

IMPROVING EDUCATIONAL QUALITY PROJECT

The Role of Assessment in the Rhythms of Reform

Individual Pupil Planning

Classroom Decision-Making

Intervention Development

Evaluating Progress

Informing National Efforts
-CRT
-Curriculum

International Knowledge Building

A PRESENTATION MADE AT THE COMPARATIVE INTERNATIONAL
EDUCATIONAL SOCIETY (CIES) MEETING
IN
WILLIAMSBURG, VIRGINIA
MARCH 6-10, 1996

Abigail Harris
Graduate School of Education
Fordham University
113 W. 60 St.
NY, NY 10023

Improving Educational Quality Project
Institute for International Research
1815 N. Fort Myer Drive
Arlington, VA 22209

The Role of Assessment in the Rhythms of Reform
Abigail Harris, Fordham University

Until recently, testing in developing countries has been an add-on, a maneuver used to control the flow of students to upper levels of schooling or to certify end of cycle performance (Lockheed & Larach, 1993). Its role in promoting quality was mostly as a motivator for those pupils with some hope of using their test performance to advance in school or work. Further, because of the high stakes involved, what's tested was what was taught. If the end of cycle tests provided a comprehensive assessment of the curriculum, quality was enhanced. If it didn't--say for example it focused on just math and reading--instructional quality in other subject areas in the curriculum (science, social studies, writing) was jeopardized (Harris & Ilon, 1992).

In IEQ, testing isn't an add-on, its an integral part of a dynamic process designed to empower educators to sustain learning progress. My intent today is to illustrate the role of assessment in improving educational quality in Ghana and to get your feedback about how best to disseminate this process and what we've learned about it.

The kind of assessment I will be referring to is curriculum-based assessment (CBA). **CBA is the practice of asking students to perform tasks that have been drawn directly from the curriculum and then using assessment results to adapt instruction to reflect the learners' needs.** CBA provides a way of linking curriculum with learning, and, in effect, adjusting instruction to fit the students (Gickling & Havertape, 1981; Hargis, 1987; Shinn, 1989; Shinn & Good, 1992). For example, students might be asked to read a passage from their textbook or compute the answers to mathematics problems sampled from their texts. "Direct observation and recording of a pupil's performance in the local curriculum [is used] as the basis for gathering data to make instructional decisions." (Deno, 1987, p. 41)

What distinguishes CBA is that the specific tasks that students are asked to perform are drawn directly from the curriculum and the tasks are selected, administered and scored using standardized procedures.

In all CBA models, instructional decisions are based on information generated by an ongoing assessment of student performance on the curriculum. The primary goal is to guide the instructional decision-making process (Blankenship, 1985; Graden, Zins, & Curtis, 1988; Marston & Magnusson, 1985) so that instruction continues to be relevant to the students' instructional needs, thereby increasing the chances of successful learning.

CBA can help improve the quality and effectiveness of the education process. Today I will describe how CBA is being used for (1) individual instructional planning, (2) improving instruction through classroom level planning, (3) developing instructional interventions, (4) evaluating educational progress and programs, (5) informing national efforts, and (6) contributing to knowledge building in the international community.

(1) Individual Instructional Planning: Creating and Using Diagnostic Profiles.

Although diagnosis is commonly taken to mean the identification of a learning problem, diagnostic profiles in CBA focus more on each pupil's strengths. The goal in CBA is to determine precisely what students know and where they fall on a hierarchically ordered learning continuum. The basis for this continuum is the curriculum. CBA assumes that there is an underlying ordering in the curriculum such that the year 2 textbook builds on what was covered in year 1, and the year 3 textbook builds on what was covered in years 1 and 2, and so on. It follows that if a year 4 student can perform a task in the year 4 text (for example, reading a passage with fluency and comprehension), this student should have comparable or better success on similar tasks taken from lower level texts. Similarly, if a year 4 student is unable to perform the reading task, it is functionally useful to determine at what level the student can perform the task. Thus, rather than stopping the assessment at the point that the child fails, the examiner continues probing downward on the curriculum continuum to the point where the child succeeds. For some preliterate students this probing extends to finding out if the student has pre-reading skills such as letter identification. When the assessment is complete, it is possible to construct diagnostic profiles of individual students and groups of students. These profiles are relevant for instructional planning and decision making at all levels of the educational system from the classroom to national policy.

