

QUIPS Impact
on
Primary School Achievement

The Mitchell Group, Inc.
Performance Monitoring and Evaluation Project

Isaac Amuah
Emma Gyamera
Richard Vormawor

with

Elizabeth Barcikowski
The Mitchell Group, Inc.

The following presents results from a recent study of the QUIPS Program impact. Staff members of The Mitchell Group, Inc. Performance Monitoring and Evaluation Project and members of the Evaluation Assessment Research Center worked collaboratively in the data collection, data analysis, and presentation of results. The current presentation focuses on three sets of findings:

1. Cohort 1 test results from July 2001 with emphasis on the sustainability of achievement gains made by QUIPS from June 1998 to December 2000.
2. Results from the Cohort 3 impact study.
3. Results from achievement performance analyses of the QUIPS achievement tests.

Cohort 1 and 2: Analysis of the sustainability of Cohort 1 achievement gains.

Review of Cohort 1 and 2 Achievement Gains, December 1998 to December 2000.

The QUIPS interventions in Cohort 1 schools began in June 1998 and the intervention program was introduced in Cohort 2 schools in January 1999. Achievement performance monitoring for Cohort 1 and Cohort 2 incorporated a cross-sectional study of Grade 3, Grade 4, and Grade 6 achievement in Mathematics and English. In this study ten pupils were randomly selected from each of three classes (i.e., P3, P4, and P6) to participate in the achievement testing program at baseline and annually through the 2-year cycle of intervention (i.e., December 1998, December 1999, and December 2000). Beginning in December 1999 ten randomly selected pupils from a set of comparison schools were also tested. This latter situation allowed for post-intervention comparisons of pupil achievement between the partnership schools and a set of comparable control schools. The initial study of QUIPS impact on pupil achievement for Cohort 1 and Cohort 2 schools was conducted in the spring of 2001 following the final achievement testing cycle which took place in December 2000. The impact study focused on two primary questions:

1. For the QUIPS partnership schools, was there a significant and meaningful improvement in the average class achievement for English and mathematics from baseline (i.e., achievement measures taken in December 1998) to post-intervention data collection (i.e., achievement measures taken in December 2000)?
2. Following the 2-year cycle of intervention is English and mathematics achievement higher on the average for pupils in QUIPS partnership schools than for pupils attending a set of comparable *control* schools?

Table 1 on page 3 summarizes the gains in achievement made for the Cohort 1 and 2 QUIPS partnership schools from baseline to the end of the 2-year intervention cycle. It can be seen in Table 1 that on the average a 50% improvement in mathematics and a 24% improvement in English was made in the QUIPS schools from baseline to the end of the 2-year intervention in December 2000 (i.e., see the compiled results for Grade 3, Grade 4, and Grade 6 given in the row labeled "Total" for mathematics and English). That is, on the average, classes in the QUIPS schools obtained 11.30% more of the mathematics test items correct at the end of the intervention than at baseline (i.e., an increase from 22.53% at baseline to 33.83%) and 10.51% more of the English test items (i.e., an increase from 44.32% at baseline to 54.83%). These changes from baseline to post-intervention testing were statistically significant ($p < 0.025$) and considered to be a meaningful gain in overall achievement. The positive linear trends in mathematics and English achievement are presented in the line graphs in Figure 1 and Figure 2.

Table 1. Averaged Class Means for 45 Cohort 1 and Cohort 2 OUIPS Partnership Schools (Averaged Class Mean Percentage Correct Scores)

	Mathematics*		
	December 1998	December 1999	December 2000
Grade 3	32.80	44.72	49.45
Grade 4	21.60	27.82	29.45
Grade 6	13.18	16.35	22.60
TOTAL	22.53	29.63	33.83

	English*		
	December 1998	December 1999	December 2000
Grade 3	39.50	49.25	51.64
Grade 4	40.22	48.37	49.23
Grade 6	53.24	56.13	63.63
TOTAL	44.32	51.25	54.83

*Significant positive linear trends (p < 0.01)

