A study was conducted to find ways to identify Wake County, North Carolina, public school students who need additional assistance in the early grades before they would be taking the North Carolina end-of-grade (EOG) tests. Also explored was the focus of assistance, since the EOG tests do not provide diagnostic information at the individual level and are not even very reliable statistically at the individual level. The apparent solution appeared to be the use of classroom profiles developed in the Wake County Public Schools to measure reading, mathematics, and writing. These profiles had been developed with teacher input but without reliability and validity checks, yet their correlation with third-grade EOG tests was determined. The third grade EOG tests were a preliminary to the full EOG tests that follow in later years. Correlations improved with additional teacher training in the profile system. For the 1999-2000 school year, schools were given a combined list of students who might need additional assistance based on the EOG early tests and the profiles. The impression of evaluators is that multiple methods of assessing student mastery of the material that will be on the EOG tests has been a positive, although expensive, addition to student evaluation. Attachments contain the literacy assessment profile, a mathematics observation matrix, and the assessment data capture form developed for the profile system. (SLD)
Background

The North Carolina End of Grade (EOG) tests have been given each spring at grades 3-8. Third graders are also given a short pretest in the fall. The stakes connected with this testing have risen dramatically since they were first instituted in 1993, with school ratings, teacher bonuses, and student retention screening now tied to the results. In addition, the Wake County Public School System (WCPSS) has adopted a goal that 95% of our students in grades 3 and 8 will score "on grade level" (Levels III or IV) by 2003. About 77-87% currently score at those levels in grades 3 and 8. A major instructional assistance effort (with a budget of about $7 million) has been designed to provide additional assistance for students.

Given this context, the initial challenges for Evaluation and Research (E&R) were:

♦ How can we identify students who need additional assistance at the early grades (1-3) before the full EOG tests are given?

♦ What should teachers focus on when providing assistance, since EOG does not provide diagnostic information at the individual student level (and is not even very statistically reliable at the individual level)?

The obvious solution appeared to be the use of results from another measure used in WCPSS--classroom profiles that had been developed in reading, mathematics, and writing by the Curriculum and Instruction department staff with input from teachers. These profiles were based on the state curriculum and profile models (although WCPSS profiles are far more comprehensive than the state's model). While the WCPSS profiles have a great deal of face validity, one limitation is that E&R was not asked to be involved until after the instruments were developed. Reliability and validity checks were not considered in the developmental process. E&R was asked after the system was being implemented to design data capture scan forms, collect and analyze the data, and
help set cut scores for targets at the early grades. We had concerns based on the first year's data about whether teachers within and across schools were truly implementing the system in a consistent way, even after three days of training. Some of the analyses reported here provide data related to that concern.

A second more specific set of questions arose once the decision was made to try to use the profile results. Our results address these questions.

1. Do spring results on the grade 3 classroom profile data correlate well with the results of the third grade EOG tests?

2. Incoming third graders are a critical group for accomplishment of the school system achievement goal. Is the state's EOG fall third grade assessment or the spring of second grade classroom profile data a better measure to use in identifying students who may need help?

3. Should another EOG measure be provided to schools to identify progress mid-year in terms of the success of assistance provided?

Do spring results on the grade 3 classroom profile data correlate well with the results of the third grade EOG tests?

While EOG testing data have been available since 1993, the systematic use of standardized classroom literacy performance assessments began in a group of pilot schools in WCPSS in 1996-97. During 1997-98, training in the use of the assessment process was provided for primary grade teachers system-wide, although some schools did not fully participate in training until 1998-99. The form is included as Attachment 1.

A key component of the K-3 literacy assessment system is a system of "leveled" books and the use of "running records" to identify the level of book most appropriate for each student. Running records are conducted for each student at least quarterly with several pieces of data recorded on a profile card that is kept with a student's cumulative record. During the primary grades, most students move from Level 1 to Level 32. For purposes of screening for at or below grade level status, the key piece of information was identified as the book level at which a student could read with 90% accuracy and a "retelling" score of 3 or 4 on a four point scale.
In math, an observational matrix is provided for each grade that represents seven strands in the curriculum. Students who show on grade level performance in at least 6 of the 7 strands are considered to be “on grade level.” A sample from grade 2 is included as Attachment 2.

