

Improving Educational Quality (IEQ) II Project

**DEMONSTRATING IMPACT ON THE GROUND
THROUGH ASSESSMENT OF LEARNING**

IEQII undertaken by:

American Institutes for Research

in collaboration with

The Academy for Educational Development

Education Development Center, Inc.

Juárez and Associates, Inc.

The University of Pittsburgh

Report prepared by:

Abigail Harris

January 2001

Table of Contents

INTRODUCTION	1
THE ROLE OF ASSESSMENT IN IEQ – GHANA	6
Background	6
CBA in IEQ – Ghana	7
Instructional Value	12
Monitor Impact--Regular Monitoring of Growth	12
CBA and the IEQ Process in Ghana	18
CBA AS A TOOL FOR USE IN DEVELOPING COUNTRIES	20
RECOMMENDATIONS	22
NOTES	23

Introduction

This manuscript describes a “rhythm of reform”--a cycle of assessment-assimilation-action occurring throughout the life of a reform effort--designed to assess and enhance the impact of interventions within the reform. Although multiple sources and kinds of assessment are useful for analyzing and enhancing impact, the focus of this paper is to demonstrate that direct assessment of student achievement is crucial because of its power to guide and motivate action.

Measured change in student learning is the most direct indicator of the effects of educational investments and the success of efforts to improve educational quality. Yet often success is estimated through proxies--counting the number of teachers who have received training, documenting the number of books that have been distributed, recording student enrollment and attendance--all indicators that money has been spent as planned but not whether the goal of the spending was achieved. Even in those projects where student achievement is assessed, it is often an activity that is external to project decision making and typically it occurs when it is too late in the project to take full advantage of the role that achievement information can play in the improvement process.

At least two kinds of interests are served by including on-going direct assessment of student learning as an integral part of educational projects. The first is the obvious: monitoring changes in student learning in order to evaluate program success as well as identify statistical relationships between inputs and improvements. The availability of valid outcome data on student learning allows policy makers and donor agencies to estimate and compare the relative contribution of different investments to desired outcomes (in this case, improvement in student learning). A useful by-product is that the early and continued attention to learning outcomes provides a reminder to all involved that the target is student learning.

Another powerful role that direct assessment can play is through the instructional process. Faced with timely, actual achievement data on what children can and cannot do--relevant to the local curriculum and stated in meaningful terms--educators (policy makers, curriculum developers, teacher trainers, district officers, and teachers) and other stakeholders have something concrete with which to work. These data serve as a powerful stimuli for discussion, problem solving, and, in many instances, corrective action.

An example illustrates this point. One of the simplest forms of direct assessment is to ask students to write their names and then to write as many words as they can in 10 minutes. They can be asked in the vernacular and words can be in any language as long as the words are spelled correctly and the child can tell you what the word is.

Figure 1.

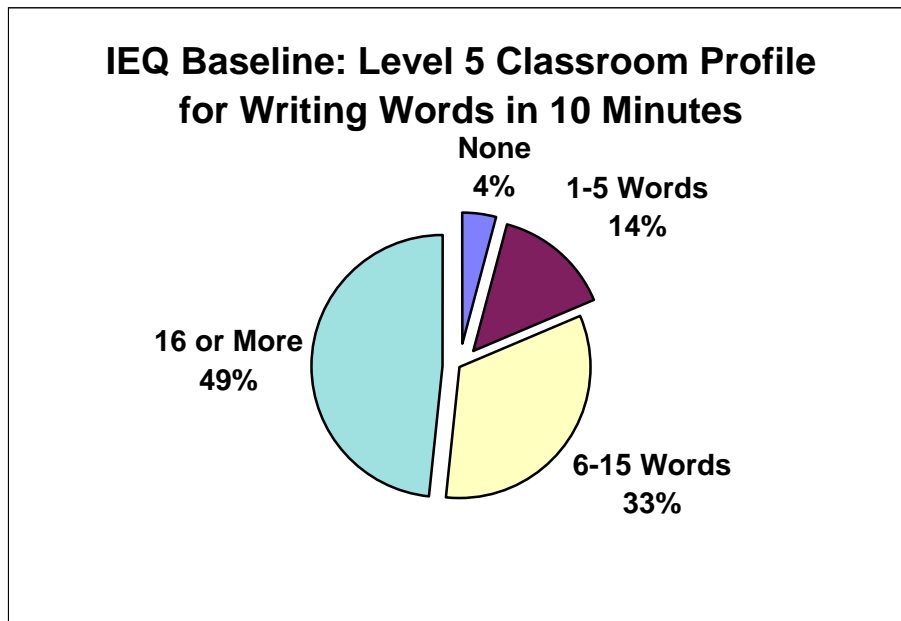


Figure 1 represents actual baseline data from a sample of fifth grade classrooms in Ghana. What would you know about the instructional needs in a 5th grade classroom with the profile illustrated in this pie chart?

The data indicate that 4% of the pupils are unable to write any words-not even their names and 14% are able to write between 1-5 words (typically their names and perhaps the name of their teacher, school or friend). About 49% of the students are able to write 16 or more words, with some writing over 100 words.

