and .95 for grades 3 to 5. Criterion-related validity with the reading tests of the Stanford Achievement Test ranged from .62 to .85.

The RC subtest assesses literal and inferential comprehension with brief reading passages presented in a multiple-choice cloze format and short passages followed by multiple-choice questions. For the RC, internal consistency reliability was .96 at grades 3 through 5. Criterion-related validity with the reading tests of the Stanford Achievement Test ranged from .68 to .90.

Structure of Instruction Rating Scale (SIRS). The SIRS (Deno, King, Skiba, Sevcik, & Wesson, 1983) was employed to measure the structure of instruction. An observer completes the scale following a 20-minute observation of teacher interaction with one student. Eleven variables (instructional grouping, teacher-directed learning, active academic responding, demonstration/prompting, controlled practice, frequency of correct answers, corrections, positive consequences, pacing, oral reading practice, and silent reading practice) constitute the scale, each selected because of its relation to academic achievement (Reith, Polsgrove, & Semmel, 1981; Stevens & Rosenshine, 1981). Internal consistency reliabilities (Cronbach's alpha) obtained in this study for the 11-item scale were between .88 and .89.

Teacher questionnaire. A teacher questionnaire (see Fuchs, Deno, & Mirkin, 1982b) designed for the study requires teachers to report (a) the adequacy of students' progress toward reading goals, (b) whether they have made changes in reading goals, and (c) descriptions of students' current levels of functioning in reading.

Student interview. An interview schedule, designed for the study (see

Fuchs et al., 1982a), questions students about their reading progress and goals and requires interviewers to judge the accuracy of student responses.

Procedure

Treatments. DBPM is described in *Procedures To Develop and Monitor Progress on IEP Goals* (Mirkin et al., 1981). Employing these procedures in the area of reading, the experimental teachers first wrote curriculum-based IEP goals and objectives. The annual goal specified the segment of the curriculum and the date on which a student would read at a certain proficiency. The objective contained supplementary information indicating the weekly rate at which the student would improve in order to meet the goal. Teachers then developed curriculum-based measurement systems to match goals. They measured students' oral reading performance at least twice weekly, for 1 minute, on a randomly selected passage from the goal material. Then, teachers graphed student performance and implemented a data-utilization rule that required them to introduce a program change whenever a student's improvement across 7 to 10 measurement points appeared to be inadequate for goal attainment.

Contrast teachers set IEP goals and monitored progress as they wished. These teachers reported that they relied predominantly on periodic teacher-made tests, nonsystematic observation, and workbook exercises to assess goal mastery (Fuchs et al., 1982b).

Training. Each of three teacher trainers (TTs) was assigned to a set of schools and to the experimental and contrast teachers within that set of schools. TTs provided training to teachers weekly in one-to-one sessions from November through May. During these sessions, TTs taught DBPM to the experimental teachers. They taught the contrast teachers strategies for diagnosing and treating learning and social behavior problems, for structuring and managing their instructional groups, and for using audio-visual equipment and paraprofessionals. TTs met individually with experimental teachers an average of 23.52 hours ($SD = 5.95$) and with contrast teachers an average of 20.60 hours ($SD = 6.22$). A t test on the difference between the training hours for the two groups revealed no statistically significant difference.

Data collection. Each student was tested individually on the PRT pre- and posttreatment and in groups of four to eight on the SA and RC at the end of the study. All examiners were trained in standard testing procedures.

Data collectors were taught and practiced the SIRS during a 5-hour session, with interobserver agreement[1] on two training tapes at .87. During

[1] Interobserver agreement was calculated using the following formula (Coulter in Thompson, White, & Morgan, 1982): Percentage agreement = agreements between observer A & observer B ÷ agreements between A & B + disagreements between A & B + omissions by A + omissions by B.

each trimester of the study, data collectors observed each teacher providing reading instruction to the same randomly selected subject and then completed the SIRS. Interobserver agreement, calculated on 11% of the observations, was .84.

Each teacher completed the questionnaire twice, once during the second and once during the third trimester of the study. The student interview was administered at the end of the study to a random sample of 30 students. Interrater agreement[1] on the last part of the interview, which required data collectors to judge the accuracy of student responses, was assessed on 33% of the interviews at .90.