In 1997-98, teachers were asked to use a scanable form to code the results of the running record conducted near the end of the school year. A sample form is included as Attachment 3. These scan forms were used to compile a system-wide summary of students’ book levels. This summary was studied by the district reading instruction specialists, and a level of 23 or higher was chosen to represent "at or above grade level" at the end of second grade. A level of 29 or higher was identified as "at or above grade level" at the end of third grade. The application of these cut-points identified approximately 25% of students as below grade level at the end of second grade and approximately 35% of students as below grade level at the end of third grade. EOG test results at the end of third grade identified 21% of students as below grade level in 1998.

The Pearson correlation between the reading book level and the EOG test scores for the 6,009 third grade students for whom we had both pieces of information from spring 1998 was .64. This was slightly lower than the correlation between the third grade fall pretest and the EOG reading test (.74).

The data collection process was repeated in 1998-99 and showed that:

- Approximately 20% of second grade students were below grade level according to the literacy assessment profile;
- Approximately 28% of third grade students were below grade level according to the literacy assessment profile;
- Approximately 20% of third grade students were below grade level according to the spring EOG results;
- A Pearson correlation of .63 existed between book level and EOG reading score for 6,944 third grade students; and
- For students scoring below grade level (Level 29) on the literacy assessment, the correlation with EOG was lower at .51.

Since teachers were still being trained in 1998-99 in the use of the literacy performance assessment system, the correlation between performance assessment and EOG multiple-choice score was considered encouraging. We hope that as teachers gain
expertise in the administration and interpretation of running records, the correlation will rise and judgements based upon the literacy profile data will be made with greater confidence.

Incoming third graders are a critical group for accomplishment of the school system achievement goal. Is the state’s EOG fall third grade assessment or the spring of second grade classroom profile data a better measure to use in identifying students who may need help?

We ran Pearson correlations between the spring 1999 grade 3 EOG results and the fall of grade 3 EOG results as well as spring 1998 grade 2 classroom profile results.

- In reading, correlations with spring grade 3 EOGs were .72 with fall 3rd grade EOGs and .67 with spring of grade 2 classroom profiles.

- In math, correlations were .76 with the fall 3rd grade EOG and .59 for the spring of 2nd grade classroom profile.

Thus, both had reasonable correlations with the third grade spring EOGs, with some evidence that the fall EOG was a better predictor, especially in math. Of course, the EOG fall and spring tests are both multiple choice and similar in structure and format, and the tests are closer together in time, so this seems reasonable from that standpoint. However, the standard error on the fall EOG test is larger than the EOG, primarily because the test is much shorter, so the results are somewhat surprising. The next two charts illustrate the percentage of students who were classified correctly and incorrectly on the spring EOG based on the fall EOG and spring classroom profile data.

Even though the correlations with the third grade EOG appeared acceptable for both instruments, about 16% of the students were misclassified with both instruments in terms of where they ultimately scored on EOG in the spring of their third grade year.

The good news is that a higher percentage of students scored on grade level after scoring low on the pretest or class profile than the reverse (in which students scored at or above grade level initially and fell below grade level on the spring of grade 3 EOG). The difference in percentages was more positive for the fall of 3rd grade pretest than the classroom profiles. We hope students who improved their status did so as a result of assistance provided, although we suspect some was due to measurement error.
We plan to re-run these analyses once spring 2000 data are in and hope correlations and student classifications improve with an extra year of use of the instruments (especially the classroom profiles).

### Spring EOG 3rd Grade Results in Relation to Fall EOG 3rd Grade Results:

#### Reading

<table>
<thead>
<tr>
<th>Fall Grade 3 Status</th>
<th>EOG Status Spring Grade 3, N=7,221</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On Grade Level: III or IV</td>
</tr>
<tr>
<td>Level III or IV</td>
<td>5,064 (70.1%)</td>
</tr>
<tr>
<td>Level I or II</td>
<td>785 (10.9%)</td>
</tr>
</tbody>
</table>

#### Mathematics

<table>
<thead>
<tr>
<th>Fall Grade 3 Status</th>
<th>EOG Status Spring Grade 3, N=7,248</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On Grade Level: III or IV</td>
</tr>
<tr>
<td>Level III or IV</td>
<td>5,117 (70.6%)</td>
</tr>
<tr>
<td>Level I or II</td>
<td>512 (7.1%)</td>
</tr>
</tbody>
</table>

### Spring EOG 3rd Grade Results in Relation to Spring Grade 2 Class Profile Results:

#### Reading (Book Levels)

<table>
<thead>
<tr>
<th>Spring Grade 2 Status</th>
<th>EOG Status Spring Grade 3, N=5,747</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On Grade Level: III or IV</td>
</tr>
<tr>
<td>Above Cut (Level 23)</td>
<td>4,109 (71.5%)</td>
</tr>
<tr>
<td>Below Cut</td>
<td>547 (9.5%)</td>
</tr>
</tbody>
</table>