Translated into instructional terms, the teacher has at least 20% of the students in her 5th grade class who are pre-literate and another third of her students have some beginning skills. About 50% of the 5th grade pupils have some written language skills. The teacher needs to have instructional skills and materials to teach at several levels and needs to be able to manage a multi-ability classroom. Often teachers in these circumstances make a choice to focus on the 50% with some skills, leaving the others idle. Focusing on the most able pupils tends to begin early and provides one explanation for the 4% of pupils in 5th grade who are unable to even write their names.

Consider the implications if these data represent a composite profile for fifth graders from a sample of public primary schools in two of the more developed regions of the country. The national syllabus indicates that in grade 5, “Pupils should be able to write short paragraphs on given topics,” as well as, “Write letters to friends and relations.” Consider as well that the 5th grade pupils who were tested were

less than 6 months away from taking the national criterion referenced test (CRT), a multiple choice exam with 100 questions each in English and Mathematics based on the national syllabi.

From an educational policy standpoint it becomes clear that the syllabus is pitched at a level far above the current skills of at least 50% of students. It should be no surprise when 50% or more of the pupils score at or below chance level on the CRT. These findings coupled with findings from other similar kinds of exercises (for example, percent of words read correctly from a textbook passage) provide the stimulus for rich discussion. Should the syllabi be revised? Should scarce resources be redirected to interventions for upper primary school non-readers (with the aim of bringing all citizenry to at least a minimum level of literacy)? What strategies and materials are needed for successful remediation? Perhaps the promotion and repetition policies should be re-examined? Meaningful findings from classroom level assessment lead to meaningful exchanges and more informed decision-making.

If assessment is so useful, why isn't it carried out more often? A variety of explanations or excuses are given for not assessing student learning:

Availability of Tests:

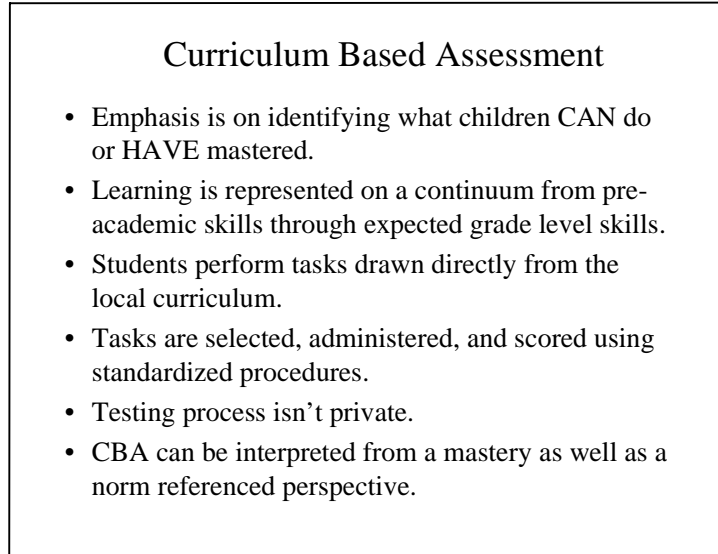
- No available tests designed for this population,
- Even if tests were available, it is unlikely that they reflect the current local curriculum,
- If we get someone to construct tests, what do the scores mean?,
- Doesn't it take a long time to develop good tests?

Who would conduct it?

- Locals aren't trained,
- It would require too much effort and organization.

These issues are not trivial but the task of assessment can become more manageable. Recent research and experience in several developing countries as well as in the US has highlighted the utility of an approach that focuses on direct assessment of student performance on tasks from the local curriculum. The approach is called curriculum-based assessment (CBA) or direct assessment and it includes some forms of performance assessment and continuous assessment. For clarity and ease of discussion, the approach will be referred to as CBA.

Figure 2.



As is evident from the above example of direct assessment of writing, CBA is not the typical set of multiple-choice questions that you find on most norm-referenced or criterion-referenced tests. Figure 2 provides a list of its characteristics.

In CBA, the emphasis is on identifying what children can do and what skills they have mastered. Many forms of traditional assessment cover a restricted range of knowledge and skills. They focus on grade level objectives and students are assumed to have mastered prerequisite skills. When students perform poorly, the failing scores indicate that the students have not achieved at the expected level but there is rarely diagnostic information about student mastery of prerequisite skills. In CBA, student learning is represented on a continuum from pre-academic skills through expected grade level skills. Sometimes goal level skills (skills associated with higher grades or levels) are included as well. Students begin with the easier tasks and continue along the continuum until their performance indicates non-mastery of subsequent skills on the continuum. Every student can be located on this continuum and their learning can be monitored.

In CBA, students perform tasks that have been drawn directly from the local curriculum. The syllabi (sometimes called “Table of Instructional Objectives” or “Scope and Sequence”) are used to define the continuum and then reading passages, math problems, writing assignments, and so on are selected from the curriculum to correspond to the national standards. The specific tasks are selected, administered, and scored using standardized procedures. Often tasks are pilot tested to create multiple, equivalent forms of the assessment “test.” Procedures for developing and using CBA are readily available.

A unique feature of CBA is that the assessment process isn't private. Teachers and other educators are encouraged to share in the process. The best way for a teacher to "teach to the test" is to teach the curriculum. When teachers realize that all students will be expected to write their names, read from the textbooks, and so on, it is not uncommon to see teachers asking students to work in pairs or small groups to practice these skills.

