Resource Consultant Training Program
Research Report No. 13

RCTP

The Portfolio Concept with Applications in Curriculum-Based Measurement

Tracey Hall
Gerald Tindal

University of Oregon, Division of Teacher Education, Special Education Area,
Eugene, Oregon, 97403-1215

Published by
Resource Consultant Training Program
Division of Teacher Education
College of Education
University of Oregon

Staff
Gerald Tindal, Program Director
Jerry Marr, Editor

Hall, Tracey
Tindal, Gerald
The Portfolio Concept with Applications in Curriculum-Based Measurement

Preparation of this document was supported in part by the U.S. Department of Education, grant number H029B80119-89. Opinions expressed herein do not necessarily reflect the position or policy of the U.S. Department of Education, and no official endorsement by the Department should be inferred.

Cover design: George Beltran

The Portfolio Concept with Applications in Curriculum-Based Measurement

Tracey Hall
Gerald Tindal

University of Oregon

Abstract

A popular response to current testing procedures, and the confusion created regarding assessment analysis and application, is the student learning portfolio. Most applications of the portfolio concept are theory-driven and lack any empirical validation. This research report describes an alternative to traditional assessment procedures in the form of a specialized learning portfolio. This portfolio applies the concept of keeping data on students over time using the tools of Curriculum-based measurement (CBM). With the standardization of CBM assessment tools in the long term portfolio, educators will be able to use this document to represent student performance over time, within and across classes, and across abilities.

Testing has become a critical component in documenting student learning today. Not only do educators have a plethora of tests available, but they are also making increasingly more decisions based on those tests. For example, schools give students a standardized, norm-referenced test each year to report their accomplishments to the general public. In many states, minimum competency tests are used yearly to ascertain who should graduate. Numerous curriculum-embedded tests are used by teachers to move students through programs. Specialists also use such tests to place students into programs. Speech therapists, school psychologists, and resource teachers all use specialized tests primarily as decision-making tools for eligibility in special services.

The Problem with Research on Testing

For the past 20 years, a considerable amount of research has focused on the use of tests in our public schools. The work of Goslin (1969) and Stetz and Beck (1979) began to provide an understanding of the testing process. In these two studies, teachers were questioned as to how they used standardized achievement measures. Both studies concluded that teachers do not rely heavily on the information standardized tests yield. Stetz and Beck (1979) found that 80% of the teachers reported making only some or little use of standardized test data. The overall findings indicate that teachers most frequently used such test results to diagnose individual student problems and report performance back to students and parents. Fewer than 20% of the teachers used test information for such activities as curriculum alterations, instructional decision-making, or teaching-method changes.

The Test Use Project conducted at The Center for the Study of Evaluation at the University of California at Los Angeles updated and expanded this line of inquiry. Findings have been summarized in research reports by Lazar-Morrison, Polin, Moy, & Burry (1980), and Yea, Herman, & Rudner (1981). Three primary questions drove this line of inquiry: (a) What is the nature of current testing practices; (b) what factors influence the use of test results; and (c) what costs are associated with testing? Using a far broader methodology than in previous studies,

Requests for reprints should be addressed to Tracey Hall, 175 Education, University of Oregon, Eugene, OR, 97403-1215.

schools from large, small, rural, and urban districts were sampled; teachers from the fourth, sixth, and tenth grades were surveyed; and finally, interviews and record reviews were considered. The authors reported the following results:

1. Teachers tend not to value standardized published tests.
2. Teachers don't want to eliminate standard -ized tests completely.
3. Teachers place greater reliance on the results of their own judgements of students' performance.

Although the level of intensity, quality of research, and school personnel surveyed has varied and improved throughout the years, the basic conclusion remains the same: Standardized tests do not provide teachers with information they find useful.

Ironically, what appears to have taken place over the last 20 years as an outcome of such test dissatisfaction is the development of more tests. These new tests have been added to the existing test batteries in schools and, as a consequence, instruction has been interrupted more frequently. Consequently, educators find themselves in a real dilemma. Is there any way out of this tension of continued test development and limited instructional time? Many factors play a part in the big picture of educational assessment: teachers, administration, policies, politics, history, big business, to name several. In view of such diverse factors, as well as the complexity of implementing change, we will focus on the teacher as the central figure, and therefore the key to improving the process of assessment.

Making Current Practices Functional

With the current test utilization research findings, at least three solutions are possible: Because teachers can't be expected to use or value what they don't understand, one solution would be to better train teachers on how to use the available tests. This option is consistently supported by 20 years of research. Another possibility, however, is to begin creating assessments that address teachers' needs. Rather than simply create more tests, which may not be functional for classroom instruction, we should consider the demands teachers face in their instruction, the resources available to them, and then provide alternatives that are functionally and technically adequate (this would entail throwing out most current tests). A third option would be to take an inventory of what is now being done, eliminate duplicated practices, and fill in those areas that haven't been well developed to this point. This final option is essentially a combination of the first two, utilizing some extant measures and introducing some alternative measures and/or procedures. The critical question is, which option should we take? Any answer, of course, implies the use of certain criteria for selecting an option. For example, ease or difficulty of implementation, cost, resources available and required, or any combination of the above.

The first option implies that more measurement training would make tests more useful, which would require extensive preservice or inservice training. Several issues surface with this option. As stated previously, a multitude of tests are available and in use throughout public school systems across the nation. Preservice training programs could teach trainees on basic measurement issues. However, it would be difficult to deal with specific tests, because of the large number and wide variety of tests in existence. The result most likely would be an inefficient introduction to testing issues, not meeting a criteria of true assessment understanding. An additional concern with this option is the function these tests provide for the classroom teacher. Standardized tests don't provide teachers with information they find useful.

If we take the second option and construct new tests in assessment procedures, we should not forget that educators use published achievement test results and are unwilling to get rid of them completely. Therefore, an additional selection criterion would be to take advantage of procedures that have been shown to work and make them functional in the classroom setting. This implies that teachers actually use test data to make decisions. However, many tests are being used inappropriately, so use alone can not be a critical factor. A prime illustration of inappropriate test use is the application of published achievement tests to plan instruction, even though we know the content validity of such tests is lacking (Jenkins & Pany 1988). On the other hand, published achievement tests may have established validity with schools and teachers, which makes them useful for other decisions, such as program certification or classification. Ultimately, the criterion for selection deals less with the validity of a test in an absolute sense and more with the type of inference or decision being made (Messick, 1989).

This decision-making purpose of assessment leads us to the third compromise option, in which there would be an elimination of duplicated practices and development of measures in those areas that haven't been well developed to this point. The most important criterion for this option must be a clear purpose of assessment at the district, building, and classroom level. Additionally, this level of the educational system must also clarify the procedures

for distributing information, so that reports, scores, etc. are clearly presented and usable for teacher and public consumption. Simultaneously, assessment must also assist teachers with instructional planning and evaluation of the instructional group and individuals. Before focusing on the third option, we first analyze the first two options.

If we assume that the major problem with testing in our schools is less with the tests and primarily with their use, then we need to train educators in a range of data utilization practices. They need to know (a) criteria for judging tests' technical adequacy (reliability and validity), (b) what kinds of inferences are appropriate, and (c) when to use particular measures?

Knowledge of testing technical adequacy is an area that has plagued classroom teachers for years. Training programs have little or no emphasis on test use, interpretation, or evaluation. Generally, technical adequacy of assessment has focused on two issues: reliability and validity. Reliability is defined as the extent to which a test provides a consistent set of results over time, and within similar test situations. For a test in education to be useful it must have this consistency or reliability. Validity is an evaluation of a test's truth in what it purports to measure. Does the tool actually measure what it is constructed to measure? A test must be valid for a purpose when designed, and used for that same purpose when administered and results are analyzed. Validity is dependent on appropriate test use as well as construction.

The concept of technical adequacy recently has been reformulated to focus more on decision-making issues in education (Messick, 1989), which leads to questions regarding the types of inferences that are appropriate with particular measures, and the decision of when to use particular measures.

In practice, judgements regarding test results are highly misunderstood not only at the classroom level, but are additionally misunderstood by administration and the public at a high rate as well. We continue to give standardized test scores great weight for reasons outside a test's design or original intent. Test data have been used to determine the fate of educators (Shepard 1989). Merit pay, school bonuses, and school rankings, provide an impetus to increase student scores on standardized tests. Our education systems (i.e. school boards) often demand that we raise test scores rather than look at what skills are being learned in the school curriculum. Districts are ranked within a county or state, states are rated throughout the nation, and neighborhoods are considered prime real estate based on test results from local schools (Brandt, 1989; Wiggins, 1989). Concern over the misuse of test scores in school systems is quite appropriate (Shepard, 1989, Brandt, 1989; Rogers & Stevenson, 1988). We should carefully evaluate the utilization of test results at all levels.

Most of the previous research on test results and use has been done in relation to teachers' perceptions of standardized achievement measures. Yet this may be the wrong group for that focus; administrators and program evaluators would be a more appropriate audience given the test design. Many decisions regarding testing in schools is away from direct service personnel (i.e. teachers and principals). Those making decisions about what to test or what test to use don't apply the results at the classroom level, nor do they show teachers how the information is used at the building, district, or state level. Questions regarding the reliability and validity of tests appear to be addressed at another level, out of the teacher's hands. The statements that appear to be made to classroom teachers are something like, "Trust this because I said so," and, "This test looks good, and we all know we need to assess students, so..."

Although teachers are frequently assigned with the role of test administration, they are inadequately informed as to why a particular test may have been selected and for what purpose results will be used. Yet, they are expected to administer the test and make use of the test results several months later. Unfortunately, many such tests with a high rate of use are not technically adequate (Salvia & Ysselkyke, 1989), and teachers do not have the skills in making judgements as to when or how to use the results.

We must ask ourselves; Why is this the state of testing? How has such a set of misunderstandings and useless numerals been created? When each of these issues is taken into account, the findings from twenty years of research are still relevant: A clear need exists for training educators on appropriate test use. If teachers are the target group, the issue must revolve around overall program deployment, not specific instructional decision-making. Teachers make many diverse decisions using assessment results. For example, teachers determine if placing students in specialized programs (i.e., Chapter 1, TAG, and special education) is appropriate based on published achievement test scores. Furthermore, standardized achievement tests are being used to make many different decisions. In special education they are being used for individual educational programs, while, in regular education, they are often used to make program placement decisions (Haller & Waterman, 1986). In all educational settings, these tests are being used as measures of growth. The

result of this overuse is that many inappropriate inferences are being made: Students aren't mastering goals; programs are designed on the basis of test outcomes; and programs aren't effective.

Educators need to appreciate the limitations of tests and not endorse inappropriate applications when they occur. Ranking a school based on a standardized test score does not take into consideration such factors as instructional method, students, curriculum, or coaching (Brandt, 1989). "Teaching to the test cheapens instruction and undermines the authenticity of scores as measures of what children really know" (Shepard, 1989, p.7).

Frequently, the focus of standardized testing is on test scores, reports, and outcome decisions, neglecting the student. A shift to the focus on student behavior is important since the test results infer actual student behavior. We should, therefore, attend to the student requirements as they demonstrate skills on tests. In evaluating students, educators need to consider the actual response style by the student, in addition to the stimulus material that student's actually read and work on to demonstrate competence on the test.

Furthermore, educators must limit their interpretations of learning and achievement to fit within the context of the testing conducted and the actual constructs of the test. For example, reading passages on standardized tests often fail to approximate the contexts that students encounter in their classes (Valencia, Pearson, Peters, & Wixson, 1989). Spelling assessment procedures in the majority of standardized measures do not, in fact, assess spelling ability, but instead, test a student's ability to proofread words. The multiple-choice test format is convenient for rapid scoring. However, as described above, the activity tested is transformed into another skill area altogether. Because of such issues, there is a concern that what test-makers are measuring for students is fairly unrelated to the skills to which we generalize their use (Jervis, 1989). The multiple-choice format only shows us what students don't know, not what strategies may be completely or partially in place (Tindal, 1989). If we are to use extant measures, then we must at least appreciate their limitations and/or generalizations that can be made to classroom performance.

The discrepancy between format of tests and contexts of classrooms is important not only in the inferences derived, but also in the changes created in making them consistent. since the use of selection response formats in many testing procedures has become popular, there has been an impact on classroom instruction and assessment. Teachers have reported eliminating essay tests in their own assessment of students, because such tests are inefficient in preparing students for multiple-choice tests, the preferred format for standardized tests (Shepard, 1989). This format requires a selection response for a correct answer. Multiple-choice formats, the most common procedure used in assessment of large numbers of students, don't measure the student's ability to organize information and present a coherent and effective argument (Shepard, 1989). This procedure has very little flexibility and does not allow credit for partially correct answers. Few academic skills utilize multiple-choice format, in fact, the dominant behavior in classrooms is some form of extended writing. We need to consider the response requirements by students on assessment procedures and reflect on format when making important educational decisions.

Educators need to discriminate when the response mode used in testing is, or is not, appropriate in monitoring instructional achievement. In some cases, multiple-choice tests may be useful for measuring learning (Nitko, 1989). For example, "higher-order thinking skills" may best be tapped using such a format (e.g., concept discrimination). However, other kinds of skills depend upon production responses. In the latest report for the NAEP, students were found to have insufficient reading and writing skills. That is, students were not proficient in elaborating upon information they had read. Additionally, multiple-choice may be appropriate for certain decisions, (i.e. screening students into specialized programs).

In summary, the option of accepting extant measures requires that education as a field must substantially improve knowledge of tests, including their technical adequacy, appropriate administration, application, and interpretation. Educators need specific and sufficient training on the functional application of certain testing procedures for making particular educational decisions.

Creating a New Measurement Technology.

The problems noted above may be so serious that an entirely new technology is necessary. That is, the problem with inappropriate test use may lie in the tests themselves, and any training on appropriate utilization (analysis of technical adequacy, usage, and decision-making) is solving the wrong problem. The format of the student response and the limitations of the student materials indeed are so constraining that they can not functionally be fixed.

Published standardized achievement tests were developed during an era when they served a particular purpose. Although educational testing has a long and diverse history (Levine, 1976), it has a fairly recent and narrow application in the United States.

In the early 1900s testing focused on measurement of individual differences; yet, since World War II, such tests have been increasingly used to group and label individuals. Today, our educational programs are delivered not to individuals, but to groups according to comparable test performance. For example, students are labeled Learning Disabled (LD) and placed in a resource room that offers a common alternative; other students are labeled talented and gifted (TAG) and offered a common enriched alternative. In both cases, instruction is delivered to individuals in groups on the basis of perceived commonalties, as measured by standardized achievement tests.