Data Analysis

Student achievement data. Because teachers, rather than students, were assigned randomly to treatments, "teacher" was designated the experimental unit of analysis. An F test (see Table I) computed on the pretest error scores revealed a statistically significant difference, $F(1,37) = 4.12, p < .05$. The reading posttest scores therefore were analyzed with a multivariate and follow-up univariate two-way analyses of covariance. The experimental factor was measurement/evaluation treatment (experimental vs. contrast), the blocking factor was teacher trainer (1 vs. 2 vs. 3), and the covariate was the pretest error score. Before using the analysis of covariance, the assumption of homogeneous regression coefficients was tested and scattergrams were inspected; the assumptions of equal regression slopes and linearity of Y on X appeared tenable.

Teacher decisionmaking data. Teacher responses to survey items were analyzed with chi-square statistics. The two dimensions of each contin-

TABLE I

Means, Standard Deviations, Adjusted Means, ANCOVA Results, and Binomial Effect Size Displays (r) on Pre- and Posttest Variables

	Measurement/evaluation treatment						F test		
	Experimental			Contrast					
	$\bar{X}$	SD	Adjusted mean	$\bar{X}$	SD	Adjusted mean	F	df	r
Pretest									
Words correct per minute	41.58	33.72		51.51	40.33		.689	(1,37)	
Errors per minute	8.19	3.26		6.24	2.74		4.121*	(1,37)	
Posttest									
Words correct per minute	70.23	45.75	78.93	51.29	38.18	43.81	4.221*	(1,32)	.34
Errors per minute	5.63	2.08	5.21	5.64	2.75	5.99	.001	(1,32)	.01
SA	39.79	12.08	40.74	29.65	15.34	28.84	7.194*	(1,32)	.43
RC	43.95	10.52	45.37	33.02	15.39	31.81	4.223*	(1,32)	.34

* $p < .05$.

gency table were (a) measurement/evaluation treatment with two levels and (b) the item on the survey with two to three levels (depending on the response format). Composite SIRS scores were subjected to a one between (measurement/evaluation treatment), one within (trials on SIRS) analysis of variance. Then the average SIRS change scores and the student achievement posttest data were entered into partial correlations with pretest error scores as the control variable.

Student knowledge data. Responses on the end-of-year student interview were analyzed with chi-square tests. Measurement/evaluation factor and the item on the survey were the two dimensions of each contingency table.

RESULTS

Student Achievement

The multivariate two-way analysis of covariance was conducted on the four posttest reading variables. Wilks' lambda criterion was used to test for equality of group centroids. The values calculated with the Wilks' lambda procedure were transformed into F values through Rao's approximation (Cooley & Lohnes, 1962). The tests for lambda produced significant F values for the measurement/evaluation factor, $F(4,29) = 3.80$, $p < .02$, and the teacher trainer factor, $F(8,58) = 1{,}300$, $p < .001$, and a nonsignificant F value for the measurement/evaluation × teacher trainer interaction, $F(8,58) = 1.86$, *ns.*

Follow-up univariate two-way analyses of covariance (ANCOVAs) revealed statistically significant differences between the measurement/evaluation conditions on three dependent variables. Table I displays relevant descriptive data, ANCOVA results, and binomial effect size displays (Rosenthal & Rubin, 1982).

The ANCOVAs included teacher trainer as a blocking factor only to control a known source of error. Given the absence of a statistically significant interaction between the measurement/evaluation and teacher trainer factors on the multivariate analysis, further discussion of the teacher trainer conditions would be extraneous to the purpose of this paper.

Teacher Decisionmaking

SIRS. The one between, one within analysis of variance on the composite SIRS scores revealed no significant main effects, but a significant measurement/evaluation treatment × trials interaction, $F(2,69) = 6.57$, $p < .01$. Across the study trimesters, mean composite scores increased (2.31, 2.76, 2.98) for the experimental teachers and decreased (2.82, 2.52, 2.34) for the contrast teachers. When the average increase score between subsequent SIRS trials and the mean student achievement posttest data for each teacher were correlated with the pretest reading error scores serving as the control variable, two correlations were statistically significant: the change score with SA, $r = .37$, $p < .02$, and with RC, $r = .27$, $p < .05$.

