7-1-1985

Learning from Experience to Improve Outcomes in Reading: A Case Study

Annette B. Weinshank
Michigan State University

Ruth M. Polin
Michigan State University

Christian C. Wagner
Oakland University

Follow this and additional works at: https://scholarworks.wmich.edu/reading_horizons

Part of the Education Commons

W
WESTERN MICHIGAN UNIVERSITY
Dorothy J. McGinnis
Reading Center and Clinic

LEARNING FROM EXPERIENCE TO IMPROVE OUTCOMES IN READING: A CASE STUDY

Annette B. Weinshank and Ruth M. Polin
Michigan State University

Christian C. Wagner
Oakland University

How can teachers learn more effectively from their own practice? Experience may indeed be a valuable teacher but it has a serious limitation: classroom events happen too quickly for the practitioner reliably to aggregate the results of instruction in order to profit from them. Isolation and overreliance on personal experience further restrict the range of necessary information available to a teacher that can be used to evaluate instructional decisions and actions in light of student achievement outcomes.

In the teaching of reading, the absence of specific information linking student outcomes back to reliable diagnosis and instruction has presented a major obstacle to learning systematically from one's own experience. Improving student outcomes in reading requires that the teacher receive information that reliably classifies student performance into the appropriate diagnostic categories. Assurance of diagnostic reliability is central to evaluating one's instructional practices. If the diagnostic categorizations are unreliable, it is not possible to determine whether instruction was inadequate for the student's problem or whether student performance was incorrectly diagnosed and good instruction was wasted on the wrong problem.

Unfortunately, diagnostic unreliability has been shown to be a pervasive problem in the field of reading. Only recently has it become possible to remedy the situation and to provide teachers with reliable diagnostic information. A series of studies on the diagnostic reliability of educational practitioners (Vinsonhaler, Weinshank, Wagner & Polin 1982; Weinshank 1982; Vinsonhaler, Weinshank, Wagner & Polin 1983; Weinshank & Vinsonhaler 1983) showed that reading and learning disability specialists and classroom teachers did not agree with themselves or with one another in their diagnostic judgements about simulated cases of children with reading problems.

A second series of studies (Vinsonhaler, Weinshank, Polin & Wagner 1983) showed that the reliability of diagnostic decision-making could be improved dramatically through the use of a specific type of training program. Two important products of the training studies series were 1) a reading diagnostic test battery based primarily on performance in instant word recognition, decoded word recognition, oral reading, silent reading comprehension and listening comprehension, and 2) a computer program for generating a reliable diagnosis given the data from student performance on the battery. Thus, it became possible to provide teachers with reliable diagnostic information about their students' reading performance.

This paper presents a case study of a junior high school remedial reading teacher who was provided with reliable diagnostic information about her students' reading performance and was thus able, for the first time, to learn systematically from her own experience and make changes for the better in her instruction.

A Case Study of Diagnostic Validity

The ability to train practitioners to be more reliable and/or have reliable diagnoses provided them became the basis for a series of outcome studies in the primary author's seventh grade pull-out remedial reading program. (The author is a collaborating teacher with the Institute for Research on Teaching, conducting research half-time and teaching half-time in the public schools.) The studies were designed to test the methods needed to move from diagnostic reliability to issues of diagnostic validity: verifying linkages that exist between diagnosis, instruction and student achievement. That is, given reliable diagnostic information, what instructional plans do teachers carry out and with what effect on students?

During the first year of the study (Wagner 1982), the diagnostic battery was administered to each student in the program as a pre and posttest. Using diagnostic decision aids developed in the training series, the teacher wrote a diagnosis for each student. The same data were used by the computer program to generate a computer-aided diagnosis. Agreement between teacher and computer was better than 90%. Thus, with the help of decision aids and the computer program, this teacher had become reliable in making diagnostic judgments about her students' reading problems. The remedial instruction which the teacher implemented with her students was docu-

mented by her daily notes and supported by interview and observational data collected by the project staff. At year's end, gain scores across the diagnostic categories were computed for all students.

