

PN-ACK-268



Improving Educational Quality (IEQ)

Curriculum-Based Assessment
And Improving
The Quality of Primary Education In
GHANA



CURRICULUM-BASED ASSESSMENT
AND IMPROVING THE QUALITY OF PRIMARY EDUCATION IN GHANA

Contributors to this manuscript include: Abigail M. Harris (Fordham University), Beatrice Okyere (CRIQPEG), Aida Passigna and Jane Schubert (Institute for International Research).

This manuscript was produced as part of the Improving Educational Quality Project which is funded by the Center for Human Capacity Development/U. S. Agency for International Development (Contract # DPE-5836-C-00-1042-00).

TABLE OF CONTENTS

1. Introduction

2. Background in Ghana--Need for CBA

3. Developing the Assessment Materials

- * Assessing English Proficiency
- * Principles of Instrument Development
- * Desirable Qualities in Assessment Instruments
- * Developing the Instruments
 - Oral Language
 - Reading
 - Writing
- * Pilot Testing and Revising the Instruments

4. Assessment and Data Management Process

- * Sampling Pupils
- * Testing Process
- * Recording and Handling Data
- * Reviewing Data for Accuracy

5. Summarizing Results

- * Hand Tabulation of Results
- * Computer Entry
- * Quality Control
- * Data Analysis

6. Using Results and Providing Feedback

- * Individual Instructional Planning: Creating and Using Diagnostic Profiles
- * Improving Instruction Through Classroom Level Planning
- * Informing National Efforts to Improve Quality: Understanding the CRT
- * Intervention Development
- * Evaluating Progress and the Effectiveness of Interventions
- * Knowledge Building in International Development

7. Reflections for Future Research

8. References and Resource List

9. Appendices (not included in Biennial Report)

Sample CBA Manual

Data Record Forms

Pupil Performance Summary Forms

Class Summary Sheets

School Summary Sheets

Data Layout for computer entry

Excel data record sheet

Table 1: Curriculum-based assessment instruments used in Ghana

Table 2: Letters and Sounds Data Record Sheet

Table 3: Pupil Performance Class Record Sheet

Table 4: Classroom Profile for Reading

Table 5: School Profile for Reading

Table 6: Overall Profile Baseline Profile for Reading

Table 7: Individual Profile: Mary Appiah from an urban school in the Central Region of Ghana

Table 8: Classroom Summary Form for Writing

Figure 1: Pie chart summarizing baseline data for Level 5 on writing words in ten minutes.

Figure 2: Pie chart summarizing baseline performance of Level 5 pupils on reading decoding tasks.

Figure 3: Pie chart summarizing baseline performance of Level 5 pupils on reading comprehension tasks.

Figure 4: Pie chart summarizing August 1995 Follow-up for Decoding of Reading Passage from the Level 5 Text (Intervention Schools only)

Figure 5: Pie chart summarizing August 1995 Follow-up for Level 5 Passage Reading Comprehension (Intervention Schools only)

Figure 6: Bar graph summarizing IEQ Baseline Performance in Ghana--Average Number of Words Produced in 10 Minutes

Figure 7: Bar graph summarizing IEQ August 1995 Follow-up in Ghana--Average Number of Words Produced in 10 Minutes

CURRICULUM-BASED ASSESSMENT AND IMPROVING THE QUALITY OF PRIMARY EDUCATION IN GHANA

I. Introduction

More children are learning to speak, read and write in English, teachers and pupils are coming to school more often, and teachers, pupils, and parents describe real changes in how children are learning. And, the excitement doesn't stop there! Circuit Supervisors, policy makers, teacher training college directors, education officers, advisors to donor agencies--representatives from throughout the educational system have been participating and providing support as the changes occur.

In the meantime, word about the work in Ghana has spread. This publication was developed in response to repeated questions from educators within Ghana and internationally, questions such as, "How did you decide where to begin in improving instruction?", "Which skills are Ghanaian pupils learning and which skills still need to be learned?", "How did you know what strategies to use to improve instruction?", and "How do you know whether these strategies are working to improve pupil learning?" Once educators and policy makers heard about the exciting changes that were happening in selected Ghanaian classrooms, they wanted to know more about how this improvement came about.

A core element of the improvement process has been curriculum-based assessment (CBA). CBA is the practice of asking students to perform tasks that have been drawn directly from the curriculum. The specific tasks are selected, administered, and scored using standardized procedures and the assessment results are used to adapt instruction to reflect the learners' needs. The primary goal is to guide the instructional decision making process (for example, the selection of appropriate materials, the level and type of instruction, and so on) so that instruction continues to be relevant to the students' instructional needs, thereby increasing the chances of successful learning. (For an overview of the value of curriculum based assessment in developing educational systems, see Harris & Pasigna, 1995.)

Driving the Improving Educational Quality (IEQ) process are two fundamental beliefs: (1) effective change is an inclusive process and only if stakeholders are invested in the process of change is the change likely to occur and be sustained, and (2) educational decisions are improved by the use of relevant and meaningful data. Using CBA is consistent with these beliefs. CBA begins with the local curriculum, the assessment process is easily understood by local educators, and the results are meaningful for decision-making at all levels of the educational system from the individual child to the national level and beyond.

2. Background in Ghana--Need for CBA

In Ghana, IEQ formed a partnership with the University of Cape Coast, USAID/Ghana, and the Ghana Ministry of Education to conduct innovative classroom-based research

focused on improving pupil learning. The effort was designed to complement the Ghana Primary Education Program (PREP), a USAID supported initiative to improve primary education. To conduct the research, the Centre for Research on Improving Quality of Primary Education in Ghana (CRIQPEG) was established at the University of Cape Coast. The first phase of CRIQPEG/IEQ research examined the availability, sources, and use of instructional materials in classrooms in six primary schools in the Central Region of Ghana. Phases II and III focused on assessing and improving language learning.

In Ghana, most children enter school speaking little or no English. Throughout the country there are at least 44 indigenous languages (most falling into one of 13 major language groups) and there are about 1700 different languages spoken in sub-Saharan Africa, giving rise to a complex linguistic situation (Hall, 1983). To facilitate communication and commerce, Ghanaian government policy prescribes that by fourth grade all instruction must be in English. Pupils are expected to have mastered sufficient English to be able to (1) comprehend lessons taught using oral English, (2) respond in class using oral and written English, and (3) read from textbooks written in English in core subject areas such as English, Mathematics, Science, and Social Studies.

Results from CRIQPEG's Phase I study, launched in February, 1993, suggested that many Ghanaian pupils were experiencing difficulty meeting these expectations and that children may not have had the opportunity to acquire even basic language skills. Classroom observations revealed that children were generally not interacting with the teacher, classmates, or written materials in ways that would promote language fluency and literacy.

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. As part of PREP, 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 very disappointing and hard to interpret. The Ministry of Education and USAID asked CRIQPEG to collect data to explore these results.

Through discussions with educational leaders and policy makers CRIQPEG identified a growing consensus. It became clear that language learning was key to improving the quality of education in Ghana and there was a critical need for research that would (1) shed light on the current status of English language proficiency and instruction, and (2) stimulate and guide efforts at improvement. It was this charge that led to the development and use of CBA instruments as part of the improvement process.

Initial Phase II research activities focused on collecting data that would provide Ghanaian educators with instructionally relevant profiles of the English oral language, reading and writing proficiency levels of children in grades 2-5. In the Fall of 1994, CRIQPEG in collaboration with IEQ consultants developed and pilot tested assessment instruments drawn directly from the Ghanaian syllabi and textbooks. Using techniques associated with curriculum-based assessment, three parallel test forms were created to measure pupil

proficiency levels in oral language, reading, and writing. In each area, the range of skills assessed spanned from very basic (for example, copying letters and responding correctly to simple oral questions) to grade level appropriate (for example, reading an average difficulty passage from the English textbook with at least 70% accuracy). Task groups composed of CRIQPEG researchers and IEQ consultants collaborated in the development, pilot testing, and revision of these instruments.

In February, 1994, using the newly developed and pilot tested instruments, CRIQPEG teams collected baseline data in 14 schools from two regions of Ghana. Selected pupils in grades 2-5 were tested individually, teacher ratings of pupil performance were collected, and all data were compiled onto summary sheets for data processing and analysis. During the next three months, results were summarized and shared with Ghanaian educational leaders.

Committed to expanding the research agenda to active participation in improving primary education, CRIQPEG team members also shared the results of the research findings with local educators and together they identified and implemented strategies (or interventions) for improving English language learning. Classroom interventions in Ghana focused on increasing pupil exposure to print, encouraging practice with oral English, and adjusting instructional practice so that all pupils become successful learners. Circuit Supervisors, headteachers, teachers, parents, and CRIQPEG team members collaborated periodically over several months to promote these improvements in 7 of the 14 participating schools. Educators and parents from the other 7 schools only participated in the research data collection activities.

All 14 schools in the sample received the textbooks that were being developed and distributed as part of the PREP initiative. The rationale for all schools receiving the textbooks was that it is already well documented that performance in schools with textbooks is better than performance in schools without textbooks (Fuller, 1987; Fuller & Clarke, 1994). Furthermore, USAID intended to supply textbooks throughout Ghana, so documenting what happens in "no textbook" schools would not be necessary or useful. The more important question for USAID/PREP and reform in Ghana was what can be done to increase the effectiveness of the innovation (supplying textbooks), that is, comparing "textbooks only" with "textbooks plus training in using them effectively".

At the end of the following school year (August, 1995), follow-up data were collected on pupils with baseline data. Of the original 1032 pupils, over 800 were located and reassessed. In addition, replacement pupils were tested as well. All pupils were tested using one of the parallel forms of instruments that had been developed prior to baseline data collection. Instruments for level 6 were developed (most pupils who had been in level 5 for the baseline were in level 6 the following year). Also, in response to an increased emphasis on pupil writing, several new measures were added: Spelling, Dictation of Sentences, and Written Expression (Story and Letter Writing; Level 6 only). In addition, the Oral Proficiency measures were revised and expanded.

The process of CBA development, use, feedback to educators, and revision is described below. Our intent is to document what was done and in so doing to share with other educators the rationales for decisions that were made and the lessons that were learned in the process of improving the quality of primary education in Ghana.

3. Developing the Assessment Materials

Assessing English Proficiency

Language forms and levels. When assessing proficiency in English, as in any other language, one has to recognize that language has two forms--oral and written--and that the communication process involves the use of these two forms at two levels--receptive and productive/expressive. The Ghanaian English language curriculum for basic education reflects these forms and levels of usage by providing experiences in the following four language modes: *listening* (oral language at the receptive level), *speaking* (oral language at the productive/expressive level), *reading* (written language at the receptive level), and *writing* (written language at the productive/expressive level).

English as a foreign language. For native speakers of English, receptive (listening) and expressive (speaking) oral language skills are acquired early at home, developing these skills "naturally" in the sense that they are acquired without formal instruction. When children enter school, they already have an oral language proficiency which can be used as the foundation for receptive (reading) and expressive (writing) written language development in school (Harris & Sipay, 1990). This is not so for the Ghanaian learners who are learning English as a foreign language. Ghanaian children, particularly in the rural areas, were found to have little exposure to English in their homes or community. Even in the urban areas, English is a "second language" in only those exceptional cases where the learner comes from a middle class urban family exposed to the use of the English language at home and through media such as billboards, newspapers, radio, and television.

The assessment instruments developed for the primary schools participating in the IEQ initiative in Ghana are based on the finding from the Phase 1 research that the learners in these schools are learning English as a foreign language and that, particularly in the rural areas, young children's use of English is limited to the school environment. The performance assessment, therefore, includes items representing the four language modes--listening (comprehension), oral expression (speaking), reading (decoding and comprehension), and writing--guided by the curricular expectations for each grade level as indicated in the English syllabi, teachers' guides, and pupils' textbooks.

Prerequisites and emergent literacy concepts. In the determination of English proficiency through the performance assessment instruments, care was taken to include items on prerequisite skills that affect one's performance in the more complex tasks. This was done to increase the probability of accurately identifying the level at which a child is able to perform and thus provide diagnostic information to teachers for purposes of planning remedial work when needed. Also included are items pertinent to "emergent literacy"--

that is, how children learn to read and write (Tompkins & Hoskisson, 1995). Examples of these are print awareness and concepts about print such as book orientation, direction concepts, and letter and word concepts (Clay, 1972, 1979).