#### Mathematics

<table>
<thead>
<tr>
<th>Spring Grade 2 Status</th>
<th>EOG Status Spring Grade 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On Grade Level: III or IV</td>
</tr>
<tr>
<td>Above Cut*</td>
<td>4,463 (69.5%)</td>
</tr>
<tr>
<td>Below Cut</td>
<td>575 (9.0%)</td>
</tr>
</tbody>
</table>

*Level III or IV on 6 or 7 of 7 instructional strands

We ran correlations of the fall of grade 3 EOG results for 1999 and the spring of grade 2 data capture results for 1999 and found the two measures (fall 3rd grade EOG and spring 2nd classroom profiles) had a moderate positive correlation with each other in reading (.63) and in math (.54). As shown in the next figures, about 17%-18% of the students were low on only one or the other of the two measures (about 1,000 students in reading and 1,200 in math).
Fall EOG 3rd Grade Results in Relation to Spring Grade 2 Class Profile Results:

<table>
<thead>
<tr>
<th>Spring Grade 2 Status</th>
<th>EOG Status Fall Grade 3, N=6,473</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On Grade Level: III or IV</td>
</tr>
<tr>
<td>Above Cut (Level 23)</td>
<td>4,732 (73.1%)</td>
</tr>
<tr>
<td>Below Cut</td>
<td>406 (6.3%)</td>
</tr>
</tbody>
</table>

Mathematics

<table>
<thead>
<tr>
<th>Spring Grade 2 Status</th>
<th>EOG Status Fall Grade 3, N=6,642</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>On Grade Level: III or IV</td>
</tr>
<tr>
<td>Above Cut (Level 23)</td>
<td>4,699 (70.8%)</td>
</tr>
<tr>
<td>Below Cut</td>
<td>762 (11.5%)</td>
</tr>
</tbody>
</table>

However, when we looked at only those students identified as low by either instrument to see how much they overlapped, the same students were identified as low by both instruments for only 39% of the cases in reading and 38% in math.

STUDENTS IDENTIFIED AS LOW ACHIEVERS IN MATH, FALL 1999

Total Students 1,943

Low on Both Profile and EOG
745 Students

Low on Math Profile
Spring 1999
762 Students

Low on Math EOG Only
Fall 1999
436 Students
Our decision for this year at least was to provide schools with the combined list of students who might need additional assistance this year and ask them to assess whether the students truly needed additional help to reach grade level on the EOG in the spring of third grade. We did provide recommendations based on whether students were low on both instruments (definitely provide assistance) or just one (consider class work from fall plus additional classroom profile information that was current). This appeared to be the only prudent decision given the 95% goal. We thought it would be better to have too many students considered for help than too few. However, this was an expensive decision, since schools were given allocations for the new Accelerated Learning Program (ALP) based on students who scored low on the assessments. The number of students potentially eligible almost doubled for grade 3 compared to selecting one instrument or the other.
Should another measure be provided to schools to determine mid-year status, progress made, and areas for further assistance?

This year WCPSS added yet another measure of performance--mid-year optional testing using "secure for local use" forms of the EOG tests. The optional testing was provided in response to multiple requests over several years, and was only possible because the state released some forms of the state tests used in prior years for local use on a secure basis. We had concerns about mis-use of data from administration of one form of the EOG, since the form would not cover the full curriculum in any great depth, but training highlighted appropriate and inappropriate uses of the test. Results provided an item analysis for groups of students and rosters of student responses to test items, but no total score.

The primary benefit of the optional mid-year testing appears to be increased teacher familiarity with the design of NC EOG test items and the structure of the EOG tests. Principals report that teachers have a new understanding of the expectations of the test and the need to target instructional time to the goals and objectives of the North Carolina Standard Course of Study. Hopefully, students have benefited from an opportunity to discuss specific test items in the classroom.

The item analyses and student rosters do provide some diagnostic information to teachers based upon patterns of questions missed and incorrect answers chosen by students. However, it is questionable whether additional testing provides greater incentive and/or ability to identify low-performing students in need of special assistance.

Conclusions

What do our results suggest about whether multiple measures are a boon, bane, or baloney? Or, to throw in another "b", are multiple measures a boondoggle—a pointless, unnecessary waste of time and money?