Mary A. recently completed level 6, the end of the primary cycle at an urban school in the Central Region of Ghana. She was tested by CRIQPEG midyear when she was in Level 5 and again near the end of Level 6. Table 1 describes selected aspects of her performance at these two points in time.

Table 1: Mary A. from an Urban School in the Central Region of Ghana

Task	Mid-Year Level 5 (3/94)	End of Year Level 6 (8/95)
Writing Words	21 correctly spelled words	61 correctly spelled words
Spelling-Correct words		48% correct
Spelling-Correct letters		68% correct
Letter/Story Fluency (Ave)		31 words
Letter/Story Words (Ave)		24 correctly spelled words
Letter/Story Correct Writing Sequence (Ave)		8 correct writing sequences (spelling, punctuation, grammar)
Aided Reading	64 % correct	96% correct
Reading Most Used Words	24% correct	76% correct
Reading-Words/Min. (Ave)	11.5 words per minute	15.6 words per minute
Reading-Decoding Words Percent Correct (Ave)	44.25 % correct	66.55 % correct
Reading Comprehension (Ave)	27.5 % correct	52.5 % correct

Consider the instructional implications of her 5th grade performance. She was able to write 21 English words in 10 minutes--her name and 19 other words. When asked to read a list of the most frequently used words in her text, words such as "has", "also", "is", "will", and "very", she was only able to read about 1/4 of these words. When these same words were read to her (Aided Reading), she was able to point to almost 2/3 of them.

When asked to read passages from 2nd through 5th grade texts she averaged 11.5 words per minute and she read less than half of the words correctly. With regard to reading comprehension, she answered about 1/4 of the questions correctly.

If you were her teacher, how could you use this information? It is clear from these results that Mary had limited reading and writing vocabulary. She had some decoding skills but the passages in the textbooks were at a frustrational level for her. At this point, if she tried to read passages in textbooks for levels 2-5 independently, she would struggle with every other word. (Worth noting is that research suggests that efficient learning takes place when the reader reads nine out of every ten words without assistance or prolonged hesitation). Given her difficulties in decoding, such a low level of reading comprehension is not surprising but it does reinforce the need for vocabulary building.

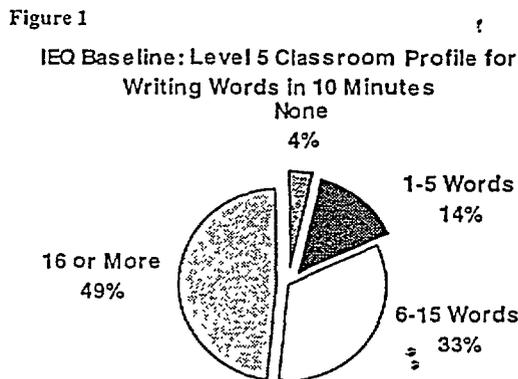
Mary's performance at the end of level 6 shows some improvement. She has a larger writing and reading vocabulary and her reading comprehension has improved. While her performance suggests that the available instructional materials are still too difficult for independent work, she does have some of the basic literacy skills. With some preparatory bridging activities such as using flashcards or word games to provide relevant practice with new vocabulary, Mary could be helped to use the texts effectively.

One of the strengths of curriculum based assessment is that it isn't "private". Teachers who observed the process learned strategies they could use to monitor pupil progress. With curriculum-based assessment there is no need to worry that the teachers will coach the pupils for future testing. To do this teachers would need to have the children practice reading the passages in their textbooks or writing letters--both skills that are part of the curriculum. Because different tasks (e.g., different passages and different writing prompts) are used for each assessment, the children would need to master the skill rather than simply memorizing one passage or letter. By having children read multiple passages and produce a variety of writing samples (writing words, dictation, expressive writing), it is possible to obtain a reliable estimate of each child's skill levels. In this way, the assessment supports the curriculum and provides a means of monitoring pupil progress.