Principles of Instrument Development

In Ghana, as in many countries, there are no published tests available to educators for assessing individual pupil performance on the national curriculum. Criterion-referenced tests developed for national monitoring purposes focus on the end of cycle performance of pupils and were never intended for monitoring individual performance or for providing immediate feedback to local educators. Standardized achievement tests from other countries such as the United States have not been validated for use in Ghana, nor are they likely to be appropriate. There is the obvious likelihood of content or cultural bias: not many children in Ghana are likely to know about hair dryers, circus clowns, or rocket ships (all found in questions in recently published US tests). Also, as was discussed earlier, pupils in Ghana are learning English as a foreign language. The sequence of curriculum objectives in the Ghanaian syllabus and curriculum is not likely to be reflected in tests from outside of Ghana, and as a result, these tests would be unfair because they would not accurately reflect what the children have learned. Furthermore, knowing a child's score or the average score for a classroom or group of children on a test based on a curriculum designed for monolingual English students would not provide the specific information needed to inform instructional intervention on the Ghanaian national curriculum. Consequently, the first step in analyzing the instructional needs of Ghanaian primary pupils was to develop assessment instruments that were directly relevant to the Ghanaian curriculum.

Desirable Qualities in the Assessment Instruments

First and foremost, assessment instruments should allow the user to answer the questions for which they are being developed and used, that is, they should be valid for the intended purpose. IEQ/CRIQPEG wanted to know what children CAN do as well as what they still need to learn. The assessment instruments needed to provide this information with accuracy and dependability. CRIQPEG also wanted the results to be meaningful to teachers, administrators and policy makers. Finally, the assessment instruments needed to be practical.

Validity--Answering the Questions That Are Being Asked: When an instrument effectively answers the questions for which it is being used, it is thought to have validity for that use. In the past, a test was thought to be either valid or not valid. More recently, educators have realized that a test that is valid for one purpose may not be valid for other purposes. For example, a test that is valid for selecting the highest performing pupils may not be valid for assessing the skills of average or low performing pupils.

IEQ/CRIQPEG wanted to know what children in Ghanaian primary schools CAN do as well as what they still need to learn. This meant that the scope of the assessment instruments needed to include skills associated with the pupils' current level as well as skills that were prerequisite to this level. Thus, a pupil in level 5 was assessed on level 5 skill objectives as well as the skill objectives of earlier levels. This strategy was useful for several reasons. First, most pupils had not mastered the skills associated with their current level in school. If the test had assessed only these skills, the results would have provided information on what most of the pupils can't do but not what they CAN do. Knowing what pupils can do was critical to determining where to begin instruction and remediation. It was the basis for developing instructional interventions that would be used for improving educational quality in Ghanaian primary schools. Secondly, for research purposes, CRIQPEG wanted to follow a group of children for several years and to monitor their learning progress. To do this, it was important to have instruments that accurately measured a broad range of skills.

Part of Accuracy is Reliability. To be useful, test scores need to accurately and dependably represent a child's skills within the domain being assessed. There are several ways to enhance the dependability of assessment. One way is to extend the assessment, that is, make the test longer, covering more of the domain. A short, limited sampling of a child's performance is subject to chance variations and likely to fluctuate with repeated testing. On the other hand, sampling from a variety of tasks and levels within the domain increases the likelihood that a child's score accurately represents the child's current skill level.

To enhance the reliability of their assessment, CRIQPEG decided to sample from a variety of levels (Level 2--the period when a child is just beginning to learn English-- through to the child's current level) on a variety of related tasks, using tasks there were representative of the level and curriculum. For example, in the reading domain, CRIQPEG identified tasks to represent the range from pre-reading skills to grade appropriate skills. In addition, several steps were taken to increase the likelihood that tasks at each level were representative of that level. For example, when selecting which reading passages from the textbooks to use to represent a grade level, it was important to make sure that the selected reading passages were of average difficulty for the levels they represented--not the hardest or the easiest.

Making the Results Meaningful: To be meaningful, test results must be linked to some criteria or reference point. This allows the responses to be interpreted and understood by people who want to use the results. For example, consider the statements below:

Ama answered correctly 65% of the 20 reading questions.

Ama read more words correctly than 65% of the children in her third grade classroom.

Ama read correctly 65% of the words in a typical 200 word passage from the third grade textbook.

The first statement sounds like it is meaningful when in fact it tells you very little. Where did the reading questions come from? How hard are they? How much reading was required to answer the questions? The second statement tells you how Ama performed as compared with other pupils in her class, but, without knowing the skill levels of pupils in her class, you still don't have the kind of information you need to make decisions about Ama's instructional needs. The third statement links Ama's performance directly to Ghanaian grade level expectations. It provides convincing evidence that Ama can in fact read some English words. Also, it conveys that Ama is likely to struggle when she reads from the available third grade textbook, but typical passages are close enough to her skill level to be within her grasp if she is provided with learning activities that help her bridge the gap.

Below are 3 statements about the performance of the children in a classroom:

The average reading score for children in the Level 3 classroom in this school was 14.

The average score for children in this Level 3 classroom was higher than the average scores from 70% of the Level 3 classrooms in the schools that were tested.

In this Level 3 classroom, 18 (45%) of the children could read most of the words in a typical passage from their textbook, another 10 (25%) of the children could read some of the words but the passage was too difficult for them to read it independently, and the remaining 8 (20%) children were preliterate: they could not read the words in the passage nor had they mastered even the most common words in their textbook such as "the" and "look".

Again, the first statement provides no meaningful information because it is not linked to any meaningful criteria. The second statement links average pupil performance to the average performance in other schools where Level 3 pupils were tested. This kind of statement informs the reader about the relative standing of this classroom when compared with other classrooms. It may be useful for program evaluation purposes but it doesn't link performance to specific skills. Perhaps literacy in all the classrooms is very low and this classroom's average was artificially bolstered by one very high scoring child. Conceivably, all of the tested classrooms, including this one, are composed mostly of children who are pre-literate.

The third statement describes performance in terms of specific skills that are relevant to the local curriculum. This description is meaningful to the classroom teacher who plans instruction for children in the classroom. It is also useful to local educators as well as

policy makers who wish to monitor academic progress of children in the schools and improve educational decision-making. It communicates to these educators that almost half of the pupils are meeting grade level expectations and available materials are appropriate. However, it also indicates that at this time the available textbooks are too difficult for over half of the pupils in the classroom. Twenty percent of the children need intensive beginning reading instruction. The implications for instructional materials, teacher training and in-service support, are clear. Finally, this kind of information is useful for international comparisons in that the reading difficulty of Ghanaian passages at each level can be linked through research to the reading difficulty of passages from other countries (see for example, Kazas & Harris, 1996).

In constructing the tasks for the assessment, CRIQPEG and its IEQ collaborators wanted to be able to link pupil performance to mastery of specific skills. To do this, the English syllabus was used to identify meaningful levels of performance and the specific enabling skills that could be used to represent each level. In some instances, the relevance of the skills was self-evident: writing one's name or recognizing upper and lower case letters. In other instances, mastery levels were established based on international standards and/or research. For example, the "percentage of words read correctly in a passage" has been used to identify optimal efficiency for various kinds of instruction (Glicking & Armstrong, 1978; Glicking & Havertape, 1981; Powell, 1971). A passage is thought to be at the instructional level if the child is able to read at least 85-95% of the words correctly. For instructional drill the corresponding percentage is 70%, whereas for independent work the percentage is 95% or higher.

Instruments must be Practical: IEQ/CRIQPEG had two specific uses for the assessment instruments: (1) there was an urgent need for useful diagnostic information about the language skills of primary school pupils, and (2) the instruments were needed as part of the on-going Improving Educational Quality Project process of conducting classroom research. Practical considerations such as ease and economics influenced the decisions made about instrument development, administration, scoring and use.

For test development, CRIQPEG chose to rely on established procedures for developing curriculum-based assessment instruments. Devoting years to test development was not feasible within the time and budgetary constraints of the IEQ project. Also using expensive or sophisticated test materials or apparatus was not appropriate. The intent was for the materials to take advantage of everyday objects and materials found in the classroom (e.g., textbooks, pencils, tables, and so on) rather than to import unfamiliar or costly alternatives. Timepieces or watches were the only piece of equipment that researchers required.

Several other practicalities were considered as well. Many of the researchers were not experienced in testing children and the amount of time available for training was limited. Also, it was hoped that the instruments would be used by teachers and other educators in the future. Therefore, the instruments needed to be fairly easy to learn to administer and the administration directions needed to be described in sufficient detail to allow the

researchers (and other users later on) to refer to them after the training and to minimize questions or confusion once the researchers were in the field.

Ease of scoring was another important factor. An attempt was made to develop assessment instruments that could be scored objectively, that is, the need for the researcher/user to make subjective judgments was minimized. For the most part, all of the scoring was done in the field. This had the advantage that scores could be recorded on summary sheets promptly and efficiently. This also reduced problems associated with keeping track of all of the different pieces of data for each child. This did mean however that some of the more subjective strategies available for assessing language skills were not feasible. For example, holistic scoring of expressive writing is a commonly used method for evaluating writing skills. However, considerable training is needed for adequate inter-scoring reliability and typically "table scoring" (getting all the tests and all the scorers together for the scoring exercise) is preferred. Consequently, IEQ/CRIQPEG opted for a more objective method of scoring the expressive writing tasks.

Also, the assessment data needed to be easily and quickly scored and recorded. For example, sending computer scannable answer sheets to a test publisher was neither practical nor economical. Besides, this is not something a teacher or Circuit Supervisor would want to do. Also, the data from the instruments needed to be easily quantifiable. Analysis of data at the classrooms level, school level, and across 14 different schools is greatly simplified when the data can be hand tallied or entered into a desktop or laptop computer and analyses performed using commonly available software.

One final bit of practicality and efficiency had to do with using answer sheets for recording the pupil responses. Initially, there were lengthy test booklets for some of the tasks. The researchers recorded pupil responses in the test booklets, requiring sufficient duplication of the test booklets so that there was a booklet for each pupil who was tested. This represented costly duplication and time consuming scoring and recording of responses. Following the initial use of these instruments, researchers began recording pupil responses on summary sheets; reusing one test booklet for administering the questions and one answer sheet per class or pupil depending on the length of responses. Worth noting was that the original use of separate test booklets for each pupil allowed researchers sufficient room to record helpful comments and to note responses that were difficult to score. Thus, it seems that sometimes for a first administration, it may be useful to have separate booklets for recording each pupil's responses and any difficulties that the researcher encounters. For later administrations, once examiners are experienced and difficulties with the test are resolved, alternative more efficient strategies can be explored. Also, this was not a situation in which pupils were actually using and handling the test booklets. In conducting research or in instances when scores are being compared from one assessment to the next, it is not a good idea to switch midstream from a situation in which pupils write directly on their own test booklets to one in which they use an answer sheet. Using an answer sheet involves an additional step and it alters the complexity of the task for the pupils.

Developing the Instruments

The process began with a review of the Ghanaian English language curriculum. Using the English syllabus, the teacher guides, and the textbooks, it was possible to identify the objectives and related skills that each pupil was expected to master at each level of primary education. This formed the basis for deciding what to measure. Major skills were identified and listed in the order in which they were represented in the instructional materials.

For each language mode, oral language (listening and speaking), reading, and writing, specific tasks were developed to measure the skills identified in the Ghanaian language curriculum. Table 1 provides a listing of these tasks. For each task, there is a description of what it is intended to measure, sample questions to illustrate how the skill is measured, a brief statement of how it is scored, and illustrative examples of how the scores can be interpreted. In providing the sample interpretations, an attempt was made to show the range of possible ways the data can be used. Thus, some of the interpretations describe the performance of an individual pupil, some profile a class or group of children, and some refer to interpretations such as are used in program evaluations.

Work Group Process

In developing the assessment materials a work group of ten researchers was formed. The group represented a collaboration between CRIQPEG and IEQ consultants and included two language specialists, and three others with expertise in assessment. Subgroups were formed to be responsible for preparing the assessment instruments in each area and for each grade level. Before the subgroups started working on the instruments, the whole group met to agree on the time/work schedule and to identify the following sub tasks to be performed:

1. define the skills to be measured in each area (Oral Language, Reading, and Writing);
2. identify the tasks to be performed by pupils to measure their performance in each of the skills being measured;
3. construct and review test items in each skill area;
4. develop administration and scoring procedures;
5. assemble draft forms of the assessment instruments;
6. pilot test the instruments and the administration and scoring procedures;
7. analyze test items to determine item relevance and consistency;
8. revise tests and administration procedures;
9. prepare forms for recording data as it is collected
10. compile an administration manual that clearly identifies: materials needed, procedures for preparing to test, procedures for administration of the instruments, directions for scoring, directions for recording each pupil's responses and total score, and procedures for storing of data;
11. train test administrators (including practice sessions).