The information on student gains confirmed some of the teacher's subjective evaluations and disconfirmed others. Much lower than anticipated gains in oral reading led the teacher to alter instructional procedures for a new cohort of students the following year. With the help of an instructional aide, the teacher instituted two major changes in her practice. First, she decided to abandon drill on common word families; the expectation that the patterns would be generalized to unfamiliar words encountered in text was not supported by the performance scores. Instead, vocabulary words would be drawn exclusively from the materials which the students would be reading. These words had to be mastered to a speed and accuracy criterion before text reading was begun. Syllabication strategies and sound-symbol association drill would be embedded within the vocabulary words. Second, instead of group oral reading for a portion of each session, daily oral reading on an individual basis was established.

The teacher's stated instructional goals emphasized improvement in word recognition and analysis skills within the context of extended, supervised oral reading. Five groups of four to six students came to the reading room from their respective classrooms for a 25-minute session four days a week between October and May, for a total of 40 instructional hours. The teacher had two major goals: (1) automate analysis and recognition skills in a framework of contextual reading and (2) give the students confidence in their ability to control the flow of print visually and orally. The teacher specifically stated that direct instruction in comprehension ran second to these two instructional goals. In her judgment, instructional time constraints coupled with serious deficits in word analysis and fluency dictated a primary emphasis on bringing those skills up to grade placement. Comprehension probes were used informally within the oral reading context.

The students moved through a fixed sequence of materials, entering at a level determined by the results of their performance on the word recognition and oral reading subtests of the battery. While sometimes allowed to skip certain steps, students generally moved through the sequence at their own pace. The program functioned

more as a clinical than as a classroom setting. Students read orally twice during each session. Vocabulary lists preceded various subsets of text, and those words had to be mastered to a speed and accuracy criterion before oral reading began.

The teacher described her program as a staged one. In Stage 1 students used a linguistically based reading program. This type of program typically starts by presenting single-syllable words made up of highly regular sound-symbol correspondences and then proceeds slowly to more irregular ones. Students were placed in reading texts appropriate to their levels of performance. No worksheets or isolated phonics drills were used. In the second stage, a small number of syllabication heuristics were presented so that the previously mastered sound/symbol associations could be transferred to two- and three-syllable words. Content-based texts were used during this phase. The goal of the third stage of instruction was to move students toward fluency and automaticity with texts of increasing difficulty. A variety of basal texts at grade levels 6 and 7 were used for contextual practice.

The teacher recorded oral reading performance all through the year. Every time a student read, the date, page(s) read, elapsed time and words miscalled were entered on a sheet devoted to the particular text the student was working with at that time. Text read at a very slow rate and/or with excessive miscalls was repeated and documented so that students could both hear and see changes in their own performance. After completing an oral reading session, the student continued reading silently until called back 5-10 minutes later to read orally again.

The procedures for this second year's study were to (11) collect reliable diagnostic data on each student using the diagnostic battery before the year's instruction began; (2) return the performance information to the teacher; (3) repeat diagnostic testing at the end of the year and return the performance information to the teacher; (4) determine student outcomes and return that information to the teacher, and (5) examine the links, between instruction and outcomes.

Collecting Reliable Diagnostic Data

The diagnostic battery was administered individually by trained testers at the beginning and end of the year

to the 23 students in the reading program. The students admitted to the program were reading two or more years below grade placement. The battery included measures of known reliability and provided data on seven variables of interest: grade level for word recognition and percent of words correctly identified by decoding (Slosson Oral Reading Test); grade level for oral reading of paragraphs based on number of miscalls and mean words per minute (Gray Oral Reading Paragraphs); grade level for silent reading comprehension based on percent of memories recalled and mean words per minute (Durrell Silent Reading Comprehension); and grade level for listening comprehension based on percent of memories recalled (Durrell Listening Comprehension).

Returning Pretest Information to the Teacher

The teacher received pretest diagnostic information in two different forms: (1) numerically, as individual and mean group scores for each of the seven diagnostic categories and (2) graphically, as performance charts. The four charts allowed the teacher to see group configurations before and after instruction and to track changes for individual students. Figure 1 shows a pretest performance chart for oral reading.

RATE: WORDS PER MINUTE

0-60	61-70	71-80	81-90	91-100	101-110	111-120	121-130	
				30				6th
		03 21 27						5th
		12 32	24 05	09 33 17		28	34	4th
23	15 11	02 04 14	29	26	16			3rd
		20						2nd

GRADE EQUIVALENTS

In this class, 87% of the students entered reading orally at three or more years below grade placement. By the end of the year, that figure had dropped to 57%. A trend toward increased reading rate can also be noted. On the individual level, student 14, for example, entered the program reading at a third grade level and exited reading at an eighth grade level.