2. Improving Instruction Through Classroom Level Planning

This kind of analysis can be used for classroom level planning as well. Effectiveness of student learning has been shown to be closely related to academic learning time (Denham & Lieberman, 1980; Levin & Lockheed, 1993). When instruction is too easy, students become bored and assignments are not taken seriously. When instruction is too difficult and students do not have the necessary prerequisite skills, they become frustrated and discouraged. Academic learning time occurs when the student is motivated by a task that is appropriately challenging. Thus, it is critically important to identify what skills each student possesses and to use instruction to progressively build upon this foundation. (Block, 1971; Hargis, 1987; Popham & Baker, 1970).

Figure 1 is a pie chart that summarizes Ghana baseline data for level 5 on writing words in ten minutes. Note that 4% of the class were not able to write even one word and that includes producing their names; this is even when they were asked in the vernacular to perform this task. About 1/5 wrote fewer than 6 words. On the other hand, about half of the class wrote 16 or more words. In fact, some pupils were able to write over 100 words correctly.



The next two charts (Figures 2 and 3) illustrate the baseline performance of level 5 pupils on reading decoding and comprehension tasks. Over half of the level 5 pupils at midyear (6 months before CRT testing) were unable to decode (read) even 30% of a typical passage in their textbook and 85% got less than 1/3 of the comprehension questions correct.

Figure 2
IEQ Baseline: Level 5 Classroom Profile for Decoding of Reading Passage from the Level 5 Text

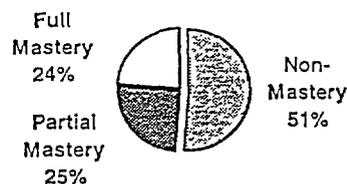
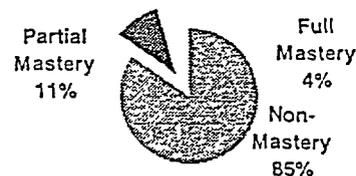


Figure 3
IEQ Baseline: Level 5 Class Profile for Reading Comprehension of Passage from Level 5 Text



CRIQPEG shared this information with classroom teachers and local education officers. Initially, some teachers were defensive, and CRIGPEG team members and Circuit Supervisors had to reassure teachers that the information was not unique to their school and that there were many reasons for low scores (for example, textbooks had not been available in past years). The next step was to turn to the teachers as classroom experts and ask them what could be done to improve learning. Together the classroom teachers, head teachers, local education officers, and CRIQPEG team members collaborated to devise instructional strategies or interventions. Over the next several months, CRIQPEG team members, head teachers and teachers tried different strategies and shared feedback on the effectiveness of these strategies.

3. Informing National Efforts to Improve Quality: Understanding the CRT

At the same time that IEQ was getting off the ground in Ghana, criterion referenced tests (CRT) were being developed for national use to monitor the end of cycle performance of primary school pupils. Multiple forms of multiple choice tests in reading and mathematics were developed and administered to a large carefully selected sample of entering level 6 pupils throughout Ghana. Performance on the test was disappointing and hard to interpret. The Ministry of Education and USAID asked CRIQPEG to collect data to explore these results.

CRIQPEG findings shed light on why the pupils experienced so much difficulty with the CRT. Figures 2 and 3 are based on mid year data for level 5 pupils; CRT testing was conducted at the beginning of level 6 (just a few months later). Reading and comprehending multiple choice questions was beyond the reach of all but about 15% of the level 5 pupils tested in the 14 participating schools.

4. Intervention Development

About the same time as CRIQPEG was sharing the results with local educators, specialists from IEQ were studying the results as well. In response to the performance patterns, specialists identified 3 goals for intervention development:

1. Constant exposure to print: This goal was developed because so many children were unable to write even a few words. The intent was to provide more exposure to print so that pupils could be learning all the time, not just when the teacher wrote something on the chalk board.
2. Frequent practice with oral English: This goal was a reaction to low performance on oral language and reading comprehension assessment tasks and was intended to build pupil vocabulary.
3. Using instructional strategies that help every pupil to be a successful learner. This third goal was designed to encourage teachers to assess pupil learning and to adapt instruction to promote efficient learning for all pupils.

In professional development seminars and in the participating schools, Circuit Supervisors, head teachers and CRIQPEG team members learned and shared specific strategies for achieving these goals.