Figure 1. Math Trends Cohort 1 & 2 Combined

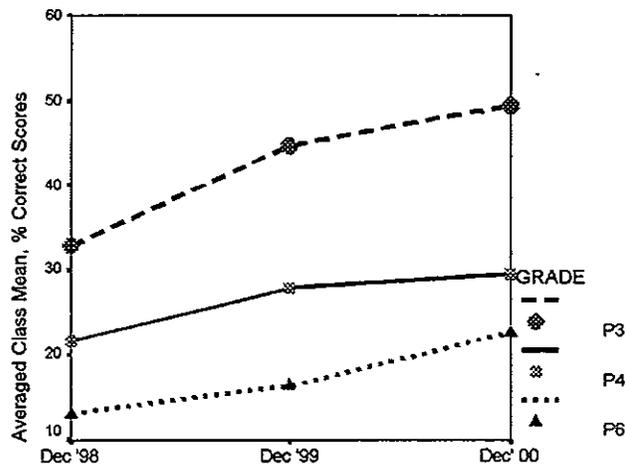
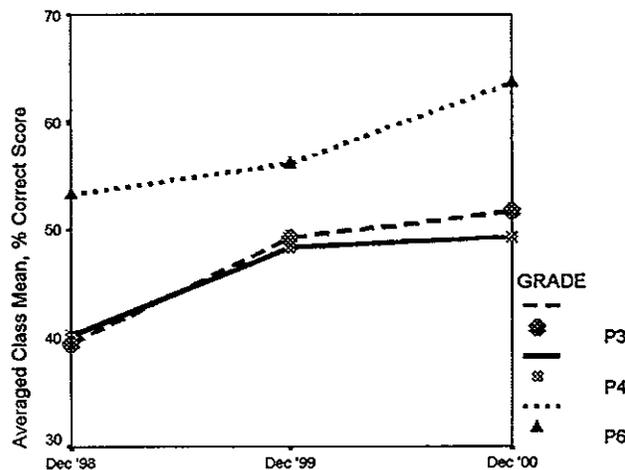


Figure 2. English Trends Cohort 1 & 2 Combined



Post-intervention group comparisons between Cohort 1 and Cohort 2 QUIPS partnership schools and the set of Cohort 1 and 2 comparison schools indicated that, on the average, pupils in the QUIPS schools outperformed pupils in the control schools for both mathematics and English. These results are presented in detail in Appendix 1 and summarized in Table 2 below. It can be seen in Table 2 that when the grades are considered together, the average class performance in mathematics for the QUIPS Partnership Schools was 6.23 percentage points higher than the average class performance for the set of comparison schools and 9.89 percentage points higher for English. Post-intervention group differences between the QUIPS partnership and comparison schools was statistically significant for both mathematics and English ($p < 0.025$).

Table 2. Averaged Class Performances for QUIPS Partnership and Comparison at the End of the Two-Year Intervention Cycle, December 2000. (Averaged Class Mean Percent Correct Scores)

	<u>Mathematics*</u>	
	<u>QUIPS Partnership Schools</u>	<u>Comparison Schools</u>
Grade 3	49.45	42.54
Grade 4	29.45	23.46
Grade 6	22.60	16.80
TOTAL	33.83	27.60
	<u>English*</u>	
Grade 3	51.64	43.45
Grade 4	49.23	39.54
Grade 6	63.63	51.81
TOTAL	54.83	44.94

*Overall group differences significant ($P < 0.025$)

Investigating the Sustainability of the Achievement Gains in Cohort 1

In September 2000 the QUIPS incorporated an evaluation component into the Performance Monitoring and Evaluation System (QUIPS PME) that would address the degree to which schools supported by the QUIPS Program were successful in maintaining or even increasing further the quality improvements made after direct QUIPS Program support is withdrawn.

The first follow-on achievement testing was conducted in July 2001 in a sample of 14 Cohort 1 and Cohort 2 schools (i.e., a 30% sample). The follow-on testing program for Cohort 1 and Cohort 2 is in some ways a special case in that the achievement test instrument in these earlier cohorts was modified for the later cohorts, Cohort 3 through Cohort 6. This presented a small problem when considering the design of the follow-on achievement test program, at least for these earlier cohorts. The follow-on testing program serves two purposes. The first purpose is to assess the sustainability of the gains made as a result of the QUIPS program and the second is to provide the annual performance data required for USAID's annual R4 performance review. As the reporting requirements of the R4 involve the aggregation of performance data across cohorts it is important that a common achievement test instrument be used in the follow-on testing program. It was therefore decided to use the new achievement test instrument for all follow-on achievement testing, even for that conducted in the Cohort 1 and Cohort 2 schools. In order to assure that the data collected from the new instrument could be used to investigate Cohort 1 and 2 trends in achievement over time both the earlier achievement tests and the modified ones were administered to the same children in

December 2000. This provided the data needed to convert one test to the other for purposes exploring class achievement trends in Cohort 1 and 2 schools from baseline testing through the follow-on years.