One other strength of CBA is that results can be interpreted from a mastery as well as a norm referenced perspective. Students can be described in terms of percent mastery such as was done above. In fact, student performance on some kinds of tasks (such as reading rate) can be compared to student performance internationally. In addition, local and national norms for performance can be developed and individual, classroom, or grade level data can be compared and monitored over time.

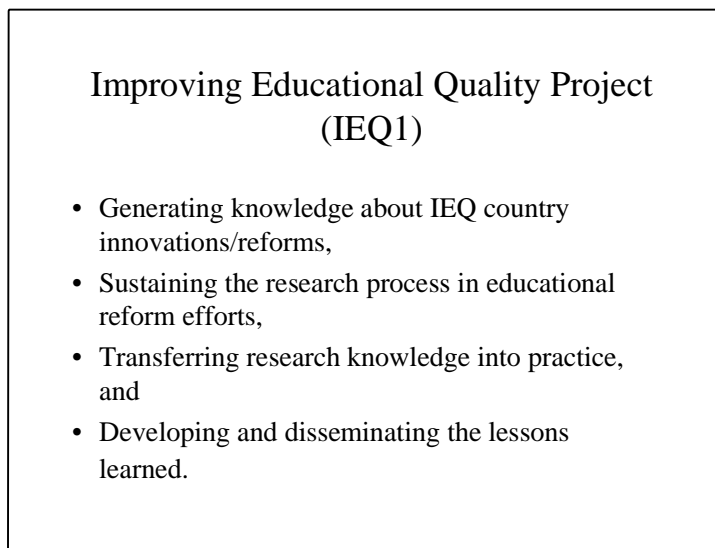
What follows is a description of this process in action. During IEQ I, CBA was used throughout the project for both monitoring and instructional purposes.

The Role of Assessment in IEQ - Ghana

BACKGROUND

In 1991, the Improving Educational Quality Project (IEQ)ⁱ was launched with the purpose of conducting classroom level research to guide the generation and adaptation of innovations that hold promise for improving the quality of primary education. To do this, IEQ formed partnerships with host country colleagues and collaborated with them in efforts to learn about and improve the school and classroom experiences of educators and pupils.

Figure 3.



At about the same time the Government of Ghana with support from USAID launched an initiative to strengthen the policy and institutional framework required to insure a quality, effective, equitable and financially sustainable primary education system in Ghana by the year 2000 (Executive Summary, Ghana Primary Education Program (PREP), September 10, 1992). Major activities included (a) distributing textbooks and guides to schools in Ghana for 1.8 million school children, (b) organizing a comprehensive in-service training program for approximately 30,000 primary teachers, (c) developing end of primary cycle criterion-referenced testing (CRT) for grade 6 pupils in English and Mathematics, and (d) preparing and implementing an Equity Improvement Plan. Although strong support for primary education had been available in the schools in the 1960's and early 1970's, in the recent past, it had become necessary to hire many minimally prepared or untrained teachers, educational materials were scarce, and teachers were left to their own devices in their efforts to help children acquire basic

literacy and numeracy (King, Glewwe, & Alberts, 1992; Ministry of Education, 1994). For example, in 1990-91, base-year figures showed that only 10% of pupils in primary schools had the basic textbooks (Kraft, 1994). One goal of the Ghana/USAID initiative was to reverse this trend and put sufficient textbooks into the schools for every child to have textbooks in English, Mathematics, Science, and Social Studies. Another goal was to provide training to teachers in instructional content and processes in each of the subject areas.

There was a natural fit between the primary school initiative and the goals of IEQ. Through a partnership between IEQ and the University of Cape Coast, the Centre for Research on Improving Quality of Primary Education in Ghana (CRIQPEG)ⁱⁱ was formed to conduct research designed to support and complement PREP innovations. Over a four year period, researchers followed a cycle of assessment-assimilation-action designed to assess factors affecting learning, share findings with stakeholders, collaborate with stakeholders on improvements, re-assess, and so on. CBA--direct assessment of student learning--played a central role in this cycle.

CBA IN IEQ-GHANA

To begin the assessment process, 14 schools representing urban, rural, and semi-urban, several language groups, two regions, and including equity schools, were selected. Seven of the schools were designated as intensive intervention and seven as comparison/non- intensive intervention. Table 3 compares the intervention activities for intensive and comparison/non-intensive schools.

In retrospect, it would have been useful to identify some true “control” schools as well. As is evident from Table 1, research activities in the comparison schools likely created a variant of the Hawthorne effect, hence the labeling “non-intensive” intervention schools.

In order to assess pupils skill levels, curriculum-based assessment instruments were developed for grades 2-6 in oral language (functional expressions, listening comprehension and speaking), reading (pre-reading, decoding, and comprehension), and writing (copying, writing words, dictation, and written expression). Faculty from the University of Cape Coast collaborated with consultants and CBA tasks were developed, pilot tested, and three parallel forms were created for grade levels 2-6. For in-depth information on the development and use of the assessment instruments, see Harris, Okyere, Pasiona, & Schubert (1997). Table 2 provides a brief description of the instruments, sample questions, and sample interpretations of the scores.

Table 1. Overview of intervention activities for intensive intervention and comparison schools.