Table 1:
Curriculum-Based Assessment Instruments Used in Ghana

TASK	LEVELS	WHAT IT MEASURES	SAMPLE QUESTIONS	SCORING	SAMPLE WAYS SCORES CAN BE INTERPRETTED
ORAL LANGUAGE					
Oral-Functional Language	P2-P6	Proficiency with everyday functional oral English	"What's your name?" "How old are you?" "What is the name of your school?"	Extra credit for complete answers.	When children are questioned using everyday English, 30% of the children respond correctly to most of the questions.
Listening Comp: P2-P6	P2-P6	Following oral directions. Comprehension of oral English. Drawn from English syllabus.	"Push the table." "Open to page 13 and point to the monkey "	% Correct	When P6 children are questioned using oral English associated with the syllabus for each level, 60% respond correctly to P2 questions but only 20% of the children respond correctly to P6 questions.
Listening Comp: P6 Passage	P6	Understanding a passage from the textbook that is read to the pupil	Comprehension questions based on the passage, e.g., "What did Dede find out?"	% Correct	When a passage from the P6 textbook is read to the child, she demonstrates that she understood by responding correctly to more than 75% of the comprehension questions.
Oral Expression. P2-P6	P2-P6	Speaking English appropriately in response to questions drawn from the English syllabus.	"Name 2 things we use water for in the house?" "Who are pounding nuts in this picture?"	Extra credit for complete answers.	Although 40% of the level 6 pupils were able to demonstrate understanding of oral English, far fewer pupils were able to demonstrate competence in speaking English.
PRE-READING/READING					
Concepts about Print*	P2-P6	Hands on exposure to print.	Questions asked in English and vernacular, e.g., "Turn to page 5."	# Correct	By level 5, all but a few of the pupils demonstrated mastery of basic skills associated with using printed materials such as finding a page or turning to a specific unit.
Letter/Sound Recognition	P2-P6	Alphabet recognition/discrimination	Upper and Lower case letters	# Correct	While very few of the level 2 pupils recognized a majority of printed letters, by level 5 this skill was mastered by most of the pupils.
Aided Reading	P2-P6	Pointing to words that are read.	same as below	% Correct	There were 20% of the pupils who couldn't pronounce the words but they were able to locate the words when the words were read aloud.
Reading Most Used Words	P2-P6	Reading of most commonly used words in the P2-P6 textbooks	Word lists with words such as and, the, for, one, they, etc.	% Correct	75% of the pupils demonstrated that they were able to read all or almost all of the most commonly used words in their textbook.
Reading Passage from Textbook	P2-P6	Decoding accuracy	Passages selected from P2-P6 textbooks	Words % Correct	For 25% of the pupils in this classroom, the passage is too difficult and consequently pupils will become frustrated and the learning will be inefficient.
		Decoding speed	Speed in first minute of reading the above passages	Words/Minute	Pupils in level 3 average about 18 words per minute whereas level 6 pupils are able to average about 42 words per minute.
Passage Comprehension	P2-P6	Reading Comprehension	Questions based on the above passages	% Correct	Reading quickly and accurately was associated with reading comprehension. Pupils who read slowly also had more difficulty with the comprehension questions.

Table 1:
Curriculum-Based Assessment Instruments Used in Ghana

WRITING	LEVELS	WHAT IT MEASURES	SAMPLE QUESTIONS	SCORING	SAMPLE WAYS SCORES CAN BE INTERPRETTED
Copying Letters	P2-P6	Copying letters using a pencil and paper	Pupil is asked in vernacular to copy his/her name.	Pass/Fail	About 3/4 of all level 2 pupils could copy letters. By level 3, more than 9 out of every 10 children were able to copy letters.
Writing Name	P2-P6	Writing name correctly without help	Pupil is asked in vernacular to write his/her name.	Pass/Fail	By level 6, all pupils could write their names without assistance.
Writing Words	P2-P6	Writing vocabulary	Pupils are asked to write as many words as they can within 10 minutes.	# of correctly spelled words	Most pupils in level 2 and 3 experienced difficulty writing more than a few words whereas by level 6 most pupils wrote more than 10 words or more.
Spelling	P4-P6	Approximate spelling of commonly used words	Words taken from most commonly used words in English textbooks	Letters % correct (must be in correct sequence)	While only 20% of level 4 pupils spelled most of the words correctly, about 40% of the pupils were able to approximate the spelling by identifying most of the correct letters in these words.
		Spelling of commonly used words.	same as above	Words % Correct	By level 6, 40% of the pupils were able to spell correctly (with at least 75% accuracy) the most commonly used words in their textbooks.
Dictation	P4-P6	Writing words that are dictated	Dictated sentences from the P2-P4 English textbooks	Words % Correct	When sentences from the P2-P4 textbooks are dictated, pupils are able to spell 30 % of the words correctly.
		Correct use of capital letters	same as above	Capital Letters % Correct	On the average, pupils correctly capitalized about half of the letters that should have been written in upper case.
		Correct spelling, punctuation, etc.	same as above	Correct Writing Sequences % Correct	When sentences from the P2-P4 textbook are dictated and spelling, punctuation, and capitals are considered, pupils in the intervention schools performed significantly better than pupils in the comparison schools.
Writing Story	P6	Fluency in written expression	Pupils are given a topic from English syllabus and asked to write story.	Words produced - # (spelling not considered)	When asked to write a story, Akua wrote 125 words.
		Words spelled correctly in written expression	Exa: Most children in Ghana know Anansi stories. Write an Anansi story or some other kind of story.	Words spelled correctly (#)	Of the 125 words that Akua wrote, 119 (95%) were spelled correctly.
		Correct word use (syntax and semantics), spelling, punctuation, etc. in written expression	same as above	Correct writing sequences (#)	When word use, spelling, punctuation, and capitalization are considered, Akua was able to produce 70 correctly written sequences.
Writing Letter	P6	Fluency in written expression	Pupils are given a topic from English syllabus and asked to write a letter.	Words produced - # (spelling not considered)	The median number of words produced by level 6 pupils who were asked to write a letter was 38. Some pupils were not able to express any written words whereas the longest letter contained 182 words.
		Words spelled correctly in written expression	Exa: Imagine a friend gave you a gift. Write a letter to the friend thanking the friend for the gift. Include something you like about the gift.	Words spelled correctly (#)	The average number of correctly spelled words in letters written by level 6 pupils was 34.
		Correct spelling, punctuation, etc.	same as above	Correct writing sequences (#)	When correct word use, spelling, and punctuation are considered, the average number of correctly written sequences in the intervention schools was 27 whereas the average in comparison schools was 21.

15

Developing Oral Language Instruments:

Consistent with foreign language learning and the Ghana curriculum, three aspects of oral language were identified for assessment: functional English, listening comprehension, and speaking or oral expression. In developing the oral proficiency test, curriculum materials such as the English textbooks, the teacher's handbooks or guides, and the English syllabi for the appropriate grade levels were consulted. Each of the subgroups worked on the materials for a particular grade level. The groups started working by first familiarizing themselves with the textbooks, syllabi, and teachers' guides. Using this information, they constructed the test items (i.e., questions).

Functional English referred to items that assessed pupils' understanding of functional or everyday English such as, "Good morning," "How are you?", "What is your name?" and so on. This section was used to test pupils' skills in routine English expressions. When someone is learning a foreign language, these are often the first expressions that they learn and practice. In the classroom, teachers often begin English class instruction with these expressions and children gain practice in choraling or individually providing the responses.

With regard to listening comprehension, the instrument was developed to test the child's understanding of simple commands and instructions. In this section, pupils needed to demonstrate understanding of oral questions by pointing to objects in pictures or following simple directions. The pictures that were used were from the textbooks and the directions were drawn from typical exercises in the curriculum or classroom. Also, the Level 6 version included a task in which a passage from the Level 6 English textbook was read aloud to the pupil and the pupil was asked to respond to oral comprehension questions.

The oral expression questions tested pupils' ability to express themselves in English. Pupils needed to be able to understand the questions and to respond appropriately using spoken English. Children were instructed to provide a complete answer and extra credit was given to children who used a sentence in responding to the questions. When children provided one or two word answers, they were reminded and encouraged to provide a complete answer. The criteria for judging correctness of responses were based on communication--that is, that the pupil used the correct expressions or sentences and was able to communicate this to the examiner. More importance was given to whether the pupils' responses were within the phonemic range of the English language, rather than on whether or not the pupil was able to approximate native-like enunciation, pronunciation, or intonation.

Developing Reading Instruments:

In the area of reading, IEQ/Ghana wanted to know answers to questions such as: In a primary classroom, how many children can read and comprehend the available textbook? When children can't read the textbook, what reading and prereading skills do they possess? How many children in the classroom are non-readers? The aim was to assess children's skills on a continuum from pre-reading skills (letter/sound recognition, concepts about printed material) to the skills required for the assigned English textbook.

Pre-Reading and Beginning Reading Tasks

Pre-Reading Tasks (Letters and Sounds; Print Concepts): For the prereading skills, work by Maria Clay served as a framework for measuring the skills needed for beginner readers. For the Letters and Sounds task, children were asked to identify upper and lower case printed letters. The intent of the task was to determine if the child could discriminate between different printed letter shapes. Thus, children were given credit for recognizing the letter if they could say the letter name, make the sound the letter makes when used in a word, OR say a word that begins with the letter. Another pre-reading task measured children's hands-on familiarity with books and print. As noted by Clay (1979a, b), some of the preliminary steps in learning to read are knowing how to orient the book, knowing that reading of English proceeds from left to right, and so on. While Clay developed her own materials for testing print concepts (Clay, 1979a, b), the Ghana version makes use of the local textbooks to measure children's familiarity with books and print.

Beginning Reading (Aided Reading; Reading Most Used Words): Once children begin to read simple words and sentences, they begin to recognize the most commonly used words. Words such as "the", "and", "she", and "you" are found more often in print than words such as "table" or "family". When children read stories it is important that they recognize the common words so that they become less frustrated by uncommon or new words. There is research to suggest that children learn best when they know most of the words they encounter and there are only a few new or unfamiliar words (Glicking & Armstrong, 1978; Powell, 1971). For this reason, it seemed useful to determine pupils' familiarity with the most commonly used words in the English textbooks. To do this it was necessary to determine the most commonly used words in each of the English textbooks. In Ghana, there is one English textbook for each level. The number of times each word in the textbook was used was tallied and the 60 most used words were identified. These 60 words were ranked from most used to least used and distributed to one of 3 lists in order to create 3 parallel forms, each with 20 words. The most used word was placed on list A, the next most used word was placed on list B, the third on list C, the fourth on list C, the fifth on list B, the sixth on list A, the next on list A again and the process was repeated (ABCCBA...) until all 60 words had been distributed. One of these forms was used for baseline data collection, another for the August 1995 follow-up assessment, and the third is available for use in a final assessment.

Each child was asked to read the words on the Most Used Word list for their level. If they were unable to read the first or second word, they were encouraged to find and point to any words they could read. They were given credit for any words they were able to read correctly.

When this task was first discussed, some team members felt that pupils may not be able to read the words but they may recognize them. This seemed particularly likely since the pupils were learning English as a foreign language. To assess whether pupils would recognize the words even if they couldn't pronounce them, an aided reading task was

added to the measurement process. Words from the Most Used Word list were dictated to the pupil in random order and the pupil was asked to point to the word. This Aided Reading task was viewed as another way of making sure that each pupil's progress in learning to read was captured in the assessment process.

Reading: Decoding and Comprehension

Selecting Reading Passages : In order to evaluate the match between available curriculum materials and children's reading skills, children were asked to read typical passages from their English textbook and the textbooks from grade levels below their own. Several factors were considered in selecting these reading passages. In Ghana, there is one English textbook per grade level and 30 units in each textbook. What was needed for the CBA were passages for each primary grade level that were representative of other passages in the textbook.

Initially 5 passages were selected from each textbook. These five passages met the following criteria: (1) they were not drawn from the first 10 units in the textbooks (it is not unusual for teachers to devote most of the year to the passages at the beginning of the textbooks; pupils may memorize these passages and performance on passages from this section are less likely to be representative of performance on the book as a whole); (2) they were not drawn from the last 10 units (there may be legitimate reasons why teachers do not reach the last third of the book); (3) they were similar in style to the other units and passages; and (4) they contained sufficient narrative to yield a long enough sample of the child's reading (about 70 words at level 2; 170 words in level 3 and at least 200 words in upper primary.) During pilot testing, all five passages were administered to each of the children and, for each passage, the total number of correctly read words was tallied. Based on the pilot test data, the easiest passage and the hardest passage at each level were eliminated. Visual inspection of pilot test data was used to assure that pupil performance on the remaining three passages was not wildly discrepant.

