

CHAPTER SIX: FOCUS GROUP INTERVIEW ANALYSIS

6.1 Introduction

This chapter addresses those indicators that are described as qualitative and whose scores have been derived through focus group interviews and inter-rater scoring. Two levels of surveys were conducted. The first addresses indicators that measure processes and outcomes as perceived by managers and aggregated to derive a district-level score. The second set of measures were used to aggregate scores related to the different groups, including control groups, to measure pilot activities. Table 6-1 provides results for the district-level surveys while other tables (6-3 to 6-6) provide results used in evaluating pilot group responses. These responses were provided by students, teachers, school-level managers, and community members. The tables indicate which stakeholders responded to which indicators.

This chapter attempts to focus on quality. Whereas quantitative measures discussed in earlier chapters focus on specific inputs, outputs and outcomes in terms of quantity, such as number of desks purchased (an output), current enrollment (input), student to classroom ratio or student achievement (outcomes), and many others, the qualitative dimension looks at how inputs are treated. Such measures as quality of construction or in-service programs (processes) or attitude towards science (outcome) add a new dimension to the research and may better help the reader understand the perceptions of the human dimension in this research. Although different approaches may be used to evaluate qualitative measures, the focus group was chosen to secure responses to different standardized questions representing different constructs of quality, and rating those group responses, thus converting them to measurable indicators.

The second part of this chapter dealing with the different groups is crucial in trying to understand the impact of various processes on specific outcomes and attributing them to specific interventions. These interventions are the similar pilot activities chosen by a group of schools forming a Group or sample. A control group is important in this type of analysis and control groups for primary and secondary levels were created based on the fact that no interventions were provided under the project. Also important is the relationship between processes and outcomes. Outcomes are both quantitative and qualitative. To complete the analysis in this chapter, it is necessary to introduce results from selected quantitative outcomes such as test scores and absence rates in order to examine the possible relationship to various processes.

Most research of this type introduces experimental rigor in order to control for outside influences. This was not possible since the main purpose of the project was to integrate pilot activities into the normal flow of schools from selected districts. It was not possible, therefore, to establish a study based on appropriate measurement principals. For example, validity and reliability of instrumentation was not rigorously developed. Interventions from other donors could not be controlled. Certain statistical applications could not be used such as significant difference as were applied in Chapter three in the use of ANOVA to analyze achievement test results. Thus, many of the results suggested in this chapter are left open to challenge. Nevertheless, there is a large body of data that can offer some insight into how schools may or may not have changed as a result of pilot interventions.

6.2 District Level Evaluation

School level managers including representatives from the school committees participated in the focus groups. A total of 14 indicators formed the basis for evaluating process-type variables reflecting quality. A total of 124 primary schools and 28 secondary schools participated across the six districts. This represents a sample size of 14 percent and 24 percent respectively. Data were collected through focus group interviews and scoring by raters after the interviews were completed. An evaluation scale ranging from 0 to 5, where zero is the lowest score and five is the highest score, was used. Results reflected the perceived quality of processes seen as important in educational effectiveness.

The following table (table 4-1) shows the results of the surveys across the six districts. Column one contains the 14 indicators and that managers were surveyed. Results for each district were achieved by averaging responses from all focus groups and are shown in three columns where the first contains the results of the baseline survey, the second containing the results of the mid point survey and the third containing results from the post pilot survey. The timeframe from baseline to post pilot represents a span of 22 months. Scores within each box generally represent an integer; however, some contain tenth decimal points. The decimal points occurred when raters could not agree on an integer and their respective scores had to be averaged.

The last three rows provide some results. In the third row from the bottom, the two numerals for each district illustrate how many indicator scores increased by one or more (range being 0 to 5) between the baseline and post pilot surveys. The second number reflects the same change but in a negative direction. Due to the lack of a rigorous statistical analysis, this represented an attempt to standardize an approach to defining significant difference. In two cases, where baseline scores were at four, an increase of more than half a point was considered significant. Thus a score changing from 4 to 4.7 was included. Where significant upward changes occurred, the numeric values are shown in bold. If there were a significant decrease, then the numeric values are shown in bold and italic.

The next-to-last row shows results in a different way. A median score for any indicator is 2.5. Any score of 3.0 or above is considered high average to above average. The next-to-last row shows the number of indicators that were scored at three or above at each survey. It would be predicted that the project interventions would lead to increases in scores and that the total number of scores at 3 or above would increase as well. This row provides the results of the change across the three surveys.

The last row shows the number of times the raters selected a score of zero. The zero score indicates that there is no evidence that the indicator is present in the pilot. For example, a zero for in-service would indicate that no in-service was conducted. A zero for use of DEPs would indicate that DEPs were not being used at the school level. A zero for communication links would indicate there was no communication taking place between two levels of the system. The analysis of how many times the zero rating was used may reflect on the indicator or on the raters and may prove to be important. (These last three rows are also used in analysis of the group evaluations in tables 6-3 to 6-6)

Analysis may be found directly after the Table 6-1.

**Table 6-1
Pilot Project Evaluation by District – Post Pilot Survey**

*(Aggregated data based on 2 outlier primary schools per zone and for secondary 1 CDSS and 1 CSS school per cluster)
(Primary school sample = 124 schools; secondary sample = 28 schools)*

Indicator	Nsanje			Thyolo			Mchinji			Nkhata Bay			Machinga			Ntchisi		
	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot
Teacher use of lesson plans	2.5	2.0	2.5	3.0	3.0	3.0	2.0	2.0	3.5	3.0	3.0	3.0	3.0	3.8	3.0	2.0	3.0	3.0
Teachers use of record keeping	2.4	3.0	3.6	2.0	3.0	3.4	2.0	2.0	4.0	2.0	3.6	3.0	3.0	3.3	3.0	2.0	3.6	3.2
Quality head teacher in-service	3.0	3.0	0.0	4.0	3.0	0.0	3.0	3.0	4.7	3.0	3.6	3.7	0.0	3.0	4.7	3.0	3.7	0.0
Active school committee	3.8	3.0	4.0	4.0	4.0	3.7	4.3	2.0	4.0	3.0	4.0	4.0	2.0	4.3	5.0	2.0	5.0	3.5
Community engagement	3.8	2.0	3.5	2.5	3.0	2.0	3.0	2.0	5.0	3.0	4.0	0.0	0.0	2.8	5.0	3.0	4.0	3.5
Quality of construction	2.3	1.0	3.0	1.0	3.0	2.5	2.0	2.0	0.0	2.0	3.0	3.0	2.0	1.8	0.0	2.0	1.7	2.7
Quality PEA/cluster in-service	2.3	4.0	0.0	3.0	3.0	0.0	2.5	1.0	5.0	3.0	4.3	3.7	2.3	2.8	5.0	0.0	5.0	3.7
Effective use of TDCs	2.5	3.0	4.0	1.0	4.0	4.0	3.0	2.0	5.0	3.0	4.0	4.0	2.0	2.8	5.0	5.0	5.0	4.0
District/zone communication	3.0	3.0	4.3	3.0	5.0	4.3	2.7	3.0	4.7	3.0	3.6	4.0	4.0	4.0	4.7	2.0	4.0	3.3
Stakeholder communication	3.3	4.0	4.8	2.0	4.0	4.8	2.5	2.0	4.8	4.0	3.5	3.8	3.0	4.0	5.0	3.0	5.0	3.3
District /division communication	4.0	4.0	5.0	4.0	5.0	4.8	1.0	4.0	4.8	3.0	3.3	4.0	4.0	3.5	5.0	4.0	5.0	4.0
Accuracy-data collection and record keeping\	3.5	2.0	4.0	4.0	2.0	4.0	2.0	1.0	4.5	2.0	4.0	4.0	2.0	3.0	3.0	1.0	4.5	3.5
Use of DEPs a. quality from 0 to 5 (managers)	2.5	1.0	4.0	4.0	2.0	3.5	1.0	1.0	5.0	2.0	3.0	4.0	2.0	3.2	5.0	4.0	3.5	3.0
Transparency-finances & procurement	2.5	2.0	5.0	3.0	3.0	4.0	1.0	3.0	4.5	4.0	2.5	4.0	3.0	3.2	5.0	4.0	5.0	3.0
# of indicators up/indicators down (bold=up <i>bold italics =down</i>)	7 up / 2 down			7 up / 2 down			12 up / 1 down			8 up / 1 down			11 up / 1 down			6 up / 3 down		
N = 14 Number of indicators at 3 or above	7	8	11	9	12	10	4	4	13	10	13	13	6	10	13	7	13	12
Number of Indicators scored as Zero	0	0	0	0	0	2	0	0	1	0	0	1	1	0	1	1	0	1

By aggregating manager scores across the system, a gross measurement is acquired. Since standard deviations were not calculated, it is not possible to determine the dispersion from the mean and project the possible error in measurement. Based on this limitation, the following points were observed:

6.2.1 Primary and Secondary Education

- In all cases, the number of indicators that increased by one or more between baseline and post pilot were two times or more as many than the number of decreases. This shows that managers perceive that the quality of their schools as measured by changes in processes has improved in many more ways than has decreased.
- The number of indicators that were ranked as high average or above average (score of 3 or more in a range from 0 to 5) in all cases increased from baseline to post pilot (Nsanje = 7 to 11; Thyolo = 9 to 10; Mchinji = 4 to 13; Machinga = 6 to 13; and Ntchisi = 7 to 12). With N equaling 14, all managers indicate that most of their school processes are performing at above average level.
- Almost no indicators were ranked as zero at anytime.
- In-service for teachers, head teachers and PEAs demonstrated the most erratic behavior. This is probably related to pilot activities assuming that target audiences received no in-service from other donors; therefore, the zero rating was used when no pilot activities were instituted for in-service, and high rankings given by those where pilot activities included in-service.
- One group of indicators with the most consistent increases are those indicators related to communication. In almost every case post pilot results showed above average communications across the system, when at baseline communication was considered average or below. This demonstrates a significant improvement in vertical linkages as measured by communication.
- In all cases, transparency was rated high average to above average, and in four cases, this represented a significant increase from the baseline scores. In one case, the indicator value stayed the same and in a final case, the value decreased.
- Use of DEPs was also erratic. In four cases, the indicator scores increased significantly to well above average. In one case the score remained the same at above average while in the final case the score fell from above average to high average. It is unclear how schools and other managers were using the DEPs, however.
- Teachers use of data collection tools increased significantly in most cases to high average and above, which may account for why data accuracy scores all increased to high average and above. This may also relate to why DEP use is high.
- Active school committees increased significantly to above average and much higher. Community engagement scores, however, were much more erratic. While some increased significantly, one decreased to zero which seems to be a significant piece of data and should be investigated further.
- In all but one case, use of the TDCs increased significantly, probably since pilot activities were often held at TDCs. In all cases the post pilot evaluation was well above average.
- As suspected, scores for quality of construction were erratic and probably related to those districts that received pilot funding for different construction activities.