This logic cannot be vindicated. If programs need to be developed for individual students, then measures need to be sensitive to these individuals. Standardized, published achievement tests cannot accomplish this outcome. Rather a measurement technology needs to be established that is based on specific curriculum and instruction delivered to individual students. In part, the criterion referenced testing (CRT) technology that was developed in the late 1960s (Popham 1968) was an effort to move past group summary scores and relative rankings (i.e. percentile ranks, standard scores, etc.). CRTs were anchored to specific curricular domains and focused on well-defined student skills.

Although criterion referenced measures have been proffered as an appropriate testing alternative, these have often failed in helping make many educational decisions (Glass, 1978). Frequently, interpretation of these tests is limited with use of mastery scores; comparability across domains is entirely lacking; and, finally, little flexibility exists for adjusting for instructional differences in the curriculum. For example, two students may score 85% on mastery or criterion referenced tests, but, can these tests be compared? Not without answering questions such as the following: What is the specific domain? How were items sampled? What was the average difficulty? Did the students take the same test? How many problems were on the test? When were the tests taken? What score is considered mastery? What scores did other peer group members obtain on the same test(s)?

Lurking behind all of these problems is the issue of convenience of assessment. To accommodate large numbers of students taking many skill-specific tests, assessment has become routinized, with students tested using multiple choice formats and group administration. The selection response, as described earlier, has documented problems with respect to measuring actual student behavior.

As a consequence, increasing attention has been devoted to developing alternatives that acknowledge the unique focus on the individual student without the contrived formats that accompany mass-production assessment. In recent issues of *Educational Leadership* (Costa, 1989), and *The Reading Teacher* (Flood & Lapp, 1989), learning portfolios are described. Learning portfolios are highly diverse and may incorporate any performance dimension in any subject area. Using the idea of an artisan or craftsman, self-created products are assembled to reflect both the diversity and proficiency of skills. For example, a photographer may use a range of photos, slides, and prints, to display craftsmanship. An architect may develop a portfolio that includes the variety of structures drafted, demonstrating proficiency in design. This concept is now being adopted in education.

The notion of a learning portfolio has tremendous transportability. Products within a portfolio can accommodate a wide range of academic behaviors that are useful for making a number of educational decisions. They can also be used by a wide range of professionals across general education and special education environments. Because the products are classroom-focused, within the school, year-long portfolios have a great potential for showing progress over time. Furthermore, they can be useful across grade levels. Finally, a great diversity exists in the manner in which portfolios are generated and/or maintained. Following are some examples of portfolios that reflect these opportunities.

In education, the portfolio concept would transfer into a folder of products reflecting academic proficiency. In an academic area such as writing, a folder may include samples of student poetry, journal entries, essay works or reports. Over the course of the year, writing samples should reflect the improvement that the student has made as a function of instruction. We would hope to see samples near the end of the year that are better organized, develop a unique statement, and reflect more consistent language and mechanics. Such entries can be used both to display the diversity of a student's skill, as well as improvement over time.

In the evaluation of student performance, learning portfolios can be used as work samples directly or summarized in the form of progress charts. With a slight variation in administration, many products generated in the classroom could be included within a portfolio. Academic areas, such as math, science, written expression, reading, and health, may be easily incorporated into such a system. An interesting dilemma occurs in the area of reading, because there is some difficulty capturing a

permanent product for display. The most efficient option would be an audio taping of the student or a written retell. Equally useful products could be included within a portfolio, reflecting other instructional outcomes such as math computations and story writing. In the end, a teacher could have access to student products that encompass the entire range of classroom behaviors for which they have focused their instruction.

As teachers become familiar with students in their classes and come to understand the improvements children are making (or not making), they should be better able to make instructional decisions. For some students, this might mean simply providing more structure, more practice, and more guidance. For others, more intensive interventions might be necessary, such as working one to one with peers and counselors, or providing guided practice. There may be evidence of the need for extreme interventions requiring assessments for special education placement. In this scenario, learning portfolios are very transportable across the general education and special education environments. They provide the focus from which professionals can interact in developing individualized instruction.

In a larger sense, portfolios are quite timeless. Any particular work sample within the portfolio simply represents a 35 millimeter snapshot of what a student has done at one point in time. Portfolios need little explanation, other than the time and context in which they were constructed. To illustrate, parents take pictures of their children and explain that this is what they looked like when at age five. As such, portfolios provide a powerful reporting function. Teachers can generate learning products any time during the year and show them to parents to say this is what your child "looked like" (how she performed) in November . . .

> "Now that it's the end of the year, look at the great improvement that has occurred. Notice how her handwriting has improved, she's using higher level vocabulary ... "
>
> "Your son/daughter has clearly mastered several language and writing skills."

The very same reporting function can be accomplished within a school across teachers. As students move through the grades, teachers can inform the next grade level teacher about specific student tendencies (i.e. skill specific strengths and weaknesses) as illustrated through the portfolio.

While it seems as if learning portfolios will solve all of our assessment needs, caution is in order. In the examples above, several measurement problems are present. With products developed in many different academic domains, inherent non-comparability is present. In contrast to the published tests that use standard scores for comparing performance across academic areas, no common diagnostic profile exists. A student could be very strong in math, but very weak in reading or writing, yet, this outcome would be difficult to ascertain unless portfolios are developed in all academic areas. In other words, without information in all academic areas, we would not want to draw conclusions about overall student performance with portfolio examples in only one skill area.

Although a normative classroom performance range is present with the portfolio, comparisons are difficult to explicate. Student performance—how students use their time, the amount of time each has to complete a task, how much assistance they may receive from the teacher or peers— can vary greatly within an activity. Additionally, teachers may inadvertently score different students' work on varying criteria, comparing the student to their previous work. This procedure greatly affects the comparison of scores across students. Because of the absence of systematic procedures, teachers must be cautious when rating students against one another on such measures.

Portfolios may appear to provide a common language that professionals from both general and special education can use, yet comparability is deceptive. The portfolio speaks only of student performance. However, we also know some curriculum specific features may greatly influence student performance. For example, if worksheets provide the content of a portfolio, any interpretations of performance must be couched within the content of that curriculum, as well as the contents of the worksheet itself. Furthermore, professionals come with different knowledge bases for interpreting performance. General education teachers may be focusing more on average developmental skill levels. Special education teachers may come predisposed for extremely low or prerequisite skills. School psychologists may focus on quantifiable aspects of behavior or interpretations relative to aptitude. This scenario is very much like the phone game, wherein a message is whispered from one individual to another in a circle until it reaches its originator; the message can become very distorted by the time it completes the circuit. Likewise, everyone assessing performance might have a different interpretation, even though the product that informed the inferences was the same. Unless we structure the language to use with the portfolio concept, we may lose the dialogue across professionals.

Growth is a very elusive concept. While change is occurs daily in classrooms, the anchors used to reflect those changes are in motion themselves. Teachers are teaching many different skills over the course of a year, and capturing growth implies some consistency. We have two choices. Either we vary the tasks to fit instruction (use a criterion-referenced view), or we establish a common task and watch for improvement over time. In the former approach, growth is implied less by performance improvements than by the qualitatively different behaviors that are generated. In the latter approach, the behaviors show the change in student performance on a task or set of skills that has been carefully controlled. The products, therefore, represent student improvement on that specific skill, the change occurring as a result of instruction and maturation.

Furthermore, little has been said about how the products were created, how much prompting or teacher student interaction was present. If teachers differentially help students (i.e., some students more than others), then the products become difficult to compare. To the degree that the products reflecting learning are different, nothing can be said about improvement. Variations that are problematic include the following: (a) stimulus formats for student responses (b) administration procedures, such as directions by the teacher, written direction on forms, timing, etc., (c) scoring and marking procedures, and (d) summarizing and reporting techniques. Such variations can occur within a teacher, within and across students over time, across products, or across all three—teachers, students and products.

Learning portfolios appear to be a good idea without specifics regarding their development. Until the problems noted above are resolved, what we may have is simply a permanent product version of anecdotal information. To be sure, this information can be used in communicating with parents and helping teachers inform instruction. It is unlikely, though, that it could be useful in formatively and substantively evaluating instructional programs. The question then, is whether we can capture the essential features of a portfolio assessment and infuse some kind of standardization in their creation.

While most of the literature on learning portfolios has come from the general education environment, a fair amount of research and development has actually occurred in special education. In the following section these research and development efforts are described under the rubric of curriculum-based measurement (CBM).

Combining Extant and Alternative Measures

An alternative measure that has undergone extensive empirical validation is Curriculum-based measurement. In the late 1970s Congress funded a number of institutes to conduct research on programs for the mildly handicapped. One of these, the Institute for Research on Learning Disabilities (IRLD) at the University of Minnesota, focused on developing appropriate measurement systems for teachers to use in improving instruction. Begun under the guise of data-based program modification, a number of formats were systematically investigated.

We assert that this system holds promise in combining the advantages of the current psychometric measures with the appropriateness of alternative portfolio measures. It makes no sense to simply train educators to better use outdated measures; nor is the technology of portfolio adequate to replace current systems. Rather, a middle ground needs to be created that captures the advantages of both without incorporating their limitations. Specifically we propose that a curriculum -based learning portfolio be developed with the following features: It should (a) generate production responses instead of selection response, (b) use both qualitative and quantitative scoring procedures that are reliable and valid reflections of behavior, and (c) employ efficient scoring and have results reported within an appropriate context that includes diagnostic information, which is both process and product oriented. When all of these criteria are in place, instruction can be truly informed and data-based.

First and foremost, teachers need measurement systems that reflect those behaviors that they teach. In the basic skill areas, this includes production responses like reading passages, spelling words, writing words in sentences to communicate new ideas, and calculating math answers to computation and story problems. CBM has an emphasis on production response, in contrast to selection response typical of most criterion and norm referenced tests. The tasks that were initially investigated at the IRLD include oral reading fluency, correct letter sequences, correct word sequences, and digits and problems correct (Deno,1985). These responses appear regularly in elementary schools throughout the school day; as such, they become grist for the mill.

Not only are such measures production oriented, but they create a lasting product that can be analyzed using both qualitative and quantitative scoring procedures. Qualitative procedures often involve an interpretive look at students' performance; judgements of quality are often made using subjective

criteria. For example, in reading, teachers often listen closely to *how* a student reads, taking note of the prosodic features (i.e., expression and voice quality) that are present. In writing, teachers critically consider their general impression to a composition; such judgements represent critical thresholds with implicit criteria of quality. If left at that level, such products would be largely inaccessible. However, quantitative scoring systems can also be used to buttress and corroborate the more ill defined qualitative interpretation. A hidden advantage to this dual interpretation is that behavior can be based on both ordinal and interval scales.

Production responses and flexible scoring systems may be important, but they pale in comparison to the need for reliable and valid reflections of behavior. All measurement systems must generate data that are consistent (i.e., reliable) and are useful in helping make decisions (i.e., valid). Traditionally, the former is considered a necessary but not sufficient condition for the latter. We first need to capture behavior that occurs similarly across two close occasions (test/retest), different versions of the same measure (alternate forms), interpretations by different scorers (inter judge) , or within a measure over items (internal consistency). Once such consistency is in place, our attention turns to the use of data to make decisions (Messick, 1989). The CBM research conducted to date has focused on all educational decisions, ranging from initial screening and eligibility decisions to formative and summative evaluations (Tindal, 1991).

Finally, in the developmental research for CBM, a key criterion for any measurement system was that it be user friendly. That is, teachers needed information that was cheap to produce, easy to collect, and quick to score. Anything less would simply not be used. The CBM measures which evolved and were eventually produced all required: (a) little teacher preparation time (i.e. can be based directly out of the curriculum), (b) brief administration duration (one to five minutes) and standardized administration procedures, and (c) clear scoring rules. Most systems were easy enough for paraprofessionals to become proficient in their implementation.

In summary, we have identified three options for educators, given the utilization of tests and current research. The first was to continue with current practices, with the addition of teacher training to understand and appropriately use assessment measures. Secondly, we reviewed the option of creating new assessment procedures that more appropriately address teacher's needs. The third option of developing procedures containing the best qualities of current practices and newly-developed, empirically validated products, and eliminate duplicated practices to create a curriculum-based measurement (CBM) portfolio. In the remainder of this paper we emphasize a specific methodology to attain this last option.

Method

To illustrate the portfolio concept as a document representing progress or change in student performance over time, within and across classes, and across abilities, we implemented curriculum-based measurement with the students in the Fall and Spring of the 1989-90 school year. The two testing periods were early October and late May. Our methodology deals with three issues: subjects, instrumentation, and data analysis. All students were tested in four basic skill subject areas, described below. Results are reported on selected students, with a complete listing of student results for each assessment reported in Appendix 1.

Subjects

Twelve students, representing five grades and three ability levels were selected for this study. All students were from a small, urban, Pacific Northwest school district of approximately five thousand students. Subjects were selected by their teachers to represent a range of ages and skills. The age and ability levels across subjects were as follows: three regular education students of high ability, (Grade 1 = 2, Grade 5 = 1); three regular education low ability students, (Grade 1=2, Grade 5=1); two students receiving services in Chapter One reading (Grade 3 = 2); four identified special education learning disabled students (Grade 2 = 2, Grade 4 = 2).
Results are reported using fictitious student names. Reading results are reported for two third grade Chapter One students, Jenny, and Chad, a second grader, Emily, identified as mildly handicapped and receiving specialized education, and two fourth graders receiving special education services in reading, spelling, and math, Greg and Tom. Finally, in reading, a high performing fifth grade student, Carl, is discussed. In written expression, we demonstrate results with Matt, a second grade mildly handicapped student receiving services in reading, language arts, and math, and one of the third grade chapter one students, Jenny. Emily is a second grade student identified as learning disabled, and currently receives special education services in reading, math, handwriting, spelling, and oral language. We present her math performance. An additional set of tests in math are reported on the third grader, Matt.

Instrumentation

Students were tested in four basic academic areas: reading, written expression, spelling, and math. The measures require a minimal amount of

time to administer and are representative of the instructional curriculum. Following is an explanation of each measure.

Reading

All students were asked to orally read passages selected from the school adopted reading series. Readability levels in this text varied greatly, sometimes as much as five grade levels within the text. Therefore, only passages with readability levels within + 1.5 grade-level range were identified, and a random selection from this pool was chosen for the portfolio assessment.

Students were tested individually, away from classroom distractions in an office or lobby area outside the classroom. Standard administration procedures were used for each test. Each student's oral reading was timed for 1 minute on the passage. Identical procedures were used in the fall and spring testing periods.

Writing

For the expressive writing samples, an incomplete sentence (story starter) was used to generate writing, and students were asked to write what happened. Administrative procedures for written expression required the tester to read the story starter to students and allow a 1 minute period to think about what they would write and then a 3 minute writing period.