	Intensive Intervention Schools	Comparison/Non-intensive Intervention Schools
Instructional Materials	Supply of English and Mathematics textbooks for 1:1 ratio in most classrooms; Periodic re-supply.	Supply of English and Mathematics textbooks for 1:1 ratio in most classrooms; Periodic re-supply.
	Supplementary materials such as labels, puppets, sample instructional posters, sample flash cards	
Assessment/ Data Collection Activities	Interviews provided opportunities for reflection on the instructional process and factors affecting language learning. Involved administrators, teachers, community, and pupils.	Interviews provided opportunities for reflection on the instructional process and factors affecting language learning. Involved administrators, teachers, community, and pupils.
	Pupil Assessment and Teacher Ratings of pupil skill level offered teachers and pupils opportunities to consider pupil assessment and learning.	Pupil Assessment and Teacher Ratings of pupil skill level offered teachers and pupils opportunities to consider pupil assessment and learning.
	Classroom and pupil (high and low achieving boys and girls) observations heightened teachers' awareness of instructional process.	Classroom and pupil (high and low achieving girls and boys) observations heightened teachers' awareness of instructional process.
Professional Development	Circuit Supervisors, Headteachers, and CRIQPEG participated in professional development seminars focusing on strategies for (1) increasing oral practice, (2) increasing exposure to print, and (3) teaching to make every pupil a successful learner. Follow-up seminars focused on managing a multi-ability classroom.	
	Circuit Supervisors, Headteachers, and CRIQPEG led on-site workshops for teachers (6/94, 11/94)	
On-going Instructional Support	Headteachers and CRIQPEG observed classrooms and provided support for teachers.	
Collaboration	Circuit Supervisors, Headteachers, and teachers participated in periodic feedback sessions providing insights on improving instruction.	

Table 2.

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.

Table 2 (continued).

TASK	LEVELS	WHAT IT MEASURES	SAMPLE QUESTIONS	SCORING	SAMPLE WAYS SCORES CAN BE INTERPRETTED
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.
WRITING					
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.

Table 2 (continued).

TASK	LEVELS	WHAT IT MEASURES	SAMPLE QUESTIONS	SCORING	SAMPLE WAYS SCORES CAN BE INTERPRETTED
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.

INSTRUCTIONAL VALUE

Table 3 provides profiles of student performance at each grade level. When profiles such as these were shared with educators, they provided a stimulus for discussion and problem solving. In this way, the results served as a motivator. The Minister of Education at the time promised his support. Circuit Supervisors and headteachers willingly participated with CRIQPEG in intervention workshops; they then joined in leading the school level training sessions. At each level, actual CBA results were the stimulus for shared problem solving, decisions about what instructional materials were needed, and decisions about what additional training was needed.

Worth emphasizing about these profiles is that they clearly state what children CAN do. Tasks were drawn from Ghanaian instructional materials. In addition, the results were presented by Ghanaians. Results presented by “outsiders,” using non-local tests, emphasizing what students can not do or how poorly students measure up to expectations in the syllabus, would likely have been received with defensiveness and resistance.

MONITOR IMPACT--REGULAR MONITORING OF GROWTH

CBA is a useful tool for monitoring student learning and evaluating program impact. In IEQ Ghana, results of longitudinal data on approximately 2000 pupils were used to compare all 14 participating schools at the beginning of the project and then to compare findings during the implementation process. Selected findings from this research are presented below.

FINDING: WITH A FEW EXCEPTIONS, PUPIL PERFORMANCE IN INTENSIVE AND NON-INTENSIVE INTERVENTION SCHOOLS WAS COMPARABLE AT BASELINE.

Using the baseline data, analyses of variance were computed to determine whether there were significant differences between the Intensive and Non-intensive schools prior to the implementation of the interventions. Table 4 provides a summary of the results of these analyses.

In grade levels 2, 3, and 4, most comparisons were non-significant and the few significant differences were equally distributed such that half indicated a significant difference favoring the Intensive schools and half indicated a significant difference favoring the Non-intensive or comparison schools. At grade 5, there were 15 non-significant comparisons (excluding composite scores), one significant comparison favoring the Non-intensive schools, and six significant comparisons favoring the Intensive schools.

With consultant support, data were collected and analyzed by teams of local educators from the University of Cape Coast. Within 3 months of data collection, results were disseminated by this group to the Ministry, Circuit Supervisors, Headteachers, and teachers.

Table 3. Pupil profiles derived from 1994 baseline data collection.