6.3 Post Pilot Survey Of Schools By Groups

Samples of schools implementing similar pilots were grouped together. There are five distinctly different school groups and a sixth group representing the control group. Pilot groups at the primary level are different than pilot groups at the secondary level. The four tables below provide the results of the focus group interviews with the different pilot groups. Tables 6-3 and 6-4 show total of all results. There are a total of 41 indicator measures to be analyzed for primary education and 36 for secondary education. More importantly tables 6-5 and 6-6 contain results for each of the six primary school groups and secondary school groups. This is discussed later. Table 6-3 shows the results of the four primary schools selected for each of the six groups that chose similar pilot activities, while table 6-4 shows the results for the two pilot secondary schools (one CSS and one CDSS). The first column

shows the list of indicators and which stakeholder groups responded to the raters' questions. The balance of 18 columns contain the scores, ranging from 0 to 5 to one decimal point. Each group of similar pilot schools has three columns. The three columns represent baseline, mid term and post pilot results rounded to one decimal place. Those indicators that changed by one point or more between the baseline and post pilot (mid term scores were ignored) are highlighted in bold (up) or bold and italics (down). The last three rows contain the summaries of various results.

Although schools with similar pilots were grouped together, the actual pilot activities varied. For example, while schools within a group chose public awareness campaigns, the nature of those campaigns might be different. One school may have focused on HIV/AIDS, another on dropout prevention, while a third on gender. Such differences will confound results to an uncertain degree. Also, while teacher in-service was focused on instructional methodology, some programs may have emphasized content as well or introducing HIV/AIDS into the curriculum. As discussed earlier, it was not possible to control for these differences, which will have an unspecified impact on the results. A further confounding factor involves the timing of the pilots. Some pilots were introduced in the first year of the project while the others introduced during the second year. In-service might, for example, be conducted in the first year while public awareness in the second year. This might account for why mid term evaluation scores for in-service spiked during the mid term evaluation and then returned to zero at the time of post pilot evaluation. The following table 6-2 shows the types of pilot activities selected for each group and the schools that were included on the evaluations:

6-2 Pilot Activities and Schools for Primary and Secondary Education

Group	Pilot Activities	Participating Schools	
Primary Level			
6	<ul style="list-style-type: none"> • Classroom construction • Data accuracy in-service 	1. Chikale, Nkhata Bay 2. Lombwa, Mchinji	3. Mlare, Nkhata Bay 4. Sunama, Mchinji
5	<ol style="list-style-type: none"> 1. Teacher in-service 2. Educational manager in-service 	1. Kalinganya, Ntchisi 2. Matuwamba, Mchinji	3. Mtuwanjovu, 2, Ntchisi 4. Nkwali, Nkhata Bay
4	<ul style="list-style-type: none"> • Public awareness campaign • Sanitation construction • Teacher in-service 	1. Bandawe, Nkhata Bay 2. Luchenza, Thyolo	3. Konzalendo, Thyolo 4. Kayuni, Machinga
3	<ul style="list-style-type: none"> • Furniture procurement • Teacher in-service 	1. Chigumukirre, Nsanje 2. Chinduzi, Machinga	3. Mtsiransembe, Ntchisi 4. Mberenga, Thyolo
2	<ul style="list-style-type: none"> • Public awareness campaign • Sanitary construction 	1. Liwonde, Machinga 2. Bua, Mchinji	3. Mikachu, Machinga 4. Mpinji, Thyolo
1	• No pilot activities (control group)	1. Bangula, Nsanje 2. Nyamadzere, Nsanje	3. Nyanga, Ntchisi 4. Mtawira, Nsanje
Secondary level			
6	<ul style="list-style-type: none"> • School construction • Classroom furniture procurement 	1. Maula, Nkhata Bay – CDSS 2. Tukombo, Nkhata Bay – CDSS	
5	<ul style="list-style-type: none"> • Instructional materials procurement • Teacher in-service • Classroom furniture procurement 	1. Mtambanyama, Thyolo – CDSS 2. Bvumbwe, Thyolo – CDSS	
4	<ul style="list-style-type: none"> • Data collection & reporting • Classroom furniture procurement 	1. Mawiri, Ntchisi – CDSS 2. Bua, Mchinji – CDSS	
3	<ul style="list-style-type: none"> • Science Kits and supplies procured • Classroom furniture procurement 	1. Chinkwezule, Machinga – CDSS 2. Kayoyo, Ntchisi – CDSS	
2	• Assistance to educational managers	1. Mtowe, Nsanje – CDSS 2. Magoti, Nsanje – CDSS	
1	• No pilot activities (control group)	1. Puteya, Machinga – CSS 2. Ludzi Girls', Mchinji – CSS	

Although tables 6-3 and 6-4 provide a significant amount of information, the importance of the results rests with their relationship to specific pilot activities. A brief analysis of these tables follows, but the major part of the analysis will follow tables 6-5 and 6-6:

Table 6-3
Pilot Project Evaluation by Group (Primary) – Post Pilot Survey

(Represent average score for the four pilot primary schools in each of six pilot groupings)

Indicators	Group 6			Group 5			Group 4			Group 3			Group 2			Group 1		
	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot
26. Teaching/learning process																		
a. 0 to 5 students	1.7	2.4	1.5	1.7	2.2	2.9	1.8	1.9	2.7	2.3	2.4	1.8	2.1	2.3	2.7	2.1	2.2	3.5
b. 0 to 5 teacher	1.7	2.3	1.5	1.4	2.5	2.9	2.0	2.5	3.3	2.5	2.9	2.9	3.3	2.4	2.5	2.3	2.5	3.0
27. Instructional materials																		
a. 0 to 5 students	1.1	0.9	1.0	1.3	2.2	1.7	2.0	1.1	2.8	1.8	1.7	1.9	2.1	1.9	2.1	2.0	2.2	2.5
b. 0 to 5 teacher	1.7	0.9	1.1	1.6	2.2	2.4	1.9	2.0	2.6	2.1	1.9	2.1	2.6	2.3	2.6	2.3	1.6	2.4
28. Continuous assessment																		
a. 0 to 5 students	2.3	1.9	1.0	2.3	2.1	2.0	3.0	2.7	2.0	3.1	3.2	1.9	3.2	2.4	2.2	2.8	2.8	2.4
b. 0 to 5 teacher	2.0	3.4	1.8	2.9	3.2	3.3	2.7	3.8	3.5	3.6	4.1	1.9	3.7	3.4	3.8	3.4	3.8	2.4
29. Use Lesson plans																		
a. 0 to 5 managers	2.7	2.5	2.3	2.1	3.6	3.0	2.6	2.8	2.9	3.1	4.1	2.8	2.9	3.5	3.8	2.9	2.8	3.4
30. Teacher in-service																		
a. 0 to 5 teachers	0.0	1.6	2.1	1.3	3.5	3.5	2.3	2.9	3.9	1.1	2.2	3.5	1.2	2.4	4.0	2.6	1.5	3.8
31. Use of Record keeping tools																		
a. 0 to 5 teachers	2.0	1.3	1.8	3.3	2.6	2.4	3.7	3.2	3.1	3.5	2.3	2.8	3.5	3.3	3.3	3.8	2.4	3.1
b. 0 to 5 teachers	2.5	1.7	2.4	2.0	3.1	3.0	3.3	3.2	3.6	3.5	0.0	0.0	4.0	0.0	0.0	3.2	0.0	0.0
33. HIV awareness																		
a. 0 to 5 student	1.3	1.5	0.0	2.5	1.3	0.0	3.0	3.0	0.0	2.8	2.0	0.0	2.8	2.3	0.0	2.5	1.5	0.0
b. 0 to 5 teacher	2.7	0.0	0.0	2.3	2.3	2.0	2.7	3.3	3.3	3.5	3.0	2.3	3.3	3.2	2.3	2.5	2.3	2.8
c. 0 to 5 community	1.3	1.5	0.0	2.3	2.8	2.5	3.3	1.8	3.8	3.0	2.4	2.8	3.5	3.3	3.5	1.8	2.0	2.5
34. Gender awareness																		
a. 0 to 5 students	2.8	2.5	0.0	2.6	2.8	0.0	3.0	3.5	0.0	2.6	2.5	0.0	2.4	3.3	0.0	3.0	2.5	0.0
b. 0 to 5 teacher	2.8	3.3	2.6	3.1	3.6	3.2	2.7	3.6	3.5	2.4	2.8	3.0	2.8	3.4	3.1	2.6	3.1	2.7
c. 0 to 5 community	3.0	3.5	2.1	3.6	3.5	3.3	3.3	3.5	3.8	3.3	3.6	3.3	3.4	3.2	3.8	3.3	3.3	3.1
35. Quality head teacher in-service																		
a. 0 to 5 managers	1.0	1.8	0.0	2.8	3.6	0.0	3.6	3.3	0.0	0.8	3.4	0.0	2.0	3.3	0.0	2.3	3.2	0.0
36. School environment																		
a. 0 to 5 student	3.2	2.5	0.0	3.0	3.0	0.0	3.2	3.3	0.0	3.5	3.1	0.0	3.1	2.1	0.0	3.1	2.6	0.0
b. 0 to 5 teacher	3.4	2.7	2.5	2.9	3.1	2.8	2.9	3.1	3.0	2.7	3.1	2.3	2.3	2.4	2.8	3.3	3.2	2.9
37. Parent Support																		
a. 0 to 5 student	3.3	3.0	2.6	0.0	3.1	2.8	0.0	2.7	3.6	0.0	3.3	2.2	3.3	0.0	3.3	2.9	0.0	3.3
b. 0 to 5 teacher	2.8	1.6	0.0	2.9	3.1	0.0	3.2	2.0	0.0	2.0	1.8	0.0	1.8	2.6	0.0	2.0	2.7	0.0
c. 0 to 5 community	3.0	2.1	2.6	2.5	3.1	3.6	2.9	3.0	3.5	1.3	3.3	2.3	3.3	2.5	3.7	2.8	2.3	1.8
38. Active/relevant school committee																		
a. 0 to 5 managers	3.7	0.0	2.2	2.1	0.0	3.4	3.3	0.0	2.3	3.5	0.0	2.3	3.3	0.0	4.2	2.4	0.0	3.6
b. 0 to 5 community	3.8	2.5	0.0	3.0	2.3	0.0	3.8	3.8	0.0	3.6	3.0	0.0	3.1	3.5	0.0	3.8	3.3	0.0
39. Community engagement by school																		
	3.3	1.5	0.0	2.6	2.2	0.0	2.8	3.1	0.0	2.4	4.0	0.0	2.8	2.7	0.0	2.6	2.9	0.0