Spelling

For the spelling assessment, students were instructed to write words dictated from a list. The spelling test was developed by obtaining a random selection of words from the grade level spelling curriculum. The examiner used a "rolling dictation" to present 16 to 18 words to the students, depending on grade level. The tester dictated each spelling word three times, following standardized procedures: The word was stated by itself, repeated in a pre-planned phrase, and, then stated by itself again. The words were dictated with 8 to 10 second intervals, depending on grade level. The tests were administered individually because of the small number of participating students.

Math

Math tests varied by grade level. We restricted the assessment to numeral identification, dictation, counting and computation problems, versus problem solving applications such as story problems. The following descriptions are of procedures used for this portfolio concept with first through fifth grade students.

First grade: Number dictation. Students were asked to write the number on a line as dictated on a "10-second roll." They were told what line to write on, and the number was stated twice. Items ranged from single digit to three digit numerals. The tester assisted students in finding the correct line on which to write the numeral.

First grade: Object counting. Students were given one minute to count object shapes on their papers and write the number of the objects having a specific shape. The number of objects represented a random sample of 1-digit to 2-digit numerals less than 20.

First grade: Operations. Students were given a random sample of basic addition problems selected from the district curriculum of instruction. Problems were written using both horizontal and vertical formats. The students were allowed two minutes to complete as many problems as they could. They were also told to cross out any problem they did not know how to do. This same procedure was applied to subtraction operations in the spring for the first grade students taking the CBM tests.

Second through fifth grade: Operations. Students were tested in operations representative of the curriculum. Mixed probes consisted of a random sample of math operations from the district curriculum text representative of the students' grade level. This test was given to each student with directions to complete as many computation problems as possible in 2 minutes. Students were reminded to look carefully at the sign for each problem, as there was a mix of operation types in this test. also, they were told to cross out any problem they did not know how to solve and go on to the next problem, showing their best work.

Data Analysis

Initially, all tests were scored using quantitative procedures yielding numerical scores. Students were evaluated on individual performance using measures of rate, percent correct, and percent improvement. A final method of quantitative evaluation used was norm referenced. In this investigation, we used box plots to describe the normative data. Box plots graphically represent the distribution of scores (see Figure 1). The box represents the middle 50% of the population, the top and bottom "Ts" representing the 10th and 90th percentiles, respectively. Within the box, the bottom line represents the 25th percentile, the box top the 75th percentile, and the line mid-box represents the 50th percentile. The vertical axis contains numerals representing performance by rate correct, number of occurrences, and so on. Performance at a percentile level can be read easily using the box plot.

The second type of analysis we employed was qualitative, with student performance evaluated on variables that are less countable and somewhat subjective. Qualitative procedures are described for

reading and written expression tests. There are few, if any, qualitative measures used in spelling and math scoring. One might be able to evaluate the student's neatness in relation to handwriting or alignment of numerals or letters; however, these skills are less informative or relevant in math and spelling. We will not be describing any qualitative measures for those subject areas.

Reading

The following quantitative error types were counted in scoring reading: mis-identification—the student said the wrong word; substitution:—the student said a synonymous word instead of the printed word (e.g., 'house' for 'home'); omission—the student completely skipped a word or words printed in the text; reversals—the student transposed a word or words (e.g., 'was' for 'saw,' 'said she,' for 'she said'); three-second hesitations—the student did not decode the word within 3 seconds of reading the previous word. Repetitions (a word or words reread), insertions (the addition of words not on the page), and self-corrections (the student corrected a decoding error within 2-3 seconds), were not considered errors. The number of errors was subtracted from the total number of words read to determine the number of words read correctly per minute.

Qualitative analysis was based on audio recordings of each student's reading, which was reviewed for voice quality, reading expressiveness, attention to end marks and other punctuation.

Written Expression

In written expression we counted total number of words written, number of correct word sequences, and percent of correct word sequences.

Total words. This measure is a count of the total number of words written during the three-minute period regardless of spelling, grammar, or punctuation. The only items not counted in this measure are numerals (e.g., 23, 118) and special symbols (&,$, @).

Correct word sequences (CWS). A word sequence occurs when a student appropriately joins two words that are spelled correctly and are syntactically correct. This index takes account of grammar, punctuation (capitals and end marks), and spelling. Words that begin a sentence must be capitalized to be considered correct; sentences must have an ending mark, as well. In the case of run-on sentences, scorers "forced" an ending to the sentence, which lowers the count because capitals and end marks would be absent.

Percent of correct word sequences. This measure is simply the number of correct word sequences divided by total word sequences. Recent research has revealed this measure to be quite sensitive to growth and representative of student performance (Parker, Hasbrouck, Tindal, 1989).

Written expression is an area where a multitude of variables occur, which must be evaluated qualitatively as well as quantitatively. In the examples selected for this study, we looked at qualities such as handwriting, overall neatness, semantics, descriptive vocabulary, transitional phrases, sentence development, and content. These more subjective judgements lend another dimension to the evaluation procedure.

Spelling

Spelling tests were analyzed using two scoring procedures. First we counted a word correct when the entire word was written using the appropriate letters in the proper order. Second, we counted *Correct Letter Sequences* (CLS), i.e., pairs of letters within the word that are in the appropriate order. For example, if a student were to spell the word 'myself,' 'miself,' the CLS would be counted as 5, (one count for starting the word correctly with 'm,' one for the 's' to 'e,' one for the 'e' to 'l,' one for the 'l' to 'f,' and a final count for ending the word with an 'f.' (The word myself would receive 6 counts if spelled entirely correct).

Math

The number of digits in the correct place value was counted for scoring procedures in math. Rather than counting an entire problem as correct (or incorrect), this scoring procedure attends to the digits the student writes for each calculation within a problem; all digits in the work shown and in the answer are counted. For example:

```
 27          5          7
2)54        3          1)7          22
 4         +9          7           x43
 14        17 [2]       [2]         66
 14                                 88
  [7]                              946 [7]
```

Instead of a student receiving the same score for completing a math facts problem (e.g. 5 + 4 = 9) and a multiplication problem (e.g. 96 x 82 = 7,872), this procedure of counting correct digits within the problem "gives credit" for more complex tasks. Numerals in brackets indicate the total digits possible in the problem.

Results

The results of this study illustrate the flexibility in using a CBM portfolio. Student scores can be evaluated in relation to individual reference, growth or improvement over time, as well as in comparison of an individual's performance to a peer group population (in this case, a district normative group). Student performance in reading and written expression have been both quantitatively and qualitatively judged. We will be reporting several cases from the entire study, but not all results on each student tested are addressed (see Appendix 1 for complete results tables). Select cases have been prepared to exemplify the CBM portfolio.

Reading

There are a number of variables we can count, hear, and observe as a student reads. The following selections of case studies illustrate the use of CBM measures in reading while evaluating both quantitative and qualitative data. The students' reading has been made a permanent product by audio taping the assessment.

Chad and Jenny

Chad and Jenny were third-grade students receiving services in Chapter One reading. These students are from the same school and their instruction for Chapter One is basically the same supplementary pull-out reading instruction to the third grade reading program.

Chad averaged reading 13 words correctly read in the fall testing period, and 45 words read correctly in the spring testing period. This individual growth is quite impressive at 51% improvement (see Table 1). To illustrate a student's relative ranking within the entire third grade for the local district we compared their individual performances to the grade-level norm. When we compared to the norm group for third grade, Chad, (marked in Figure 1 with a •), ranked substantially below the average performance of his peers, in fact below the 10th percentile of the group in both testing periods. So, while Chad improved in words read correctly per minute over the course of the year, this performance increase was not large enough to catch up with the performance of the local third grade population.

Table 1. Third-Grade Chapter One Student Reading: Chad & Jenny

	Fall		Spring		Percent
Words Read:	Correct	Incorrect	Correct	Incorrect	Improvement
Chad	13	20	45	5	51
Jenny	23	8	56	0	26

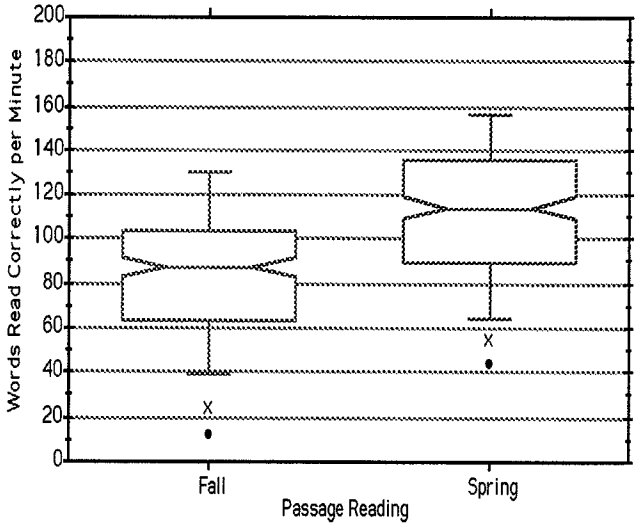

Figure 1. Norm Testing Box Plot for Third-Grade Reading

Jenny's scores were somewhat higher than Chad's. She also improved in reading rate and accuracy over the school year. However, once again, when compared to the normative group, Jenny fell below the average performance range of third-grade students (noted in Figure 1 with an x.)

To help reference this improvement, we can look at the number of words read correctly by the mean of the norm group made in the same period; there was an increase of 27 words read correctly per minute, from the fall to spring. Both Jenny and Chad improved 32 and 33 words respectively from fall to spring.

Some of Chad and Jenny's change is reflected in the number of words read incorrectly, which for both students drops from fall to spring. Chad read 20 words incorrectly in the fall, and only 5 words incorrect in the spring. Jenny made an 8-word change in words read incorrectly.

Emily

We also evaluated student performance qualitatively. Emily was a second-grade special education student receiving services in several areas, including reading. Second graders read the following from a passage in the fall:

> The next afternoon, Hasan closed his shop very early. Nodding to this one and smiling at that one, he strolled home. Everyone was feeling the happiness of the holidays.
>
> His wife met him at the door.
>
> "Come in," she said. "Your mother and our daughter are here."

We have phonetically transcribed a sample of Emily's one-minute reading of this passage. The periods between words indicate a time lapse as she read each word in the passage; more periods represent more time. Examiner-assisted, and self-correct errors are also noted in the transcription. A key is provided below to explain each symbol used.

KEY

.... = time between words, greater with each period
] = end of one minute time period
( ) = examiner said word
/ = quickly self-corrected

Emily's one-minute timed reading, fall, 1989, is presented below. As can be seen, her reading was very choppy, each word separated from the previous word, making the relationship of words in phrases and sentences very difficult to understand. She sounded out most words, slowing her rate, and she continually omitted words. This behavior was observed across several lines of text. In some cases, she omitted a word or two from the line; in other cases, she omitted the entire line. In fact, tracking could have been a problem for this student. She also omitted any punctuation or inflection related to sentence phrasing.

the ...n..e.xxt...next..and this.. nn.
....shshopped.. with...then... theeiis.. .to...
his..one..aa..dolls..at...that..one...his...the...has/
his...w..ii..ththth..
with went...him at... the.k..t..d..o.oo (door)
door....Come in she said....you..
mother..and...farver..this..this...do..r..ar..home...
for...]

When we compared her fall reading to spring reading along these dimensions, word omissions did not happen at all. The actual passage was as follows:

All night long it snowed and snowed and snowed.
In the morning Bud and Beth Page looked out the window. The yard was covered with snow. Near the yard was a patch of woods. All the trees and branches were frosted with lacy white.
Bud and Beth were excited.]

Emily's spring reading of this passage is presented in the next column. She attempted each word; words or lines were not skipped. However, she continued to have problems with word accuracy, missing 17 words in the 49 attempted.

All..night..long.it..shsh..shshwoned.and.shown ed.and.showned....Inthe.morning.. bad. and.BBrurn.began.t...looking.out.the.window. In. the.yard.was.c.cold with.... no..Near..the yard.was.pa.pach.ed...panched.of.woods. All.the. thing/things.and br. brrrbrrabrasshhshsh.brached.. was/were..fast.t.ed and.then.l..like..white. Brad and.b .breathe .. were.next...they.had...]

Greg and Tom

Two fourth-grade special education students, Greg and Tom, read passages from the grade-level material used in the mainstream classroom. Scores from these students' fall and spring reading are listed in Table 2.

Table 2. Fourth-Grade Special Education Student Reading: Greg & Tom

	Fall		Spring		Percent
Words Read:	Correct	Incorrect	Correct	Incorrect	Improvement
Greg	49	4	57	2	14.1
Tom	77	7	101	3	33.8

Tom's reading rate was substantially higher in both fall and spring than Greg's. When compared to the norm group, Greg performed significantly lower than his peer group (below the 10th percentile). Tom, however, performed between the 25th and 50th percentiles in the fall and his spring scores were quite close to the mean of 106.2 for the norm group. Why, then, was this student receiving special services for reading?

When we look beyond the numbers, we can get a better understanding of why Tom was identified as needing help in reading. A more qualitative view was obtained by listening to these students actually read. There were changes detected in performance beyond that which was quantifiable. The fall passage is depicted below:

Morning. Katie John opened her eyes and looked at the strange room. Yes, they were here, all right. She pulled on her blouse and shorts and ran out of the house. Did it really look as horrible as it had when they arrived last night?

Oh, worse. Katie John groaned. It was nothing by an ugly old brick house, squatting in the sunlight. Square as a box, flat roof, not even a bit of ivy on the walls to soften the sharp corners. It was three]

The next example is a phonetically written sample of how Tom read the fall passage:

> Morning.....Kathy.....John.....o..pen.ed.....her...
> eyes... aand...looked...at the....ststrange .....room
> .....yes.....there...were... her/here..... all... all.....
> right..... .she... pulled... on..her blouse...and
> shirt...and ran...out of the.....house did...
> it...really...look as horhorible... as...it...had
> ...when...thththey...arrriveed? last...night
> ...oh...worse.....KatieJohns....groaned
> it...was...nothing...but a...ugly...old...brick...house
> squating...in the.....sunlight...square...as a
> box...flat...roof...not even...a bite/bit...of...
> ivy...on...the...waaalls..to soften...the..sharp
> ...corrrners...it...was...three]

In the fall attempt, Tom read each word in complete isolation of the other words in a sentence, resulting in a choppy reading style. There was little or no connection of phrases or sentences. He did not stop for breath or change tone when reading words at the end of a sentence and never used inflection to differentiate words, phrases, and sentences. Although the passage contained a question, and he attempted to use a questioning tone (raising the voice), each word was so disconnected from the other within the sentence that the question was not clearly audible.