Grade Level	Oral Language	Reading	Writing
<p>Primary Grade 2 (P2) (n=261)</p>	<p>The typical pupil in grade 2 is able to respond correctly to very basic questions like “What is your name?” and “How old are you?” (Functional English)</p> <p>The typical pupil is able to follow a few simple oral directions spoken in English.</p> <p>The typical pupil is unable to communicate in simple conversational English or respond to simple oral questions using words and pictures from the P2 Pupils English textbook.</p>	<p>The typical pupil recognizes some of the letters of the alphabet but can name less than half of the upper and lower case letters.</p> <p>The typical pupil shows some familiarity with printed material (for example, responds correctly when asked in the vernacular to show the front of a P2 textbook) but is unable to point to where to begin reading or to follow along as someone else reads.</p> <p>With rare exceptions, grade 2 pupils are unable to recognize or read the words in the grade 2 English text.</p>	<p>Most grade 2 pupils could copy letters in the alphabet.</p> <p>About half of the pupils could write their names. About half could not write their names or any other words.</p> <p>About 1/4 of the pupils could write 6 or more words.</p>
<p>Primary Grade 3 (P3) (n=259)</p>	<p>The typical pupil in grade 3 is able to respond correctly to very basic questions like “What is your name?” and “How old are you?”</p> <p>The typical pupil performs slightly better than the grade 2 pupil in listening comprehension (i.e., following simple oral directions spoken in English.)</p> <p>The typical pupil is unable to communicate in simple conversational English or respond to simple oral questions using words and pictures from the P3 Pupils English textbook.</p>	<p>The typical pupil is able to recognize most or all of the letters of the alphabet.</p> <p>Most of the pupils show some familiarity with printed materials and over half show signs of beginning literacy.</p> <p>While most pupils can’t recognize or read the most commonly used words in their textbooks, about 1/3 have at least some success at this task.</p> <p>With rare exceptions, pupils can’t read or comprehend passages taken from the English textbooks for grades 2 and 3</p>	<p>The typical pupil can copy letters and write his/her name without mistakes.</p> <p>Some of the pupils cannot yet write their names without help (about 1/4).</p> <p>Most pupils are unable to write more than 5 words.</p>

Table 3 (continued).

Grade Level	Oral Language	Reading	Writing
<p>Primary Grade 4 (P4) (n=265)</p>	<p>The pupil responds more readily than P2 and P3 pupils to the standard identification questions. (Functional English)</p> <p>Listening comprehension skills are comparable to those of younger pupils with the typical grade 4 pupil experiencing difficulties responding to oral directions involving grade 4 materials.</p>	<p>Most pupils demonstrate familiarity with printed materials.</p> <p>The typical pupil is unable to read many of the most frequently used words in the grade 4 textbook.</p> <p>The typical pupil cannot read passages from the English textbooks for grades 2-4.</p> <p>Even those children who experience success in reading the passages (about 1/5 of the pupils) have difficulties comprehending what they read.</p>	<p>Everyone can copy letters and the great majority (80%) can write their names without assistance.</p> <p>The typical pupil can write more than 5 words but has difficulty writing more than 15 words.</p>
<p>Primary Grade 5 (P5) (n=254)</p>	<p>The grade 5 pupil can satisfactorily answer most of the standard identification questions and comprehends simple oral instructions such as “Sit down,” “Stand up,” and “Walk to the door,” but is unable to respond correctly to approximately half of the listening comprehension tasks and oral directions requiring the use of vocabulary and pictures found in the textbooks.</p> <p>The pupil cannot express himself/herself in English--oftentimes responding in the local language to questions asked in English.</p>	<p>Children at this level demonstrate familiarity with using printed materials.</p> <p>They can read many but not all of the most frequently used words in their texts.</p> <p>The typical pupil is able to read (decode) about half of the words in passages taken from P2-P5 textbooks.</p> <p>Reading comprehension continues to lag behind with only a small percentage of pupils comprehending what they read.</p>	<p>With rare exceptions, everyone can write their names unassisted.</p> <p>Most pupils can write more than 5 words and almost half can write more than 15 words.</p>

Table 4. Summary of baseline ANOVA results for achievement measures for overall sample (Intensity x Grade Level).

	Overall--Baseline		Level 2	Level 3	Level 4	Level 5
Functional Lang.	C>I		C>I	NS	NS	NS
Listening Comp	C>I		C>I	C>I	NS	NS
Oral Expression	I>C + Intens X Lvl		I>C	I>C	NS	NS
Copy Letters	C>I		C>I	C>I	NS	NS
Write Name	I>C		I>C	NS	NS	NS
Write Words	NS		--	--	--	--
Concepts-Print	Intens X Lvl		NS	I>C	NS	C>I
Letters/Sounds	NS		--	--	--	--
Aided Reading	NS		--	--	--	--
Most Used Words	NS		--	--	--	--
TRD-1 Minute	I>C		NS	NS	NS	I>C
TRD-% Whole	NS		--	--	--	--
TRD-Comprehension.	NS		--	--	--	--
Read P2/minute	I>C		NS	NS	NS	I>C
Read P2 %	NS		--	--	--	--
P2 Read Comp	NS		--	--	--	--
Read P3/minute	I>C			NS	NS	I>C
Read P3 %	I>C			NS	NS	I>C
P3 Read Comp	NS			--	--	--
Read P4/minute	I>C				NS	I>C
Read P4 %	NS				--	--
P4 Read Comp	NS				--	--
Read P5/minute						I>C
Read P5 %						I>C
P5 Read Comp						NS

NS indicates that the differences in average performance were not significant at the .05 level.

I>C indicates a significant difference ($p < .05$) with pupil performance in Intensive schools greater than pupil performance in the Comparison (Non-intensive) schools.

C>I indicates a significant difference ($p < .05$) with pupil performance in the Comparison schools greater than pupil performance in the Intensive schools.

TRD 1 Minute, TRD % Whole, and TRD Comprehension are composite variables combining performance from multiple passages.