a. 0 to 5 managers b. 0 to 5 community	3.0	2.8	2.0	1.3	2.0	3.3	2.0	2.5	3.9	2.6	3.8	3.4	3.0	2.6	3.4	2.4	3.5	3.1
40. Quality of construction																		
a. 0 to 5 teacher	1.1	1.9	0.0	1.3	1.7	0.0	1.9	2.1	0.0	1.9	1.7	0.0	2.3	1.8	0.0	2.5	2.3	0.0
b. 0 to 5 managers	0.8	1.2	1.8	1.3	2.2	2.4	1.4	2.3	2.7	2.0	2.1	2.4	2.2	2.0	3.0	2.1	2.4	2.9
41. Quality cluster in-service																		
a. 0 to 5 managers	0.0	0.0	0.0	0.0	0.8	0.0	1.3	1.8	0.0	0.2	2.2	0.0	0.6	2.2	0.0	0.8	2.3	0.0
42. Effective use of TDCs																		
a. 0 to 5 teachers	1.0	2.0	0.0	2.0	2.7	0.0	2.3	1.8	0.0	2.5	2.3	0.0	1.9	2.8	0.0	3.0	2.3	0.0
b. 0 to 5 managers	2.0	2.4	2.3	3.1	2.8	3.1	3.3	0.0	0.9	3.6	2.7	3.4	2.9	3.2	3.4	2.9	1.8	3.3
43. District/zone communication																		
a. 0 to 5 managers	0.3	0.9	0.0	0.5	1.5	0.0	1.7	2.3	0.0	0.9	2.8	0.0	1.4	2.5	0.0	2.9	1.8	0.0
44. Stakeholder communication																		
a. 0 to 5 managers	0.7	1.3	0.0	0.3	2.3	0.0	2.3	1.8	0.0	0.4	2.6	0.0	1.3	2.5	0.0	1.8	1.8	0.0
b. 0 to 5 community	2.0	2.0	0.0	1.3	1.5	1.1	1.3	1.3	0.0	0.6	1.9	0.0	0.3	1.8	1.0	1.3	0.5	0.0
45. District/division communication																		
a. 0 to 5 managers	0.3	0.7	0.0	0.8	1.3	0.0	2.0	2.0	0.0	0.5	2.4	0.0	0.8	3.8	0.0	2.1	2.3	0.0
46. Accuracy-data collection/records																		
a. 0 to 5 teachers																		
b. 0 to 5 managers	0.3	0.9	0.0	0.4	1.5	0.0	1.2	1.6	0.0	0.8	2.5	0.0	0.4	2.0	0.0	1.4	2.6	0.0
	1.3	2.0	1.0	1.6	2.8	1.7	2.5	2.9	0.4	1.3	2.5	1.0	1.6	3.3	0.6	2.4	2.8	0.5
48. Use of DEPs																		
a. 0 to 5 managers	0.7	0.0	0.0	1.5	0.0	0.0	1.5	0.8	0.0	0.5	0.8	0.0	0.5	2.4	0.0	2.0	0.0	0.0
49. Transparency-finance & procurement																		
a. 0 to 5 managers	2.0	0.5	0.0	2.3	0.0	0.0	3.0	0.3	0.0	1.3	0.8	0.0	0.0	3.1	0.0	1.5	0.3	0.0
b. 0 to 5 community	3.0	1.3	0.0	1.5	0.0	3.3	3.0	1.1	1.8	0.3	0.0	0.0	4.5	1.3	0.0	0.8	0.3	0.0
54. Student positive attitude																		
a. 0 to 5 teacher	2.3	3.0	4.4	2.5	2.8	3.0	3.7	3.0	4.0	3.0	3.0	3.5	2.8	3.0	3.0	3.1	3.0	3.3
# of indicators up/indicators down (bold=up bold italics =down)	4 up / 16 down			9 up / 12 down			5 up / 24 down			3 up / 18 down			1 up / 20 down			3 up / 24 down		
N = 41 Median = 2.5	10	5	1	6	12	12	16	15	14	13	14	6	13	15	13	9	8	10
Number of indicators at 3 or above																		
Number of Indicators scored as zero	2	4	20	2	4	16	1	2	17	1	3	19	1	3	18	0	4	19

Table 6-4
Pilot Project Evaluation by Group (Secondary) – Post Pilot Survey

(Scores represent average score for the two pilot secondary schools in each pilot grouping)

Indicators	Group 6			Group 5			Group 4			Group 3			Group 2			Group 1		
	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot	Base Line	Mid Point	Post Pilot
26. Teaching/learning																		
a. 0 to 5 student	1.5	1.8	2.7	2.8	2.0	3.8	2.4	2.4	2.3	1.2	2.0	2.0	2.7	2.5	2.6	2.7	2.0	2.6
b. 0 to teacher	2.0	2.5	2.7	2.4	3.5	4.2	1.8	2.9	2.9	1.9	2.3	2.6	2.3	2.5	3.0	2.8	1.5	1.5
27. Instructional mtrl																		
a. 0 to 5 student	1.5	2.0	2.1	2.5	1.5	1.2	1.2	1.0	2.1	1.5	1.5	2.1	1.8	1.5	2.0	2.5	1.5	2.8
b. 0 to teacher	1.0	1.5	2.0	1.9	1.5	2.3	1.7	1.1	2.4	1.0	1.4	2.8	2.2	1.5	2.2	2.4	1.5	1.7
28. Continuous assess																		
a. 0 to 5 student	2.0	2.1	2.5	3.5	3.5	1.7	2.7	1.6	2.0	2.5	1.6	2.0	2.4	3.5	3.7	3.9	1.5	2.0
b. 0 to teacher	2.0	2.9	3.4	3.0	3.0	3.2	3.2	3.9	3.7	2.2	2.5	2.7	3.5	3.5	4.0	3.7	2.5	2.0
29. Use of lesson plan																		
a. 0 to 5 managers	3.0	2.0	1.5	1.5	1.5	3.3	2.0	2.9	2.7	0.0	2.5	3.0	4.0	2.5	3.8	0.0	3.5	0.0
30. Teacher in-service																		
a. 0 to 5 teachers	1.5	1.2	3.5	3.2	3.0	2.7	2.9	0.5	1.7	0.0	2.7	3.9	2.0	2.5	2.0	1.7	0.0	2.5
31. Use of recording tools																		
a. 0 to 5 teachers	2.5	3.1	2.8	4.0	2.0	3.5	3.5	2.6	3.7	3.8	1.5	2.7	4.0	2.0	3.1	3.8	2.5	1.9
b. 0 to managers	2.5	2.9	3.3	2.9	3.0	3.3	3.0	3.5	3.8	3.0	2.5	3.0	3.9	2.5	4.3	2.0	2.5	4.1
33. HIV awareness																		
a. 0 to 5 student	1.5	2.5	2.0	2.0	3.5	3.5	3.0	2.5	1.9	2.5	2.5	1.7	3.0	3.5	5.0	4.0	2.0	3.5
b. 0 to teacher	1.5	4.5	1.5	3.5	2.5	4.5	3.0	3.5	4.5	3.0	2.0	2.5	3.0	2.5	3.0	4.0	3.0	2.0
c. 0 to 5 community	1.5	1.5	2.5	3.0	3.0	4.0	3.5	1.5	3.0	2.5	1.5	2.3	4.0	3.0	4.3	3.0	3.0	1.0
34. Gender awareness																		
a. 0 to 5 student	3.5	3.8	3.2	2.5	3.0	2.8	2.8	3.6	2.8	3.0	2.0	2.5	3.5	4.0	3.3	3.5	3.5	1.8
b. 0 to teacher	2.5	3.3	3.5	3.3	3.5	3.8	3.0	3.8	4.0	2.5	3.0	2.4	3.5	3.5	3.5	3.5	3.0	0.0
c. 0 to 5 community	3.0	3.5	4.0	3.0	4.5	3.8	3.0	3.4	3.8	3.3	3.3	3.9	2.8	3.5	4.1	3.0	3.0	3.0
35. Quality head teacher in-service																		
a. 0 to 5 managers	3.0	3.4	3.9	3.3	2.5	3.7	2.9	1.5	3.0	2.0	3.2	3.4	1.5	2.5	3.9	1.7	3.0	2.4
36. School environment																		
a. 0 to 5 student	2.5	3.0	3.2	2.7	3.0	3.2	2.7	3.2	3.9	3.0	3.0	3.5	3.4	3.5	4.0	3.7	3.0	3.7
b. 0 to teacher	1.5	3.2	3.5	3.2	3.0	4.0	3.2	3.5	1.5	3.2	3.2	3.4	3.5	3.5	4.0	3.8	3.0	2.4
37. Parent support																		
a. 0 to 5 student	2.5	3.1	2.9	2.9	4.0	3.2	2.2	2.5	3.3	2.1	3.2	2.4	3.1	3.0	3.7	3.5	2.0	2.8
b. 0 to 5 teacher	2.0	2.7	3.1	2.3	2.0	2.7	2.1	1.7	0.5	2.4	2.0	2.3	2.5	3.0	3.7	1.2	2.0	1.3
c. 0 to 5 community	3.0	3.2	3.5	2.7	3.5	3.3	2.6	3.0	4.0	2.4	3.2	3.1	3.2	4.0	4.0	3.2	2.0	2.6
38. Active/relevant school committee																		
a. 0 to 5 managers	3.5	4.4	4.0	3.3	3.0	3.7	2.7	1.5	4.3	3.0	2.7	4.2	4.0	3.5	4.3	0.0	3.0	3.2
b. 0 to 5 community	3.5	4.4	3.7	2.4	3.4	2.8	2.9	2.9	4.2	2.3	2.9	4.0	4.2	4.0	4.3	3.8	2.0	2.5