Decoding: Pupils were provided with copies of their English textbooks and asked to read the selected passages. The researchers used directions provided in the administration manuals for each grade level (see the sample taken from the Level 4 Manual). As the child read aloud the words in the passage, the researchers followed along on a copy of the passage and noted which words the child was and was not able to decode correctly. Reading speed was determined by noting how many words the pupil read correctly within one minute. Reading accuracy was determined by calculating the percentage of words in the passage that the pupil read correctly.

Reading Comprehension: For each passage, reading comprehension questions were developed to assess whether the child understood the passage. For baseline data collection, only 5 questions were used for each passage. For the second pupil performance data collection, questions were added so that each passage had 8 questions. Typically, a few questions were recall questions about facts in the passage (e.g., "What colour is the cap?", a few required some manipulation of available information from the story (e.g.,

"How many animals are mentioned in the story?"), and one question for each story required the child to make an inference based on the data provided in the story (e.g., "Why would the boy wear different clothes in the garden than he wears to town?").

SAMPLE TAKEN FROM THE LEVEL 4 MANUAL

Reading Passages

Directions for Administration, Scoring, and Recording of Scores

Materials Needed:

- *English Pupil's Books for P2, P3, and P4
- *Individual Record Sheets (Numbered copy of the passages for P2, P3 and P4; 1 copy per pupil of each; used by the examiner for recording the known and unknown words)
- *Pencil or pen (for the examiner)
- *Timer
- *Clipboard
- *Straight edge for helping the child to focus on one line (e.g., this could be a rectangular piece of plain cardboard or heavy paper)
- *Pupil Performance Class Summary Sheet

Directions for Administering:

BEGIN WITH THE P2 PASSAGE AND THEN REPEAT THE ADMINISTRATION PROCEDURES FOR THE P3 AND P4 PASSAGES. P4 PUPILS WILL HAVE READING SCORES FOR P2, P3 AND P4 PASSAGES!

1. Open the textbook to the selected passage. Hand the opened textbook to the pupil. Provide the pupil with the straight edge and demonstrate how it could be used. Using the straight edge is optional.
2. Place the Individual Record Sheet for the passage in front of you but shielded so the pupil cannot see what you record. A clipboard is useful for this purpose.
3. Say these specific directions to the pupil for each passage:

When I say 'begin,' start reading aloud at the top of this page. Read across the page

(DEMONSTRATE BY POINTING TO THE PLACE WHERE YOU WANT THE PUPIL TO BEGIN).

Try to read each word. If you come to a word you don't know, I'll tell it to you. Be sure to do your best reading. Do you understand what I want you to do?

4. If you are uncertain as to whether the pupil understands what is expected, the directions may be repeated in the mother tongue.
5. Say 'begin' and start your timer when the pupil says the first word. If the pupil fails to say the first word of the passage after 3 seconds, tell them the word and mark it as incorrect, then start your stopwatch.

IF THE CHILD STOPS READING BEFORE THE END OF THE PASSAGE, TELL THE CHILD TO KEEP READING

Show the child where you mean if necessary.

6. Follow along on your copy. Put a slash (/) through words read incorrectly (see scoring procedures).
7. At the end of 1 minute, place a bracket (]) after the last word. Allow the pupil to finish the passage.

When the child has finished the last sentence on the numbered copy, say,

Please stop. Thank you.

8. If at the end of one minute a pupil has read less than 4 words correctly and the child is struggling, stop the child, place a straight edge under the first line and say, **Now I want you to look at the rest of this line. Can you read something from this line? Do you see any words you know? Look and say the words you know.**

After the child responds, move the straight edge to the next line and repeat this prompt. Continue for each line in the story.

On the Individual Record Sheet circle the words that the child is able to read (the child must point to the word and say what it is). If the child knows a word that is repeated in the line, ask if the child sees that word anywhere else.

9. Next say: **Now I am going to ask you some questions about what you have read.** Then ask the comprehension questions. The pupil can look back over the story.

Directions for Scoring:

This task is scored by counting the number of correctly read words. The manual provides rules for determining if a word has been read correctly. [These rules are adapted from Tilly, W. D., & Carlson, S. (1992). Administration and scoring. In M. R. Shinn, N. Knutson, and W.D. Tilly (Eds.). Curriculum Based Assessment: Training Modules (3rd ed.) (pp. 8-12). Eugene, OR: University of Oregon.]

Directions for Recording the child's scores: Record the child's name and identifying information on the P2, P3, AND P4 Individual Record Sheets (the numbered copy of each of the reading passages). Count up the total number of words read or attempted within one minute [use the numbers at the end of each line as a guide]. Then count the number of errors and the number of words read correctly [number of errors + number of correctly read words = total words read/attempted]. Record these numbers in the box that has been provided on the Individual Record Sheet. Record the number of words read correctly in one minute on the Pupil Performance Class Summary Sheet.

Then count the number of words read correctly for the whole passage. Even if the child didn't finish the passage (i.e., if the passage was too hard and the child read fewer than 4 words correctly in one minute) the Total Words Read for the whole passage is still reported as the total number of words in the passage [i.e., the number following the last line of the passage]. Also, words not attempted are scored as errors. Count up the number of errors. The number of errors plus the number of words that were read correctly should be the same as the total number of words in the passage.

Then calculate the percentage correct for the entire passage (percentage correct=words read correctly/total number of words in the passage). Record this number in the Individual Record Sheet box and on the Pupil Performance Class Summary Sheet.

Individual Record Sheet (Examiner's Copy)

Level 4: (Text taken from Unit 16 of An English Course for Ghanaian Schools: Pupil's Book 4, p. 67.)

HOW CRAB LOST HIS HEAD	5
A long time ago Crab was a	7
very handsome man. He was	12
tall. He had fine hair. He was	19
also very kind. He had many	25
friends. Crab's friends knew that	30
he was a kind man. So they	37
asked him to do many things for	44
them. Crab was able to do	50
everything for them	53
One day Scorpion said, "I want to take my house to a	65
new place. Can you carry it for me?" Crab was able to	77
carry the house for Scorpion.	82
On another day, Snail said, I am going to my farm.	93
The river is very full of water. I cannot cross it. Can you	106
drink up the water for me?" Crab was able to drink up all	119
the water, and Snail crossed to his farm.	127
A few days later Lobster	132
came to Crab. He said "My	138
friend, I am going to marry.	144
"Please give me a kente cloth to	151
wear." This was the least difficult	157
thing for Crab to do. He gave	164
Lobster the kente cloth; he also	170
gave him a pair of sandals.	176
One day Spider came. He said, "Crab, I'm in trouble. I	187
must go to the chief's house now. I must think and speak	199
well to the chief.	203

Name of pupil: _____	P1	P2	P3	P4	P5	P6	School: _____
	1 Minute		Total			Pupil ID# _____	
Words Read Correctly: _____	_____		_____				
Errors: _____	_____		_____				
Total Words Read: _____	_____		_____				
Comprehension (Percent correct)	_____		_____				

Developing Writing Assessment Instruments

For most children, the most basic level of writing is to be able to copy forms and shapes. Often children begin by copying the letters in their names. Once children are able to copy letters, a priority in the Ghanaian curriculum is to learn to produce their name from memory. After learning to write their name, pupils begin learning to write other words. This continuum formed the basis for the writing assessment process. Initially, pupils of all levels were asked to write their names. When pupils were unable to write their names, they were asked to copy letters from their name. Pupils who were able to write their names without assistance were automatically given credit for being able to copy letters. Next, pupils were asked to write as many words as they can. They were encouraged by prompts such as, "Can you write 'the' or 'and'?" ... "Can you write the names of your friends?" "Can you write the names of animals: dog, cat, ...?" After pupils finished writing the words, they were asked to read what they had written. Only correctly spelled and read words were counted. This avoided the possibility that children had memorized copying exercises without comprehending what they were doing. While asking pupils to write any words they know doesn't capture their ability to formulate sentences or to demonstrate comprehension of the meaning of these words, it does permit them a wider range of response. Not knowing the correct verb or grammar doesn't inhibit their performance.

After the first assessment, it was clear that some pupils were able to go beyond writing words. For follow-up assessments (August 1995), additional writing exercises were included to assess dictation skills as well as expressive writing skills. Spelling and dictation skills were assessed in levels 4-6. The words that were used in the spelling component came from the most used words list from the Ghanaian English textbooks and the sentences used for dictation came from passages from the textbooks. In addition, level 6 pupils were asked to write a letter and a short story and these were used to assess their written expression skills. The prompts for the letter and short story were based on objectives in the English syllabus. The scoring of the spelling, dictation, and the open ended writing experiences was based on work by Shinn and his collaborators (Shinn et al., 1990). For example, in addition to recording the number of correctly spelled words, the researchers calculated the number of correctly identified and sequenced letters. This additional scoring strategy provides a rough estimate as to whether the pupils are approximating the correct spelling. For example, a pupil, Agnes, may only spell 25% of the words correctly but if she gets 65% of the letters correct and in the correct order, this suggests that she has an understanding of how to spell words. By contrast if she only gets 25% of the letters correct, she may have memorized a few words but not really understood the principles of spelling. For the written expression, writing samples were scored with regard to fluency (number of words attempted), number of correctly spelled words, and number of correct writing sequences, an objective scoring procedure that takes into consideration punctuation, grammar, and so on. The researchers opted not to score the samples for characteristics such as creativity or style because this kind of scoring takes considerable training and practice in order for the resulting scores to be reliable.

Pilot Testing and Revising the Instruments

Pilot Testing: After the initial instrument development, the instruments were tried out with pupils from two primary schools representing rural and urban schools. In order to prevent schools participating in the research from having an opportunity to preview and practice on the assessment instruments, the schools where the instruments were pilot tested were not schools that were participating in the research. Pupils from four classes, that is levels 2 through 5 were used to pilot test the instruments. To insure that the pilot testing included pupils with varying skill levels, class teachers were asked to group the pupils in each class into three categories, namely very good, average, and below average. Then from these three groups, nine pupils were selected randomly, three from each group.

Each pupil was assessed individually. A quiet convenient location away from the classroom block was selected for the assessment in order to minimize distractions. Each pupil was escorted to the testing location and then tested using all of the instruments appropriate for their level. In addition, pupils were administered all of the oral proficiency questions and they read all five of the reading passages at each level. During the testing, pupils were given brief breaks as needed. Oral proficiency tasks were administered first, reading second, and writing third.

Revising the Instruments: After the pilot testing, the researchers involved with the pilot testing discussed all the test items, analyzing them to determine their suitability and validity. Some directions for pupils were clarified to ensure that pupils understood what they were being asked to do. Also, some of the scoring rules were reviewed and discussed to improve the consistency of scoring. After analyzing the test items, some of them were revised and others dropped based on the results obtained from the pilot testing. The oral proficiency items were ranked from easiest to hardest based on the pilot test performance and then distributed into three parallel forms. Pilot test performance on the reading passages was used to eliminate the easiest and the hardest at each level. The remaining three passages at each level constituted the three parallel forms of the test.

Training CRIQPEG Researchers to Administer and Score the Tasks: A training workshop was organized for the CRIQPEG researchers to prepare them for the administration and scoring of the CBA instruments. The workshop was led by a consultant with expertise in assessment. The researchers were first taken through strategies in preparing for the testing. This included discussions on topics such as preparation of the tester, preparation of the test environment, preparation of the pupil to be tested (for example, establishing rapport before beginning the test; using the vernacular to maintain rapport and encourage pupil effort), and the test administration which included topics like standardization of administration procedures, scoring and recording pupil responses, handling of pupil performance records, and observation of the test behavior of the pupil during the testing. Many of these topics are reiterated in the pupil performance manuals. Also, participants were taken through the test items in order to seek clarification on any of the items or directions for administering or scoring that were unclear. Finally, the participants were

encouraged to practice administering the tests on primary school children in the community before going out to administer them in the project schools.

Recording Data on Summary Forms as It Is Collected: Keeping track of over 1000 pupils and their scores was an important challenge for the teams, particularly because CRIQPEG wanted to maintain a longitudinal data set that would allow researchers to follow the progress of each pupil over multiple school years. To facilitate this process, data summary forms (e.g., Most Used Words Data Record Sheet, Pupil Performance Class Summary Sheets, etc.) were generated prior to collection of the baseline data and directions for using these forms were incorporated into the test administration manuals and training. For subsequent CBA data collection, data summary sheets for each class with pre-printed pupil names and IDs were generated using the computerized baseline data set. In addition to helping to keep track of pupils, these forms were set up so that the data were ready for quick hand tallies as well as computer entry and analysis.