Tom's spring reading score improved by 34% when we evaluated his performance quantitatively. His reading was also much smoother and easier to listen to in the spring evaluation. Not only did he read whole words rather than word parts or symbols in isolation, he demonstrated better phrasing, inflection, and expression. In talking to Tom's special education teacher, we found that these very characteristics had been the goals established for Tom that school year in oral reading.

Carl

Our final illustration in reading is Carl, a student receiving regular classroom instruction in the fifth grade. He made substantial individual growth from the fall to spring testing. Carl's fall reading score of 95 words read correctly per minute was well within an acceptable reading performance level when compared to his peers. When tested in the spring, Carl had increased his oral reading fluency to 151 correct words per minute. The spring reading level for Carl was between the 50th and 75th percentiles, as compared to his peers.

Individually, Carl improved over 60 words per minute. The most typical change in individual performance in a school year for this grade level is a 15 to 20 word gain (Hasbrouck & Tindal, 1991). This performance growth is double the change of the peer group when comparing performance to the norm group. Scores from Carl's fall and spring reading are listed in Table 3.

Table 3. Fifth-Grade Special Education Student Reading: Carl

	Fall		Spring		Percent
Words Read:	Correct	Incorrect	Correct	Incorrect	Improvement
Carl	95	10	151	1	.09

Writing

The written expression assessment results in a permanent product to evaluate and include in a portfolio, allowing scoring and re-scoring along different dimensions as desired. Here, we have selected student tests to illustrate scoring and evaluation issues, and the relationship of the student's product along individual or group comparisons.

Matt

The fall story starter for second graders was as follows:

> "I was walking down the street and found a key in front of a big gold door. I took the key, opened the door and..."

Matt, a second grader, scored 33% correct word sequences on this fall writing measure. When tested in the spring with the story starter:

> "One very dark and spooky night I was camping in the woods. When I heard a strange noise..."

Matt wrote 66% correct word sequences. Note in Table 4 that the number of words written changed very little (1 word). Although Matt did not necessarily learn to write more over the course of the school year, the percent of correct word sequences in his writing showed substantial improvement (in word spelling, grammar and punctuation).

Table 4. Third-Grade Chapter One Student Written Expression: Matt

	Fall	Spring
Total Words Written	8	9
Correct Word Sequences	3	6
Incorrect Word Sequences	6	3
Percent Correct Word Sequences	33	66

In addition to evaluating Matt's writing "by the numbers," as shown in Table 4, the CBM portfolio allows a teacher to look directly at the student's work via the permanent product. Over the course of the year when evaluating these writing products and "eyeballing" the student's skills, variables such as handwriting height and alignment did not change much. In both samples, Matt wrote letters off the lines, and they were incorrectly formed. Spacing of words and letters improved somewhat from fall to spring. It was easier to discriminate one word from another in the spring sample, but, judgement of handwriting skills does not alter greatly from fall to spring.

Another variable we are able to evaluate when looking at the permanent product is the choice of words the student uses to tell (or write) the story. In both cases, Matt selected basic, short (3-to 6-letter words to complete the story. The major change from fall to spring for Matt was the correctness of spelling words used for each story.

Each of these factors can be taken into account as the teacher(s) evaluates a student's progress over the school year. If a teacher adopts the CBM portfolio concept for written expression, he or she may opt to collect samples under repeatable conditions more frequently. Teachers may also look across students in their class to compare writing in terms of the quantitative measures, dealing with numbers of words and correctness, as well as word usage, handwriting skills, grammar, syntax and semantics at the qualitative level.

Jenny

In another written expression example we see a student quantitatively make little change or even look worse from fall to spring (see Table 5), yet, qualitatively demonstrate clear improvement (see Figure 2).

Table 5. Third-Grade Chapter One Student Written Expression: Jenny

	Fall	Spring
Total Words Written	24	21
Correct Word Sequences	18	10
Incorrect Word Sequences	7	11
Percent Correct Word Sequences	72	47

When evaluating Jenny's writing samples, the quantitative variables, percentage of correct word sequences, and number of words written when individually referenced, and compared to the norm group, appear to be slipping. Nearly all numbers shown on Table 5 are lower in spring than fall. This quantitative look at her performance is not very encouraging.

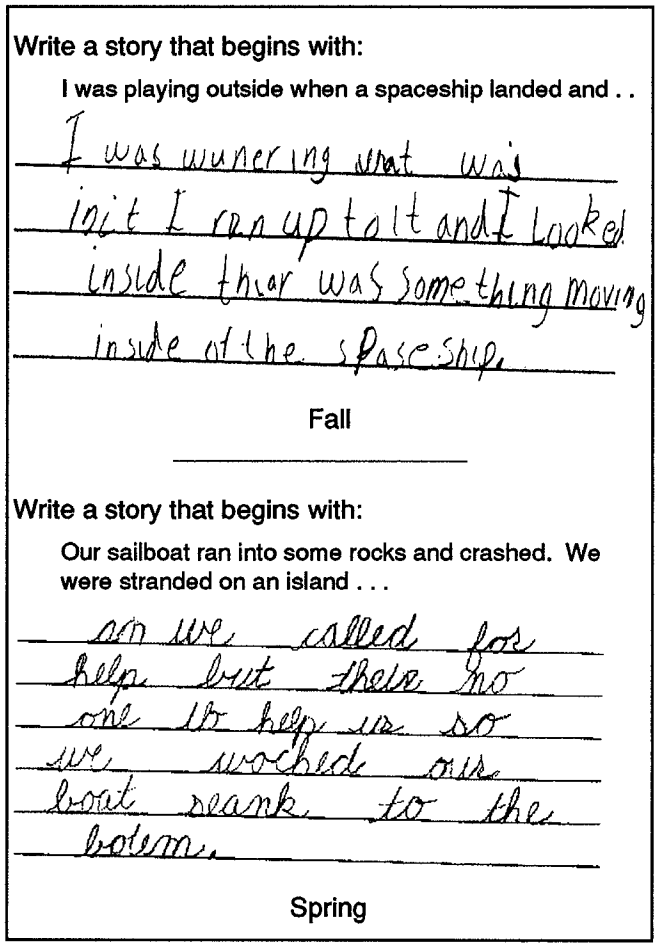

Figure 2. Jenny's Writing Samples, Fall and Spring

Jenny's case demonstrates the value of retaining the written product. Qualitative variables such as handwriting, word usage and action words written in the sample greatly improve from fall to spring. Jenny has made the transition over the course of the school year from using a manuscript handwriting style that is legible, but not considered very neat for a third grader, to a good quality cursive writing style that is easy for the reader to decode.

Another variable that can be evaluated as we analyze the permanent product is the student's use of words. In the spring sample, Jenny's thoughts were well connected and she used transitional words and phrases such as but and so, demonstrating a more advanced writing style than we observed in the fall writing sample.

Math

We will now review three students' scores in math. Students were given 2 minutes to complete math calculations. These measures are easily scored and evaluated given the permanent product from the student's work.

Emily

Emily, a second grader, had been in special services for 2 years and had received specialized instruction in reading, math, spelling, handwriting, and oral language. In the fall, Emily was only able to calculate one digit accurately (see Table 6). She did not even attempt five problems and several digits were incorrect for the math calculations. In contrast, her spring scores were significantly higher. When we compared Emily's scores over time, the most sensitive scoring procedure, correct digits, increased 14 digits (from 1 to 15) reflecting an improvement of 44 %. As can be seen in Table 6, Emily also attempted more problems in spring than fall, with a larger number of correct digits.

Table 6. Second-Grade Special Education Student Math Calculations: Emily

Mixed probe	Fall	Spring
Correct Digits	1	15
Incorrect Digits	6	11
Problems Not Attempted	5	0

Another evaluation of Emily's performance was obtained by comparing fall and spring test results to the normative group. There is dramatic improvement in her performance in relation to peers. Emily's score of one digit correct in the fall when compared to the norm group was well below the tenth percentile. Students performing at the 25th, 50th and 75th percentile in the norm group obtained scores of 5, 7 and 10 digits correct, respectively.

When we looked at Emily's spring score of 15 and compared that to the spring norm group, we found she was above the 50th percentile. Any student with a score of 15 correct digits at this testing, was performing between the 50th and 75th percentile. With box plots such as those presented in Figure 3, teachers can plot a student's performance (seen here with an x) and actually see the relationship of Emily's math performance to her regular-education peers over time.

To summarize, Emily made great individual improvement over the course of the year as demonstrated by the growth of 1 to 15 digits correct. Additionally, she moved on the normative distribution from below the 10th percentile to between the 50th and 75th percentiles. This is a substantial gain in relation to her peer group.

Matt

The scores in Table 7 are those of another second grader, Matt, who received special services. We can evaluate him using individual comparison and norm-group comparison, as above, and contrast the results.

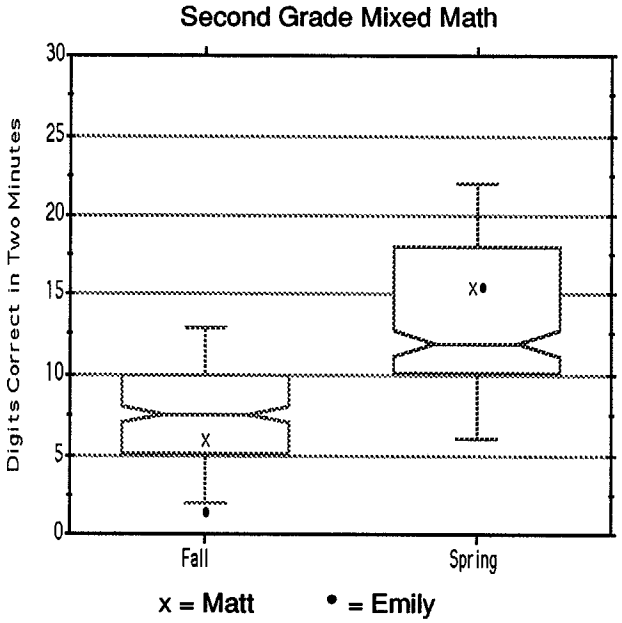

Figure 3. District Box Plots CBM Norms of Second-Grade Math Computations Second Grade Mixed Math

Matt appeared to have made some reasonable growth when the scores from fall to spring were compared; his correct digits increased from 5 to 15, this is an individual improvement of 40%. When these scores were then compared to the norm group, we could see that this growth was a movement from approximately the 25th percentile in the fall to above the 50th percentile in the spring. The apparent growth, when viewing the individual, was confirmed when evaluating Matt in relationship to his peers (see Figure 3).

Table 7. Second-Grade Special Education Student Math Calculations: Matt

Mixed probe	Fall	Spring
Correct Digits	5	15
Incorrect Digits	8	4
Problems Not Attempted	0	0

Discussion

There has been a call for reform in educational assessment over the past few years. From the inception of testing in education, two fundamental goals have remained the same: (a) to predict the success of students in the academic setting and (b) to inform the development of instructional programs that can facilitate the performance of the learner (Campione, 1989).

We initially presented three options for assessment reform in education. The first developed from the premise that teachers don't use or value what

they cannot understand. Therefore, this option included providing training for teachers on how to use the tests that are available. Our second option evolved from a research perspective that has shown a lack of connection between the design and function of tests and teachers' instruction. This option is to create new assessments that address teachers' needs in the classroom, providing alternatives that are functionally and technically adequate, and throwing out most current tests. The final option we presented was essentially a combination of the first two: make use of available and functional tests in current practices, eliminate duplicated practices, and fill in empty assessment needs with newly developed measures that provide a technically adequate, operational service for teachers.

The need continues for teachers to be accountable for their instruction in the form of increased knowledge and performance ability of their students. An example of this reform can be traced to the mid-1970's when statewide or state-regulated testing became mandated in 46 states (Valencia, Pearson, Peters & Wixson, 1989). Instead of instructional accountability in education, the reactions to reform demands have resulted in more testing. Yet, such procedures and tests do not tap the issue at hand, that of very little (if any) linkage between instruction and assessment. As a result, many researchers and educators advocate for change in the way we assess. Resulting research has demonstrated that such assessment practices in these mandated tests are at odds with instructional theory (Valencia & Pearson, 1987; Farr & Carney, 1986).

Teachers look to tests to help them make curricular and instructional decisions. Therefore, the conflict between curriculum practices and assessment creates a discrepancy. "Teachers should not have to set aside good instruction to prepare students to take a test; instead, good instruction itself should be the best preparation." (Valencia, Pearson, Peters & Wixson, 1989, p. 62). An additional issue facing such testing is that paper-and-pencil tasks, the main assessment structure, is too far removed and narrow in scope from the teaching domain (Hodgkins & McKenna, 1982).

Educational literature has documented the public and educational community desire to have student test scores exceed the national average. Naturally no classroom, school, district, or state would want the test results of their population to be below the national average. Most in fact boast the opposite, that they are indeed above average on the standard measures. We need to look closely at what that actually means with regard to groups of students and individuals. High-stakes testing such as standardized achievement tests are under investigation regarding claims brought to light by Cannell's study (1988), in which most states indicated that their schools are scoring above average on nationally normed elementary achievement tests, a phenomenon affectionately labeled the "Lake Wobegon Effect" (" . . . where all the children are above average.").

Several questions are raised with regard to performance above the national average. Do test publishers and educators have adequate information themselves and provided to the public as to the meaning of "national averages?" How up-to-date are the national norms to which we currently compare students? Because of the age of normative data, the population which norms were developed and "teaching to the tests," it may be the case that test gains are real. In other words, student scores are literally above average. We must therefore look carefully at the conclusions drawn and the decisions made with results of such measures. The quality of being "above average" does not necessarily mean that instruction and learning are adequate. While these questions and concerns are quite specific and focus on the tests themselves, we want to look at and evaluate the larger picture. Are standardized testing procedures valid in their representation of teacher and student performance in the classroom?

In the end, we can't lose sight of the real purpose of testing: to improve instructional programs for all students. We would argue that the need exists for testing practices that enhance instructional linkage. This does not exist with current procedures or reactional reforms. Training teachers in test use, application, or interpretation won't improve the central problem, that of testing procedures that do not reflect the needed connection between instruction and student performance.

Option two addresses new alternatives in educational assessment. Throw out the old procedures of nationally standardized tests and develop new measures at the state and local levels to assess school and district level achievement of mandated instructional goals. State-regulated testing has been adopted in 46 states with mixed acceptance and success (Valencia, Pearson, Peters, & Wixson, 1989). The notion of connecting instruction and assessment is the driving theme of such programs. Examples include the California Assessment Program (CAP), the Oregon Statewide Writing Assessment Program (OSWAP), and the Illinois and Michigan Statewide assessment projects. Each of these attempts to develop tests that reflect current or adopted instructional theory in a variety of subjects, usually basic academics (i.e., reading, math, writing, and, science).