Teacher survey. In the second and third trimesters of the study, measurement/evaluation treatment was related to teacher judgments that their students had made sufficient progress to reach goals, $\chi^2(2) = 6.70$, $p < .05$ for April, and $\chi^2(2) = 7.47$, $p < .025$ for June. Greater percentages of contrast teachers reported that their students would meet goals or that they were uncertain, whereas a greater percentage of experimental teachers reported that their students would not meet goals.

With respect to whether teachers had adjusted students' reading goals, in April a statistically significant relation existed with measurement/evaluation treatment, $\chi^2(1) = 4.14$, $p < .05$; experimental group teachers tended to change goals more than the contrast teachers. By June, the relation no longer existed.

Teachers were asked to write precise statements to describe students' current functioning levels. In April there was no relation between the specificity of their statements and the measurement/evaluation treatment. By June, however, greater percentages of experimental teachers' descriptions were rated as "highly" or "somewhat" specific and greater percentages of contrast teachers' descriptions were rated as "not at all" specific, $\chi^2(2) = 7.01$, $p < .05$.

Student Knowledge About Learning

There was a statistically significant relation between measurement/evaluation treatment and (a) students saying they knew their goals, $\chi^2(1) = 4.89$, $p < .05$; (b) students actually stating their goals, $\chi^2(1) = 4.89$, $p < .05$; and (c) students' accuracy in judging whether they would meet their goals, $\chi^2(1) = 4.89$, $p < .05$. Greater percentages of experimental pupils said they knew their goals, actually stated their goals, and were correct when they judged whether they would attain their goals. In addition, when asked how they knew whether they would meet their goals, experimental students tended to say that they relied on their graphed data, whereas contrast students tended to say that they "just thought so," $\chi^2(1) = 9.18$, $p < .005$.

DISCUSSION

Children whose teachers employed the ongoing measurement and evaluation system, DBPM, achieved better than students whose teachers used conventional monitoring methods, such as periodic teacher-made tests, informal observation, and workbook samples. Superior reading progress was evidenced by the experimental students not only on the passage reading test, a measure similar to the task that had been practiced during the study, but also on decoding and comprehension measures. This suggests that teachers who repeatedly employ a simple passage reading test as an index of student progress might interpret gains as representing more general achievement, including fluency, decoding, and comprehension.

Other investigations of comprehensive repeated measurement and evaluation systems typically have failed to evidence superior student growth on measures more global than those used for testing during the investigation. The generalized growth demonstrated in this study may have been either a function of the relatively long treatment or due to dimensions of DBPM that differentiate it from other systems, including (a) student measurement on the long-term goal behavior, (b) a focus on psychometrically acceptable testing, and (c) a comparatively prescriptive and edumetrically sound data-utilization rule.

An alternative explanation for the differential student achievement, as well as for results to be discussed below, is the overrepresentation of brain-injured children in the experimental group. However, the plausibility of this alternative interpretation is limited, because evidence (Reed, Rabe, & Mankinen, 1968) suggests that the category of brain injury (a) is defined and identified inconsistently, and (b) provides no implications for teaching reading. A second competing hypothesis is the difference in training hours allotted to the treatment groups. Although of no statistical difference, the mean 2.92 hours of extra training received by experimental teachers may, at least partially, explain results.

Concurrent with better student achievement, results suggested that DBPM affected pedagogy. Whereas contrast teachers' structure decreased, experimental teachers increased their structure. Such increases were related to posttest scores on decoding and comprehension tests, a finding that corroborates previous work suggesting that increased structure contributes to handicapped children's achievement (Reith et al., 1981; Stevens & Rosenshine, 1981).

In addition, experimental teachers appeared to be more realistic about and responsive to student progress. In judging student growth, contrast teachers were more uncertain, which may have been due to a relatively scant data base, and they were more optimistic. Given their students' relatively poor progress, such optimism would appear to be unfounded and may corroborate work (Einhorn & Hogarth, 1978; Fuchs & Fuchs, in press) suggesting a tendency for clinicians and educators to overestimate their effects without systematic assessment. Contrast teachers also tended to be less specific in describing students' current performance levels and to maintain established goals. Given the difficulty of establishing initially appropriate goals (Fuchs & Deno, 1982), experimental teachers' more frequent goal revision may have been necessary and may have signaled more accurate assessment of and greater responsiveness to students' current performance levels.