Posttest and Outcome Information

The diagnostic battery was again administered to all students after the year's instruction was over. Values for each student (posttest minus pretest) for each of the diagnostic categories were calculated. Posttest and outcome information were returned to the teacher in numeric and graphic forms. The posttest and outcome charts for oral reading are shown in Figures 2 and 3.

RATE: WORDS PER MINUTE

	0--60	61--70	71--80	81--90	91--100	101--110	111--120	121--130	
10th gr.							30		G R A D E
8th gr.			14						
6th gr.			16 15	21	05 03 24	27			E Q U I V A L E N T S
5th gr.		02	32 29	26	34 28 09 33 17				
4th gr.		04	12	20	23				
3rd gr.			11						

Two digit numbers plotted = student I. D.

Fig. 3 - CHANGE IN RATE: WPM

	-30	-20	-10	0	+10	+20	+30
+5							
+4					30		
+3		16			15		
+2			02 26	29	24 20 05		
+1	34 28			04 09 32 17	33	03 21	27 23
0				12	11		
-1							

Losses and gains given in years.

Outcome results show the differential effects of the teacher's instruction, at least as measured by the diagnostic tests that were used. The students showed strong gains in oral reading accuracy with some signs of increased rate. Their word recognition skills improved, and their use of decoding as a strategy for word attack was increased. Performance on silent reading comprehension paragraphs was not as encouraging. A trend toward increased silent reading rate suggests that the students may well have been reading both more fluently and more accurately given their oral reading and word recognition performances, but these subskills did not appear to impact comprehension of the test paragraphs. Listening comprehension showed the least change. In general, the students seemed to find the subject matter

of the paragraphs as obscure at the end of the year as the beginning.

To summarize, the outcome results were as follows:

1. Word Recognition. The mean gain was 1.3 years in grade level (range = .3-2.3); mean percent of words decoded rose from 6.5% on the pretest to 16% on the posttest.
2. Oral Reading. The mean gain, as measured by number of miscalls, was two grade levels; mean words per minute remained essentially unchanged (84 w.p.m. on the pretest and 87 w.p.m. on the posttest.)
3. Silent Reading Comprehension. The mean gain was 0.5 grade level (range = 0-3.0); mean words per minute increased 7.4% (100 w.p.m. on the pretest; 108 on the posttest).
4. Listening Comprehension. There was essentially no change in mean performance (grade level 5.0 on the pretest; grade level 5.1 on the posttest).

Linking Instructional Practices to Outcomes

The overall outcome results confirmed the relative efficacy of the teacher's instruction in word recognition and oral reading and the relative lack of impact of her instruction on silent reading comprehension and listening comprehension. A more fine-grained examination of linkages connecting diagnosis, instruction, and outcomes is made possible by observing the effects of the instructional program on students in any given pretest cell of a performance chart. If each cell does, in fact, represent a diagnostic category that correctly separates like from unlike students, then it should be possible to pinpoint differential effects of instruction across categories.

Outcome results for both word recognition and oral reading for students with identical pretest scores were analyzed. Results show that the teacher was most effective with the most disabled of the students in both word recognition and oral reading. In word recognition, the strongest gains were made by the students who entered three or more years below grade placement. In oral reading, the strongest gains were made by students who entered four years below grade placement. Of the eleven students in this category, nine made gains ranging from one to four years. The effects of instruction on oral reading rate are ambiguous. Strong gains in grade-level placement were as often associated with increases in

rate as with decreases.

Analyzing word recognition performance on the dimension of percent decoding showed that all students in the 5% decoding category increased their level of decoding irrespective of grade-level placement. However, of the students in the 15% category, only those who entered three or more years below grade placement increased the number of words arrived at via decoding. This appears to further confirm the efficacy of the program for the weakest readers.