At issue with these assessment procedures is the lack of empirical validation. These procedures are theory driven, and therefore we must question the validity of both the new-wave instructional practices and the assessment procedures developed to test individual student mastery of skills.

This study, an initial investigation of our third option, combined extant practices with some new alternatives for assessment. The alternative studied through the concept of a learning portfolio was curriculum-based measurement (CBM). The benefits of using CBM in a portfolio appear to be quite promising. Teachers have many options for data collection, scoring, interpretation of data both quantitatively and qualitatively, diagnostic evaluation, scaling, and communication across classrooms and grades.

Curriculum-based measurement, with its great flexibility in a learning portfolio, offers a solution to some of the problems identified in the first and second options discussed above. Combining the successful components of each option with the foundation and empirically validated measures of CBM provides us with a tool that can accommodate a wide range of academic behaviors useful for making educational decisions. Unlike standardized tests, they are curriculum- or objective-linked, similar to the attempts at statewide assessment.

A CBM portfolio is a formative evaluation procedure for measuring and evaluating student growth. Unlike standardized measures, in which the scores represent performance at a single point in time, or criterion-referenced testing, where students are measured on material recently taught, the CBM portfolio measures students in goal material, perhaps several times throughout the year. The teacher, therefore, can evaluate a students' progress toward an instructional goal and make programmatic, instructional, or placement changes during the school year.

Educators are also able to interpret the types of data collected,both quantitative and qualitative, using curriculum-based measures. The advantage of this flexibility in assessment is complete evaluation of the student—at the individual level, in comparison to groups and across time. The teacher has data to support instructional or placement changes for a student based on CBM measures in the portfolio within a school year versus waiting for costly individual assessment and/or annual tests. This formative evaluation and multiple use of data collected in a portfolio is possible because of standardized measures and procedures of CBM. Measures of this type applied to a portfolio, provide a picture of what the student has learned at the end of an instructional year.

Additional advantages of the CBM portfolio are illustrated in the study examples. Interpretations of student performance is multifaceted. As an example, students' reading performance was scored both quantitatively and qualitatively. The objective measures were countable and observable skills such as the number of words read in one minute. These scores were used in an individual level, in which the student's performance was compared over time and evaluated for improvement. Additionally, we were able to make group comparisons. In other words we could compare one students' performance with peers using the data provided by local norms, which use the same assessment procedures as the portfolio.

In the case of Chad, our third-grade Chapter One student, individual improvement was outstanding. Chad increased the number of words read correctly from 13 to 45, an individual improvement of 51%. However, as stated in the results section, Chad's performance was not within a normal range as compared to his age level peers. He was performing more like a second-grade reader at the end of his third-grade year. An interesting analysis, beyond the scope of this study would be to evaluate the learning curve of the individual student and that of his peers and estimate how much actual improvement over time would be necessary for the student to effectively catch up in reading fluency.

Another case worth discussion is that of Carl, a regular-education fifth-grade student with an individual reading improvement of over 60 words per minute from fall to spring testing. As stated earlier, research has shown normal growth for intermediate level students to be about 20 words in a school year. Several variables could contribute to such outstanding growth in one year. Carl could have had a significant amount of prior knowledge in the topic area of the randomly selected passages read for spring assessment. However, the likelihood of the two passages being within a high level of knowledge for this student would have been unusual. Although of improbable, another possible explanation of this great performance growth could be that Carl had read both passages previously. This would increase his familiarity with the material, and, if he had practiced the passages, would have helped increase his reading fluency in the spring testing period.

We seriously doubt that the above possibilities explain Carl's extraordinary growth over the school year. Rather, this appears to be a case of the "rich getting richer." Carl appears to be a student who has been able to take full advantage of the educational program put before him and never having needed specialized instruction.

Our writing assessment focused on illustrating various scoring procedures, evaluation attributes, and the relationship between individual performance to over time and individual performance as compared to the group. Test scores clearly showed how diagnosis is possible with a permanent product. Students may show change in attributes that are not reflected in the numerical representation of their writing. This is clear in the case of Jenny, where the quantitative scores were very similar from fall to spring, but qualitatively, she made very pleasing growth in handwriting, word usage and sentence structure.

Apart from the scoring procedures explained here, a few other issues should be addressed. Some of the variation from fall to spring in the writing could be attributed to the writing stimulus (story starter) given students. Even though the assessment procedures are standard, a variety of stimuli are available and, realistically, these may have triggered variations in creativity and quality of students' work based on personal interests, reading experience, and prior knowledge. To determine what is actually happening for a particular student, the teacher could obtain several (more than two) samples to represent the student's writing performance and change over the course of the school year.

The ability to communicate clearly the portfolio concept with CBM is illustrated in several of these selected assessment cases. With the use of box plots, such as Figure 1, it is quite easy for teachers to visually display a student's performance for parents, teachers, as well as the student. The score can be plotted to describe the relationship of one student to the group. The group might be the classroom, grade level in the building, or district grade level peers. A teacher's report to parents, an eligibility team, administrator, or the student might sound something like the following:

"We have tested the whole grade on the same type of test that Joe took. With this information we are able to chart all of the students in the class (or grade) and compare an individual to the group. This chart is called a box plot, it is a picture of how the students in my class did in math on the same type of 2-minute test.

This box represents where most of the children are scoring and everything above the box is very high performance. Those scores below the box indicate performance that is below most of the students in this grade.

Joe's test score was here (teacher points to the chart indicating the Joe's score on the box plot). His score is (above or below) most of the students in my class. That indicates to me that I need to ...."

Teachers are also able to evaluate the permanent product at a diagnostic level. The errors students make on any product are informative for instructional design issues. For example, if a student continually omits end marks in a writing sample, that can indicate some instructional remedies at the mechanics level in language arts. In math, the teacher is able to evaluate factual as well as operational errors.

These diagnostic and evaluation examples are no different than procedures used in specialized diagnostic tests. They don't differ from the information that a teacher may collect on a daily basis when correcting student's work. The benefit of these CBM portfolio tests is the number of evaluation procedures that can be applied to a single assessment which takes very little time to administer. The conditions under which the data is collect is also uniform, so a teacher, or administrator can legitimately compare student performance. Curriculum-based measures are reliable, valid, "user friendly," easy, fast, and inexpensive.

The notion of a learning portfolio has tremendous transportability. Products within a portfolio can accommodate a wide range of academic behaviors that are useful for making a number of educational decisions. They can also be used by a wide range of professionals across general education and special education environments. Because they are classroom-focused, year-long portfolios have a great potential for showing progress over time. Furthermore, they can be useful across all grade levels. Finally, a great diversity exists in the manner in which portfolios are generated and/or maintained.

As teachers become familiar with students in their classes and come to understand the improvements children are making (or not making), they should be better able to make instructional decisions. For some students this might mean simply providing more structure, more practice, more guidance. For others, more intensive interventions may be needed, such as working one-to-one with peers, counselors, or providing guided-practice. There may be evidence for needing extreme interventions requiring assessments for special education placement.

In this scenario, CBM portfolios are very transportable across the general education and special education environments. They provide the focus with which professionals can interact in developing individualized instruction.

In a larger sense, CBM portfolios are timeless. Any particular work sample included within the portfolio simply represents a 35-millimeter snapshot of what a student has done at one point in time. They need little explanation, other than the time and

context in which they were constructed. The very same reporting function can be accomplished within a school among teachers. As students move across the grades, teachers can inform the next grade-level instructor about specific student tendencies (i.e. skill-specific strengths and weaknesses) as illustrated through a CBM portfolio.

Although this study demonstrated the flexibility of the CBM portfolio, it is by no means designed to be prescriptive. In other words, there are no magic numbers for how many times a teacher should collect samples of student works to add to their portfolios.

A number of issues remain to be examined regarding the CBM portfolio. For example, is there an ideal number of samples to collect within a school year? Should the collection of work samples under standardized conditions be restricted to reading, math, spelling, and written expression? How much is enough to capture change over time? Should every child have a portfolio in all areas? Before schools jump ship from current assessment practices and adopt a portfolio concept using CBM, questions like these must be researched and answered empirically to ensure good practices for educational decision-making.

References

Brandt, R. (1989). On misuse of testing: A conversation with George Madaus. *Educational Leadership, 46* (7), 26-29.

Cannell, J.J. (1988). The Lake Wobegon effect revisited. *Educational Measurement: Issues and Practices, 7* (4), 12-15.

Cannell, J.J. (1988). Nationally normed elementary achievement testing in America's public schools: How all 50 states are above the national average. *Educational Measurement: Issues and Practices , 7* (2), 5-9.

Costa, A. L. (Ed.) (1989). Redirecting assessment [Special issue]. *Educational Leadership, 46*(7).

Deno, S.L., (1985). Curriculum-based measurement: The emerging alternative. *Exceptional Children, 52* (3), 219-232.

Deno,S.L., (1989). Curriculum-based measurement and special education services: A fundamental and direct relationship. In M. R. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp. 1-17). New York: Guilford.

Farr, R., & Carney, R. F. (1986). *Reading: What can be measured?* (2nd Ed.). Newark, DE: International Reading Association, Inc.

Flood, J., & Lapp, D. (1989). Reporting reading progress: A comparison portfolio for parents. *The Reading Teacher, 42*(7), 508-515.

Glass, G. V. (1978). Standards and Criteria. *Journal of Educational Measurement, 15*, 237-261.

Goslin, D.A. (1969) *Guidelines for the collection., maintenance, and dissemination of pupil records.* Troy, N.Y.: Russell Sage Foundation.

Hall, T., & Tindal, G. (1989). Using curriculum-based measures to group students in reading. In G. Tindal, K. Essick, N. George, & M. George (Eds.), *The Oregon Conference '89 Monograph.* Eugene, OR: College of Education, University of Oregon.

Haller, E. J., & Waterman, M. (1985). The criteria of reading group assignments. *The Reading Teacher, 38*, 772-781.

Hasbrouck, J., Tindal, G., & Parker, R. (1991). *Forming reading groups using curriculum-based measures: The maze and oral reading fluency.* Manuscript submitted for publication.

Hodgkins, R., & McKenna, B. (1982). Testing and teacher certification: An explosive combination. *Educational Measurement: Issues and Practices, 1*(2), 10-12, 26.

Jenkins, J.R., & Pany, D. (1988). Research on teaching reading comprehension: Instructional variables. In J. Guthrie (Ed.) *Reading comprehension and education..* Newark, Delaware: International Reading Association.

Jervis, K. (1989). Daryl takes a test. *Educational Leadership. 46* (7), 10-15.

Lazar-Morrison, C., Polin, L., Moy, R., & Burry, J. (1980). A Review of the Literature on Test Use. *CSE Report No. 144*, Center for the Study of Evaluation. University of California, Los Angeles.

Levine, M. (1976). The academic achievement test: Its historical context and social functions. *American Psychologist, 31*, 228-238

Martinez, M.E., & Lipson, J.I. (1989). Assessment for Learning. *Educational Leadership, 46*, 73-75.

Messick, S. (1989). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher. 18*(2), 5-11.

Neill, D.M., & Medina, J.J. (1989). Standardized testing: Harmful to educational health. *Phi Delta Kappan, 70*, 688-697.

Nitko, A. J. (1989). Designing tests that are integrated with instruction. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.) (pp.447-474). New York: McMillan.

Popham, W. J. (1978). *Criterion-referenced measurement.* Englewood Cliffs, N.J. : Prentice-Hall.

Rodman, B. (1985). Hundreds in New Jersey seek alternative teacher certification *Education Week, 4,* 39.

Rogers, V.R., & Stevenson, C. (1988). How do we know what kids are learning in school? *Educational Leadership. 45*(5), 68-75.

Salmon-Cox, L. (1981). Teachers and Standardized Achievement Tests: What's really happening? *Phi Delta Kappan , 62,* 631-634.

Salvia, J., & Ysseldyke, J.E. (1988). *Assessment in special and remedial education.* (3rd Ed.). Boston: Houghton-Mifflin Co.

Sheperd, L.A. (1989). Why we need better assessments. *Educational Leadership. 46*(7) 4-9.

Stetz, F., & Beck, M. (1979). *Comments from the classroom: Teachers' and students' opinions of achievement tests.* Paper presented at the Annual Meeting of the American Educational Research Association, San Fransisco.

Tindal, G. (1988). Curriculum-based assessment. In J. Graden, J. Zins, & M. Curtis (Eds.), *Alternative educational delivery systems: Enhancing instructional options for all students* (pp 111-135). Washington, DC: National Association of School Psychologists.

Tindal, G. (1989). Evaluating the effectiveness of educational programs at the systems level using curriculum-based measurement. In M. Shinn (Ed.), *Curriculum-based measurement: Assessing special children* (pp 202-238). New York: Guilford Press.

Tindal, G. (1991). A review of curriculum-based assessment procedures on nine assessment components. In J. Kramer (Ed.), *Buros Symposium,* Lincoln, NE: Buros Mental Measurements Institute.

Valencia, S.W., & Pearson, D. (1987). Reading assessment: Time for a change. *The Reading Teacher. 40*(8), 726-732.

Valencia, S.W., Pearson, D., Peters, C.W., & Wixson, K.K. (1989). Theory and practice in statewide reading assessment: Closing the gap. *Educational Leadership, 46,* 57-63.

Wesson, C.L., Viertaler, J.M., & Haubrich, P.A. (1989). An efficient technique for establishing reading groups. *The Reading Teacher, 42,* 466-469.

Wiggins, G. (1989). Teaching to the (authentic) test. *Educational Leadership, 46,* 41-47.

Wolf, D. P., (1989). Portfolio assessment: Sampling student work. *Educational Leadership. 46,* 35-39.

Yea, J.P., Herman, J.L., & Rudner, L.M. (1981). *Teachers and testing: A survey of test use.* Center for the Study of Evaluation. Los Angeles: University of California: .