Table 2 provides a sample of a completed Letters and Sounds Data Record Sheets.¹ Notice that the children's names are written in a column on the left side of the page and there is a place to record each child's performance on each upper and lower case letter. A "1" indicates that the child received credit for the letter; a "0" means the child was not able to identify the name of the letter or the sound it makes. As you can see from this sheet, most children in this classroom recognized the most commonly used letters. They had more difficulty with letters that occur less frequently such as "x" and "v". In the column on the far right side of the page is a total score for each child for the Letters and Sounds task.

Table 3 provides a sample of a completed Pupil Performance Class Summary Sheet.¹ Notice that the children's names are again written in the column on the left side of the page. They are in the exact order as they were written on the Letters and Sounds Data Record Sheets and all other record sheets. Also, the children's total scores for the Letters and Sounds task have been copied onto this summary sheet. Similarly, the children's total scores from other tasks have been recorded.

4. Assessment and data management process

Sampling pupils: Following the training, the research teams (each composed of a leader and 3 other members) went to the 14 participating schools to administer the measures to pupils selected for the research sample. Teachers in participating classrooms were asked to provide a class list and to group the children in their class into three categories: good, average or poor students. Then, on a rotating basis, pupils from each category were randomly selected. In all, 25 pupils were selected from each of the grade levels 2-5 in the intensive schools, while 15 pupils were selected from each of the grade levels 2-5 in the seven non-intensive schools. These class lists and teacher groupings were kept by the team and used when replacement pupils were needed. The intent was to reduce the chances that pupils with better attendance (and possibly better skills) would be overselected because of their availability.

Page <u>2</u> of <u>2</u>		Part 2 Upper Case Letters																				Total (0-26)	TOTAL Parts 1 & 2 (0-52)						
ID #	Pupil's Name (Last, First)	A	N	C	I	H	E	S	Q	D	K	O	T	U	P	W	J	V	B	F	Z			L	G	M	R	Y	X
01	Hagan, Raymond	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
02	Kingsley, Eric	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
03	Mensah, Angelina	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	25	51
04	Ahmed, Ishek	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
05	Aborgye, Agnes	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
06	Akka, Ekua	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
07	Debrah, Grace	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24	48
08	Mendis, Naomi	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	51
09	Sarsah, Jason	1	1	1	1	1	1	1	1	0	1	1	1	0	1	0	0	0	0	1	0	1	1	1	1	0	0	18	28
10	Lea, Theophilus	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	46
11	Diauwu, Sylvia	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24	50
12	Hayford, Janet	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	1	0	1	1	1	1	0	20	45
13	Gyatey, Augustine	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
14	Abaka, Samuel	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	1	1	0	1	1	1	1	1	1	22	48
15	Bardoo, Victoria	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	25	51
16	Appiah, Mary	1	1	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	23	45
17	Denayah, John	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26	52
18	Musah, Ishek	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	1	0	1	1	1	1	1	0	21	46
19	Kufoe, Urey	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	0	0	1	0	1	0	0	1	1	1	18	32
20	Ofori, Augustina	1	1	0	1	0	1	1	0	0	1	0	1	0	0	0	0	0	0	1	0	1	0	1	1	1	1	14	28

School Mpeasem Primary
 School Code 16
 Teacher Brown

Letters/Sounds
 Data Record Sheet*

Level 5
 Date 25 Feb. '14
 No. on Roll 48

Page <u>1</u> of <u>2</u>		Part 1 Lower Case Letters																				Total (0-26)							
ID #	Pupil's Name (Last, First)	ID Info	f	k	p	w	z	b	h	o	j	u	a	c	y	l	m	d	n	s	x		i	e	q	r	v	t	g
01	Hayan, Raymond		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
02	Kingsley, Eric		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
03	Mensch, Angelina		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
04	Ahmed, Ishak		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
05	Aboagye, Agnes		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
06	Acton, Ekua		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	24
07	Debrah, Grace		1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	25
08	Mends, Naomi		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	10
09	Sarsah, Jason		0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20
10	Lee, Theophilus		1	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1	0	1	0	1	26
11	Drauwie, Sylvia		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
12	Hayford, Janet		1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
13	Gyatey, Augustine		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
14	Abaka, Samuel		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
15	Baidoo, Victoria		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	22
16	Appiah, Mary		1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	26
17	Denayah, John		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	25
18	Musah, Ishak		1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	14
19	Kuohie, Urey		1	1	0	0	0	1	1	1	0	1	0	0	0	0	1	1	0	1	0	1	1	0	1	0	1	0	14
20	Ofori, Augustina		0	1	0	0	1	0	1	0	1	1	0	0	0	0	0	0	0	1	1	0	1	1	0	1	0	1	14

Page	of	Pupil's Name (Last, First)	Gender (M or F)	Oral Language			Writing			Reading															Comments	
				Functional	List. Comprehension	Oral Expression	Copy Letters (0 or 1)	Write Name (0 or 1)	Writing Words (# correct)	Print Concepts- Ghana	Letters & Sounds (0-52)	Aided Reading (% correct)	Most Used Words (% correct)	P2 Passage			P3 Passage			P4 Passage			P5 Passage			
														1 Minute (# correct)	Whole Passage (% correct)	Comprehension (% correct)	1 Minute (# correct)	Whole Passage (% correct)	Comprehension (% correct)	1 Minute (# correct)	Whole Passage (% correct)	Comprehension (% correct)	1 Minute (# correct)	Whole Passage (% correct)		Comprehension (% correct)
01		Hagan, Raymond	M	21	4	24	1	1	44	10	52	100	100	74	99	100	96	99	100	73	100	100	72	99	60	
02		Kingsley, Eric	M	21	3	4	1	1	27	9	52	100	100	51	100	50	81	99	20	77	97	20	60	96	20	
03		Mensah, Angelina	F	21	3	5	1	1	33	10	51	100	80	37	96	50	56	99	20	49	100	20	21	93	20	
04		Ahmed, Ishak	M	21	4	7	1	1	25	10	52	100	96	56	96	50	42	97	40	63	94	20	29	91	20	
05		Abnogye, Agnes	F	21	4	14	1	1	17	10	52	100	84	14	61	25	10	75	20	14	68	20	14	53	0	
06		Ackon, Ekua	F	21	4	10	1	1	35	10	52	96	92	46	97	75	42	98	40	59	96	40	42	91	60	
07		Debrah, Grace	F	21	6	9	1	1	21	8	48	100	96	26	75	75	24	79	75	23	79	40	16	63	20	
08		Mends, Naomi	F	21	3	14	1	1	37	8	51	92	88	57	96	100	56	98	60	77	98	40	58	90	0	
09		Sarsah, Jason	M	6	1	0	1	0	3	3	28	20	0	2	8	0	2	6	0	3	8	0	2	4	0	
10		Loa, Theophilus	M	18	2	12	1	1	12	6	46	76	40	6	56	0	11	50	40	23	70	80	11	45	20	
11		Drauo, Sylvia	F	18	1	0	1	1	15	8	50	56	84	2	24	0	2	23	0	2	42	0	2	15	0	
12		Hayford, Janet	F	12	1	4	1	0	4	5	45	40	12	3	21	20	4	20	0	3	21	0	6	16	0	
13		Gyatey, Augustine	M	21	4	8	1	1	45	9	52	96	96	57	97	75	88	94	100	85	96	80	57	92	60	
14		Abaka, Samuel	M	12	2	5	1	1	15	8	48	92	40	14	60	25	11	58	0	8	40	0	10	44	0	
15		Baidoo, Victoria	F	20	4	18	1	1	26	10	51	100	96	56	97	100	80	98	80	62	96	60	47	87	40	
16		Appiah, Mary	F	18	3	6	1	1	21	10	45	64	24	11	40	50	13	51	0	10	48	20	12	38	40	
17		Denanyah, John	M	20	2	16	1	1	23	9	52	96	100	10	70	25	9	63	25	7	68	0	9	25	0	
18		Musah, Ishak	M	18	3	7	1	1	16	5	46	88	80	5	35	0	8	31	0	7	50	25	10	54	0	
19		Kuofre, Tracy	M	12	1	3	1	0	2	8	32	20	12	0	12	0	11	11	0	2	15	20	1	8	0	
20		Ofori, Augustine	F	6	1	0	1	1	6	9	28	16	12	2	17	0	5	17	0	2	21	0	6	24	0	

Testing Process: Baseline testing in all 14 schools took place over a period of about 1 month. Each team visited and tested the pupils in two schools. Every team had at least two members who were fluent in the local languages and the other members were at least conversant in the local languages. This was crucial for establishing rapport with the children and asking CBA questions that could be posed in the local language. Research team members who were less fluent in the vernacular carried out other data collection activities (e.g., recording of data, interviews with Circuit Supervisor and headteachers, etc.). Teams were in each school for approximately 1 week.

During the training, researchers discussed several test administration strategies. They considered dividing up the assessment tasks and having each tester administer a particular task or cluster of tasks to all of the children. This would mean that each tester would become expert in their task and only need the materials associated with the tasks they administered. Alternately they considered having a tester administer all tasks to a child before moving on to assess the next child. Based on their experiences, the researchers determined that it was best to have one person administer all the subtests to a pupil, and if possible, to do it at one or two sittings. This gave the tester a chance to establish rapport with the child and for the child to become more comfortable with the person doing the testing. This meant that each tester needed to become familiar with all of the tasks and needed to have a complete set of the testing materials but the end result was better data.

Recording and Handling Data: As soon as pupils were selected for inclusion in the sample, their names were listed on the data record forms that were described earlier and each pupil was assigned an ID number. Thereafter, this number followed the pupil on all subsequent data collection activities. As the researchers tested pupils, they recorded and scored pupil responses. All pupil scores were recorded on the Pupil Performance Class Summary Sheet prior to leaving the school. Processing the data in this way greatly facilitated data management.

Reviewing data for accuracy: When the teams returned from the field, data were organized and brought to the CRIQPEG office for vetting. One person was hired to check each set of materials as it was submitted and to follow-up on any missing pieces or misrecorded information. Also, because seven different teams were working independently, there was greater opportunity for inconsistencies and error. To monitor consistency, the Coordinator randomly selected files of data and reviewed them for accuracy in scoring and recording across teams.

5. Summarizing results

Hand Tabulation of Results: Shortly after the teams returned from collecting the CBA data, they summarized the findings by doing some simple hand tallies using the Pupil Performance Class Summary Sheet. They began by creating class profiles. These were then summarized to create school profiles and summary profiles for all schools in the IEQ project.

School Mpeasem #16 Teacher Brown

Classroom Summary
Reading

Baseline X
New Sample _____

Date 25 Feb. 1994 Level 5

	Non-Mastery		Partial Mastery		Full Mastery		Number	Number	Number
	Number	% of tested	Number	% of tested	Number	% of tested	tested	lost	in class
Letter Recognition	0	0	3	15	17	85	20	-	48
Concepts About Print	1	5	3	15	16	80	20	-	48
Aided Reading	3	15	3	15	14	70	20	-	48
Reading Most Used Words	5	25	2	10	13	65	20	-	48
Reading Passage from P2 Text	5	25	5	25	10	50	20	-	48
P2 Passage Comprehension	10	50	4	20	6	30	20	-	48
Reading Passage from P3 Text	5	25	5	25	10	50	20	-	48
P3 Passage Comprehension	12	60	4	20	4	20	20	-	48
Reading Passage from P4 Text	4	20	6	30	10	50	20	-	48
P4 Passage Comprehension	13	65	4	20	3	15	20	-	48
Reading Passage from P5 Text	6	30	6	30	8	40	20	-	48
P5 Passage Comprehension	15	75	5	25	0	0	20	-	48

Non-Mastery = Less than 30% correct

Partial Mastery = Between 30% and 69% correct

Full Mastery = More than 70% correct

School Mpeasem Primary #16

School Summary
Reading

Date 25 Feb. 1994

	Level 2--Percent Mastery			Level 3--Percent Mastery			Level 4--Percent Mastery			Level 5--Percent Mastery			
	Non	Partial	Full										
Letter Recognition										0	15	85	
Concepts About Print										5	15	80	
Aided Reading										15	15	70	
Reading Most Used Words										25	10	65	
Reading Passage from P2 Text										25	25	50	
P2 Passage Comprehension										50	20	30	
Reading Passage from P3 Text										25	25	50	
P3 Passage Comprehension											60	20	20
Reading Passage from P4 Text											20	30	50
P4 Passage Comprehension											65	20	15
Reading Passage from P5 Text											30	30	40
P5 Passage Comprehension											75	25	0

3/ Non-Mastery = Less than 30% correct
 Partial Mastery = Between 30% and 69% correct
 Full Mastery = More than 70% correct

Table 6

Baseline-February 1994
Reading

	Level 2--Percent Mastery			Level 3--Percent Mastery			Level 4--Percent Mastery			Level 5--Percent Mastery		
	Non	Partial	Full									
Letter Recognition	41	36	23	19	27	55	10	28	62	1	14	85
Concepts About Print	30	42	28	9	34	58	11	39	50	9	21	71
Aided Reading	88	8	3	60	22	19	51	17	32	37	23	40
Reading Most Used Words	96	4	0	70	16	14	59	20	21	43	23	34
Reading Passage from P2 Text	86	13	1	73	20	7	65	20	16	41	28	32
P2 Passage Comprehension	91	7	2	85	11	4	79	14	7	73	17	10
Reading Passage from P3 Text				69	19	12	61	22	17	43	24	34
P3 Passage Comprehension				92	7	1	83	11	5	83	13	5
Reading Passage from P4 Text							66	18	16	48	17	35
P4 Passage Comprehension							90	5	5	87	7	7
Reading Passage from P5 Text										51	25	24
P5 Passage Comprehension										85	11	4

Non-Mastery = Less than 30% correct
 Partial Mastery = Between 30% and 69% correct
 Full Mastery = More than 70% correct

N=1039

Tables 4-6 are sample completed profiles for CBA reading data.¹ Table 4 is a Classroom Profile that was constructed using the data from the Pupil Performance Class Summary Sheet in Table 3. Table 5 illustrates how the data from one Classroom Profile (Table 4) can be combined with similar data from other classrooms in the school in order to create a School Profile. Table 6 includes actual data from the 1994 baseline data collection. All of these tables were constructed quickly and easily using hand tallies of data found on the Pupil Performance Summary Sheets.