In July 2001 TMG administered the new achievement tests in a random sample of Cohort 1 and Cohort 2 QUIPS partnership schools. Using the conversion formula developed in December 2000, class performances converted to scores comparable with the earlier set of achievement tests were used in a sustainability study focusing on Cohort 1 findings. At the time of the follow-on achievement testing, July 2001, QUIPS support had been discontinued in the Cohort 1 schools for 1 year (i.e., since July 2000) and had been discontinued in the Cohort 2 schools for 6 months (i.e., since December 2000). The reason for focusing on the Cohort 1 schools is because achievement testing was conducted in the Cohort 1 schools immediately after the final interventions, in July 2000. Achievement testing in July 2000 was not conducted in the Cohort 2 schools because the interventions were not scheduled for completion in Cohort 2 until December 2000. The July 2000 testing in Cohort 1 provides for a meaningful comparison of post-intervention achievement levels with the achievement levels one-year after the withdrawal of QUIPS support in July 2001. As achievement testing was not appropriate and not conducted in the Cohort 2 schools in July 2000 there was no data from Cohort 2 that could be meaningfully compared to the July 2001 achievement results. Investigation of the sustainability of achievement gains made in Cohort 2 will need to be postponed until after the second follow-on achievement tests are administered.

Table 3 on page 6 shows the trends in mathematics and English for Cohort 1 only through the 2-year intervention cycle and one year after QUIPS support had been withdrawn. All 18 Cohort 1 partnership schools participated in all of the December test occasions and the July 2000 testing. The results presented for July 2001, however, are estimated results based on data collected in six randomly selected Cohort 1 partnership schools. The achievement gains made in the Cohort 1 schools can be seen by reviewing the increases in the averaged class performances from December 1998 to December 2000. Considering results from Grade 3, Grade 4, and Grade 6 together (see Table 3, rows labeled *TOTAL*), a 38% increase in mathematics achievement was made over the two-year intervention cycle, an improvement of 9.82 percentage points. A 15% increase was made in English from baseline to December 2000, an improvement of 7.46 percentage points. An trend analysis of achievement change from December 1998 to December 2000 within the Cohort 1 schools resulted in a significant positive linear trend in both subject areas for all three grades tested ($p < 0.01$).

Analyses conducted immediately following the completion of the QUIPS Cohort 1 interventions (July 2000) and 6 months after (December 2000) indicated that class averages in mathematics and English achievement were significantly higher, on the average, in the QUIPS partnership schools than in the set of comparison or *control* schools ($p < .01$). Compiling P3, P4, and P6 results, the average class performance for the Cohort 1 QUIPS schools was approximately 10 percentage points higher than that of the control schools in mathematics and 15-18 percentage points higher in English (i.e., 15.5 points higher in July 2000 and 18.7 percentage points higher in December 2000). No Control schools were tested in July 2001 and therefore post-intervention group comparisons one year following the withdrawal of QUIPS support could not be conducted.

Table 3. Averaged Class Means for Cohort 1 Schools

	Mathematics							
	Annual Assessments, Dec 1998- Dec 2000*				Follow-on Study Results***			
	Dec '98	Dec '99	Dec '00**		July '00**	Control	July '01	
	(n=18)	(n=18)	QUIPS (n=18)	Control (n=18)	(n=18)	(n=18)	(n=6)	
Grade 3	37.84	40.33	51.42	40.78	53.73	42.38	59.38	
Grade 4	24.57	28.68	31.10	22.32	34.38	24.81	35.84	
Grade 6	14.89	17.78	24.24	13.74	24.18	16.22	22.85	
TOTAL	25.77	28.93	35.59	25.61	37.43	27.80	39.35	
			English					
Grade 3	46.31	48.17	53.55	38.67	60.51	44.64	63.17	
Grade 4	46.41	51.76	53.77	33.62	59.28	44.35	59.76	
Grade 6	59.84	61.29	67.61	46.45	71.94	56.23	65.67	
TOTAL	50.85	53.74	58.31	39.58	63.91	48.41	62.87	

*Significant positive linear trends in mathematics and English from December 1998 to December 2000.
 **Grades considered together, post-intervention group differences between QUIPS and Control Schools were significant ($p < 0.01$)
 ***Class achievement levels for the QUIPS schools in July 2000 were not statistically different from the achievement levels obtained in July 2001 ($p > 0.05$).

To investigate the success of the Cohort 1 partnership schools in maintaining the achievement gains made as a result of QUIPS, the average achievement levels of the Cohort 1 partnership schools in July 2000 were compared to the estimated achievement levels obtained one year later, in July 2001. These results are presented in Table 3 above in the column labeled "Follow-on Study Results". Inspection of the group means given for the QUIPS schools in July 2000 in comparison to the estimated performance of the QUIPS schools in July 2001 demonstrated that, with the exception of Grade 6 English, the achievement levels for the QUIPS schools immediately following the interventions in July 2000 were similar to the estimated achievement levels one year after direct QUIPS support had been withdrawn, in July 2001. In Grade 3 and Grade 4, the estimated mathematics and English performances in July 2001 were slightly higher than those obtained one year before, though in no situation were the differences from July 2000 to July 2001 statistically significant. Figure 3 and Figure 4 on page 7 present the mathematics and English trends for Cohort 1 from baseline in December 1998 to one year after the QUIPS interventions were discontinued, in July 2001. The data to generate these graphs were taken from Table 3.