Important sign: During a feedback session after several months of collaboration, teachers in participating schools requested assistance on managing classrooms with diverse achievement levels. This was a clear sign that teachers (1) recognized the diversity of achievement levels in their classrooms, and (2) were motivated to try to achieve goal 3 of adjusting instruction to pupil needs. Subsequent professional development seminars targeted goal 3.

5. Evaluating Progress/Effectiveness of Interventions

Naturally, everyone wanted to know: Did it work? Are efforts to improve educational quality having an impact on pupil performance? In August of 1995, 18 months after baseline data collection and 14 months after the initial sharing of data with local educators, CRIQPEG returned to all 14 schools to collect another round of CBA data. Of the original baseline group of 1032, about 800 were located and retested. Replacements for missing children were randomly selected from baseline class lists.

Figures 4 and 5 illustrate the August 95 follow-up performance of level 5 pupils from the intervention schools. Recall the earlier baseline charts for level 5 pupils (used to help interpret CRT findings). For decoding, 51% were performing at the non-mastery level at baseline as compared with 19% at follow-up: low performing pupils benefited. Full mastery went from 24% at baseline to 63% at follow-up. In the more troublesome area of reading comprehension where 85% of the pupils were at the non-mastery level at the time of baseline data collection: at follow-up this number was 56%. Still not good, but a definite improvement. Full mastery levels went from 4% to 21%.

Figure 4

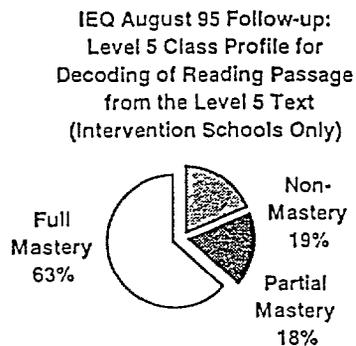
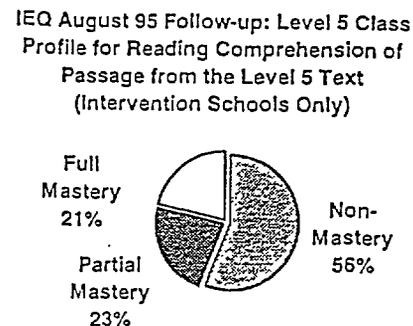


Figure 5



Switching over to writing words and looking at more grade levels: Figure 6 provides a summary of baseline performance on writing words at each level tested for comparison and intervention schools. Note that at each level, comparison schools performed slightly higher on average than intervention (intensive) schools. Figure 7 provides comparable data for August, 1995. There are two points I'd like to make. First, both intensive and comparison schools performed better at the follow-up data collection (Hawthorne strikes again). (Actually, from the start, CRIQPEG termed the intervention schools, "intensive schools" and the comparison schools, "non-intensive". This is actually more accurate than calling them comparison schools--data collection such as classroom observations, interviews, as well as pupil performance assessment constituted a less intensive intervention.) The second point to make with this figure is that at follow-up, pupils in the intensive schools have significantly greater writing vocabulary than pupils in the non-

intensive or comparison schools. These are just 2 examples of how the assessment process contributes to the improvement process by allowing us to monitor pupil performance changes.