39. Community engagement by school																			
a. 0 to 5 managers	3.0	2.8	3.5	3.5	3.0	3.3	1.8	1.5	3.0	2.3	2.3	3.8	3.0	3.5	3.8	1.3	2.5	3.5	
b. 0 to 5 community	2.0	1.8	4.0	2.0	3.0	2.3	1.5	2.0	3.8	3.0	3.0	2.6	4.2	4.0	4.3	1.8	0.5	0.5	
40. Quality of construction																			
a. 0 to 5 teacher	2.0	2.9	3.2	2.4	3.0	3.0	2.7	2.2	3.3	2.3	1.5	2.9	2.2	2.5	3.3	3.5	3.0	1.9	
b. 0 to 5 managers	2.0	2.0	3.2	2.7	2.5	2.5	2.7	3.2	3.4	2.7	2.4	2.7	2.2	2.5	2.7	2.0	3.0	3.7	
46. Accuracy-data collection/records																			
a. 0 to 5 teachers	0.0	2.5	4.3	1.5	1.5	0.0	0.3	1.5	1.5	1.3	1.5	1.5	1.8	2.5	0.8	0.8	1.5	0.0	
b. 0 to 5 managers	1.0	2.0	2.8	3.0	2.5	1.3	0.5	3.3	4.0	2.0	2.5	2.8	2.3	1.5	1.5	0.5	2.0	2.5	
48. Use of DEPs																			
a. 0 to 5 managers	1.0	0.0	0.0	0.5	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	
49. Transparency-finance/procurement																			
a. 0 to 5 managers	4.0	0.0	2.0	1.0	0.0	3.0	4.0	1.0	0.0	3.0	1.0	2.0	2.8	0.5	3.5	1.5	3.5	0.0	
b. 0 to 5 community	3.0	0.0	3.3	0.5	0.0	2.0	3.0	0.5	2.0	2.0	1.5	1.9	2.3	1.0	4.0	n/a	1.0	0.0	
54. Student positive attitude																			
a.. 0 to 5 teacher	3.5	4.3	2.0	3.3	3.5	3.8	3.5	3.5	3.8	3.5	2.5	3.8	3.0	3.5	4.0	3.5	3.0	2.0	
57. Attitude-science																			
a. 0 to 5 student	2.5	3.5	3.0	4.0	3.0	1.5	2.5	3.7	2.5	3.0	3.3	3.3	2.0	2.5	3.5	4.0	3.5	3.5	
b. 0 to 5 teacher	2.5	3.0	3.5	2.5	4.0	3.5	3.0	3.0	3.5	3.0	3.0	3.0	1.5	3.0	3.0	4.0	3.5	2.0	
# of indicators up/down (bold=up bold italics =down)	15 up / 4 down			7 up / 5 down			12 up / 6 down			7 up / 3 down			9 up / 1 down			5 up / 15 down			
N = 36 Median = 2.5																			
Number of indicators at 3 or above	11	15	21	15	22	22	13	13	20	13	10	14	18	18	28	18	16	8	
Number of Indicators scored as zero	1	3	1	0	1	2	1	1	1	2	1	1	1	1	1	2	2	6	

Whereas Table 6-1 aggregated data by district, tables 6-3 and 6-4 aggregated data according to pilot programs. More stakeholder groups were involved in the evaluation and the number of indicators increased due to the involvement of more stakeholder groups (N = 41 for primary, and N = 36 for secondary). The control group was introduced in this analysis. Results are quite different when compared to the district level analysis:

6.3.1 Analysis of Primary Education Results

- In all cases, including the control group, the number of indicators that increased significantly was far less than the number that decreased between baseline and post pilot. For example, Group six had four indicators increase by at least one point while 16 decreased by at least one point between baseline and post pilot. The Group One (control group) had three increase and 24 decrease). This may suggest that different raters were used during post pilot than during baseline.
- The number of score at 3 or higher shows no pattern. In some cases, the number decreased, while in others they increased, and others stayed the same. This is possible since the groups are reflecting on the specific pilots they selected. Therefore, the control should be about the same which it is (9 at baseline and 10 at post pilot). The largest decrease was in Group 6 decreasing from 10 to 1 which is also predictable given the pilot configuration. Given that the balance of pilots contain some form of in-service and/or public awareness, indicator changes will vary accordingly. Group 5 focuses on in-service for teachers and managers and one would predict the greatest increase in indicator scores for this Group. Again, this is the case with the number of indicators having score of 3 or above increasing from 6 to 12.
- The biggest surprise, one which supports the prediction that different raters were used, is that the number of zeros used in scoring indicators increased dramatically between baseline and post pilot. The number of zeros used to evaluate all six groups during baseline was 7 while the number of zeros used in the post pilot increased to 107. There is clearly something wrong with this evaluation and further investigation is necessary.

6.3.2 Analysis of Secondary Education Results

- The pattern of up and down indicators is much less dramatic at the secondary level. In fact, all experimental groups (groups 2 to 6) had more indicators increase by one or more points than decreased. The control group was the only group where the reverse is true. Some changes were large such as in group six where 15 indicators increased by one or more points between baseline and post pilot while only four decreased (see table 4-4 for results).
- The number of indicators with scores of 3 or higher followed the same pattern where all experimental groups increased between baseline and post pilot and the control group decreased (group 6 = 11 to 21, group 5 = 15 to 22, group 4 = 13 to 20, group 3 = 13 to 14, group 2 = 18 to 28, and group 1 18 to 8).
- The zero problem identified in primary education does not exist in secondary education. At baseline, seven zeros were used for all six groups while at post pilot 12 zeros were used, the largest number of six zeros used for group one. This is in line with previous comments.

A more detailed analysis of groups is provided in the following pages.

6.4 Analysis Of Primary And Secondary Education By Groups

Table 6-5 contains results related to primary education by groups (the following explanation also applies to table 6-6 and will not be repeated). The first column contains the group number; the categories of activities in relation to PIF categories; the types of pilots selected and the specific outputs of the pilot activities. The second column contains a list of indicators. The selection of these indicators is based on a prediction of what process and outcome indicators would be impacted by the pilot activities. If, for example, classroom furniture were a selected pilot activity, what processes and outcomes would be achieved by this intervention? The column reflects this selection based on the evaluator's understanding of education systems. This accounts for why different process and outcome indicators have been selected for each of the different groups.

The next four columns contain average results of the surveys of baseline and post pilot data for the schools participating in the pilot (see table 6-2 for participating schools). Results of the control groups for the same indicators at the same points in time are also provided. Certain outcomes, such as absence rate and student achievement, were identified that are not qualitative but essential to the analysis of how processes might impact on outcomes. Although quantitative data appear in another chapter, results were imported on to these tables to assess the potential relationships between processes and outcomes. Recall that the systems model suggests a complex relationship between inputs, processes, outputs and outcomes. A visual analysis (as opposed to a statistical analysis) is being used to assess these relationships. The last column contains a narrative description of the prediction explaining why certain indicators were selected, and a brief analysis either supporting or contravening the prediction.

Since the analysis lack statistical rigor, it is possible to challenge the analyses and conclusions. Although true, there is still much that can be gleaned from the results and will suggest topics for further research using appropriate measurement techniques.

Table 6-5: Predictions for Each Pilot Group for Primary Educational Indicators

Primary Level						
Groups and Pilot Activities	Affected Indicators by Statement from Survey Tools	Pilot Schools		Control Schools		Prediction and Analysis
		Base	Post	Base	Post	
Group 6	Use of data collection tools-teachers (process)	2.0	1.8	3.8	3.1	<p>PREDICTION:</p> <ul style="list-style-type: none"> Not only will there be an increase in the number of classrooms (output), but the quality of the construction should be seen by stakeholders as improved. With training in data collection and reporting, schools will increase the use of school registers while data accuracy will increase. Data collection and reporting will improve communication links from school through division Less crowding in classrooms will increase motivation resulting in improved attendance by students and teachers. <p>ANALYSIS:</p> <ul style="list-style-type: none"> While teachers perception of quality construction decreased, managers increased significantly. Control group followed a similar pattern. Use of registers remained static while accuracy was seen as decreasing slightly in experimental group and significantly in the control group. Communications at different levels all decreased and more so in the control group. Classroom crowding was reduced but still remains too high. In the control group reduction in crowding due to reduction in enrollment (32%) while in the experimental group new construction only partially offset the dramatic increase (42%) in enrollment continuing overcrowding in the classroom. Still student attitude and behavior improved in the experimental group but not the control group. The absence rate, however, increased for the experimental group and decreased for the control group <p>In general, the predictions were not supported by the data except for the first and last predictions and in these cases only partially.</p>
Access and Planning	Use of data collection tools-managers (process)	2.5	2.4	3.2	0.0	
	Quality of school construction-teachers (process)	<i>1.1</i>	0.0	2.5	0.0	
Classroom construction	Quality of school construction-managers (process)	0.8	1.8	2.1	2.9	
	Improved communic. Links-zone/district (process)	0.3	0.0	2.9	0.0	
Data accuracy in-service	Improved communic. Links-district/stkholders (process)	<i>0.7</i>	0.0	1.8	0.0	
	Improved communic. Links-district/division (process)	0.3	0.0	2.2	0.0	
Classroom construction	Accuracy of data collection/reporting-teacher (process)	0.3	0.0	1.4	0.0	
	Accuracy of data collection/reporting-mnnger (process)	1.3	0.9	2.4	0.5	
Data accuracy in-service	Student absence rate (outcome)	1.8	3.0	3.4	2.8	
	Attitude/behavior change (outcome)	2.3	4.0	3.1	3.3	
Data accuracy in-service	Teacher absence rate (outcome)	NA	NA	NA	NA	
	Student to classroom ratio (outcome)* enrollment data	199 497	166 703	108 2062	92 1392	
<p>Bold = significant increase</p> <p><i>Bold Italics = significant decrease</i></p> <p><i>*mid point data used since baseline not available</i></p> <p><i>NA = not Available</i></p>						