Finally, along with better student achievement and pedagogy, students were more knowledgeable about their own learning as a result of the systematic measurement and evaluation treatment. In comparison to pupils whose progress was measured and evaluated by conventional special education practice, experimental students (a) more frequently claimed they knew their goals, (b) more often stated their goals, (c) were more accurate in their estimates of whether they would meet their goals, and (d) more typically reported that they relied on data to formulate estimates of whether they would meet goals. These findings are theoretically and socially important. On the one hand, they support the hypothesis of educational psychologists (Bandura, 1982; Crow & Crow, 1963; Farnham-Diggory, 1972) and accumulating evidence (Locke, 1968; Rosswork, 1977) that student knowledge of goals may affect school performance. On the other hand, increased participation by students in their own education may itself be an important educational goal.

REFERENCES

BANDURA, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist, 37*(2), 122–147.

BLOOM, B. S., MADAUS, G. F., & HASTINGS, J. T. (1981). *Evaluation to improve learning.* New York: McGraw-Hill.

BOHANNON, R. (1975). *Direct and daily measurement procedures in the identification and treatment of reading behaviors of children in special education.* Unpublished doctoral dissertation, University of Washington.

COOLEY, W. W., & LOHNES, R. R. (1962). *Multivariate procedures for behavioral sciences.* New York: John Wiley & Sons.

CROW, L. D., & CROW, A. (1963). *Readings in human learning.* New York: David McKay.

DENO, S. L., KING, R., SKIBA, R., SEVCIK, B., & WESSON, C. (1983). *The structure of instruction rating scale (SIRS): Development and technical characteristics* (Research Rep. No. 107). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

DENO, S. L., & MIRKIN, P. K. (1977). *Data-based program modification: A manual.* Reston, VA: Council for Exceptional Children.

EINHORN, H. J., & HOGARTH, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. *Psychological Review, 85*(5), 395–416.

FARNHAM-DIGGORY, S. (1972). *Cognitive processes in education: A psychological preparation for teaching and curriculum development.* New York: Harper & Row.

FUCHS, L. S. (1981). *The concurrent validity of progress measures of basal reading material.* Unpublished doctoral dissertation, University of Minnesota.

FUCHS, L. S., & DENO, S. L. (1982) *Developing goals and objectives for educational programs.* Washington, DC: American Association of Colleges for Teacher Education.

FUCHS, L. S., DENO, S. L., & MARSTON, D. (in press). Improving the reliability of curriculum-based measures of academic achievement for psychoeducational decision making. *Diagnostique.*

FUCHS, L. S., DENO, S. L., & MIRKIN, P. K. (1982a). *Direct and frequent measurement and evaluation: Effects on instruction and estimates of student progress (Research Rep. No. 97). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.*

FUCHS, L. S., DENO, S. L., & MIRKIN, P. K. (1982b). *Effects of frequent curriculum-based measurement and evaluation on student achievement and knowledge of performance: An experimental study* (Research Rep. No. 96). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

FUCHS, L. S., & FUCHS, D. (in press). Criterion-referenced assessment without measurement: How accurate for special education? *Remedial and Special Education.*

GINN AND COMPANY. (1976). *Reading 720.* Lexington, MA: Ginn (Xerox Corp.).

GLASER, R., & NITKO, J. (1971). Measurement in learning and instruction. In R. Thorndike (Ed.), *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.

KARLSEN, B., MADDEN, R., & GARDNER, E. F. (1976). *Stanford Diagnostic Reading Test.* New York: Harcourt Brace Jovanovich.

KELLER, F. S. (1968). Goodbye, teacher.... *Journal of Behavior Analysis, 1,* 79–89.

LINDSLEY, O. R. (1971). Precision teaching in perspective: An interview with Ogden R. Lindsley. *Teaching Exceptional Children, 3*(3), 114–119.