Appendix 1

Results of Assessment for All Students in Study

First-Grade Student # C-11 Reg. Ed. High(Tchr. Report)

Test	Fall	Spring
Number Dictation	19 * *=topped out	19* *=topped out
Object Counting	9	16
Addition Comp.	Corr Dig 14 Inc Dig 1	Corr Dig 41 Inc Dig 0
Subt. Comp.	Not given Fall	Corr Dig 14 Inc Dig 1
Letters	7	12
Numbers	16	20
Reading	Corr Wds 13 Errors 4	Corr Wds 146 Errors 0

First-Grade Student # C-12 Low Reg. Ed (Tchr Report)

Test	Fall	Spring
Number Dictation	5	19 * *=topped out
Object Counting	6	17
Addition Comp.	Corr Dig 9 Inc Dig 9	Corr Dig 41 Inc Dig 1
Subt. Comp.	Not given Fall	Corr Dig 19 Inc Dig 1
Letters	12	26
Numbers	13	23
Reading	Corr Wds 4 Errors 10	Corr Wds 100 Errors 1

First-Grade Student # I-11 Reg. Ed High (Tchr Report)

Test	Fall	Spring
Number Dictation	19 * *=topped out	19 * *=topped out
Object Counting	10	14
Addition Comp.	Corr Dig 23 Inc Dig 0	Corr Dig 41 Inc Dig 2
Subt. Comp.	Not given Fall	Corr Dig 20 Inc Dig 0
Letters	11	12
Numbers	data missing	29
Reading	Corr Wds 106 Errors 1	Corr Wds 114 Errors 4

First Grade Student # I-12 Reg Ed LOW (Tchr Report)

Test	Fall	Spring
Number Dictation	missing data	19 * *=topped out
Object Counting	5	13
Addition Comp.	Corr Dig 0 Inc Dig 15	Corr Dig 18 Inc Dig 9
Subt. Comp.	Not given Fall	Corr Dig 3 Inc Dig 33
Letters	13	13
Numbers	13	29
Reading	Corr Wds 2 Errors 8	Corr Wds 12 Errors 5

Second-Grade Student #C-23: Special Education

IEP areas Reading Spelling
Math Oral Lang
Handwriting

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 1	Inc. Dig. 6	Xed Out 5	Corr. Dig 15	Inc. Dig 11	Xed Out 0
Spelling	CLS 32	Tot Wds 3		CLS 47	Tot Wds 10	
Written Expresn	Tot Wds 13	% Corr. 30%	CWS 4	Tot Wds 26	% Corr. 44%	CWS 12
Reading	Corr. Wds 24	Inc. Wds 23		Corr. Wds 32	Inc. Wds 17	

Second-Grade Student #I-23 Special Education

IEP Areas: Reading Math
Spelling

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 5	Inc. Dig. 8	Xed Out 0	Corr. Dig 15	Inc. Dig 4	Xed Out 0
Spelling	CLS 33	Tot Wds 2		CLS 65	Tot Wds 8	
Written Expresn	Tot Wds 9	% Corr. 33%	CWS 3	Tot Wds 9	% Corr. 66%	CWS 6
Reading	Corr. Wds 9	Inc. Wds 6		Corr. Wds 44	Inc. Wds 3	

Third-Grade Student #C-36 Chapter 1 (Reading)

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 28	Inc. Dig. 0	Xed Out 2	Corr. Dig 24	Inc. Dig 0	Xed Out 3
Spelling	CLS 70	Tot Wds 4		CLS 94	Tot Wds 11	
Written Expresn	Tot Wds 30	% Corr. 71%	CWS 22	Tot Wds 20	% Corr. 77%	CWS 17
Reading	Corr. Wds 13	Inc. Wds 20		Corr. Wds 45	Inc. Wds 5	

Third-Grade Student #C-35 Chapter 1 (Reading)

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 20	Inc. Dig. 2	Xed Out 1	Corr. Dig 8	Inc. Dig 11	Xed Out 0
Spelling	CLS 76	Tot Wds 6		CLS 92	Tot Wds 11	
Written Expresn	Tot Wds 24	% Corr. 72%	CWS 18	Tot Wds 21	% Corr. 47%	CWS 10
Reading	Corr. Wds 23	Inc. Wds 8		Corr. Wds 56	Inc. Wds 0	

Fourth-Grade Student # C-43: Special Education

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 4	Inc. Dig. 9	Xed Out 2	Corr. Dig 21	Inc. Dig 16	Xed Out 0
Spelling	CLS 70	Tot Wds 3		CLS 82	Tot Wds 6	
Written Expresn	Tot Wds 23	% Corr. 52%	CWS 11	Tot Wds 27	% Corr. 69%	CWS 18
Reading	Corr. Wds 49	Inc. Wds 4		Corr. Wds 57	Inc. Wds 2	

Fourth-Grade student #I-44: Special Education

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 30	Inc. Dig. 6	Xed Out 8	Corr. Dig 8	Inc. Dig 4	Xed Out 0
Spelling	CLS 56	Tot Wds 2		CLS 89	Tot Wds 6	
Written Expresn	Tot Wds 37	% Corr. 41%	CWS 16	Tot Wds 35	% Corr. 63%	CWS 23
Reading	Corr. Wds 77	Inc. Wds 7		Corr. Wds 101	Inc. Wds 3	

Fifth-Grade Student #I-51 Reg. Ed., High Level (Tchr. Report)

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 55	Inc. Dig. 0	Xed Out 0	Corr. Dig 37	Inc. Dig 1	Xed Out 1
Spelling	CLS 103	Tot Wds 8		CLS 110	Tot Wds 11	
Written Expresn	Tot Wds 58	% Corr. 78%	CWS 13	Tot Wds 58	% Corr. 96%	CWS 63
Reading	Corr. Wds 95	Inc. Wds 10		Corr. Wds 157	Inc. Wds 1	

Fifth-Grade Student #I-52 Reg. Ed., Low Level (Tchr. Report)

Test	FALL	FALL	FALL	SPRING	SPRING	SPRING
Mixed Probe	Corr. Dig 10	Inc. Dig. 31	Xed Out 0	Corr. Dig 18	Inc. Dig 8	Xed Out 2
Spelling	CLS 64	Tot Wds 3		CLS 78	Tot Wds 4	
Written Expresn	Tot Wds 34	% Corr. 71%	CWS 25	Tot Wds 35	% Corr. 78%	CWS 30
Reading	Corr. Wds 64	Inc. Wds 6		Corr. Wds 84	Inc. Wds 2	

Appendix 2

Scored Protocols for Students Discussed in Report

Written Expression

The case of Matt Fall

Write a story that begins with:

I was walking down the street and found a key in front of a big gold door. I took the ~~key, opened~~ the door and ther Was • Was^a^man Sahing l^the w

TW 9
CWS 3
IWS 6

Second grade
Special Ed. Student Sample no. I-23
Written Expression
Spring (5/90)

Written Expression

The case of Matt Spring

Write a story that begins with:

One very dark and spooky night I was camping in the woods. When I heard a strange noise....

^a^bean^ a^gun^ shot
a^pr_sht^ a^dog.

TW 9
CWS 6
TWS 3

Second grade
Special Ed. Student Sample no. I-23
Mixed addition and subtraction
Fall (10/89)

Curriculum-Based Measurement

Math 2

Fall

The case of Matt

22 + 43 65	4 3 + 8 49	48 - 8 40	2 0 + 2 4
872 - 9 91	67 + 32 100	70 + 9 7	305 - 6
447 - 51	9 8 + 4	2 2 + 4	12 + 6
212 + 769	678 - 46	782 - 8	393 + 790
13 + 70	142 - 7	33 + 26	712 - 590

Total digits correct

5

Second grade
Special Ed. Student Sample no. I-23
Mixed addition and subtraction
Spring (5/90)

Curriculum-Based Measurement

Math 2

The case of Matt Spring

11 + 6 17	28 + 50 78	50 + 31 81	502 + 326 828	9
11 + 57 02	46 - 45 01	2 9 + 5 15 -5	322 + 453 11151	4/3
42 + 54	91 + 1	10 + 22	551 - 6 555	2/1
605 - 71 6	737 - 3	73 - 37	767 - 368	
368 + 110	468 - 5	17 + 48	251 + 448	

Total digits correct

15

Second grade
Special Ed. Student Sample no. C-23
Mixed addition and subtraction
Fall (10/89)

Curriculum-Based Measurement

Math 2

The case of Emily

Fall

11 + 6	28 + 50	50 + 31 57	502 + 326
11 + 57 58	46 - 45	9 + 5	322 + 453 36
42 + 54	91 + 1	10 + 22	551 - 6
605 - 71	737 - 3	73 - 37	767 - 368
368 + 110	468 - 5	17 + 48	251 + 448

Total digits correct

1

Second grade
Special Ed. Student Sample no. C-23
Mixed addition and subtraction
Spring (5/90)

Curriculum-Based Measurement

Math 2

The case of Emily Spring

22 + 43 16	4 3 + 8 14	48 - 8 40	2 0 + 2 4
872 - 9 877	67 + 32 99	70 + 9 79	305 - 6 301
447 - 51 416	9 8 + 4 21	2 2 + 4 8	12 + 6 18
212 + 769 1	678 - 46	782 - 8	393 + 790
13 + 70	142 - 7	33 + 26	712 - 590

Total digits correct

15

Written Expression

The case of Jenny Fall

Write a story that begins with:

I was playing outside when a spaceship landed and....

^I^was•wunering•wat•was^
^iot^I^ran^up^to^it^and^I^Looked
^inside•thiar•was^something^moving
^inside^of^the•Pase.Ship.

TW 24
CWS 18
IWS 7

Written Expression

The case of Jenny Spring

Write a story that begins with:

Our sailboat ran into some rocks and crashed. We were stranded on an island ....

• an • we ^ called ^ for ^ help ^ but • their • no one • to ^ help ^ us ^ so ^ we • watched • our ^ boat • sank • to ^ the • bolem.

TW 21
^ CWS 10
• IWS 11

Appendix 3

Unscored Protocols for All Students in Study

Second grade
Special Ed. Student Sample no. C-23
Mixed addition and subtraction
Fall (10/89)

Curriculum-Based Measurement

Math 2

11 + 6	28 + 50	50 + 31 57	502 + 326
11 + 57 58	46 - 45 1	9 + 5	322 + 453 36
42 + 54	91 + 1	10 + 22	551 - 6
605 - 71	737 - 3	73 - 37	767 - 368
368 + 110	468 - 5	17 + 48	251 + 448

Total digits correct

_____

Second grade
Special Ed. Student Sample no. C-23
Mixed addition and subtraction
Spring (5/90)

Curriculum-Based Measurement

Math 2

22 + 43 V 6	4 3 + 8 14	48 - 8 40	2 0 + 2 4
872 - 9 877	67 + 32 99	70 + 9 79	305 - 6 301
447 - 51 416	9 8 + 4 21	2 2 + 4 8	12 + 6 18
212 + 769 7	678 - 46	782 - 8	393 + 790
13 + 70	142 - 7	33 + 26	712 - 590

Total digits correct

_____

Written Expression

Write a story that begins with:

I was playing outside when a spaceship landed and....

I Sol a, Spaceship
I Wadr a bat et
Hesed Iland his

Second grade
Special Ed. Student Sample no. C-23
Written Expression
Spring (5/90)

Written Expression

Write a story that begins with:

Our sailboat ran into some rocks and crashed. We were stranded on an island ....

They are live on that i sa bot they ar onL to go badR to the Woter that WEs thar on Le wae home to See thaer fa mle.

Spelling

Second Grade
Special Ed. Student Sample no. C-23
Spelling
Fall (10/89)

1. bus
2. h
3. If
4. find
5. wet
6. S
7. to
8. wati
9. not
10. yu
11. b
12. shis
13. boy
14. fiv
15. men
16.
17. su

Scott-Foresman
Grade 2

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**bus**	the **bus** is late	**bus**	4	4
2. (10)	**here**	**here** are your things	**here**	5	9
3. (20)	**of**	which **of** those is mine	**of**	3	12
4. (30)	**fine**	a **fine** day	**fine**	5	17
5. (40)	**wet**	**wet** paint	**wet**	4	21
6. (50)	**sent**	he **sent** a package	**sent**	5	26
7. (60)/(1:00)	**too**	**too** far for walking	**too**	4	30
8. (70)/(1:10)	**wanted**	he **wanted** a brother	**wanted**	7	37
9. (80)/(1:20)	**not**	**not** right now	**not**	4	41
10. (90)/(1:30)	**yellow**	a **yellow** bird	**yellow**	7	48
11. (100)/(1:40)	**birthdays**	**birthdays** are great	**birthdays**	10	48
12. (110)/(1:50)	**tries**	she **tries** very hard	**tries**	6	64
13. (120)/(2:00)	**brother**	my **brother** is ten	**brother**	8	72
14. (130)/(2:10)	**five**	**five** in the family	**five**	5	77
15. (140)/(2:20)	**some**	have **some** dessert	**some**	5	82
16. (150)/(2:30)	**mean**	a **mean** dog	**mean**	5	87
17. (160)/(2:40)	**stay**	**stay** at home	**stay**	5	92

Total Words Correct	Correct letter Sequences

Spelling

Second grade
Special Ed. Student Sample no. C-23
Spelling
Spring (5/90)

1. Onigs ____
2. they ____
3. did ____
4. Pat ____
5. runen ____
6. Prity ____
7. dad ____
8. flamoly ____
9. ran y ____
10. Riry ____
11. Working ____
12. feet ____
13. foob ____
14. dress ____
15. or ____
16. fine ____
17. into ____

Scott-Foresman
Grade 2

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**orange**	the **orange** was rotten	**orange**	7	7
2. (10)	**they**	**they** will fly home	**they**	5	12
3. (20)	**did**	what **did** you do?	**did**	4	16
4. (30)	**pat**	**pat** your tummy	**pat**	4	20
5. (40)	**running**	**running** a race	**running**	8	28
6. (50)	**pretty**	a **pretty** dress	**pretty**	7	35
7. (60)/(1:00)	**dad**	his **dad** went fishing	**dad**	4	39
8. (70)/(1:10)	**family**	my **family** likes pets	**family**	7	46
9. (80)/(1:20)	**rain**	the **rain** stopped	**rain**	5	51
10. (90)/(1:30)	**kick**	let's **kick** the ball	**kick**	5	56
11. (100)/(1:40)	**working**	I'm **working** today	**working**	8	64
12. (110)/(1:50)	**feet**	my **feet** hurt	**feet**	5	69
13. (120)/(2:00)	**food**	no **food** in the gym	**food**	5	74
14. (130)/(2:10)	**dress**	**dress** warmly	**dress**	6	80
15. (140)/(2:20)	**or**	one **or** more	**or**	3	83
16. (150)/(2:30)	**fine**	a **fine** job	**fine**	5	88
17. (160)/(2:40)	**into**	get **into** the car	**into**	5	93

Total Words Correct	Correct letter Sequences

Second grade
Special Ed. Student Sample no. I-23
Mixed addition and subtraction
Spring (5/90)

Curriculum-Based Measurement

Math 2

11 + 6 17	28 + 50 78	50 + 31 81	502 + 326 828
11 + 57 68	46 - 45 01	2 9 + 5 15	322 + 453 1151
42 + 54	91 + 1	10 + 22	551 - 6 555
605 - 71 6	737 - 3	73 - 37	767 - 368
368 + 110	468 - 5	17 + 48	251 + 448