From the viewpoint of most teachers, collecting classroom profile information to the EOGs is a positive move (a boon) for several reasons.
Schools see the EOG tests as "traditional", non-diagnostic, and subtracting from instructional time. For the most part, they see the classroom profiles (performance assessments), which actually take more time, as instructional and diagnostic, therefore less wasteful of their time.

While both assessments measure students' mastery of the North Carolina curriculum, they are different in that:

- EOG is primarily multiple choice items, with some open-ended items at some grades and writing assessments at grades 3, 5, 8. Multiple choice results are returned quickly, but open-ended and writing results take months to be returned.

- The math, literacy, and writing profile cards are performance assessments primarily, and rely on teacher observation, student class work, and some standardized methods and instruments (e.g., running records for reading).

Our impression from a system-level perspective at this point is also that multiple measures are a positive addition—albeit an expensive one. If our goal is to determine whether students are truly mastering the material we desire, measures that are closer to classroom practice and which provide teachers with clues about what students do not understand provide more useful information than our statewide assessment that is not diagnostic at the individual level. The information can be quite helpful. It would clearly be nice if these measurements were a little more exact, and if they were validated before being placed in use, but they do provide more clear-cut standards than teacher judgement alone. The expensive part is that we identify more students as potentially needing help than we might if the measurements were more reliable. We would appreciate input on other ways to narrow the field of students likely to benefit from assistance.
K-3 Literacy Assessment
Reading Continuum Summative Profile

The Reading Continuum Summative Profile provides on-going documentation of student growth in reading both during the year and at the end of the year. This assessment should be completed a minimum of twice a year, once at the beginning and once at the end of each school year. However, since the purpose of this assessment is to inform instruction, on-going assessment is recommended.

Directions for Completion of the Profile

1. Write student's name, ID #, and school in the blanks provided.

2. Review the descriptive indicators for each of the stages of reading development. Indicators are not considered to be of equal value in describing reading behavior, but together provide a holistic description of a child's reading development. The teacher will check the indicators the child exhibits in reading using the color coding listed below:

<table>
<thead>
<tr>
<th></th>
<th>Kindergarten</th>
<th>First Grade</th>
<th>Second Grade</th>
<th>Third Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginning of Year</td>
<td>Pencil</td>
<td>Black Ink</td>
<td>Pencil</td>
<td>Black Ink</td>
</tr>
<tr>
<td>During/End of Year</td>
<td>Blue Ink</td>
<td>Red Ink</td>
<td>Blue Ink</td>
<td>Red Ink</td>
</tr>
</tbody>
</table>

If the child has been retained, use the same color code but circle your new checks.

3. Mark the stage on the continuum that best describes the child's present level of functioning in reading, including the date. The teacher will indicate at what point on the continuum the child best fits. If the child exhibits only a few behaviors within a stage, the teacher may place a mark further to the left within a stage. If the child exhibits most of the behaviors within a stage, the teacher will indicate this development by a mark further to the right on the continuum. (Use the same color code as in #2.)

4. The teacher will select the book level that best matches the mark made on the continuum. If the continuum has been marked at the end of the early emergent/emergent stage, book levels 3-4 would be appropriate. If the continuum has been marked in the middle of the early independent stage, book levels 19-20 would be appropriate.

5. The teacher will take a running record as the child reads. The teacher will then complete a box on the back of the folder by including the following information: date, grade, title(s) of book(s) used for running record, accuracy rate (mark the child's highest instructional level 90-94%), self-correction (SC) rate, and retelling score (at least a 3). The information in this box is to support the reading stage marked on the reading continuum.

6. For students who are early emergent/emergent readers, the box entitled Print Concepts needs to be completed at the beginning of the year.

7. In order to document growth over time, this form needs to be marked at the beginning and again at the end of each school year. However, additional assessments are encouraged.

8. The teacher should record specific observations about the child's behavior during assessments each year.
## Documentation Over Time

<table>
<thead>
<tr>
<th>Date</th>
<th>Grade</th>
<th>Book Level:</th>
<th>Book Title:</th>
<th>Running Record</th>
<th>Accuracy Rate</th>
<th>SC Rate</th>
<th>Retelling Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Typical Grade Levels:**
- Kindergarten: 1-5
- Grade 1: 6-17
- Grade 2: 15-27
- Grade 3: 24-32

**EOG Reading:**
- Grade 1: 1-17
- Grade 2: 15-27

**Books Levels:**
- Kindergarten: 1-5
- Grade 1: 1-17
- Grade 2: 15-27
- Grade 3: 24-32
Using the Mathematics Observation Matrix
Grade 2

**Purposes of the On-Going Assessment Matrix**

Purposes of the Assessment Matrix are:
- to develop consistent performance expectations among teachers across the system related to mathematics content goals,
- to provide information for the teacher to use in planning appropriate instruction for individual student learning,
- to support students and teachers in monitoring students' progress toward reaching mathematics goals,
- to talk to parents and students about expectations related to mathematics goals and objectives,
- to guide the collection of classroom evidence of student achievement, and
- to encourage teachers to continually revisit and reinforce previously taught concepts throughout the year.