Significant differences favoring the Intensive schools were associated with reading passages and indicated that grade 5 pupils in the Intensive schools were able to read more words per minute than grade 5 pupils in the Non-intensive schools. Also, in two of the four reading passages, grade 5 pupils in the Intensive schools read correctly a greater percentage of the words than did grade 5 pupils in the Non-intensive schools. There were no significant differences with regard to reading comprehension on any of the passages. It should be noted that with this large number of statistical tests, it is expected that a small number of significant findings are likely to be spurious. Thus, for all practical purposes, it appears that baseline achievement of pupils in Intensive and Non-intensive schools in grade levels 2, 3, and 4 was comparable; in grade level 5, while performance in most areas was comparable, there were some differences in reading decoding favoring pupils in the Intensive schools.

FINDING: WITH RARE EXCEPTIONS, PUPIL PERFORMANCE IN INTENSIVE INTERVENTION SCHOOLS WAS SIGNIFICANTLY BETTER THAN PUPIL PERFORMANCE IN NON-INTENSIVE INTERVENTION SCHOOLS AT 1 AND 2 YEARS FOLLOWING THE INITIAL INTERVENTION EFFORTS.

Table 5 provides the results from analyses of variance comparing pupil performance results from Intensive and Non-intensive schools at baseline (Winter 1994), 1 year after baseline results were shared with educators in Intensive Schools (Summer 1995), and 2 years after results were shared (Summer 1996). These analyses report findings based on the longitudinal data set (only those pupils who had been tested at baseline were included). Analyses completed on pupils who were no longer attending the participating school at the time of the follow-up assessments failed to identify any patterns differentiating Intensive and Non-intensive schools with regard to pupils dropping out or transferring to other schools. Due to resource limitation, only grade 6 pupils were reassessed in Summer 1996.

As is evident from Table 5, children in the Intensive Intervention schools made significantly more academic growth than children in the Non-intensive schools--that is, although children in all 14 participating schools make substantial improvement, the improvement of children in the Intensive schools was significantly greater than the improvement of children in the Non-Intensive schools. This was true for skills associated with oral language, reading, and writing. The only area for which this was not true was expressive writing. This skill was not measured a part of baseline assessment. At follow-ups, only 6th grade pupils were asked to provide samples of their expressive writing (e.g., write letters, stories, etc.). Most of the grade level 6 pupils--regardless of whether they attended Intensive or Non-intensive intervention schools--experienced considerable difficulty with this task.

FINDING: DIFFERENCES IN FAMILY BACKGROUND DO NOT SEEM TO ACCOUNT FOR INTENSIVE/NON-INTENSIVE ACHIEVEMENT DIFFERENCES. IF ANYTHING, THE PUPILS FROM NON-INTENSIVE SCHOOLS CAME FROM HOUSEHOLDS WHERE THERE WAS GREATER PERCEIVED ENGLISH COMPETENCY THAN THE PUPILS FROM THE INTENSIVE SCHOOLS. FAMILY SIZE WAS COMPARABLE. EDUCATION LEVEL OF HOUSEHOLD MEMBERS WAS COMPARABLE.

In July 1996, estimates of household English were gathered for Level 6 pupils. Members in each household were listed and pupils were asked in the vernacular to describe whether each household member could speak, read, or write English. Each member was rated as "No," "Some," or "Yes." The number of Yes's for each aspect of language (Speaking, Reading, and Writing) was tallied and pupils in Intensive schools were compared with pupils in the Non-intensive Schools. There were significantly MORE readers of English in the households of Non-Intensive school pupils than there were in the households of Intensive school pupils (significant at the .05 level). [When kind of community was taken into account, all three ratings of English competency favored pupils in Non-Intensive schools however some cell sizes were small and thus more subject to chance variation.]

Table 5. Summary of ANOVA results for selected achievement measures.

	Overall--Baseline Feb. 94 (N=1037)	Overall-- July. 95 (N=785)	Overall-- July 96 (Grade Lvl 6 only N=222)
ORAL LANGUAGE			
Functional Lang.	C>I	I>C	I>C
Listening Comprehension	C>I	I>C	NS
Listening Comp for Passage (Grade Lvl 6 only)	--	NS	I>C
Oral Expression	I>C + Intens x Lvl: I>C Lvl3 2&3; NS Lvl3 4 &5	I>C	I>C
Total Oral Language	--	I>C	I>C
READING			
Concepts-Print	Intens x Lvl Lvl3 I>C; Lvl5 C>I	NS	NS
Letters/Sounds	NS	I>C	I>C
Aided Reading	NS	I>C	I>C
Most Used Words	NS	I>C	I>C
Words Read Correctly in 1 minute (composite)	I>C	I>C	I>C
Percent of Words Read Correctly (whole passages-composite)	NS	I>C	I>C
Reading Comprehension (composite)	NS	I>C	I>C
WRITING			
Copy Letters	C>I	NS	NS (ceiling)
Write Name	I>C	I>C	NS (ceiling)
Write Words	NS	I>C	I>C
Dictation-% correct letters	--	NS	I>C
Dictation-% correct words	--	I>C	I>C
Dictation-passage (includes spelling, punctuation, etc.)	--	I>C	I>C
Expressive Writing Fluency (Grade 6 only)	--	NS	NS
Expressive Writing Correct (Grade 6 only)	--	NS	NS

NS indicates that the differences in average performance were not significant at the .05 level.

I>C indicates a significant difference ($p < .05$) with pupil performance in Intensive schools greater than pupil performance in the Comparison (Non-intensive) schools.