Group 5						
Quality and Management	Teaching/learning interaction-students (process)	1.7	2.9	2.1	3.5	PREDICTION:
	Teaching/learning interaction-teachers (process)	1.4	2.9	2.3	3.0	
Teacher in-service	Use of instructional materials-students (process)	1.3	1.7	2.0	2.5	<ul style="list-style-type: none"> • With an emphasis on in-service for teachers and education managers, the in-service is assumed to focus on improving classroom and school approach to learning but also include HIV, gender training. This will lead to changes in how teaching/learning takes place in the classroom and school heads improve instructional leadership.
	Use of instructional materials-teachers (process)	1.6	2.4	2.3	2.4	
Educational manager in-service	Use of continuous assessment-students (process)	2.3	2.0	2.8	2.4	<ul style="list-style-type: none"> • As a result of in-service for managers and to a lesser extent teachers, the school environment should improve as well as linkages with parents through school committees and their involvement with students at home.
	Use of continuous assessment-teachers (process)	2.1	3.2	3.4	2.4	
	Use of lesson plan (process)	2.1	3.0	2.9	3.4	<ul style="list-style-type: none"> • All in-service programs for teachers, and managers should be seen as improved as well as increased use of TDCs.
	Teacher in-service quality (process)	1.3	3.5	2.6	3.8	
	HIV awareness-students (process)	2.5	0.0	2.5	0.0	<ul style="list-style-type: none"> • Various administrative processes should improve including accuracy of data, use of DEPs and transparency.
	HIV awareness-teachers (process)	2.3	2.0	2.5	2.8	
	HIV awareness-community (process)	2.3	2.5	1.8	2.5	<ul style="list-style-type: none"> • Outcomes should improve. Student achievement should benefit and teacher content knowledge should also increase.
	Gender awareness-students (process)	2.6	0.0	3.0	0.0	
	Gender awareness-teachers (process)	3.1	3.2	2.7	2.7	<ul style="list-style-type: none"> • Other indicators such as absence, dropout and repetition rates should be reduced.
	Gender awareness-community (process)	3.6	3.3	3.3	3.1	
	Head teacher in-service quality (process)	2.8	0.0	2.3	0.0	<ul style="list-style-type: none"> • Inspectors should increase the number of visits they make to schools.
	School environment improved-students (process)	3.0	0.0	3.1	0.0	
	School environment improved-teachers (process)	2.9	2.8	3.3	2.9	ANALYSIS:
	Parent support at home-student (process)	0.0	2.8	2.9	3.3	
	Parent support at home-teacher (process)	2.9	0.0	2.0	0.0	<ul style="list-style-type: none"> • As reported by students and teachers, the teaching/learning process improved significantly for the experimental group but not the control group. HIV and gender awareness did not show an improvement.
	Parent support at home-community (process)	2.5	3.6	2.9	1.8	
	School committees-managers (process)	2.1	3.4	2.4	3.6	<ul style="list-style-type: none"> • While the control group shows a definite trend downward for most indicators, the results are mixed in the experimental group, with different stakeholder groups responding up or down in no pattern. When comparing experimental to control group there seems to be a pattern of improvement for the experimental group but not a conclusive one.
	School committees-community (process)	3.0	0.0	3.8	0.0	
	Community engagement-managers (process)	2.6	0.0	2.6	0.0	<ul style="list-style-type: none"> • Teacher in-service increased significantly for experimental and control groups while decreasing to zero for both groups of head teachers. PEAs indicated an increase for experimental and decrease for control groups. This may be explained by the fact that head teachers received in-service during the first year and the teachers receiving in-service the second year. TDC use was perceived as moving down which cannot be explained.
	Community engagement-community (process)	1.3	3.2	2.4	3.1	
	PEA in-service quality (process)	0.8	1.3	0.8	0.0	<ul style="list-style-type: none"> • With the exception that the community saw a significant increase in transparency, all other indicators moved down for the experimental group while all indicators moved down significantly for the control group.
	Effective use of TDCs-teachers (process)	2.0	0.0	3.0	0.0	
	Effective use of TDCs-managers (process)	3.1	3.1	2.9	3.3	
	Accuracy of data collection-teacher (process)	0.4	0.0	1.4	0.0	
	Accuracy of data collection-manger (process)	1.6	1.6	2.4	0.5	
	Use of DEPs by managers (process)	1.5	0.0	2.0	0.0	
	Transparency of finances/proc-managers (process)	2.3	0.0	1.5	0.0	
	Transparency of finances/proc-community (process)	1.5	3.3	0.8	0.0	
	Achievement-standard 4 math (outcome)	40.0	32.6	45.9	42.5	
	Achievement-standard 4 English (outcome)	8.5	9.2	23.4	23.7	
	Achievement-standard 6 math (outcome)	17.0	12.6	18.9	21.8	
	Achievement-standard 6 English (outcome)	28.1	21.4	34.5	37.5	
	Absence rate (outcome)	6.5	3.8	3.4	2.8	
	Dropout rate (outcome)	10.6	7.8	19.8	13.7	
	Repetition rate (outcome)	16.1	17.6	28.9	13.5	
	Attitude/behavior change (outcome)	2.5	3.0	3.1	3.3	
	Student discipline case rate (outcome)	NA	NA	NA	NA	
	teacher absence rate (outcome)	NA	NA	NA	NA	
	Teacher subject knowledge (outcome)	88.3	90.6	88.4	91.4	
	Teacher discipline case rate (outcome)	NA	NA	NA	NA	

	Number of inspections (outcome)	2.7	5.3	5.0	1.8	<ul style="list-style-type: none"> • For the most part achievement test scores moved down significantly while moving up significantly for the control group. Teacher competency moved up for both groups. Student achieved behaved in reverse of the prediction, and the fact that teachers scored above 90% cannot be explained to their lack of competency. • Student absence and dropout rates moved down significantly and repetition stayed about the same for the experimental group with a similar pattern for the control group. This met the prediction. • While the number of inspections almost doubled for the experimental group, it dropped significantly for the control group as predicted. <p>Results are mixed with achievement scores showing the most disappointing results. With teaching/learning process scores showing significant improvement and with high teacher competency, the low scores are inexplicable. It is assumed that HIV and gender were not included in the teacher or head teacher in-service. Increased visits by PEAs to schools could be attributed to their in-service program. The erratic nature of the balance of indicators does not support the balance of predictions.</p>
--	---------------------------------	-----	-----	-----	-----	---

Group 4		1.8	2.4	2.1	3.5	
Quality, Gender and Access	Teaching/learning interaction-student (process)	2.0	3.3	2.3	3.0	PREDICTION:
	Teaching/learning interaction-teacher (process)	2.0	2.8	2.0	2.5	
Public awareness campaign	Use of instructional materials-student (process)	1.9	2.6	2.3	2.4	<ul style="list-style-type: none"> With an emphasis on in-service for teachers, it is assumed the focus is on improving classroom approach to learning but also include HIV, gender training. This will lead to changes in how teaching/learning takes place in the classroom. Teachers should rate in-service higher. Since training normally takes place at TDCs their use should be rated higher.
	Use of instructional materials-teacher (process)	3.0	2.0	2.8	2.4	
Sanitation construction	Use of continuous assessment-student (process)	2.7	3.5	3.4	2.4	<ul style="list-style-type: none"> Public awareness should also increase participation of parents on school committees and rate HIV and gender awareness more highly.
	Use of continuous assessment-teacher (process)	2.6	2.9	2.9	3.4	
Teacher in-service	Use of lesson plan (process)	2.3	3.9	2.6	3.8	<ul style="list-style-type: none"> Student and teacher achievement should increase as a direct result of teachers changing classroom behavior.
	Teacher in-service quality (process)	3.0	0.0	2.5	0.0	
	HIV awareness-students (process)	2.7	3.3	2.5	2.8	<ul style="list-style-type: none"> As a result of changed teaching/learning process and improved sanitation, student indicators related to absence, dropout and repetition should decrease.
	HIV awareness-teachers (process)	3.3	3.8	1.8	2.5	
	HIV awareness-community (process)	3.0	0.0	3.0	0.0	
	Gender awareness-students (process)	2.7	3.5	2.6	2.7	
	Gender awareness-teachers (process)	3.3	3.8	3.3	3.1	
	Gender awareness-community (process)	0.0	3.6	2.9	3.3	
	Parent support at home-student (process)	3.2	0.0	2.0	0.0	
	Parent support at home-teacher (process)	2.9	3.5	2.8	1.8	
	Parent support at home-community (process)	3.3	2.4	2.4	3.6	
	Relevant school committees-managers (process)	3.8	0.0	3.8	0.0	
	Relevant school committees-community (process)	2.8	0.0	2.7	0.0	
	Community engagement-managers (process)	2.0	3.9	2.4	3.1	
	Community engagement-community (process)	2.3	0.0	3.0	0.0	
	Effective use of TDCs-teachers (process)	3.3	0.9	2.9	3.3	
	Effective use of TDCs-managers (process)	3.0	0.0	1.5	0.0	
	Transparency of finances/procur.-managers (process)	3.0	1.8	0.8	0.0	
	Transparency of finances/proc.-community (process)	34.7	35.6	45.9	42.5	
	Achievement-standard 4 math (outcome)	20.3	24.7	23.4	23.7	
	Achievement-standard 6 math (outcome)	17.5	18.7	18.9	21.8	
	Achievement-standard 6 English (outcome)	38.5	37.9	34.5	37.5	
Absence rate (outcome)	4.6	1.0	3.4	2.8		
Dropout rate (outcome)	7.9	6.0	19.8	13.7		
Repetition rate (outcome)	15.7	21.5	28.9	13.5		
Attitude/behavior change (outcome)	3.7	4.0	3.1	3.3		
Student discipline case rate (outcome)	NA	NA	NA	na		
Teacher absence rate (outcome)	NA	NA	NA	na		
Teacher subject knowledge (outcome)	85.7	92.2	88.4	91.4		
Teacher discipline case rate (outcome)	NA	NA	NA	NA		
					<p>ANALYSIS:</p> <ul style="list-style-type: none"> The comparison between experimental and control group is very similar and there is little significant change either up or down. Teachers did rate in-service significantly higher, but so did the control group. Use of TDCs was rated significantly lower which is inexplicable given that training took place there. Public awareness did not show a significant increase in HIV or gender scores. Scores related to various community indicators were erratic some moving up and some down for the experimental group but significantly down for most indicators for the control group. Three of the four student achievement scores increased only slightly and increased significantly for English 4. The control group showed a similar pattern. Teacher competency scores were high and increased significantly for both groups. While absence and dropout rates declined significantly, repetition rates increased for the experimental group. The control group scores were slightly better. <p>There appear to be few changes in how teachers conduct teaching/learning and there were no major increases in learning even with high teacher competency. The control group tends to support this. The public awareness campaigns seem to have had little impact on the indicators measured here. While absence and dropout rates moved in the predicted direction, the significantly increased repetition rate cannot be explained, especially since the control group rate dropped significantly.</p>	