Computer Entry: The Pupil Performance Class Summary Sheet greatly facilitated computer key entry of the CBA data. A data layout scheme that corresponded directly with the layout of the Pupil Performance Summary Sheet was developed so that data could be entered into the computer in the exact order and form that it appeared on the sheets. Using this agreed upon data layout scheme, team members key entered the data for their schools into data files in the CRIQPEG computers and to backup diskettes. To prevent any of the problems or confusion that can occur when multiple people try to use one large data set, data for each school were entered and saved in unique files.

Quality Control: After the teams entered the data, it was necessary to concatenate the data files from the 14 schools into one data set and to verify that the data had been entered correctly. Visual inspection of a printed copy of the data was a useful strategy for identifying anomalies. Pupil records that were either longer or shorter than expected were reviewed as were records that had data in predesignated "blank" columns. These could be easily verified by referring back to the Pupil Performance Class Summary Sheet and correcting the computer file as needed. For example, in a few instances the number of columns that had been allotted on the data layout scheme was not sufficient (e.g., a few of the pupils had written more than 100 words but only 2 columns had been allotted for recording this score). Another way in which the data entry was verified was by obtaining descriptive statistics on all the variables (e.g., averages, highest and lowest scores, frequencies, etc.). Data were checked whenever pupil scores were found that were suspicious or outside of the possible range of scores. In one instance, an entire class had been entered with an incorrect grade level. Researchers noted that in that school one level was missing and another level had two times the expected number of pupils. This kind of error is very typical with a large data set and it illustrates how important it is to use multiple strategies to verify that the data are accurate before proceeding with other analyses.

Data Analysis: Initially, computer analyses were used to evaluate the comparability of the intensive and non-intensive schools. Although the schools had been selected as comparable, analysis of variance techniques were used to identify any significant differences on CBA baseline performance that could confound later results.

Another important part of the data analysis process was data reduction and the creation of composite scores. Pupils had several related scores in each of the areas that were tested (oral language, reading, and writing). For analysis purposes, it was useful to combine

some of these scores to create composite scores. For example, a level five pupil had "words per minute", "words read-percent correct", and "reading comprehension" scores on 4 reading passages (i.e., one passage each from levels 2, 3, 4, and 5). From these scores, three composite scores were created: "average words per minute", "average percent correct", and "average reading comprehension". These composite scores reduced the number of scores for each pupil from 12 to 3, facilitating greater ease in the handling of the data for later analyses. The composite scores were more reliable than scores from individual passages because they represented a larger sampling of pupil performance. Also, creating composite scores reduced the impact of minor fluctuations in the difficulty of individual passages across parallel forms. This was important since CRIQPEG used the three parallel forms to compare pupil performance at baseline and two subsequent follow-up testings.

The next phase of data analysis was to evaluate each of the research questions. CRIQPEG used the CBA data to describe the English language skill levels of Ghanaian primary school pupils, to gain insight into the factors affecting language learning, and to follow and compare the learning growth of pupils in the non-intensive and intensive intervention schools. Comparisons by type of community (urban, semi-urban, and rural) and gender provided useful insight into the unique needs of sub-groups of pupils. In some instances, the hand tallies were most useful--they were simple to prepare and easy to explain. In other instances, more formal statistical analyses were used to evaluate whether differences between groups (males versus females, intensive versus non-intensive) were large enough to be indicative of real differences and not just differences that were likely attributable to chance variations. In the sections that follow, some examples of the uses of the CBA findings are described.

6. Using results and providing feedback

Data generated by the CBA instruments were used in several ways to help improve the quality and effectiveness of the education process. CBA was 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.

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 or limiting the assessment at the point that the child fails, the examiner continues probing 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.

For example, Mary Appiah (not her real last name) 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 7 describes selected aspects of her performance at these two points in time.

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 correctly read words per minute; she read less than half of the words correctly. This amounts to about 1 correctly read word every 15 seconds--a very frustrating pace! 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 the textbooks 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.

Table 7: Individual Profile: Mary Appiah 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

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 speaking English, reading the passages in their textbooks, or writing letters--all 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 phrase, passage or letter. By having children demonstrate listening comprehension and spoken English, 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.

Improving Instruction Through Classroom Level Planning

This kind of CBA 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). Academic learning time occurs when the pupil is motivated by a task that is appropriately challenging. When instruction is too easy, pupils become bored and assignments are not taken seriously. When instruction is too difficult and pupils do not have the necessary prerequisite skills, they become frustrated and discouraged. Thus, it is critically important to identify what skills each pupil possesses and to use instruction to progressively build upon this foundation (Block, 1971; Hargis, 1987; Popham & Baker, 1970).

Table 8 is an example of a way of organizing CBA writing data so that it is easy for teachers to understand and to use to better match instruction to the learner. Notice that it clearly identifies what specific pupils can do and the kinds of instructional support they are likely to need. This kind of tool is useful as a starting point for discussions with teachers. It can lead to consideration of specific activities or techniques that might be helpful or to discussions about how to address the diverse instructional needs in their classrooms.

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.

Table 8

School _____
Teacher _____

Classroom Summary
Writing

Level _____
Date _____

1. List the names of children who were not able to copy letters. These children will need help learning to form letters.

2. List the names of children who were able to copy letters but are not able to write their names. These children can practice copying their names until they are able to write the name without help. They will need help learning to form letters and short words.

3. List the names of children who wrote their name but wrote less than 15 words. These children can begin to form short sentences with the words they know. As they learn new words, they can write new sentences using the words they know.

4. List the names of children who wrote 15 or more words. These children can learn new words and use them in sentences and short compositions.

IEQ Baseline: Level 5 Classroom Profile for Writing Words in 10 Minutes

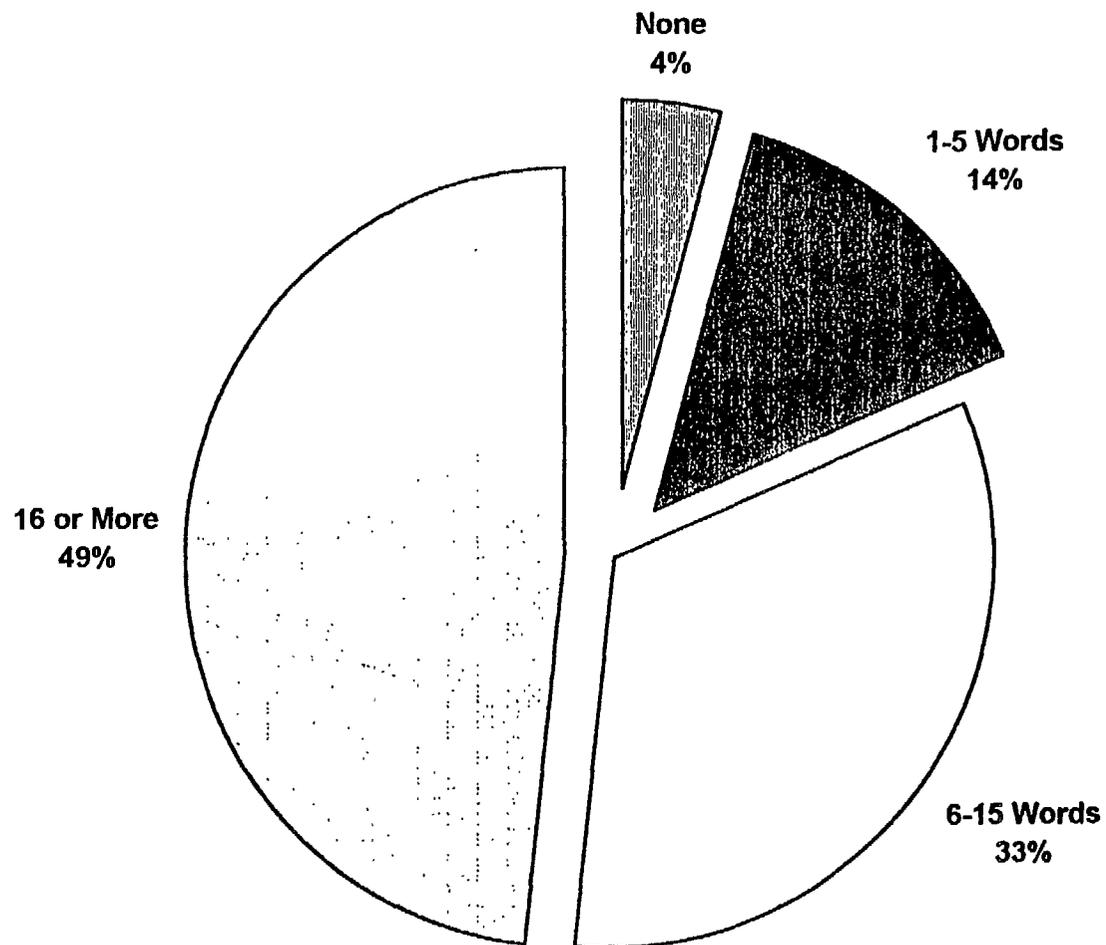


Figure 1

39



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

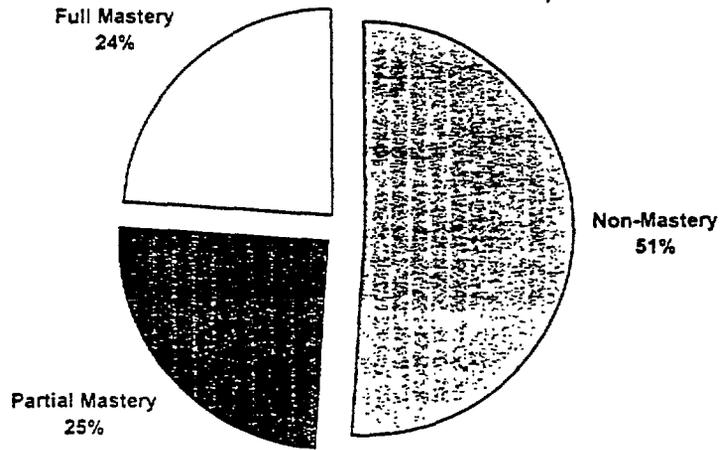
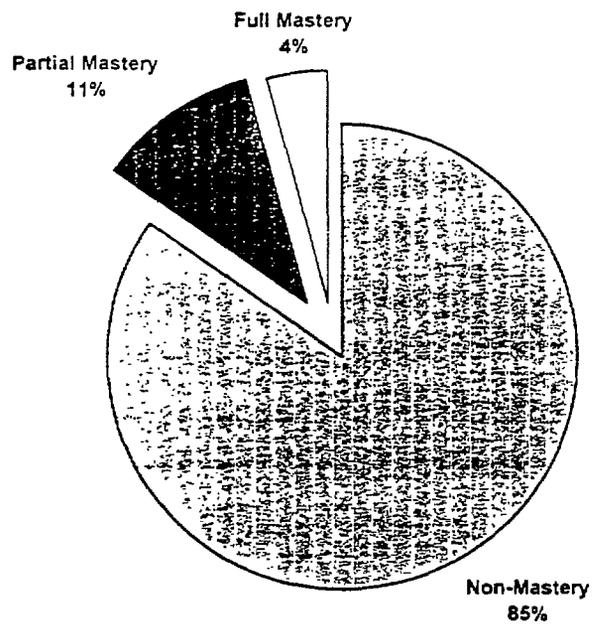


Figure 2

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



A. Harris
2/94 Ghana

Figure 3

Non-Mastery = Less than 30% Correct
Partial Mastery = 30-69 %
Full Mastery = More than 70% Correct

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.