Total digits correct

_____

Written Expression

Write a story that begins with:

I was walking down the street and found a key in front of a big gold door. I took the key, opened the door and....ther Was Wasusman Sahingl the u

Second grade
Special Ed. Student Sample no. I-23
Written Expression
Spring (5/90)

Written Expression

Write a story that begins with:

One very dark and spooky night I was camping in the woods. When I heard a strange noise....

a bear a gun shot
a prist a dog

Spelling

Second Grade
Special Ed. Student Sample no. I-23
Spelling
Fall (10/89)

1. hb
2. mer
3. hup
4. wil
5. look
6. yell
7. did
8. zir
9. Bush
10. s
11. yez
12. d
13. mel
14. few
15. wle
16. Im
17. Imr

Scott-Foresman
Grade 2

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**happy**	a **happy** day	**happy**	6	6
2. (10)	**men**	**men** were working	**men**	4	10
3. (20)	**hopped**	I **hopped** over	**hopped**	7	17
4. (30)	**wish**	**wish** upon a star	**wish**	5	22
5. (40)	**looked**	it **looked** nice	**looked**	7	29
6. (50)	**yellow**	**yellow** is a color	**yellow**	7	36
7. (60)/(1:00)	**don't**	**don't** fall down	**don't**	6	42
8. (70)/(1:10)	**stamp**	**stamp** your feet	**stamp**	6	48
9. (80)/(1:20)	**bushes**	**bushes** have stickers	**bushes**	7	55
10. (90)/(1:30)	**tries**	he **tries** hard	**tries**	6	61
11. (100)/(1:40)	**yes**	**yes** is the answer	**yes**	4	65
12. (110)/(1:50)	**desk**	the **desk** is messy	**desk**	5	70
13. (120)/(2:00)	**mail**	**mail** the letter	**mail**	5	75
14. (130)/(2:10)	**five**	**five** students	**five**	5	80
15. (140)/(2:20)	**well**	the **well** was deep	**well**	5	85
16. (150)/(2:30)	**I'm**	**I'm** almost finished	**I'm**	5	89
17. (160)/(2:40)	**inch**	an **inch** worm	**inch**	5	94

Total Words Correct

Correct letter Sequences

Second grade
Special Ed. Student Sample no. I-23
Spelling
Spring (5/90)

Spelling

1. Cool ____
2. stay ____
3. very ____
4. hapt ____
5. we're ____
6. fen ____
7. cris ____
8. tow ____
9. get ____
10. Looking ____
11. W ____
12. es cn ____
13. peht ____
14. him ____
15. dris ____
16. Bath ____
17. ther ____

Scott-Foresman
Grade 2

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**cool**	that **cool** kid	**cool**	5	5
2. (10)	**stay**	I'll **stay** home tonight	**stay**	5	10
3. (20)	**very**	it's **very** cold outside	**very**	5	15
4. (30)	**napped**	he **napped** all day	**napped**	7	22
5. (40)	**we're**	**we're** having fun	**we're**	6	28
6. (50)	**send**	**send** me a note	**send**	5	33
7. (60)/(1:00)	**cries**	she **cries** when sad	**cries**	6	39
8. (70)/(1:10)	**to**	get **to** the point	**to**	3	42
9. (80)/(1:20)	**get**	I'll **get** the car	**get**	4	46
10. (90)/(1:30)	**looking**	his **looking** upstairs	**looking**	8	54
11. (100)/(1:40)	**white**	the **white** paper	**white**	6	60
12. (110)/(1:50)	**inch**	two **inch** hem	**inch**	5	65
13. (120)/(2:00)	**paint**	**paint** the house	**paint**	6	71
14. (130)/(2:10)	**him**	give **him** a prize	**him**	4	75
15. (140)/(2:20)	**dress**	the **dress** is new	**dress**	6	81
16. (150)/(2:30)	**bath**	the **bath** felt good	**bath**	5	86
17. (160)/(2:40)	**their**	**their** party was fun	**their**	6	92

Total Words Correct	Correct letter Sequences

Third grade
Chapter 1 Student Sample no. C-36
Mixed addition and subtraction
Fall (10/89)

Curriculum-Based Measurement

Math 3

11 + 1 12	77 - 31 46	948 - 811 137	~~17~~ ~~x 3~~
~~35~~ ~~x 1~~	9 9 + 4 22	12 + 10 22	487 - 405 082
28 - 24 4	6 + 10 16	30 16 + 31 77	1 41 84 + 47 172
1 740 + 167 907	1 44 84 + 38 166	95 - 34 6	73 - 21
40 x 0	68 32 + 59	648 + 271	61 x 8

Total digits correct

_____

Curriculum-Based Measurement
Math 3

18 13 + 23 54	75 + 23 98	672 - 547 125	6 x 8	132 + 544 676
13 - 9 4	73 42 + 2 117	11 x 3	85 x 2	15 + 29 44
24 - 12 12	511 + 713 1224	49 - 22 27	33 x 3	32 + 26
48 + 39	16 - 9	1 x 2	214 + 762	3 x 8
13 + 63	973 + 334	11 x 5	70 19 + 33	6 x 6

Total Digits Correct

_____

Written Expression

Write a story that begins with:

I was playing outside when a spaceship landed and....

I is very green it is very short. it has very long hair it is a girl it has a babby named bnanas. it has three sisters their names are.

Written Expression

Write a story that begins with:

Our sailboat ran into some rocks and crashed. We were stranded on an island ....

We went to chop
down some trees so that
we could build a boat.
we found a pile of gl imer-
ings

Spelling

Third Grade
Special Ed. Student Sample no. C-36
Spelling
Fall (10/89)

1. gerges ____
2. sonethe ____
3. tired ____
4. naber ____
5. oncedce ____
6. garden ____
7. lorucr ____
8. brefest ____
9. watter ____
10. scrimimed ____
11. built ____
12. ratter ____
13. shove ____
14. streeng ____
15. sopose ____
16. hellper ____
17. may ____

Scott-Foresman
Grade 3

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	jeans	most jeans are blue	jeans	6	6
2. (10)	son	my son is asleep	son	4	10
3. (20)	tired	I'm tired of working	tired	6	16
4. (30)	neighbor	my neighbor visited	neighbor	9	25
5. (40)	ounce	one ounce of liquid	ounce	6	31
6. (50)	garden	the garden of flowers	garden	7	38
7. (60)/(1:00)	lawyer	a lawyer won her case	lawyer	7	45
8. (70)/(1:10)	breakfast	breakfast at eight	breakfast	10	55
9. (80)/(1:20)	waiter	the waiter served	waiter	7	62
10. (90)/(1:30)	scream	don't scream so loud	scream	7	69
11. (100)/(1:40)	built	they built a fort	built	6	75
12. (110)/(1:50)	writer	a writer of this book	writer	7	82
13. (120)/(2:00)	shove	shove the table	shove	6	88
14. (130)/(2:10)	string	string the beads	string	7	95
15. (140)/(2:20)	suppose	I suppose he's right	suppose	8	103
16. (150)/(2:30)	helper	your helper is here	helper	7	110
17. (160)/(2:40)	May	is May next month?	May	4	114

Total Words Correct	Correct letter Sequences

Spelling

1. explore ____
2. lisird ____
3. closed ____
4. today ____
5. beautiful ____
6. knot ____
7. table ____
8. tired ____
9. marrine ____
10. repante ____
11. navy ____
12. no ____
13. remake ____
14. Suerner ____
15. throw ____
16. canyn ____
17. kneed ____

Curriculum-Based Measurement

Math 3

11 + 1 12	77 - 31 46	948 - 811 137	~~17~~ x 3
35 x 1 35	9 9 + 4 22	12 + 10 22	487 - 405 082
28 - 24 04	6 + 10 16	30 16 + 31 77	41 84 + 47 16
740 + 167	44 84 + 38	95 - 34	73 - 21
40 x 0	68 32 + 59	648 + 271	61 x 8

Total digits correct

_____

Curriculum-Based Measurement
Math 3

18 13 + 23 43	75 + 23 98	672 - 547 134	6 x 8	132 + 544 676
13 - 9 16	1 73 42 + 2 18	11 x 3 14	85 x 2 816	15 + 29
24 - 12	511 + 713	49 - 22	33 x 3	32 + 26
48 + 39	16 - 9	1 x 2	214 + 762	3 x 8
13 + 63	973 + 334	11 x 5	70 19 + 33	6 x 6

Total Digits Correct

_____

Written Expression

Write a story that begins with:

I was playing outside when a spaceship landed and....

I was wunering wat was
init I ran up to it and I Looked
inside thar was something moving
inside of the spaceship.

Written Expression

Write a story that begins with:

Our sailboat ran into some rocks and crashed. We were stranded on an island ....

an we called for help but theis no one to help us so we wached our boat sank to the bolem.

Spelling

Third grade
Chapter 1 Student Sample no. C-35
Spelling
Spring (5/90)

1. explore
2. lizard
3. closed
4. today
5. butiful
6. knot
7. table
8. tired
9. marene
10. repaint
11. naved
12. no
13. semake
14. suvenare
15. thoughe
16. canyon
17. knew

Scott-Foresman
Grade 3

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**explore**	**explore** the island	**explore**	8	8
2. (10)	**lizard**	a **lizard** is scaley	**lizard**	7	15
3. (20)	**closed**	they **closed** the shop	**closed**	7	22
4. (30)	**today**	**today** is sunny	**today**	6	28
5. (40)	**beautiful**	a **beautiful** day	**beautiful**	10	38
6. (50)	**knot**	the **knot** won't untie	**knot**	5	43
7. (60)/(1:00)	**table**	a **table** for four	**table**	6	49
8. (70)/(1:10)	**tired**	I'm **tired** of writing	**tired**	6	55
9. (80)/(1:20)	**marine**	a **marine** scientist	**marine**	7	62
10. (90)/(1:30)	**repaint**	**repaint** the house	**repaint**	8	70
11. (100)/(1:40)	**navy**	a **navy** officer	**navy**	5	75
12. (110)/(1:50)	**no**	**no**, you can't go	**no**	3	78
13. (120)/(2:00)	**remake**	**remake** your bed	**remake**	7	85
14. (130)/(2:10)	**souvenir**	a **souvenir** t-shirt	**souvenir**	9	94
15. (140)/(2:20)	**throw**	**throw** the ball	**throw**	6	100
16. (150)/(2:30)	**canyon**	the **canyon** is huge	**canyon**	7	107
17. (160)/(2:40)	**knee**	her **knee** is bent	**knee**	5	112

Total Words Correct	Correct letter Sequences

Spelling

Third grade
Chapter 1 Student Sample no. C-35
Spelling
Fall (10/90)

1. Jeens
2. son
3. tiered
4. nober
5. unts
6. garden
7. lourer
8. bracfast
9. water
10. screem
11. biult
12. writer
13. shuve
14. string
15. supos
16. helper
17. May

Fourth grade
Special Ed. Student Sample no. C-43
Mixed math
Fall (10/89)

Curriculum-Based Measurement

Math 4

70 - 46 36	720 -499 374	974 - 14 960	94)4700
264	5 3 +9 16	5 1)7	22 x43 21
72 x 4	301 x 2	250 x 1	323 + 46
4)5	434 x 1	37 x 8	5)80
524 +412	302 x 90	111 x 9	4874 -1360

Total digits correct

_____

Fourth grade
Special Ed. Student Sample no. C-43
Mixed math
Spring (5/90)

Curriculum-Based Measurement
Math 4

$5\overline{)20}$ → 9

$958 \times 1 = 958$

$28 \times 6 = 426$

$129 \times 1 = 129$

$83 \times 11 = 83$

$180 \times 6 = 6260$

$209 \times 1 = 209$

$40 + 40 + 27 = 107$

$5849 - 1349 = 4500$

$236 \times 30 = 000 + 69180 = 69180$

174×1

$5\overline{)63}$

$3\overline{)2469}$

11×22

340×80

$5694 - 4240$

$6\overline{)42}$

210×30

$122 + 86$

394×80

Total digits correct

_____

Written Expression

Write a story that begins with:

I was playing outside when a spaceship landed and....

and tauck eve one to mars and made them work for then for two hundred yers then took them back to erth the poeple

Written Expression

Write a story that begins with:

Our sailboat ran into some rocks and crashed. We were stranded on an island ....

We starded to fix the boat but it fell down again then we Made a house then it fell down we couldod do ener thing with out a han

1. Clld hood ____
2. Clent ____
3. medl ____
4. Compooter ____
5. shodel ____
6. reptile ____
7. Preseac ____
8. meet ____
9. bach ____
10. Horse ____
11. gold finch ____
12. fichen ____
13. Ruc ____
14. Up Ses ____
15. Steel ____
16. lush ____
17. ____

Scott-Foresman
Grade 4

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**childhood**	**childhood** friends	**childhood**	10	10
2. (8)	**clarinet**	the **clarinet** sounds neat	**clarinet**	9	19
3. (16)	**metal**	the **metal** melted	**metal**	6	25
4. (24)	**computer**	my **computer** is fast	**computer**	9	34
5. (32)	**shoulder**	his **shoulder** is broken	**shoulder**	9	43
6. (40)	**reptile**	a **reptile** is cold-blooded	**reptile**	8	51
7. (48)	**parentheses**	**parentheses** for math	**parentheses**	12	63
8. (56)	**meet**	**meet** my friend	**meet**	5	68
9. (64)/(1:04)	**badge**	a **badge** for bravery	**badge**	6	74
10. (72)/(1:12)	**hoarse**	her **hoarse** voice	**hoarse**	7	81
11. (80)/1:20)	**goldfinch**	a **goldfinch** flew away	**goldfinch**	10	91
12. (88)/(1:28)	**fiction**	**fiction** is fun to read	**fiction**	8	99
13. (96)/(1:32)	**percussion**	**percussion** is a sound	**percussion**	11	110
14. (104)/(1:40)	**upstairs**	**upstairs** in my room	**upstairs**	9	119
15. (112)/(1:48)	**steel**	**steel** is a tough metal	**steel**	6	125
16. (120)/(1:56)	**thrush**	a **thrush** is a bird	**thrush**	7	132

Total Words Correct	Correct letter Sequences

Fourth grade
Special Ed. Student Sample no. C-43
Spelling
Spring (5/90)

Spelling

1. linkinny _____
2. nashin _____
3. laughter _____
4. ficken _____
5. indsry _____
6. crushes _____
7. shiver _____
8. book _____
9. crash _____
10. nisido _____
11. gbbdfinch _____
12. triming _____
13. debing _____
14. fourting _____
15. burbing _____
16. _____
17. _____