Group 3						
Access and Quality	Teaching/learning interaction-student (process)	2.3	1.8	2.1	3.5	PREDICTION:
	Teaching/learning interaction-teacher (process)	2.5	2.9	2.3	3.0	
Furniture procurement	Use of instructional materials-student (process)	1.8	1.9	2.0	2.5	<ul style="list-style-type: none"> With an emphasis on in-service for teachers, it is assumed the focus is on improving classroom approach to learning. This will lead to changes in how teaching/learning takes place in the classroom. Teachers should rate in-service higher. Since training normally takes place at TDCs their use should be rated higher.
	Use of instructional materials-teacher (process)	2.1	2.5	2.3	2.4	
Teacher in-service	Use of continuous assessment-student (process)	3.1	1.9	2.8	2.4	<ul style="list-style-type: none"> More and possibly better quality furniture will create a better classroom learning environment, leading to improved student attendance indicators. Student and teacher achievement should increase as a direct result of teachers changing classroom behavior. As a result of changed teaching/learning process, student indicators related to absence, dropout and repetition should decrease.
	Use of continuous assessment-teacher (process)	3.6	1.9	3.4	2.4	
	Use of lesson plan (process)	3.1	2.8	2.9	3.4	
	Teacher in-service quality (process)	1.1	3.5	2.6	3.8	
	Use of TDCs-teachers (process)	2.5	0.0	3.0	0.0	
	Use of TDCs-manager (process)	3.6	3.4	2.9	3.3	
	Achievement-standard 4 math (outcome)	31.2	35.6	45.9	42.5	
	Achievement-standard 4 English (outcome)	9.4	12.1	23.4	23.7	
	Achievement-standard 6 math (outcome)	15.6	17.2	18.9	21.8	
	Achievement-standard 6 English (outcome)	30.3	23.4	34.5	37.5	
	Absence rate (outcome)	1.8	3.7	3.4	2.8	
	Dropout rate (outcome)	9.2	24.1	19.8	13.7	
	Repetition rate (outcome)	13.7	33.1	28.9	13.5	
	Attitude/behavior change (outcome)	3.0	2.8	3.1	3.3	
Student discipline case rate (outcome)	NA	NA	NA	NA		
Teacher absence rate (outcome)	NA	NA	NA	NA		
Teacher subject knowledge (outcome)	93.7	94.4	88.4	91.4		
Teacher discipline case rate (outcome)	NA	NA	NA	NA		
					<p>ANALYSIS:</p> <ul style="list-style-type: none"> Little change is noted related to classroom practices. While two indicators for the experimental group decreased significantly and none moved up, the control group reported one up and one down at a significant level. Rating for TDC use was also disappointing. Student achievement followed a similar pattern between the experimental and control groups where one indicator moved up and one down. Teacher competency scores were very high, but given teachers and students reported little change in classroom behavior, one would not expect an increase in achievement. Absence, dropout and repetition rates actually increased significantly for the experimental group, with the opposite condition existing for the control group, completely opposite to predications. <p>In-service appears not to have led to behavior change in the classroom and furniture did not appear to have an impact on absence rates or other student outcomes. The control group performed much better overall than did the experimental group suggesting no relationship between interventions and improvement of processes and outcomes.</p>	

Group 2						
Quality, Gender and Access	HIV awareness-students (process)	2.8	0.0	2.5	0.0	PREDICTION:
	HIV awareness-teachers (process)	3.3	2.3	2.5	2.8	
Public awareness campaign	HIV awareness-community (process)	3.5	3.5	1.8	2.5	<ul style="list-style-type: none"> The public awareness campaigns should involve communities in more facets of school activities including improved school committee activities. Construction of latrines should reduce negative indicators involving students, especially dropout rate for girls. Although less of an impact on achievement it is expected that parent participation in more school activities will lead to improved student achievement.
	Gender awareness-students (process)	2.4	0.0	3.0	0.0	
	Gender awareness-teachers (process)	2.8	3.1	2.6	2.7	
	Gender awareness-community (process)	3.4	3.8	3.3	3.1	
Sanitary construction	Parent support at home-student (process)	3.3	3.3	2.9	3.3	ANALYSIS: <ul style="list-style-type: none"> The only significant change among the experimental group is downward for HIV, gender, parent support, school committees, community engagement, and transparency. This suggests that no public awareness activities impacted on students, teachers, managers or community. Absence rate dropped significantly but no data were available for the post pilot results for dropout or repetition rates. Control group indicators improved. Achievement improved in three cases, one significantly, and decreased significantly in one case for the experimental group. The control group demonstrated a similar pattern. <p>Results of the public awareness campaign are disappointing. Improvement of sanitation may have a relationship to absence rates, and this is the only positive prediction supported by the data. It is possible that the interventions have only a modest relationship with achievement which may be explained by the results.</p>
	Parent support at home-teacher (process)	1.8	0.0	2.0	0.0	
	Parent support at home-community (process)	3.3	3.7	2.8	1.8	
	School committees-managers (process)	3.3	4.2	2.4	3.6	
	School committees-community (process)	3.1	0.0	3.8	0.0	
	Community engagement-managers (process)	2.8	0.0	2.7	0.0	
	Community engagement-community (process)	3.0	3.4	2.4	3.1	
	Transparency of finances/proc-managers (process)	0.0	0.0	1.5	0.0	
	Transparency of finances/proc-community (process)	4.5	0.0	0.8	0.0	
	Achievement-standard 4 math (outcome)	28.4	31.2	45.9	42.5	
	Achievement-standard 4 English (outcome)	21.8	22.5	23.4	23.7	
	Achievement-standard 6 math (outcome)	16.8	21.1	18.9	21.8	
	Achievement-standard 6 English (outcome)	50.4	45.1	34.5	37.5	
	Absence rate (outcome)	10.0	1.8	3.4	2.8	
Dropout rate (outcome)	5.5	NA	19.8	13.7		
Repetition rate (outcome)	11.1	NA	28.9	13.5		
Attitude/behavior change (outcome)	2.8	3.0	3.1	3.3		
Student discipline case rate (outcome)	NA	NA	NA	NA		
Teacher absence rate (outcome)	NA	NA	NA	NA		

6.4.1 Primary Level Results

Results are mixed. As will be discussed at the end of the evaluation, questions have arisen related to the research methodology and accuracy of data. Since this is fundamental in having confidence in the results, all results are now brought into question. One cause may be related to the fact that pilots were implemented in two phases across two years. The baseline study may be accurate, but when comparing results to the post pilot a problem arises if in-service were conducted during the first phase and not the second. The focus group may have responded in the year when in-service was not conducted creating a downward trend in evaluation. With this caveat, the following observations can be made about the five pilot primary school groups:

- Three pilots included teacher in-service. The results are not consistent. While in one case pilot process indicators moved up, in another they moved down, and in the third, they remained stable. The control groups were similarly erratic. These processes are critical in improving learning so the impact on achievement should be negligible. In fact, for the pilot where the process indicators moved up, achievement scores moved down. These results suggest no relationship between teacher in-service and student achievement.
- A second set of process indicators relate to in-service of managers. This was piloted by only one group. By training managers, especially school managers, there should be a number of impacts on other processes and outcomes related to teachers, community, the school environment, and certain administrative functions. This demonstrates how important the role is of the school leader in structuring an environment that improves learning. In this pilot, many of the process and outcome measures moved downward; however, there were some significant changes upward. It was in this pilot, coupled with teacher in-service, that classroom practices improved significantly, some community indicators moved upward, and student absence and dropout rates were reduced. This may suggest that the combination of in-service for teachers and school heads of the same schools could have multiple impacts on a number of processes and educational outcomes.
- A third in-service, data collection and reporting, was selected by one pilot region. The results of the relationship between this intervention and various processes such as increase in accuracy of data and teachers' use of tools such as registers, does not appear to exist.
- Public awareness campaigning was selected by two pilot groups. Since there was a range of campaigns from dropout prevention to HIV and gender awareness, results could be diluted since sample schools within a group might have chosen different campaigns. In almost every case where other process and outcome indicators should have been impacted, the data suggest that there was no impact of the campaigns. This may be due to the variation within a group as to the types of campaigns selected.
- While the above pilots are all process-related, three other pilot activities involved construction of classroom, construction of sanitation facilities, and procurement of furniture. These interventions focused on access rather than quality and should have impacted on indicators related to access and student motivation. In two of the three pilots, there appear to be a high correlation between these inputs and student attitude, absence and dropout rates but not repetition rates. This is one area of pilot activities where there appears to be a consistent correlation.

Secondary level results follow:

Table 6-6: Predictions for Each Pilot Group for Secondary Educational Indicators

		Secondary Level				
Group 6 <u>Access</u> School construction Classroom furniture procurement	Quality of school construction-teachers (process)	2.0	3.2	1.3	3.5	PREDICTION: <ul style="list-style-type: none"> • Not only will there be an increase in the number of classrooms, but the quality of the construction will be seen by stakeholders as improved. • More and possibly better quality furniture will create a better classroom learning environment, leading to improved student attendance indicators. Less crowding in classrooms will increase motivation resulting in improved attendance by students and teachers. ANALYSIS: <ul style="list-style-type: none"> • The four indicators for which measures existed showed a significant increase. The student to classroom ratio shows a decrease by half in the experimental group while constant for the control group. Enrollments in both groups remained stable. • While absence rate for the experimental group reduced significantly, the control group absence rate was almost eliminated and it is difficult to explain this. There appears to be a correlation between improving the quality of the school environment and attendance.
	Quality of school construction-managers (process)	2.0	3.2	1.8	0.5	
	Student absence rate (outcome)	7.9	3.2	9.5	0.5	
	Teacher absence rate (outcome)	NA	NA	NA	NA	
	Student to classroom ratio (outcome)	74.8	37.4	16.1	16.1	
Enrollment data*	132	126	127	134		
Bold = significant increase Bold Italics = significant decrease						
*Used mid point data since baseline was not available						

Group 5		2.8	3.8	2.7	2.6	PREDICTION:
Quality and Access	Teaching/learning interaction-student (process)	2.4	4.2	2.8	1.5	
	Teaching/learning interaction-teacher (process)	2.5	1.2	2.5	2.8	
Instructional materials procurement	Use of instructional materials-student (process)	1.9	2.3	2.4	1.7	<ul style="list-style-type: none"> This combined with the purchase of materials for teachers to use in the classroom should improve instruction leading to higher achievement by students. More and possibly better quality furniture will create a better classroom learning environment, leading to improved student attendance indicators.
	Use of instructional materials-teacher (process)	3.5	1.7	3.9	2.0	
	Use of continuous assessment-student (process)	3.0	3.2	3.7	2.0	
Teacher in-service	Use of continuous assessment-teacher (process)	1.5	3.2	0.0	0.0	<p>ANALYSIS:</p> <ul style="list-style-type: none"> Of the seven indicators measuring classroom practices, five increased – three at a significant level while two decreased significantly in the experimental group. There was little variation between baseline and post pilot for the control group. Results tend to support the prediction. There was no change in teachers' perception of in-service which was unexpected. There was minimal change up or down in achievement scores for the experimental and control groups. Teacher competency for the experimental group dropped significantly suggesting that a different teacher or teachers' were involved in the baseline and post pilot. Data demonstrate that prior to interventions there were more students than textbooks while at the completion there were more books than students for the experimental group. No data were available for the control group. The student absence rate decreased the student attitude rating increased, supporting the prediction.
	Use of lesson plan (process)	3.2	2.7	1.7	2.5	
Classroom furniture procurement	Teacher in-service quality (process)	9.1	9.6	10.9	11.3	<p>Although in-service, teacher competency and an increased number of textbooks should have improved achievement scores, this did not happen. Worse yet, scores are exceedingly low suggesting there might have been something wrong with the tests themselves.</p>
	Achievement-form 1 math (outcome)	37.0	34.2	41.0	42.3	
	Achievement-form 1 English (outcome)	11.1	6.2	12.4	14.1	
	Achievement-form 3 math (outcome)	47.5	28.5	47.1	58.8	
	Achievement-form 3 English (outcome)	3.4	3.0	9.5	0.5	
	Absence rate (outcome)	NA	NA	NA	NA	
	Dropout rate (outcome)	NA	NA	NA	NA	
	Repetition rate (outcome)	3.5	2.0	3.5	2.0	
	Attitude/behavior change (outcome)	NA	NA	NA	NA	
	Student discipline case rate (outcome)	NA	NA	NA	NA	
Teacher absence rate (outcome)	82.0	66.3	81.9	88.5		
Teacher subject knowledge (outcome)	NA	NA	NA	NA		
Teacher discipline case rate (outcome)	1.5	0.9	NA	NA		
Student to textbook ratio (outcome)						