LOCKE, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance, 3,* 157–189.

LOVITT, T. C. (1977). *In spite of my resistance, I've learned from children.* Columbus, OH: Charles E. Merrill.

LOVITT, T. C., & HANSEN, C. L. (1976). The use of contingent skipping and drilling to improve oral reading and comprehension. *Journal of Learning Disabilities, 9*(8), 20–26.

MARTIN, M. A. (1980). *A comparison of variations in data-utilization procedures on the reading performance of mildly handicapped students.* Unpublished doctoral dissertation, University of Washington.

MIRKIN, P. K., & DENO, S. L. (1979). *Formative evaluation in the classroom: An approach to improving instruction* (Research Rep. No. 10). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities. (ERIC Document Reproduction Service No. ED 185 754)

MIRKIN, P. K., DENO, S. L., FUCHS, L. S., WESSON, C., TINDAL, G., MARSTON, D., & KUEHNLE, K. (1981). *Procedures tò develop and monitor progress on IEP goals.* Minneapolis: University of Minnesota.

MIRKIN, P. K., FUCHS, L. S., & DENO, S. L. (Eds.). (1982) *Considerations for designing a continuous evaluation system: An integrative review* (Monograph No. 20). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

PECKHAM, P. D., & ROE, M. D. (1977). The effects of frequent testing. *Journal of Research and Development in Education, 10*(3), 40–50.

POTTER, M., & MIRKIN, P. K. (1982). *Instructional planning and implementation* *practices of elementary and secondary resource room teachers: Is there a difference?* (Research Rep. No. 65). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

QUILLING, M., & OTTO, W. (1971). Evaluation of an objective based curriculum in reading. *Journal of Educational Research, 65*(1), 15–18.

REED, J. C., RABE, G. F., & MANKINEN, M. (1968). Teaching reading to brain-damaged children: A review. *Reading Research Quarterly, 1,* 289–298.

REITH, H. J., POLSGROVE, L., & SEMMEL, M. I. (1981). Instructional variables that make a difference: Attention to task and beyond. *Exceptional Education Quarterly, 2*(3), 61–72.

ROSENTHAL, R., & RUBIN, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology, 74*(2), 166–169.

ROSSWORK, S. G. (1977). Goal setting: The effects on an academic task with varying magnitudes of incentive. *Journal of Educational Psychology, 69*(6), 710–715.

SALVIA, J., & YSSELDYKE, J. E. (1981). *Assessment in special and remedial education* (2nd ed.). Boston: Houghton-Mifflin.

STEVENS, R., & ROSENSHINE, B. (1981). Advances in research on teaching. *Exceptional Education Quarterly, 2*(1), 1–10.

THOMPSON, R. H., WHITE, K. R., & MORGAN, D. P. (1982). Teacher-student interaction patterns in classrooms with mainstreamed mildly handicapped students. *American Educational Research Journal, 19,* 220–236.

TINDAL, G., FUCHS, L. S., CHRISTENSON, S., MIRKIN, P. K., & DENO, S. L. (1981). *The relationship between student achievement and teacher assessment of short- or long-term goals* (Research Rep. No. 61). Minneapolis: University of Minnesota, Institute for Research on Learning Disabilities.

WHITE, O. R., & HARING, N. G. (1980). *Exceptional teaching* (2nd ed.). Columbus, OH: Charles E. Merrill.

AUTHORS

LYNN S. FUCHS, Assistant Professor, Departments of Psychology and Education, Wheelock College, 45 Pilgrim Road, Boston, MA 02215. *Specializations:* Instructional design and development, special education.

STANLEY L. DENO, Professor of Educational Psychology and Special Education, College of Education, University of Minnesota, 178 Pillsbury Drive, S.E., Minneapolis, MN 55455. *Specializations:* Learning and instruction, teacher education for moderately handicapped students.

PHYLLIS K. MIRKIN. Dr. Mirkin died in May 1982. She was Associate Director, Institute for Research on Learning Disabilities, University of Minnesota, 75 E. River Road, Minneapolis, MN 55455. *Specializations:* Learning disabilities, instruction, formative evaluation. We regret her untimely death and thank her for her contribution to *AERJ.*