

TO: Don Conrad, Jenkins Cooper, Wade Robinson

FROM: Joel Aronson, Evaluation Specialist

DATE: September 28, 1998, October 18, 1998

RE: Trip Report, Quality Improvement for Primary Schools –
September 6 through September 25, 1998

The scope of work for this visit included 4 elements:

1. Complete the comparison of the baseline data for the Partnership Schools/Communities and the Comparison Schools/Communities
2. Brief USAID/MOE/GES on as many key baseline data as possible. Discuss degree of match between and among group 1 schools/communities (Partnership and Comparison)
3. Assist in the planning effort for the National Forum for PME.
4. Submit consultant trip report to USAID.

In addition, I accomplished the following:

5. Reviewed the revised achievement tests and suggested further changes to both the makeup of the tests and the testing process.
6. Produced an initial draft of the Baseline Report for tranche 1.
7. Helped staff at TMG/Accra upgrade their utilization of Office97 and Internet applications.

Baseline Comparisons

During the interim between my last trip (July/August) and this, Emma had completed all of the basic comparisons I had requested. These included:

- Side-by-side contrast of the primary results from the CSA questionnaires from Partnership Schools/Communities and Comparison Schools/Communities, item by item, for the 75 items we had identified as potential indicators.
- Side-by-side contrast of the primary results from the ILP questionnaires and the classroom observation form from Partnership Schools/Communities and Comparison Schools/Communities, item by item, for the 281 items we had identified as potential indicators.
- Second-order results (cross-tabulations) by parent's education level for most of the above.

- Second-order results (cross-tabulations) by school classification (A, B, or C) level for most of the above.

Because CSA had produced their questionnaire instruments before we revised our own versions (which are the versions CSA will use in the future), some of the questions we asked were not asked by CSA and some of the questions they asked had different response alternatives than ours. As a result, there are a number of items for which there are not exact comparisons between the two sets of data. For example, our questions about the enrollment of school-age children ask about boys and girls separately, theirs did not.

The bottom line for the comparison of the baseline data of Partnership Schools/Communities and Comparison Schools/Communities is that there are strong similarities between the two sets, though not complete convergence: the Comparison Schools ARE fully adequate as control locations. As I reviewed the small, but fairly consistent differences between the two sets, I sense that the Partnership Schools (which were selected first) were probably among the best in each of the school classification categories, leaving the Comparison Schools a little behind, on average. When we look at progress indications we will take those differences into account.

USAID/MOE/GES Briefings

We did not have a formal briefing for MOE/GES or for USAID, but we did deliver a presentation at USAID regarding our data gathering for representatives of a number of donor agencies and MOE/GES. In addition, we held a lengthy feedback meeting at USAID (with CSA and ILP attending) that was focused entirely on data collection/data analysis topics.

In general, our approach has been to identify a large number of potential indicators (based on their susceptibility to change as QUIPS interventions take place), with the intention of reducing the number as we gather data and see how they respond to change. Up to this point, we have incorporated only CSA/ILP indicators and we still need to explore the utility of information from other sources, such as GES' literacy and numeracy effort and HIID's annual school census data.

Planning for the National Forum

Following consultations with ILP, CSA, and USAID, we decided to handle the National Forum a little differently than originally planned. In order to elicit maximum buy-in, we decided to establish a planning group of 10-12 people, including representatives from Legon University, the University of Cape Coast, MOE, GES, ILP, CSA, PME, and USAID. This group will begin meeting in November and will help shape the Forum's agenda and list of invited participants. Though the first two-day Forum originally was scheduled for November, 1998, we have put it off until Summer of 1999 for two principle reasons:

- USAID has scheduled a review of the entire QUIPS initiative for the end of the year (1998) and its outcome may impact the Forum
- As yet, there will be little feedback by November regarding the impact of interventions from either tranche 1 or tranche 2 schools/communities. The later schedule will allow us to provide the Forum participants with more specific outcome information and more firmly link theory with relevant practice.

More Effective Achievement Testing

During my July/August visit, I discovered and reported on a serious methodological problem with the achievement tests being used by ILP to assess the level of learning achieved in primary grades 2, 4, and 6. At that time, I worked with ILP and CRIQPEG to plan a revised testing methodology that would assess learning achievement more realistically.