Informing National Efforts to Improve Quality: Understanding Criterion Referenced Testing results

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. The charts we have just been reviewing are based on data mid year 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.

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) Frequent practice with oral language: This goal was a reaction to low performance on oral language and reading comprehension assessment tasks and was intended to build pupil vocabulary. Since most of the children in the participating IEQ classrooms were in middle and upper primary, this meant frequent practice with English.
- (2) 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. It was intended to encourage

teachers to use the textbooks more often, to label common classroom objects, and to use instructional aids such as flash cards, old newspapers and food labels in classroom activities.

(3) Using instructional strategies that help every pupil to become 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.

An important sign that the teachers were reflecting on and trying to address the third goal occurred during a feedback session after several months of collaboration: teachers in participating schools specifically requested assistance on managing classrooms with diverse achievement levels. Teacher and researcher comments and questions indicated that they (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.

Evaluating Progress and the Effectiveness of Interventions

An important use of CBA is program evaluation: Were 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 CBA baseline and interview findings with local educators, CRIQPEG returned to all 14 schools to collect another round of CBA data. Of the original baseline group of 1032, over 800 pupils were located and retested. Replacements for missing children were randomly selected from baseline class lists.

Figures 4 and 5 illustrate the August 1995 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% at baseline to 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%.

Another example of using CBA to evaluate learning progress and efforts at improvement can be seen by looking CBA performance on the writing words task. Figure 6 provides a bar graph summarizing baseline performance on writing words at each level tested for comparison and intervention schools. Note that at each level comparison schools performed slightly higher than intervention (intensive) schools. Figure 7 provides comparable data for August of 1995. Comparing these two figures illustrates 2 points.

The first is that both intensive and comparison schools performed better at the follow-up data collection. Perhaps this is not surprising since, 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 when comparing these figures is that at follow-up, pupils in the intensive schools have significantly greater writing vocabulary than pupils in the non-intensive/comparison schools. These are just a few examples of how the assessment process contributes to the improvement process by allowing us to monitor pupil performance changes.

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. These findings were pursued: 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? From findings from the baseline data, it appeared that there may be different patterns in the 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; most of their exposure is in school in the current textbook.

However, analyses using the recently collected 1995 data for upper primary children shows 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 reach a minimal level of literacy. Explorations into this language learning process continue. This is just one example of how IEQ has been using assessment to build knowledge within country as well as contribute to a knowledge base in the international development community.

August 95 Follow-up: Level 5 Class Profile for Decoding of Reading Passage from the Level 5 Text (Intervention Schools Only)

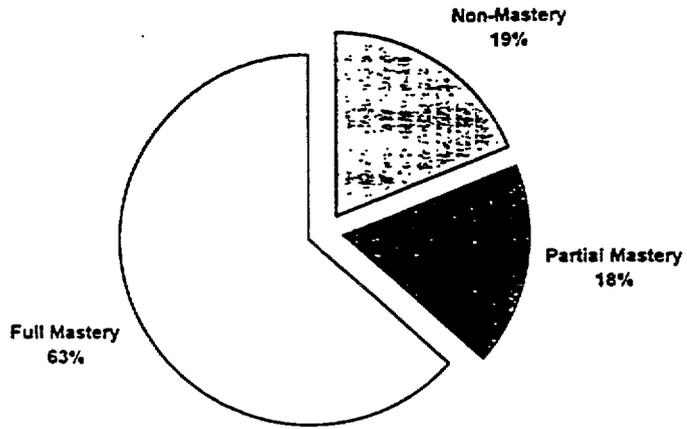
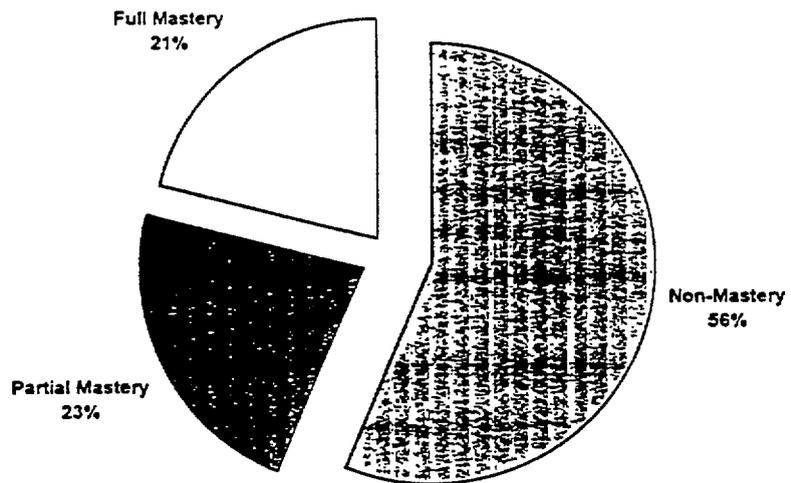


Figure 4

IEQ August 95 Follow-up: Level 5 Class Profile for Reading Comprehension of Passage from the Level 5 Text (Intervention Schools Only)



A. Harris
8/95 Ghana
wntwds5.xls

Figure 5

Non-Mastery = Less than 30% Correct
Partial Mastery = 30-69% Correct
Full Mastery = More than 70% Correct

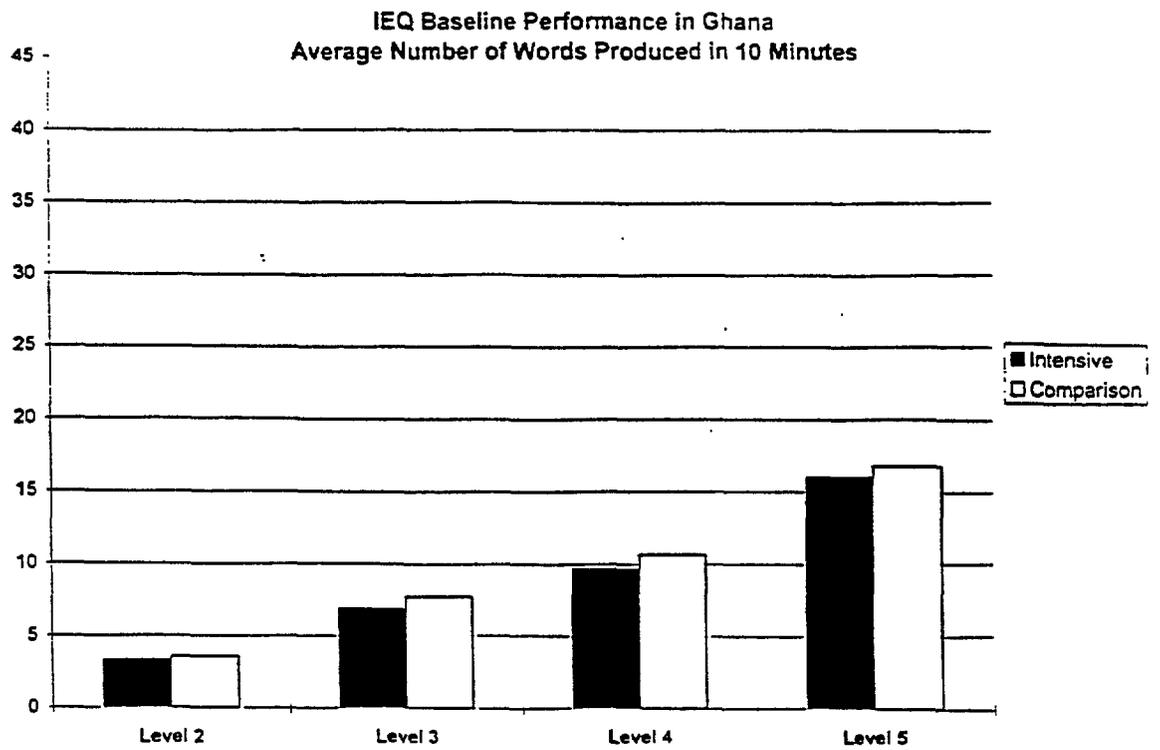


Figure 6

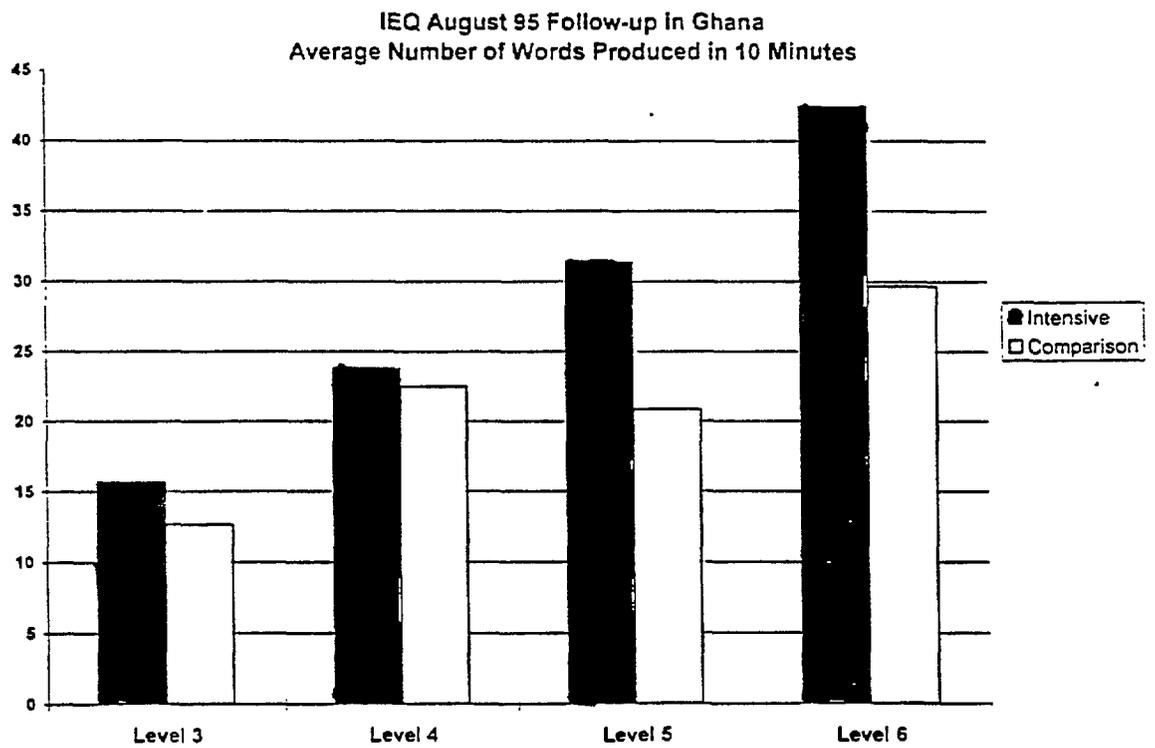


Figure 7

7. Reflections for Future Research

* **CBA instruments should lead to data that have meaning to those who will use it.** CRIQPEG found that teachers were more willing to believe test information because it was collected on their own pupils. Sharing information about the performance of children in a teacher's classroom and within the school was an important tool for motivating teachers, head teachers, and Circuit Supervisors to become involved in efforts toward improvement. It was crucial for the information to be easily understood and translated into skills that the educators could understand. Abstract scores were less useful than being able to report how many children could write their names, respond to questions such as "What is your name?", and read passages from the textbook. Also, these same kinds of illustrations were important for communicating with policy makers and national leaders.

* **Developing Good Instruments Takes Time and Revision:** Often it may take several steps to get to the final instrument: For example, the first time the reading instruments were administered there were manuals for each level. The second time, it was possible to combine the manuals into one. This saved paper and was more efficient. Also, as was mentioned earlier, the revised oral language test booklet was very long. This allowed the test administrators to note problem questions and responses that were difficult to score. This information will be incorporated into the manual and the next time it is used, the results can be recorded on an answer sheet rather than a 40 page test booklet that has to be duplicated for each child. This will save paper, take less time to prepare, and take less time to record on the data summary sheets.