Scott-Foresman
Grade 4

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**Lincoln**	**Lincoln** was president	**Lincoln**	8	8
2. (8)	**nation**	a **nation** united	**nation**	7	15
3. (16)	**laughter**	**laughter** is a medicine	**laughter**	9	24
4. (24)	**fiction**	**fiction** is not fact	**fiction**	8	32
5. (32)	**industry**	**industry** needs work	**industry**	9	41
6. (40)	**crushes**	it **crushes** the grapes	**crushes**	8	49
7. (48)	**shiver**	I **shiver** in the cold	**shiver**	7	56
8. (56)	**bookstore**	the **bookstore** closed	**bookstore**	10	66
9. (1:04)	**crash**	the **crash** was awful	**crash**	6	72
10. (1:12)	**mosquito**	a **mosquito** bit me	**mosquito**	9	81
11. (1:20)	**goldfinch**	a **goldfinch** flew by	**goldfinch**	10	91
12. (1:28)	**trimming**	he's **trimming** the tree	**trimming**	9	100
13. (1:32)	**deafening**	**deafening** thunder	**deafening**	10	110
14. (1:40)	**fourteen**	**fourteen** years old	**fourteen**	9	119
15. (1:48)	**surfing**	**surfing** is great fun	**surfing**	8	127

Total Words Correct

Correct letter Sequences

Curriculum-Based Measurement

Math 4

~~5)20~~	958 x 1 958	~~28~~ ~~x 6~~	129 x 1 129
~~83~~ ~~X 11~~	180 x 6 180	209 x 1 209	40 40 +27 107
5849 -1349 4500	236 x 30 290	174 x 1 174	~~363~~
~~3)2469~~	11 x22 22	~~340~~ ~~x 80~~	5694 -4240 1454
~~6)42~~	210 x 30 230	1 122 + 86 208	~~394~~ ~~x 80~~

Total digits correct

_____

Curriculum-Based Measurement

Math 4

70 - 46 24	520 -499 221	974 - 14 960	8 94)4700 752
2)54	5 3 +9	1)7	22 x43
72 x 4	301 x 2	250 x 1	323 + 46
4)5	434 x 1	37 x 8	5)80
524 +412	302 x 90	111 x 9	4874 -1360

Total digits correct

_____

Written Expression

Write a story that begins with:

I was walking down the street and found a key in front of a big gold door. I took the key, opened the door and....

and fond a big gold key and I like to Kupe the Key but my mom wont late me Kuep the sh. I ray I to ren tehe Key out my whinde and nast maning ~~I~~ seen. That

Written Expression

Write a story that begins with:

One very dark and spooky night I was camping in the woods. When I heard a strange noise....

it look be chonk me and
I could not see it sow
I look a gene and got
a place loote and look
it was my friend he woe
brong popcorn to me. And
then at pick

Fourth grade
Special Ed. Student Sample no. I-44
Spelling
Fall (10/90)

1. hospitol ____
2. allomana ____
3. tamberren ____
4. memro ____
5. Kalar ____
6. warid ____
7. caben ____
8. saning ____
9. lowiano ____
10. I've ____
11. doller ____
12. frie ____
13. person ____
14. sote ____
15. revrater ____
16. mary ____
17. desoge ____

Scott-Foresman
Grade 4

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**hospital**	the **hospital** is huge	**hospital**	9	9
2. (8)	**aluminum**	**aluminum** is recyclable	**aluminum**	9	18
3. (16)	**tambourine**	the **tambourine** jingles	**tambourine**	11	29
4. (24)	**memoir**	my **memoir** was written	**memoir**	7	36
5. (32)	**clarinet**	the **clarinet** played music	**clarinet**	9	45
6. (40)	**worried**	I **worried** about you	**worried**	8	53
7. (48)	**copper**	**copper** is a metal	**copper**	7	60
8. (56)	**shining**	a **shining** sun	**shining**	8	68
9. (64)/(1:04)	**Louisiana**	**Louisiana** is south	**Louisiana**	10	78
10. (72)/(1:12)	**I've**	**I've** had fun	**I've**	5	83
11. (80)/1:20)	**dollar**	a **dollar** allowance	**dollar**	7	90
12. (88)/(1:28)	**fry**	**fry** fish in butter	**fry**	4	94
13. (96)/(1:32)	**person**	this **person** is a friend	**person**	7	101
14. (104)/(1:40)	**sort**	**sort** through the toys	**sort**	5	106
15. (112)/(1:48)	**refrigerator**	a **refrigerator** is cold	**refrigerator**	13	119
16. (120)/(1:56)	**marry**	**marry** your true love	**marry**	6	125
17. (128)/(2:04)	**disagree**	we **disagree** on politics	**disagree**	9	134

_____

_____

Total Words Correct

Correct letter Sequences

Spelling

Fourth grade
Special Ed. Student Sample no. I-44
Spelling
Spring (5/90)

1. bluebird ____
2. texas ____
3. mangnet ____
4. mxdope ____
5. siting ____
6. kinold ____
7. defance ____
8. repair ____
9. ballet ____
10. sodle ____
11. lounge ____
12. loolorence ____
13. lifeboat ____
14. welcome ____
15. dozze ____
16. trees ____
17. dividing ____

Fourth grade
Spelling Test Key
Spring (5/90)

Scott-Foresman
Grade 4

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**bluebird**	a **bluebird** flew by	**bluebird**	9	9
2. (8)	**Texas**	**Texas**, the lone star state	**Texas**	6	15
3. (16)	**magnet**	a **magnet** attracts	**magnet**	7	22
4. (24)	**encyclopedia**	**encyclopedia** of dogs	**encyclopedia**	13	35
5. (32)	**sitting**	**sitting** on the porch	**sitting**	8	43
6. (40)	**chemical**	**chemical** storage bin	**chemical**	9	52
7. (48)	**defend**	**defend** your friends	**defend**	7	59
8. (56)	**repay**	**repay** your debts	**repay**	6	65
9. (1:04)	**ballot**	a **ballot** for voting	**ballot**	7	72
10. (1:12)	**soil**	this **soil** is rich	**soil**	5	77
11. (1:20)	**lounge**	**lounge** around a pool	**lounge**	7	84
12. (1:28)	**ballerina**	the **ballerina** is tiny	**ballerina**	10	94
13.(1:32)	**lifeboat**	a **lifeboat** for safety	**lifeboat**	9	103
14. (1:40)	**welcome**	**welcome** to school	**welcome**	8	111
15. (1:48)	**dust**	**dust** the furniture	**dust**	5	116
16. (1:56)	**trees**	**trees** in the forest	**trees**	6	122
17. 2:04)	**dividing**	**dividing** is more fun	**dividing**	9	131

Total Words Correct

Correct letter Sequences

Curriculum-Based Measurement

Math 5

59 x 69 = 4,071

142 x 80 = 11,360

86 x 20 = 1,720

85 x 11 = 935

40 x 5 = 200

41 + 16 = 79

93 x 11 = 1023

177 - 37 = 140

557 x 8

4) 7

90 - 33

22 x 44

95)199

3)396

422 - 65

358 x 60

199 - 10

32 x 31

45 - 45

12 x 41

Curriculum-Based Measurement
Math 5

47 x 21 47 74 0 787	215 25) 55 50 5	846 - 415 431	30.49 124) 4769 472 49	30 x 50 1500
13 36) 540 36 180 7	536 7723 + 164 8423	421 x 343 3	794 - 320	21 x 45
3265 873 + 7365	8) 216	87 x 76	125 x 51	87) 9724
9) 925	23) 69	956 - 770	4302 + 1147	523 x 48
233 x 34	9 - 5	6 x 3	1923 + 3683	36 x 5

Total Digits Correct _____

Written Expression

Write a story that begins with:

I was walking down the street and found a key in front of a big gold door. I took the key, opened the door and....

I saw a big stachue heard a big thump so I look behind another stachue. lots and lots of people passed after that I decided to go home. everybody left then I went out put the key down and ran home. I told my brother what happened he said I don't belive you so I told my parents

Written Expression

Write a story that begins with:

One very dark and spooky night I was camping in the woods. When I heard a strange noise....

It sounded like a bear growling. I got up from my bed. I went out of the tent to see what it was. I saw a giant bear eating all of our food. I couldn't stop it. I did not bring any guns or knives. I tried to get its attention but it just kept on eating. Then

Fifth grade
Reg. Ed. (High) Student Sample no. I-51
Spelling
Fall (10/90)

1. scientific ____
2. richul ____
3. Indiana ____
4. income ____
5. gaint ____
6. sample ____
7. oatmel ____
8. fisices ____
9. truthful ____
10. phrases ____
11. Oregon ____
12. atrousis ____
13. valve ____
14. dumpling ____
15. vilum ____
16. lable ____
17. incerite ____

Scott-Foresman
Grade 5

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**scientific**	a **scientific** experiment	**scientific**	11	11
2. (8)	**ritual**	the **ritual** is ancient	**ritual**	7	18
3. (16)	**Indiana**	**Indiana** is a state	**Indiana**	8	26
4. (24)	**income**	her **income** grew	**income**	7	33
5. (32)	**giant**	**giant** footsteps	**giant**	6	39
6. (40)	**sample**	a **sample** product	**sample**	7	46
7. (48)	**oatmeal**	**oatmeal** cookies	**oatmeal**	8	54
8. (56)	**physics**	my **physics** class	**physics**	8	62
9. (64)/(1:04)	**truthful**	a **truthful** answer	**truthful**	9	71
10. (72)/(1:12)	**phrase**	that **phrase** is cute	**phrase**	7	78
11. (80)/1:20)	**Oregon**	an **Oregon** fishery	**Oregon**	7	85
12. (88)/(1:28)	**atrocious**	that's **atrocious** work	**atrocious**	10	95
13. (96)/(1:32)	**valve**	the **valve** broke	**valve**	6	101
14. (104)/(1:40)	**dumpling**	a **dumpling** in soup	**dumpling**	9	110
15. (112)/(1:48)	**villain**	a **villain** in disguise	**villain**	8	118
16. (120)/(1:56)	**label**	**label** your books	**label**	6	124
17. (128)/(2:04)	**inaccurate**	an **inaccurate** answer	**inaccurate**	11	135

Total Words Correct	Correct letter Sequences

Fifth grade
Reg. Ed. (High) Student Sample no. I-51
Spelling
Spring (5/90)

1. exclude ____
2. lesson ____
3. dulphin ____
4. serisely ____
5. slumber ____
6. milatary ____
7. srudly ____
8. goal ____
9. teardrop ____
10. tribute ____
11. double ____
12. lovingly ____
13. simile ____
14. iron ____
15. main ____
16. weapon ____
17. expoducion ____

Scott-Foresman
Grade 5

Time	Word	Phrase	Word	CLS	Cum CLS
1. (0)	**exclude**	do not **exclude** him	**exclude**	8	8
2. (8)	**lesson**	**lesson** number 22	**aluminum**	7	15
3. (16)	**dolphin**	a **dolphin** does tricks	**dolphin**	8	23
4. (24)	**seriously**	I **seriously** doubt it	**seriously**	10	33
5. (32)	**slumber**	a **slumber** party	**slumber**	8	41
6. (40)	**military**	the **military** of the US	**military**	9	50
7. (48)	**shrewdly**	**shrewdly** planned	**shrewdly**	9	59
8. (56)	**goal**	a **goal** for my team	**goal**	5	64
9. (1:04)	**teardrop**	**teardrop** from the eye	**teardrop**	7	73
10. /(1:12)	**tribute**	a **tribute** to her	**tribute**	8	81
11. (1:20)	**double**	a **double** play	**double**	7	88
12. (1:28)	**lovingly**	they **lovingly** kissed	**lovingly**	9	97
13. (1:32)	**chenille**	a **chenille** bathrobe	**chenille**	9	106
14. (1:40)	**iron**	an **iron** bridge	**iron**	5	111
15.(1:48)	**main**	a **main** street in town	**main**	5	116
16.(1:56)	**weapon**	a **weapon** is dangerous	**weapon**	7	123
17.(2:04)	**expedition**	an **expedition** to Guam	**expedition**	11	134
18. (2:12)	**budget**	a **budget** for clothes	**budget**	7	141

Total Words Correct

Correct letter Sequences

Curriculum-Based Measurement

Math 5

59 x 69 736	142 x 80 000 1136 1136	86 x 20 00 172 172	85 x 11 85 85 170
40 x 5 200	22 41 + 16	93 x 11	177 - 37
557 x 8	4) 7	90 - 33	22 x 44
95)199	3)396	422 - 65	358 x 60
199 - 10	32 x 31	45 - 45	12 x 41

Fifth grade
Reg. Ed. (Low) Student Sample no. I-52
Mixed math
Spring (5/90)

Curriculum-Based Measurement
Math 5

47 x 21 47 + 94 141	2r10 25) 55 -45 10	846 - 415 431	124) 4769	30 x 50 00 +150 150
36) 540	11 536 7723 + 164 423	421 x 343	794 - 320	21 x 45
3265 873 + 7365	8) 216	87 x 76	125 x 51	87) 9724
9) 925	23) 69	956 - 770	4302 + 1147	523 x 48
233 x 34	9 - 5	6 x 3	1923 + 3683	36 x 5

Total Digits Correct _____

Written Expression

Write a story that begins with:

One very dark and spooky night I was camping in the woods. When I heard a strange noise....

Then I heard someone screming I went out to see what was a matter. Then I saw the biggest thing I ever saw it was bigfoot and it was screaming. A rock fell on its

Fifth Grade
Reg. Ed. (Low) Student Sample no. I-52
Spelling
Fall (10/89)

1. slufish _____ _____
2. read _____ _____
3. inaridane _____ _____
4. income _____ _____
5. giint _____ _____
6. sumple _____ _____
7. otmel _____ _____
8. phakic _____ _____
9. tureful _____ _____
10. frase _____ _____
11. oregon _____ _____
12. actores _____ _____
13. role _____ _____
14. dompling _____ _____
15. vile _____ _____
16. labe _____ _____
17. incoret _____ _____

Spelling

Fifth grade
Reg. Ed. (Low) Student Sample no. I-52
Spelling
Spring (5/90)

1. ers _____ ____
2. lession _____ ____
3. dopin _____ ____
4. serisly _____ ____
5. slumber _____ ____
6. miltry _____ ____
7. shrudly _____ ____
8. goal _____ ____
9. teir _____ ____
10. tripute _____ ____
11. loving _____ ____
12. sh _____ ____
13. iron _____ ____
14. _____ ____
15. main _____ ____
16. wepon _____ ____
17. exption _____ ____