Figure 6

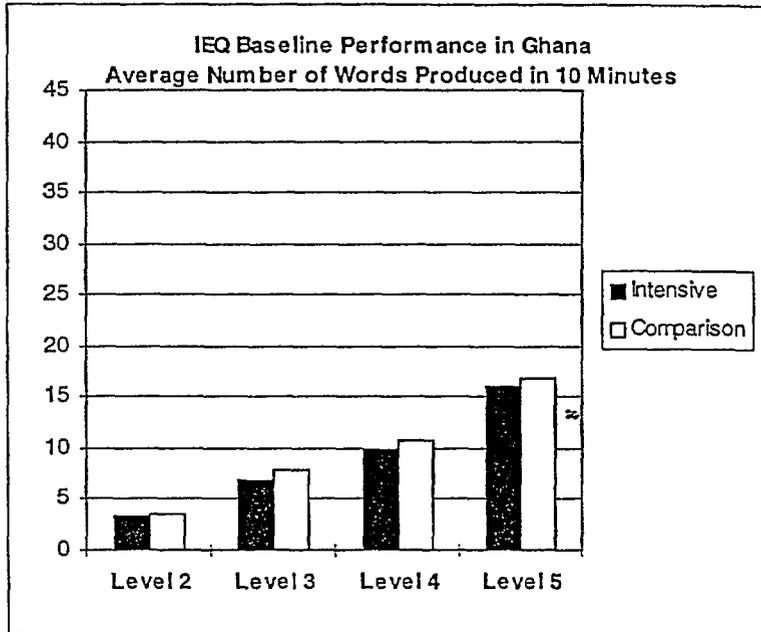
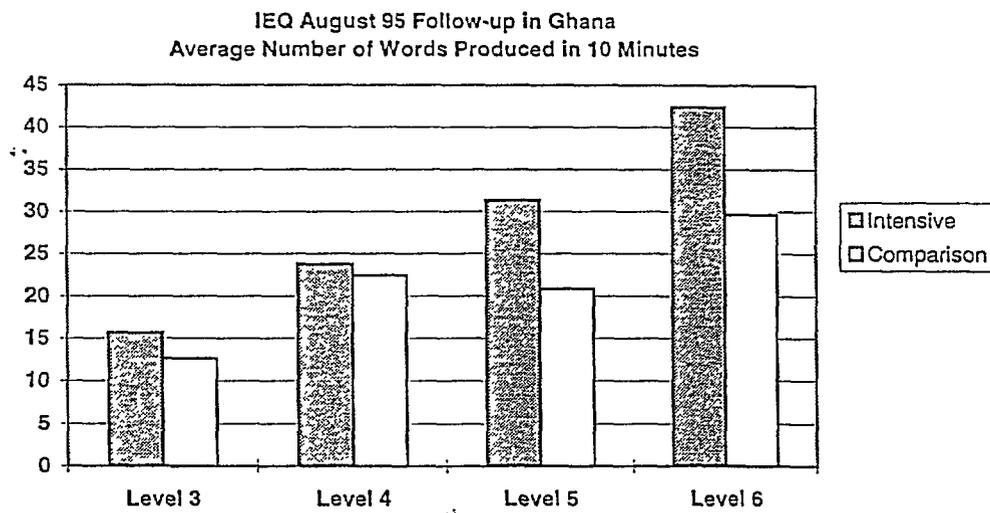


Figure 7



Knowledge Building in International Development

In reviewing the CBA baseline findings in Ghana there was an interesting pattern of reading performance. Upper level pupils read upper level passages with the same fluency and decoding success as they read lower level passages. This finding led to a follow up

investigation of the sequencing of passages in the Ghanaian textbooks in relation to actual performance. Internationally used readability formulas or indices were used to evaluate the predicted difficulty of reading passages in Ghanaian textbooks. These ratings were compared to actual pupil performance. What was interesting was that the factors typically associated with reading difficulty (e.g., word length, sentence length, etc.) didn't predict the sequencing of passages in the Ghanaian textbooks. Nor did they predict baseline pupil performance. We pursued these findings: did this mean that the readability formulas used effectively in other countries didn't apply in Ghana? Was a different mechanism for sequencing of reading materials more appropriate when children are learning English as a foreign language?

Based on the baseline data, I was prepared to propose different hypotheses for sequencing of reading materials in countries where children are learning English as a foreign language. It seemed that performance was more related to recency of exposure: pupils performed as well or better on passages that were in recently used textbooks, even though these would be upper level books, as compared with performance on passages from textbooks from the lower levels. One possible explanation for this relationship is that Ghanaian children are learning English as a foreign language; their principal exposure is in school in the current textbook.

However, analyses using the recently collected 1995 data for upper primary children show strong correlations between children's performance and text difficulty ratings based on readability indices. The explanation may be that the formula work once the children become literate. This is what we were seeing in the recent data collection.

Explorations into this language learning process continue. This is just one example of how IEQ will be using assessment to build knowledge within country as well as contribute to knowledge base in the international development community.

The Rhythm of Reform

Opening Doors Strengthening Capacity
Building Knowledge Demonstrating Potential
Developing Research Tools Stimulating Initiatives