During this present trip, I reviewed the revised tests that were being prepared for administration in October and November, just subsequent to the launch workshops for the new group of 27 Participant and 15 Comparison Schools/Communities. I found there was still a lack of clarity about what was needed, related, I believe, to two concerns that became confused:

- I had suggested the following rearrangement for year-end use of the items already written and assembled into tests for grades 1 through 6:

Grade Level	30% Items From	50% Items From	20% Items From
P6	P6 Test	P5 Test	P4 Test
P4	P4 Test	P3 Test	P2 Test
P2	50% P2 Test	50% P1 Test	

- However, The fact that testing for tranche 2 is taking place at the beginning of the school year, complicates the problem. Students sitting in P6 classrooms are really P5 students who have had a 10-week hiatus between the last of their P5 schooling and their testing of that material. To be consistent with our testing, we need to set the tests up as follows:

Grade Level	30% Items From	50% Items From	20% Items From
P6	P5 Test	P4 Test	P3 Test
P4	P3 Test	P2 Test	P1 Test
P2	100% P1 Test		

I reviewed the math tests for P4 and P6 and found they were inconsistently assembled. The P4 test included about half P3 items (based on my review of the tests prepared for tranche 1) and half P2 items. That was fine for this tranche,

but if they are to be useful for end-of-year testing at some future time, they also should include some P4 items.

The P6 test included only P5 items. In effect, that would be testing P5+ students with P5 items, which is, effectively, no change from the original testing in tranche 1 which resulted in virtually zero response. P4 and P3 items need to be added.

It remains the intention to test in P2, but the only relevant items already banked are for P1. Since beginning P2 students are really P1+ students, and since at-grade test items failed at all levels during the tranche 1 achievement testing, I don't believe we can use those items and expect meaningful results.

The achievement tests are the only direct measure of level of learning we have in QUIPS. From my perspective, it is very important that they be potent. Without the ability to identify demonstrable shifts in classroom effectiveness, we will not have much beyond descriptive data to account for all our effort. I raised these points at the USAID meeting held just before I left Ghana and there seemed to be general agreement from the group regarding Wade's suggestion that achievement testing for tranche 2 be postponed until the end of term 1. That would give everyone a chance to be sure that the tests were constructed appropriately. It would also give the teachers and students a chance during their term 1 to review the previous year's learning and to begin to engage the new work. A third benefit is that the results, though not directly comparable to results gathered at end-of-year, at least are on a continuum where the gaps can be taken into account.

Finally, there still needs to be better clarity about what the achievement tests are designed to reflect. If part of that intention is to track the progress of students at the p2, p4, and p6 levels of instruction, then testing in those grades at the beginning of the school year comes much closer to assessing p1, p3, and p5 progress. Those results can never be combined with end-of-year results. If the aim is to be able to pool tranche results and compare the effectiveness of teaching at Partnership schools with that at Comparison schools generally, the testing process needs to be more consistent, and probably should include students at all six grade levels to provide a better continuum.

Status at the Time of my Departure:

On the day of my departure, I spent several hours with Emma and Michael reviewing the analytic procedures we had developed, reviewing how to send and receive emails with attachments, and building a task list for the next several months. The task list included the following elements:

- Get together with Tom LeBlanc to work out how to incorporate the 1997 school census data (and the 1998 data when it is ready) for the QUIPS schools into our Access data base.

- Get in touch with Mr. Kanadu to obtain his Literacy and Numeracy information for the QUIPS schools. I have given him the list of participant and comparison schools for tranche 1 and the participant schools for tranche 2. At the time, I did not have the names of our 15 comparison schools for tranche 2. When we know those, ask Mr. Kanadu for their data, as well.
- Emma (and, when he is sufficiently up to speed, Michael) should plan to attend some of the launch workshops to represent TMG/PME, to learn about activities and conditions at the schools, to interact with the district education officials, and to meet the community leaders.
- As data is gathered from the tranche 2 schools/communities, the information should be entered using a system similar to the one we used for tranche 1 data. We want to keep tranche 1 data and tranche 2 data as compatible as possible, even though some of the specific questionnaire items may have changed.
- Emma should get together with Sammy (ILP) to make sure he is entering ILP data in the same way we have done it. We want to avoid having to deal with undigested text data when, with a little planning, that information can be categorized before it is entered.
- Once all data are entered, perform data analyses that parallel those that we have done for tranche 1.

As we look ahead, we need to plan for the analysis of a new round of classroom observations from term 1 for tranche 1 schools, and then for year's end, we need to prepare a report which combines results from tranches 1 and 2.