Group 4 <u>Planning and Access</u> Data collection & reporting Classroom furniture procurement	Use of data collection tools-teachers (process)	3.5	3.7	3.8	1.9	PREDICTION: <ul style="list-style-type: none"> Data collection and reporting will improve communication links from school through division, although no process measures were made at the secondary level. More and possibly better quality furniture will create a better classroom learning environment, leading to improved student attendance indicators. ANALYSIS: <ul style="list-style-type: none"> There was little change reported in the use of data collection tools in the classroom, data accuracy scores increased dramatically. Although there was little change in student attitude, student absences increased significantly, contrary to prediction. <p>Although teachers and managers perceive that accuracy of data has improved, this remains to be evaluated objectively. There was no outcome measure identified to do this. The introduction of new furniture does not necessarily account for why absence and attitude scores changed. For some reason communication indicators were not evaluated, even though evaluated at the primary level.</p>
	Use of data collection tools-managers (process)	3.0	3.8	2.0	4.1	
	Quality of school construction-teachers (process)	2.7	3.3	3.5	1.9	
	Quality of school construction-managers (process)	2.7	3.4	2.0	3.6	
	Accuracy of data collection-teacher (process)	0.3	1.5	0.8	0.0	
	Accuracy of data collection-manger (process)	0.5	4.0	0.5	2.5	
	Student absence rate (outcome)	1.5	4.2	9.5	0.5	
	Attitude/behavior change (outcome)	3.5	3.8	3.5	2.0	
	Teacher absence rate (outcome)	NA	NA	NA	NA	

Group 3		1.2	2.0	2.6	2.6	PREDICTION:
Quality and Access	Teaching/learning interaction-student (process)	1.9	2.6	3.0	1.5	<ul style="list-style-type: none"> • Science materials will lead to better equipped laboratories; however, this may not lead to better teaching unless in-service accompanies the installation of new kits.
	Teaching/learning interaction-teacher (process)	1.5	2.1	2.0	2.8	<ul style="list-style-type: none"> • This should impact on students' attitude towards science.
	Use of instructional materials-student (process)	1.0	2.8	2.2	1.7	<ul style="list-style-type: none"> • Also, math scores may improve since math is frequently used in the science curriculum.
	Use of instructional materials-teacher (process)	2.1	1.1	10.9	11.3	<ul style="list-style-type: none"> • More and possibly better quality furniture will create a better classroom learning environment, leading to improved student attendance indicators.
Science Kits and supplies procured	Achievement-form 1 math (outcome)	2.4	1.6	12.4	14.1	
	Achievement-form 3 math (outcome)	NA	NA	9.5	0.5	
	Student absence rate (outcome)	3.5	3.8	3.5	2.0	
Classroom furniture procurement	Attitude/behavior change (outcome)	3.0	3.3	4.0	3.5	
	Attitude toward science-student (outcome)	3.0	3.0	4.0	2.0	
	Attitude toward science-teacher (outcome)	NA	NA	NA	NA	
	Teacher absence rate (outcome)					

- ANALYSIS:**
- Teachers perceive a much increased use of instructional materials. Other classroom practices perceived by both students and teachers increased also but not significantly. The control group show lower scores supporting the prediction.
 - Scores for attitude toward science remained the same for the experimental group but declined for the control group.
 - Experimental group math scores remained close to zero percent, demonstrating almost no mastery in forms 1 and 3. Control group scores were very low as well, but not as bad as the experimental group.
 - No scores were available for experimental group absence rate. Attitude score did improve slightly.

Since science testing was not conducted it was not possible to determine if kits might have any impact on learning. If math scores are used as a proxy, then the situation is extremely bad with respect to scientific learning. In fact, math scores across all six groups were very low, not one student demonstrating a grade of 60% or higher. There were an excessive number of students with a score of zero.

Group 2						
Management and Planning Assistance to educational managers	Use of data collection tools-teachers (process)	4.0	3.1	3.8	1.9	PREDICTION: <ul style="list-style-type: none"> The types of in-service will cover the range of functions from financial to internal and external leadership. This should have a positive impact on all indicators where managers at school to district level are involved including such issues as parents' involvement, data collection, and use of DEPs. If specific topics such as gender and HIV are included, then there should be a positive impact on awareness. Quality of in-service should have been rated high. Changes in management practices should lead to improved indicators for teachers such as a reduction in discipline cases. Inspectors should increase the number of visits they make to schools. ANALYSIS: <ul style="list-style-type: none"> One of the three gender and one of the three HIV scores improved significantly while the other two were stable. Most control group scores moved down significantly. There was little change and possible downward movement of scores related to data tools and accuracy. All scores involving community increased, significantly. Also, school environment ratings improved. By contrast control groups were lower when compared to the experimental group and many decreased. In-service was rated as high and increased significantly while the experimental group was rated lower. DEPs were not used at all by experimental or control groups. While student attitudes improved in the experimental group it decreased significantly in the control group. There was no impact on the rate of inspections. Most indicators behaved as predicted, with control group scores moving down rather than remaining stable. In-service for managers may have had a positive impact on community indicators. Since key measures for teachers were not useable (raw data not available) there was no way to investigate the possible relationship between manager in-service and selected teacher outcomes.
	Use of data collection tools-managers (process)	3.9	4.3	2.0	4.1	
	HIV awareness-students (process)	3.0	5.0	4.0	3.5	
	HIV awareness-teachers (process)	3.0	3.0	4.0	2.0	
	HIV awareness-community (process)	4.0	4.3	3.0	1.0	
	Gender awareness-students (process)	3.5	3.3	3.5	1.8	
	Gender awareness-teachers (process)	3.5	3.5	3.5	0.0	
	Gender awareness-community (process)	2.8	4.1	3.0	3.0	
	Head teacher in-service quality (process)	1.5	3.9	1.7	2.4	
	School environment improved-student (process)	3.4	4.0	3.7	3.7	
	School environment improved-teacher (process)	3.5	4.0	3.8	2.4	
	Parent support at home-student (process)	3.1	3.7	3.5	2.8	
	Parent support at home-teacher (process)	2.5	3.7	1.2	1.3	
	Parent support at home-community (process)	3.2	4.0	3.2	2.6	
	School committees-managers (process)	4.0	4.3	0.0	3.2	
	School committees-community (process)	4.2	4.3	3.8	2.5	
	Community engagement-managers (process)	3.0	3.8	1.3	3.5	
	Community engagement-community (process)	4.2	4.3	1.8	0.5	
	Accuracy of data collection-teacher (process)	1.8	0.8	0.8	0.0	
	Accuracy of data collection-manger (process)	2.3	1.5	0.5	2.5	
	Use of DEPs by managers (process)	0.0	0.0	0.0	0.0	
	Transparency of finances/proc-managers (process)	2.8	3.5	1.5	0.0	
	Transparency of finances/proc-community (process)	2.3	4.0	0.0	0.0	
	Attitude/behavior change (outcome)	3.0	4.0	3.5	2.0	
	Student discipline case rate (outcome)	NA	NA	NA	NA	
	Teacher absence rate (outcome)	NA	NA	NA	NA	
	Teacher discipline case rate (outcome)	NA	NA	NA	NA	
Number of inspections (outcome)	1	0	0	1		

6.4.2 Secondary Level Results

The introduction for primary education also applies here. There is a major difference between primary pilots and secondary pilots. Whereas in primary groups a majority of pilot activities related to processes (in-service and public awareness) the majority of pilots under secondary education were related to inputs (construction and procurement):

- Unlike primary education, there is little evidence that correlations exist between construction and procurement and outcome indicators related to student attitude, absence, dropout and repetition rates.
- In one pilot where instructional materials and furniture procurement was included with teacher in-service, process indicators related to classroom practice showed a significant correlation. The three interventions all focused on the classroom level. This combination of interventions may be the reason that teachers and students noted a significant increase in classroom practices indicators. This did not, however, translate into increased student achievement test scores. In a separate pilot where instructional materials for science and classroom furniture but not in-service there was no significant correlation between the interventions and classroom practices or student achievement.
- The one pilot intervention where in-service was provided to managers but no other interventions were provided, a variety of positive correlations were exhibited – a total of eight. Further, the control group showed a number of negative correlations – a total of 10. This combination suggests a possible relationship between manager in-service and various indicators at different levels of the system from the school to community. Instead several other pilot interventions at the primary and secondary levels provide some indication, that various combinations of these process and procurement interventions may have a significant impact on various student and teacher quality outcomes.

6.5 Conclusions, Lessons Learned And Recommendations

The focus of this chapter has been on evaluation of process and outcome indicators and their possible correlation to input and other process indicators. Unfortunately fundamental flaws make the analysis exceedingly difficult:

- **Flawed research design:** Normally, the research design is decided before project activities begin. The design includes controlling the experimental environment, computing errors of measurement, and correlating significance levels based on varying sample sizes. In this case NIPDEP's objectives were different; however, a statistically based research design was superimposed. Although used, appropriate statistical analyses were not applied such as computing standard deviations and regression analysis except for achievement test results. This design was inappropriate for the project model. Instead, a model should have been chosen which was compatible to the type of project where it was not possible to control for outside interferences. Often a case study approach would be used involving observations as the primary data collection technique. Since the research design was not compatible with the project design, results are highly suspect.

Recommendation: Capacity does not exist within the JICA/CERT/MIE team to apply appropriate research techniques to projects of this type. One can either change how future JICA projects are designed so they conform to a research study methodology or the research methodology can adapt to the JICA approach. In either case, expertise in testing and measurement is highly specialized and not available to meet either design requirement. If scientifically evaluated results are required from this or any other projects, research designs should enable the conduct of **holistic evaluation**. Holistic evaluation encompasses the complex nature of social sector evaluation utilizing quantitative and qualitative techniques some only recently being applied. The approach can be used in any type of social sector research including the current JICA-type project. For this reason, it is recommended that donors look at the possibility of funding a long term specialist in testing and measurement to be assigned to the tertiary institution in Zomba and seconded to MIE and CERT, with the dual mission of teaching education courses at the college and building capacity at MIE and CERT to assist donors in conducting social sector evaluation. Additional funding might also be provided for CERT and MIE to hire one individual dedicated to this type of research working closely with the international consultant. In this manner, institutional capacity can be built within

two local institutions and at the same time, help introduce more rigor to educational course for undergraduate and graduate students.