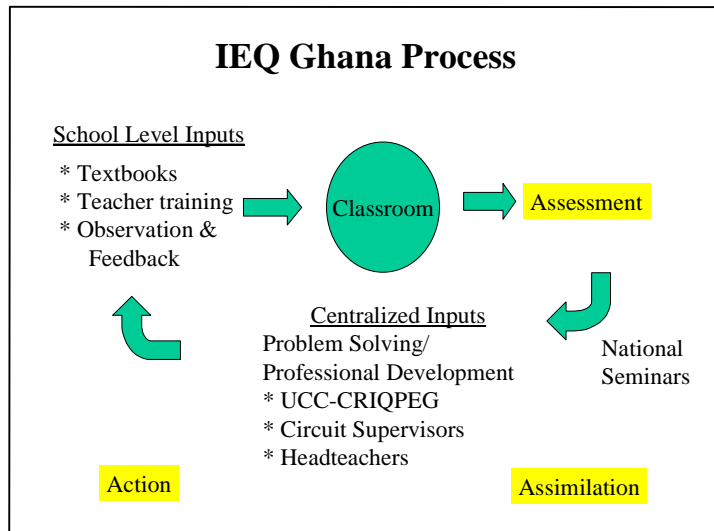
C>I indicates a significant difference ($p < .05$) with pupil performance in the Comparison schools greater than pupil performance in the Intensive schools.

Lvl=Grade level; Intens x Lvl indicates there was a significant interaction.

CBA AND THE IEQ PROCESS IN GHANA

Through these examples, it should be evident that CBA played a central role in the IEQ process in Ghana. Figure 4 illustrates the assessment-assimilation-action process.

Figure 4.



CBA was used to measure baseline performance and to catalyze local and national problem solving. This problem solving resulted in better targeted interventions in the form of teacher training and instructional materials. Classroom level assessment found its way into the instructional process. As the cycle of assessment-assimilation-action repeated itself through the project, educators saw evidence they could understand of the changes and improvement. Table 6 illustrates this interactive relationship.

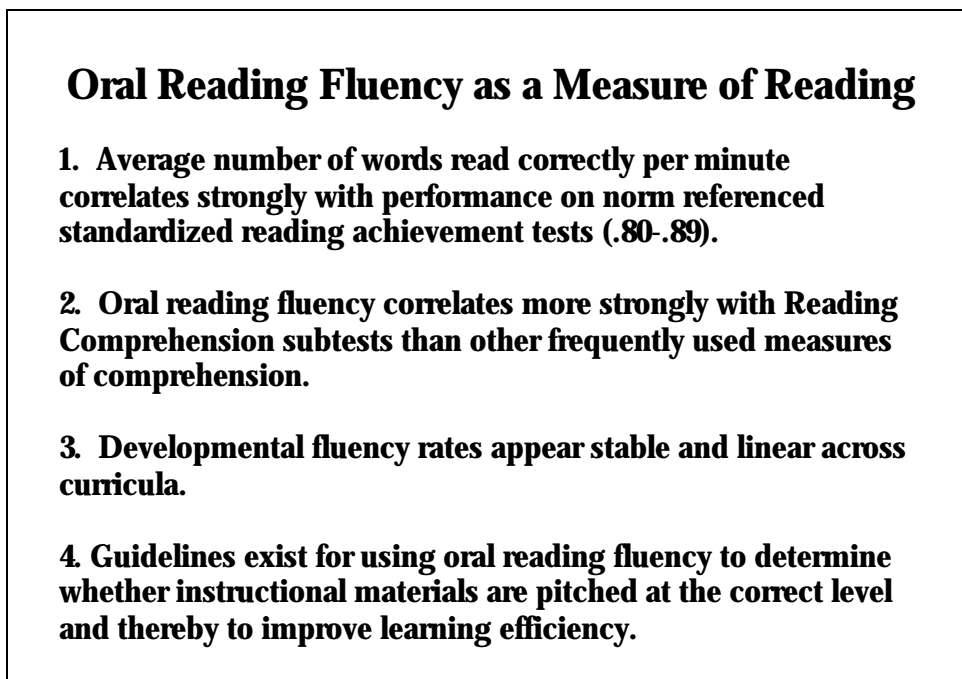
Table 6.