**When to Mark the Matrix**

Teachers may choose to mark student matrices at anytime during a grading period. Some students will demonstrate understanding long before other classmates. It is easiest to mark cards for those students at the time the teacher observes the student's mastery of objectives. Mastery is demonstrated by success on a variety of tasks. Other students will not have sufficient mastery of objectives at the end of the grading period when the teacher must make a summative judgment in order to assign a grade on the report card. Cards for these students need to be marked but may reflect characteristics of students at Levels I and II.

Teacher-student conferences and interviews may be required for some students who seem to have inconsistent understanding of particular concepts. Activities from the Mathematics Assessment Guide or the Instructional Strategies for Mathematics may be used for individual assessment.

Each grading period teachers will highlight student performance using the following colors. This will enable teachers, parents, and students to see growth at a glance.

<table>
<thead>
<tr>
<th>Grading Period</th>
<th>Color</th>
</tr>
</thead>
<tbody>
<tr>
<td>First Grading Period</td>
<td>Yellow</td>
</tr>
<tr>
<td>Second Grading Period</td>
<td>Blue</td>
</tr>
<tr>
<td>Third Grading Period</td>
<td>Pink</td>
</tr>
<tr>
<td>Fourth Grading Period</td>
<td>Green</td>
</tr>
</tbody>
</table>

**Student Work Samples**

Samples of student work which document understanding of mathematics topics should be kept and dated throughout the year. Activities from the Mathematics Assessment Guides, Strategies for Instruction in Mathematics, Testlets, or teacher made assessments may be used. It is recommended that no more than 6-8 work samples be included per year.

At the end of the year, the student matrix along with work samples will be put in the cumulative folder to go to the next year's teacher.
The Writing Continuum Summative Profile shows twelve marks divided into four overlapping stages of writing development.

For grades 1-5, please darken the bubble showing the student's end of year proficiency level in each mathematics strand, based upon mastery of objectives on the Mathematics Observation Matrix. Proficiency should be the highest level where the student has demonstrated mastery of most of the skills listed on the Matrix. For example, if the student has mastered only a few of the Level III skills under Measurement but most of the Level II skills, mark the bubble for Level II under Measurement.

<table>
<thead>
<tr>
<th>Numeration</th>
<th>Geometry</th>
<th>Patterns</th>
<th>Measurement</th>
<th>Problem</th>
<th>Solving</th>
<th>Data</th>
<th>Computation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level III</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level II</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Level I</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Reading for students in Grades K-3 (Include grades 4-5 if student is below Book Level 32)

The top of the Reading Continuum Summative Profile card shows twelve marks divided into four overlapping stages of reading development. Please darken the bubble below this item corresponding to the mark that most closely describes this child's most recent level of development. For example, if a student has demonstrated approximately one-half of the descriptive indicators under Early Independent, that child would be in the middle of Early Independent and the student's development would be an 8 on the scale.

Early Emergent/Developing

<table>
<thead>
<tr>
<th>Early Emergent</th>
<th>Developing</th>
<th>Early Independent</th>
<th>Independent</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1 2 3 4 5 6 7 8 9 10 11 12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If this child is Early Emergent/Developing how many Print Concepts have been mastered?

- 1-2
- 3-4
- 5-6
- 7-8
- 9-10
- 11-12
- 13-14
- 15-16
- 17-18
- 19-20
- 21-22
- 23-24
- 25-26
- 27-28
- 29-30
- 31-32

Writing for students in Grades K-3

The Writing Continuum Summative Profile shows twelve marks divided into four overlapping stages of writing development. Please darken the bubble above the number showing the mark that most closely indicates this child's most recent level of development. For example, if a student has demonstrated almost all of the behaviors under Emergent and a few of the behaviors under Developing, that child would be near the start of the developing stage and the student's development would be represented by a 4 on the scale.
NOTICE

Reproduction Basis

This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.

This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").