* **Most Used Words Task needs to be better understood and perhaps refined.** Most Used Words lists for different levels are very similar, with many of the same words appearing on the lists from each level. This is not surprising since the 60 most commonly used words are likely to be similar across levels. Some test developers have suggested picking words at random instead. While this is likely to produce progressively harder words at the higher grade levels, it doesn't address the intent of the measure: to determine if children are familiar with the words they encounter most often in their texts. Also, children were given the list from their current year only. This made comparisons from one year to the next difficult. In the future, it may be preferable to identify one list of the most frequently used words in the primary texts and use this list to create one set of parallel forms instead of using unique sets for each level.

In addition, it may be useful to review the words and consider their distribution in terms of parts of speech (verbs, nouns, pronouns, and so on). Typically the words that occur most frequently are function words such as "a" "and" and "the", pronouns such as "she" "we" "you" and "they" and common verbs such as "is" "are" "was" and "said". Vocabulary acquisition when learning a foreign language is likely to be depend on some combination of frequency of use and ease of remembering. Content words or words that have a direct meaning in the mother tongue may be easier to remember initially. Whereas words that are

used frequently in oral and written language have the advantage of repetition. It is these words that appear most often on writing samples. The purpose of this task is to assess whether the child has mastery of the most frequently used words so that he or she can focus on new or less frequently used words. Therefore, this task should not be construed as a measure of vocabulary.

*** Future CBA efforts may want to assess oral proficiency in the local vernacular as a readiness skill for oral English.** CRIQPEG didn't address the issue of language learning in the mother tongue. Pre-requisite to language learning in a foreign language, is adequate language development in one's mother tongue. While CRIQPEG encouraged active interaction between teachers and pupils (including encouraging lower primary teachers to converse with pupils in Ghanaian and English), the CBA instruments did not assess pupil language readiness skills in the mother tongue. There were political and practical reasons for this decision. Politically, the policy makers and researchers felt strongly that the focus should be on English. From a practical standpoint, in many instances, the teachers in Ghana are not fluent in the local mother tongue (e.g., only 43% of the teachers in participating IEQ schools reported that they were fluent in the local language). Further, it was typical in the IEQ schools for there to be children with different mother tongues and the few written Ghanaian languages often don't match the local languages or dialects. Nonetheless, it may be useful in future CBA efforts to assess pupil oral language proficiency in the mother tongue as a readiness skill for oral language learning in English. Also, in instances in which the local Ghanaian language has been transcribed, investigating the utility of developing reading proficiency in the mother tongue as a prerequisite or facilitative factor in developing English reading skills should be considered.

Summary

Exciting changes continue to happen in Ghana. IEQ findings from August 1996 reinforce earlier findings supporting the value of using classroom research to improve policy and practice. Recent CRT results obtained from the IEQ schools corroborate the IEQ/CBA findings. CRIQPEG has become a valued partner in national educational reform efforts. All of these changes provide examples of the strength of the IEQ assessment => assimilation => action process.

References and Resource List

- Blankenship, C. (1985). Using curriculum-based assessment data to make instructional decisions. Exceptional Children, 52, 233-238.
- Block, J. H. (Ed.). (1971). Mastery learning: Theory and practice. New York: Holt, Rinehart, & Winston.
- Carver, R. P. (1990). Reading rate: A review of research and theory. San Diego: Harcourt Brace Javonovich.
- Carver, R. P. (1982). Optimal rate of reading prose. Reading Research Quarterly, 18, 56-88.
- Centre for Research on Improving Quality of Primary Education in Ghana (CRIQPEG). (1993). Improving Educational Quality in Ghana--Phase I Report. Arlington, VA: Institute for International Research.
- Clay, M. M. (1972). Reading: The patterning of complex behavior. Portsmouth, NH: Heinemann
- Clay, M. M. (1979a). Sand--The Concepts About Print Test. Portsmouth, NH: Heinemann.
- Clay, M. M. (1979b). Stones--The Concepts About Print Test. Portsmouth, NH: Heinemann.
- Clay, M. M. (1979c). The early detection of reading difficulties. Portsmouth, NH: Heinemann.
- Cohen, L. G., & Spruill, J. A. (1990). A practical guide to curriculum-based assesment for special educators. Springfield, IL: Charles C. Thomas.
- Cummings, W. K. (1986). Low-cost primary education: Implementing an innovation in six nations. Ottawa, Canada: International Development Research Centre.
- Denham, C., & Lieberman, A. (Eds.). (1980). Time to learn. Washington, DC: National Institute of Education.
- Deno, S. L. (1985). Curriculum-based measurement: The emerging alternative. Exceptional Children, 52, 219-232.
- Deno, S. L. (1986). Formative evaluation of individual student programs: A new role for school psychologists. School Psychology Review, 15, 358-374.

- Deno, S. L. (1987). Curriculum-based measurement. Teaching Exceptional Children, 20, 41.
- Flores, P. (1981). Educational innovation in the Philippines: A case study of Project IMPACT. Ottawa, Canada: International Development Research Centre.
- Fuchs, L. S., & Deno, S. L. (1991). Paradigmatic distinctions between instructionally relevant measurement models. Exceptional Children, 57, 488-500.
- Fuchs, L. S., Deno, S. L., & Mirkin, P. K. (1984). The effects of frequent curriculum-based measurement and evaluation on pedagogy, student achievement and student awareness of learning. American Educational Research Journal, 21, 449-460.
- Fuchs, L. S., & Fuchs, D. (1986). Linking assessment to instructional intervention: An overview. School Psychology Review, 15, 318-323.
- Fuchs, L. S., Fuchs, D., & Deno, S. L. (1982). Reliability and validity of curriculum-based informal reading inventories. Reading Research Quarterly, 18, 6-26.
- Fuchs, L. S., Fuchs, D., & Deno, S. L. (1985). The importance of goal ambitiousness and goal mastery to student achievement. Exceptional Children, 52, 63-71.
- Fuchs, L. S., Fuchs, D., Hamlett, C. L., Walz, L., & Germann, G. (1993). Formative evaluation of academic progress: How much growth can we expect? School Psychology Review, 22, 27-48.
- Fuller, B. (1987). What school factors raise achievement in the Third World? Review of Educational Research, 57, 255-292.
- Fuller, B., & Clarke, P. (1994). Raising school effects while ignoring culture? Local conditions and the influence of classroom tools, rules, and pedagogy. Review of Educational Research, 64, 119-157.
- Gickling, E. E., & Armstrong, D. L. (1978). Levels of instructional difficulty as related to on-task behavior, task completion, and comprehension. Journal of Learning Disabilities, 11, 32-39.
- Gickling, E. E., & Havertape, J. F. (1981). Curriculum-based assessment. In J. A. Tucker (Ed.), Non-test based assessment. Minneapolis: The National School Psychology Inservice Training Network, University of Minnesota.
- Good, R. H., & Salvia, J. (1988). Curriculum bias in published norm-referenced reading tests: Demonstrable effects. School Psychology Review, 17, 51-60.

- Good, R. H., & Shinn, M. R. (1990). Forecasting accuracy of slope estimates for reading curriculum-based measurement. Empirical evidence. Behavioral Assessment, 12, 179-193.
- Graden, J. L., Zins, J. E., & Curtis, M. J. (Eds.). (1988). Alternative educational delivery systems: Enhancing instructional options for all students. Washington, DC: National Association of School Psychologists.
- Hall, E. (1983). Ghanaian languages. Accra, Ghana: Asempa Publishers.
- Hargis, C. H. (1987). Curriculum based assessment: A primer. Springfield, IL: Charles C Thomas.
- Harris, A. J., & Sipay, E. R. (1990). How to increase reading ability. (Ninth edition). New York: Longman.
- Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1986). The treatment utility of assessment: A functional approach to evaluating assessment quality. American Psychologist, 42, 963-974.
- Howell, K. (1986). Direct assessment of academic performance. School Psychology Review, 15, 324-335.
- Idol, L., Nevis, A., & Paolucci-Whitcomb, P. (1986). Models of curriculum-based assessment. Rockville, MD: Aspen Publishers.
- International Development Research Centre. (1981). Teaching yourself in primary school: Report of a seminar on self-instructional programs. Ottawa, Canada: Author.
- Kameenui, E. J., & Simmons, D. C. (1990). Designing instructional strategies: The prevention of academic learning problems. Columbus, OH: Merrill.
- Kelly, E. (1984). Evaluation of the Improved Efficiency of Learning Project. Albany, NY: Improved Efficiency of Learning Project/USAID.
- King-Sears, M. E (1994). CBA in special education. San Diego, CA: Singular Publishing Group.
- Klees, S. J., & Suparman, M. R. (1984). An evaluation of the costs of PAMONG schooling alternatives in Indonesia. McLean, VA: Project PAMONG/USAID.
- Levin, H. M., & Lockheed, M. E. (Eds.). (1993). Effective schools in developing countries. Bristol, PA: Falmer.

- Linn, R. L., Baker, E. L., & Dunbar, S. B. (1991). Complex, performance-based assessment: Expectations and validation criteria. Educational Researcher, 20, 15-21.
- Marston, D. (1988). The effectiveness of special education: A time-series analysis of reading performance in regular and special education classrooms. The Journal of Special Education, 21, 13-26.
- Marston, D. B. (1989). A curriculum-based measurement approach to assessing academic performance: What it is and why do it. In M. R. Shinn (Ed.), Curriculum-based measurement: Assessing special children (pp. 18-78). New York: Guilford.
- Marston, D., & Magnusson, D. (1985). Implementing curriculum-based measurement in special and regular education settings. Exceptional Children, 52, 266-276.
- McMaster, J. C. (1978). Cost-effectiveness analysis of Project IMPACT for the Philippines. Quezon City, Philippines: Regional Center for Innovation and Technology (INNOTECH).
- Moss, P. A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. Review of Educational Research, 62, 229-258.
- Nichols, D. (1980). Low-cost learning systems. Bethesda, MD: Institute for International Research.
- Pasigna, A. L. (1991). A post hoc analysis of Thailand's Reduced Instructional Time Project. Bloomington, IN: Author.
- Pasigna, A. L. (1993, October). Swaziland's Continuous Assessment (CA) Programme: Some questions and answers. Manzini, Swaziland: Ministry of Education and Arlington, VA: Institute for International Research.
- Popham, W. J., & Baker, E. L. (1970). Establishing instructional goals. Englewood Cliffs, NJ: Prentice-Hall.
- Powell, W. (1971). Validity of the I. R. I. reading levels. Elementary English, 48, 637-642.
- Project RIT. (1984). Project RIT (Reduced Instructional Time). Lopburi, Thailand: Muang Mai School, Ministry of Education.
- Rogers, V. (1989). Assessing the curriculum experienced by children. Phi Delta Kappan, 70, 714-717.

- Rosenfield, S., & Rubinson, F. (1985). Introducing curriculum-based assessment through consultation. Exceptional Children, 52, 282-287.
- Shepard, L. A. (1991). Psychometricians' beliefs about learning. Educational Researcher, 20, 2-16.
- Shepard, L. A. (1993). The place of testing reform in educational reform: A reply to Cizek. Educational Researcher, 22, 10-13.
- Shinn, M. R. (1988). Development of curriculum-based local norms for use in special education decision making. School Psychology Review, 17, 61-80.
- Shinn, M. R. (Ed.). (1989). Curriculum-based measurement: Assessing special children. New York: Guildford Press.
- Shinn, M. R. (Ed.). (1990). Curriculum-based measurement. CBM: Training modules. Eugene: OR: University of Oregon. (US Department of Education Grant No. H02980051).
- Shinn, M. R., & Good, R. H. (1992). CBA: An assessment of its current status and a prognosis for its future. In J. Kramer (Ed.). Curriculum-based assessment: Examining old problems, exploring new solutions. Lincoln, NE: Buros Institute of Mental Measurement, University of Nebraska.
- Shinn, M. R., Good, R. H., & Stein, S. (1989). Summarizing trends in student achievement: A comparison of evaluative models, School Psychology Review, 18, 356-370.
- Tindal, G. A., Fuchs, L. S., Fuchs, D., Shinn, M. R., Deno, S. L., & Germann, G. (1985). Empirical validation of criterion referenced tests. Journal of Educational Research, 78, 203-209.
- Tompkins, G. E., & Hoskisson, K. (1995). Language arts: Content and teaching strategies. Englewood Cliffs, NJ: Prentice-Hall.
- Windham, D. M. (1990). Economics, instructional development, and the enhancement of educational efficiency. In D. W. Chapman & C. Carrier (Eds.). Improving educational quality: A global perspective. New York: Greenwood Press.

Footnote

¹ Names and scores of children and schools provided in all tables have been changes to protect research participants. Only Table 6 (and all Figures) provides actual data. Other tables provide data that is representative of actual data but the scores have been modified to prevent readers from identifying the actual child, classroom, or school.