	Impact Assessment	Instructional Value
A. Baseline Achievement Data Collection	<ol style="list-style-type: none"> 1. Used to evaluate comparability of project schools; 2. Established a baseline of student performance; 3. Drew attention to student learning as a key indicator of program success. 4. Developed local capacity in assessment by involving University of Cape Coast faculty (CRIQPEG) in instrument development, administration, analysis, and dissemination. 	<ol style="list-style-type: none"> 1. Facilitated the development of student and classroom learning profiles; 2. Provided information on the match between available instructional materials and student skills; 3. Provided a tool for engaging stakeholders in the improvement process. 4. Demonstrated to participating educators a simple process of assessing learning. 5. Developed local capacity in instructional improvement by involving CRIQPEG in sharing assessment findings and collaborating on solutions.
B. Mid-term Achievement Data Collection	<ol style="list-style-type: none"> 1. Provided data to evaluate learning progress in non-intensive and intensive intervention schools. 2. Provided data to evaluate differential effectiveness (rural/urban; gender, etc.) and unique needs. 3. Used to generate hypotheses for further study. 	<ol style="list-style-type: none"> 1. Results presented at national conference to generate discussion on national level. 2. Raised interest of CRT developers 3. Results discussed with Circuit Supervisors and headteachers, generated new ideas & guided decisions about instructional materials and training needs
C. End of Project Achievement Data Collection	<ol style="list-style-type: none"> 1. Demonstrated significant learning improvement for all schools (Hawthorne effect?) 2. Demonstrated significant differences in learning between intensive and non-intensive intervention schools. 	<ol style="list-style-type: none"> 1. CRIQPEG researchers gained experience in data analysis and report writing. 2. Interest in project continued; requests for instruments and info within Ghana and internationally.
D. Post -project	<ol style="list-style-type: none"> 1. CRT Results (1996) indicated that Grade 6 pupils from intervention schools scored significantly better than their public school peers. 	<ol style="list-style-type: none"> 1. Current USAID funded basic education project is using the assessment instruments to monitor progress and as the basis for implementing a continuous assessment approach. 2. CRIQPEG is integrally involved in the assessment component of the project.

CBA as a Tool for Use in Developing Countries

One additional note about CBA, recent research further supports the use of this kind of assessment and its utility in monitoring (i.e., its properties as a measure of school learning). In recent years, CBA has gained popularity in the US and several studies have investigated the psychometric and statistical properties of a variety of CBA tasks. One of the most promising for consideration in developing countries is what was labeled “decoding” in IEQ-Ghana and “oral reading fluency” in the US research. Studies available from the 1980’s demonstrated strong correlations between decoding/oral reading fluency and reading achievement. However, the focus was on reading passages from basal readers or

Figure 5.



literature-based basal readers which typically include passages sequenced for difficulty. In developing countries little is known about whether or how available curriculum is sequenced. Recent investigations in the US looked at “authentic” materials which basically means stories being used in whole language programs at each level. I’d like to review some of the findings:

A recent study by Hintz, Shapiro, Conte, and Basile (1997) compared oral reading rates for children on (1) grade appropriate literature based reading series, and (2) authentic reading materials (not written for any particular reading level).

Average number of words read correctly per minute correlated with the Stanford Achievement Test subtests at an average of .89. Oral reading fluency correlated more strongly with the Reading Comprehension subtest than any of the other measures (e.g., Cloze, answering simple questions about the passage, etc.) (Fuchs, Fuchs, & Maxwell, 1988).

In a different study (Jenkins & Jewell, 1993) correlations between oral reading measures and commercially available norm-referenced reading measures (i.e., Metropolitan Achievement Test, Gates Reading) ranged from .80-.88.

Comparisons of the developmental fluency rates appeared stable and linear across curricula, using the more traditional skills and literature-based basals and “authentic” materials (Fuchs & Deno, 1982; Hintz, Shapiro, Conte, and Basile (1997).

One important caveat is that these studies were based on English as a first language and examined the sequencing of passages using readability indices. A small substudy done by Kazas (1996) using the CBA data and internationally accepted readability formulae found that the readability indices didn't correlate with actual performance until a minimal level of reading was established. This argues for careful selection of the reading passages and selection of multiple passages.

Recommendations

Improving reform efforts through on-going assessment of student learning is a powerful tool to increase program effectiveness. From our research emerged valuable insights on how to maximize this potential. Some of these insights are offered now as recommendations:

Staffing

1. Involve a broad range of local educators from the outset of the project.
2. In the interest of objectivity and as an instructional resource to the local educators, involve outside experts (perhaps build local capacity at the University level).

Assessment Strategy

1. Focus on what students CAN do--be sure your assessment process includes a full range of prerequisite skills.
2. Draw the assessment stimuli from the local syllabus and/or instructional materials. Be aware that curriculum may not be well sequenced--use extra caution in selecting passages.
3. Plan ahead--it is much easier to create several parallel tasks/forms initially than it is to create only one form and try to develop equated tasks/forms later.
4. Share assessment results with stakeholders as soon as they are available.

Notes

ⁱ The Improving Educational Quality Project was funded by the Center for Human Capacity Development/U.S. Agency for International Development (USAID). IEQ was conducted by the Institute for International Research (IIR) in collaboration with Juárez and Associates, Inc., and the Institute for International Studies in Education at the University of Pittsburgh. (Contract #DPE-5836-C-00-1042-00). IEQ's goals developed out of a focus on generating knowledge about IEQ country innovations, sustaining the research process in educational reform efforts, transferring research knowledge into practice, and developing and disseminating the lessons learned. They were: (1) understand how and why each country's classroom-based interventions influence pupil performance; (2) demonstrate a process whereby findings from classroom research on improving educational quality are utilized by the educational system; (3) create opportunities for dialogue and partnerships among researchers and educators who are seeking to improve educational quality at the local, regional, national, and international level; and (4) maintain a history of the project to document the rationales for choices made, opportunities and constraints encountered, and lessons learned. The five participating countries were Ghana, Guatemala, Mali, South Africa and Uganda.

ⁱⁱ CRIQPEG is located at the University of Cape Coast. It is coordinated and staffed by faculty, lecturers, and advanced level graduate students from the University. Although it was formed as part of the IEQ Project, the quality of its work has been widely recognized and it continues as part of the University of Cape Coast.