An analysis of the teacher's instructional document-ation for oral reading provided information, by student, on rate, number of words read orally, number of times (trials) a selection was reread, a list of all words that were miscalled and attendance data. Findings for this class include the following:

1. Mean rate for oral reading during instruction was 63 $\pm$ 7 w.p.m. This was slower than the rate for oral reading paragraphs on the pretest (84 $\pm$ 17) and the posttest (87 $\pm$ 12).
2. Mean number of words read orally was approximately 26,000. Students who made the strongest gains were below the mean; they read about 20,000.
3. Mean number of trials was 1.5. Repeated trials (more than 2) were effective in reducing miscalls but in-effective or counterproductive in increasing rate. The teacher has learned that having students read a selection more than twice is an exercise in dimin-ishing returns; students would be better off reading another selection of comparable difficulty.
4. Across all students, 2,478 words were miscalled during oral reading. Of the list words mastered prior to oral reading, only 15% were subsequently missed, sug-gesting that list mastery probably contributed to fluency.
5. Mean attendance was 82%; most of the students came to most of the sessions.

For the first time in her teaching career, this teacher received specific information which enabled her to link day to day instructional activities with student achievement in reading. On the basis of this information she made several decisions about the next year's program: (1) continue using text-based vocabulary for decoding and syllabication practice; (2) limit repeated reading of a passage to two trials only; (3) eliminate use of the linguistically-based materials except for the very

weakest students; stronger students would enter the
sequence at Stage 2, and (4) maintain comprehension
instruction at the same informal level.

Establishing Diagnostic Validity Within and Across Classrooms

The results of the study indicate that it is possible
to introduce into a classroom setting a system for estab-
lishing diagnostic validity, that is, establishing the
links that connect diagnosis, instruction, and student
achievement. The procedures are completely non-intrusive
and are sensitive tok the teacher's regular instructional
practices. Receiving information about the results of
instruction might or might not persuade a teacher to
alter practice but it does provide a rational basis for
decision making. This teacher found that overall, her
instructional program served her goals of improving word
recognition and oral reading skills. She learned that
her program is more effective for the lowest achieving
entering seventh graders than it is for those somewhat
less disabled. She has decided to alter her instructional
strategy for the stronger entering students but has not
decided how (or whether) to expand her goals to include
more direct instruction in comprehension.

The teacher can continue to learn from her own
teaching practice by repeating the documentation of the
diagnostic-treatment-outcome cycle year after year. But
the process is, in principle, very slow and there is no
comparative information available to her from other
teachers who deal with similar students. How do students
with similar entering characteristics fare in classroom
settings that are similar or different from hers? How do
they fare with similar or different materials and strat-
egies?

Diagnostic validity for students who range across
the whole spectrum of reading performance can be estab-
lished over time only by practitioners willing to learn
from their own practice and from that of their colleagues.
Learning from experience in ordaer to improve practice
in reading depends on teachers having access to reliable
information about the outcomes of their own and their
colleagues' instructional practices. In a study currently
underway, the authors are extending the procedures
described for establishing diagnostic validity to 10
fifth grade classrooms. A major goal of the research is
to document a full range of diagnostic classifications,
associated instruction and student achievement in reading.

Such documentation would provide a benchmark against which teachers and researchers could measure the consequences of instruction both within and across classrooms. Practitioners would be in a position to learn from and share their own and their colleagues' experience while simultaneously contributing to the growth of a descriptive, and ultimately prescriptive, data base that connects diagnosis with instruction and with the outcomes of instruction.

NOTE:

REFERENCES

Vinsonhaler, J., Weinshank, A., Polin, R. & Wagner, C. Improving diagnostic reliability through training. (Research Series No 126) East Lansing: Michigan State University, Institute for Research on Teaching, 1983.

Vinsonhaler, J., Weinshank, A., Wagner, C. & Polin, R. Diagnosing children with educational problems: Characteristics of reading and learning disabilities specialists and classroom teachers (Research Series 117) E. Lansing: Michigan State University, Institute for Research on Teaching, 1982.

Vinsonhaler, J., Weinshank, A., Wagner, C. & Polin, R. 1983, Winter. Diagnosing children with educational problems: Characteristics of reading and learning disabilities specialists and classroom teachers. Reading Research Quarterly, 18(2), 134-164.

Wagner, C. Learning from experience: Applying the clinical and epidemiological research paradigms to the study of the diagnosis and treatment in reading. Unpublished doctoral dissertation, Michigan State University, E. Lansing, MI, 1982.

Weinshank, A. 1982. The reliability of diagnostic and remedial decisions of reading specialists. Journal of Reading Behavior, 14(1), 33-50.