Possible Inaccurate Data Collection: An examination of various raw data tables suggests strange patterns. For example, the absence rate table containing baseline, mid point and post pilot scores varies dramatically with the mid point score spiking five times as high as the baseline and post pilot scores. Shown earlier, post pilot rater scores for process variables show an excessive use of zeros as compared to the baseline and mid point scores. These and other examples may suggest problems with instrumentation and data gatherers including focus group raters. The achievement tests, for example, were constructed based on the national curriculum; however, appropriate testing and measurement techniques were not applied to their construction. An item analysis would have determined the difficulty level for each question. This was not done. The exceedingly low scores related to secondary mathematics especially may suggest that the questions were too difficult even though they were based on the curriculum. A properly validated achievement test may have yielded very different results which provided the necessary dispersion from the mean score. The inter-rater technique used in the focus group analysis was developed as a way to reduce bias in scoring results. Nevertheless, the excessive number of zeros used in the post pilot evaluation suggests that different raters scored results quite differently, even though all raters received training. Finally, such quantitative variables as absence rates are more complicated to compute than imagined. The main problem is that inaccurate and inadequate data. Accurate students' records, especially on attendance, are not always kept at the school level (and when they are, they are not accurate) and as such computation of absence rate was not done in most cases.

Recommendation: The above recommendation applies here as well. Data collection tools require establishing validity and reliability especially if they are to be used to measure constructs such as processes used in this study. Reliability and validity require a specialized expertise. If none is available, then it is simpler to use quantitative indicators only which tends to reduce any research to a simple input-output research model. Since quality is an important issue, the more complex research methodology is required so as to measure quality. There are only two choices here: (1) use a simpler input-output model that does not rely on complicated instrumentation but which reduces the ability to assess quality; or, (2) implement more complicated instrumentation to capture appropriate quality data, requiring more demanding instrumentation. One way around this is to utilize instruments already developed for international application. These do exist and can be acquired to measure a variety of subject at different grade levels. The added advantage is that results can be compared to international counterparts. More difficult to acquire are those instruments used to measure quality. Many do exist, but their reliability and validity are established in different cultural settings and may not be appropriate for use in Malawi. This could be rectified by hiring a short term consultant specializing in testing and measurement to review instruments used for a variety of purposes in the field of education and select those that would be most appropriate for use in Malawi. Again, selected institutions such as CERT and MIE could receive training on the use, interpretation and reporting of results.

- **Possible correlations:** Even with the flaws, results may suggest certain patterns. The concept of providing in-service for teachers and managers, public awareness for community and instructional materials procurement may have the greatest impact on learning. This combination of pilots was not tested. In previous chapters results by district may have suggested certain correlations given that performance of input indicators and certain processes were high; however, when results were grouped by pilot groups these relationships appeared to be weaker. This could have resulted from the small sample size, by intervening variables (internal and external) or by the fact there was no significant relationship. Research by others suggest that certain conditions need to manifest themselves if quality is to improve. In other words, such indicators, if present, will serve as a proxy for quality. Such indicators as relevant community participation, parent/teacher meetings and parents helping with homework; teachers using certain classroom practices; maximum teacher to student ratios not to exceed 30 to 1; and a number of others, if present, will eventually lead to improvement of outcome indicators. In some cases, this impact will not be achieved for several years until different parts of the system have capacity built and then link these parts to each other. This means that teaching can be improved, school heads can be improved, and school councils can

be improved, but until they are linked to each other, school quality may be affected only on a small scale. In fact, this may what have been reflected in the data.

Recommendation: Further research is suggested using appropriate research techniques. Much more information is needed about educational quality in Malawi than currently exists or available in this report. It is recommended that a number of research topics be designed to focus on different input, process and outcome relationships. As suggested here, one might ask the question as to what impact does in-service have on learning in the classroom. One needs to embody some of the recommendations above to answer this question. First, a research design would need to operationally define indicators related to each variable in the research. Next, it is necessary to determine the size of the sample and the statistical techniques that will be employed to analyze results. Third, valid and reliable data collection instruments need to be used and trained data collectors must be available to use the instruments properly. Fourth, appropriate software may be needed to treat the data. Some of the newer software packages, such as the Statistical Package for Social Scientists (SPSS), go beyond regression analysis and may be more appropriate for use is conducting path analysis to determine the chain of relationships that exist in a complex social sector institution.

To test the impact of the systems model, future research may use the approach taken in quality school improvement research and examine different pilots. This was recommended in Chapter 5. The first may provide community building, the second, school head capacity building and the third, teacher development. The fourth would embody all three with a plan to link these elements. The fifth group would be the control group and all other variables would be controlled. The research would look at significant relationships when only one group receives development assistant for the three parts of the system that appear to have the greatest impact on learning – classroom, school and community. For this research to be successful, those conditions mentioned in the previous paragraph need to be adopted. Also, for this recommendation to be fully implemented, the first two recommendations must be considered and acted upon. Without such change in the social research arena, no meaningful research can take place.

FINAL OBSERVATIONS

The wealth of information in this study has been derived by a variety of data collection techniques and aggregated at the district and group levels. This allows the researchers to cross check results to see if they are consistent, thus improving reliability. In some cases, results of self reports by teachers and focus group interviews of teachers, managers, students and communities support each other while in other cases there is a significant difference in results. Patterns of quantitative data, for the most part, are consistent with national trends.

Both descriptive results and analysis of potential correlations among indicators do not tend to support any pattern related to effectiveness. Indeed, the measure of effectiveness itself is questioned in terms of both reliability and validity of instrumentation. So what can be surmised as a result of this momentous effort?

As was concluded, it is time to examine more closely how research should be conducted. It has been recommended that the focus should be on the school quality improvement methodology which relies on observations and the case study approach. Second, it is recommended that appropriate instrumentation be used to create a standardized approach to measuring the health of education in Malawi. This will allow comparisons to be made at different levels of the system and internationally as well as comparing changes over time.

The indicators that have been selected are many of the standard indicators used in research. Researchers attempted to broaden the range by including qualitative measures of processes and outcomes. This is dangerous ground since such indicators are intellectual constructs and their measurement is difficult. Still, by ignoring their measurement, a most important dimension of quality is ignored. As recommended, both the approach to research and instrumentation can account for these measures but more attention must be paid to the types of tools and methods used if results are considered both reliable and valid. The expense related to this approach is much higher than in more traditional approaches requiring the use of small samples and larger errors of measurement. But, this is a necessary trade-off if one is to try and understand the complicated nature of educational systems.

It was also suggested that research related to linking key indicators may reveal value added outcomes. If in-service of teachers is conducted in one sample, in service of managers in another, and in service of SMCs in a third, the impacts on outcomes may be much less than if in-service were to be conducted with all three groups in the same community. Some of the results in Chapter Six suggest that linking different elements of the system are important to increasing school effectiveness. This might include linking such inputs as instructional materials and procurement of furniture with teacher, and school manager in-service and community public awareness to yield the greatest increase in quality. These kinds of input/process relationships may be more important than isolating one indicator for the purpose of research. Future research should consider the piloting of inter-related indicators within a sample group.

The measurement of quality outcomes is suggested here as more complex than simply looking at achievement. And, given the complex nature of educational systems, the improvement in an array of outcomes at the school level may take more time than can be measured over short pilots of two years duration, since attitude and behavior change are part of the improvement equation. One cannot expect to see significant improvement in outcomes over such a short period of time. Phase 3 of JICA's project may continue over ten years, allowing sufficient time to measure such change providing that research design and instrumentation are appropriate and the capacity exists within Malawi to do the job properly.

Thus, the results of this study are important in showing the way forward in constructing an appropriate research agenda. The importance must be seen in that light – as a test of how research should be conducted more so than as a measurement of the pilots of the NIPDEP project.

REFERENCES

- Berman P. & McLaughlin M. (1981) *Federal Programs Supporting Educational Change. Implementing and Sustaining Innovations*. Santa Monica, California, Rand Corporation.
- Centre for Educational Research and Training (June 2004). *A Report for the Baseline Survey*. Report submitted to the JICA National Implementation Program for District Education Plans (NIPDEP). CERT: Malawi.
- Centre for Educational Research and Training (June 2004). *A Report for the Midpoint Survey*. Report submitted to the JICA National Implementation Program for District Education Plans (NIPDEP). CERT: Malawi.
- CERT (2005) *Issues in the Education System of Malawi: Evidence from Research*. Background Paper for the National Conference on Education in Malawi- 29th March to 1st April, 2005. Malawi Institute of Management, Lilongwe.
- Chase G. (1979) *Implementing a Human Services Program: How hard will it be?* *Public Policy* 31 (385-435).
- Chimombo J.P.G. (1999) *Implementing Educational Innovations: A study of Free Primary Education in Malawi*. Unpublished DPhil thesis, University of Sussex.
- Chimombo J.P.G. and Kadzamira E. (2001) *Community Sensitization and Mobilization in Malawi. Report Submitted to Civil Society Coalition for Quality Basic Education*. Zomba: Centre for Educational Research and Training, University of Malawi,
- Chimombo, J.P.G., Dzimazdi, C. and Kunje D. (2003). *Conditions of schooling in Malawi Primary schools*. Study commissioned by the World Bank for the Ministry of education. CERT, Zomba
- Chimombo JPG, Kunje D and Chimuzu T. (2005) *Malawi SACMEQ II Study*. Centre for Educational Research and Training. Zomba.
- JICA Project Team (May 2003). *Guide: Baseline, Mid-Point and Post-Pilot Project Evaluation, Volume 1 – Procedures*. Prepared for use by the NIPDEP Project Evaluation Team. Malawi.
- JICA Project Team (May 2003). *Guide: Baseline, Mid-Point and Post-Pilot Project Evaluation, and Volume 2 – Instruments and Report Forms*. Prepared for use by the NIPDEP Project Evaluation Team. Malawi.
- Malawi Institute for Education (June 2005). *Post Pilot Report on Student and Teacher Achievement Test Scores*. Report submitted to the JICA National Implementation Program for District Education Plans. MIE: June 2005.
- Milner, G., Chimombo, J., Banda, T., and Mchikoma, C. (2001). *The quality of education: Some policy suggestions based on a survey of schools in Malawi*. Paris: International Institute for Educational Planning.
- Ministry of Education Science and Technology (2000). *Policy Investment Framework*. MoST: Malawi.
- Pressman J. and Wildavsky A. (1973) *Implementation* Berkeley and Los Angeles: University of California Press