Figure 3. Cohort 1 Math Trends, Dec '98 to July '01

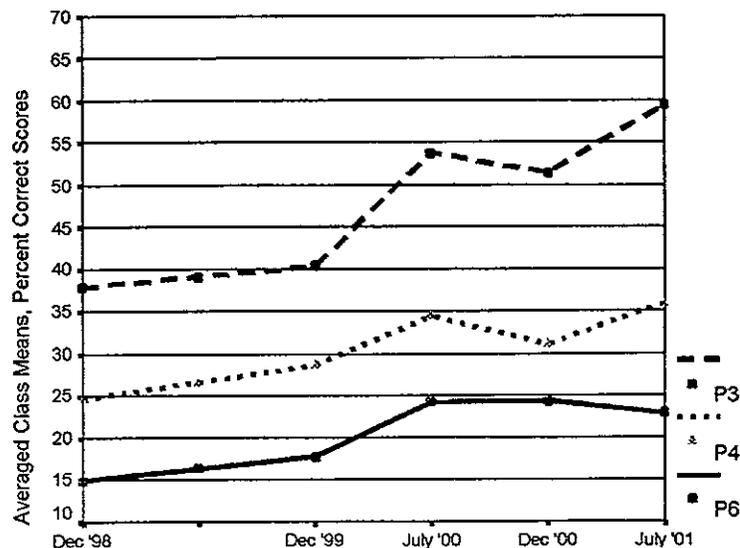
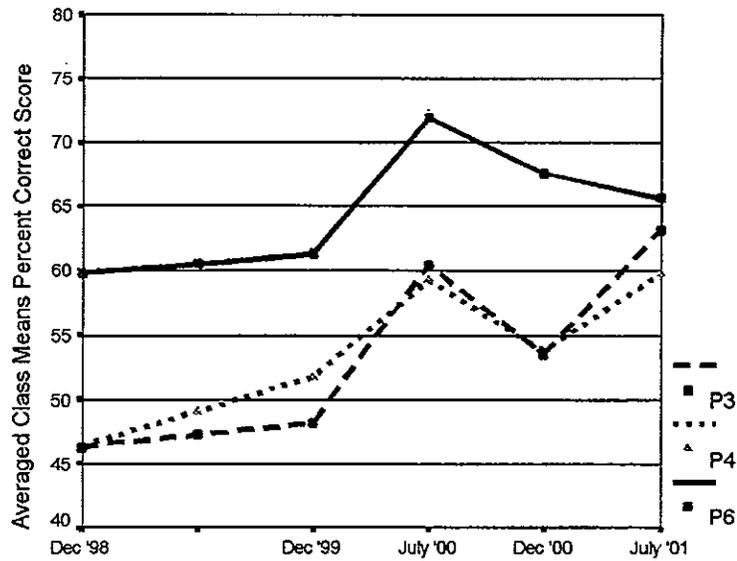


Figure 4. Cohort 1 English Trends Dec '98 to July '01



Recommendations from Cohort 1 Follow-On Study

Results from the first follow-on testing exercise demonstrated that the Cohort 1 schools are, in most situations, maintaining or even improving on the achievement gains made as a result of the QUIPS program. The latter situation was exemplified by further gains in Class 3 mathematics and English achievement that were made in the first year after QUIPS support was withdrawn. It could be that the QUIPS program has an impact on pupil achievement growth rates, particularly in the lower grades, which are reflected in positive shifts in the class average achievement levels in the years after the direct interventions have been discontinued.

One limitation in this first follow-on study was the small number of schools sampled from Cohort 1. In the future this will not be a problem because it will be possible to combine the data from the Cohort 1 and Cohort 2 schools. Another limitation is that no comparison schools were sampled during the follow-on testing. It is suggested that in the future the follow-on achievement testing program include a sample of the comparison schools in addition to the sample of partnership schools.

Cohort 3 Impact

Introduction

The Cohort 3 impact study focuses primarily on the study of group differences in the learning growth curves or *learning rates* between the QUIPS partnership schools and the set of Cohort 3 comparison or *control* schools in addition to post-intervention comparisons of the group means between the partnership and comparison schools. In general this is accomplished in three steps: 1) testing pupils on three different occasions in a sample of QUIPS partnership

schools and a sample of comparison schools; 2) estimating for each school sampled the learning growth curves or *slopes* associated with each test (i.e., mathematics, English literacy, and spoken English narrative for P3/P4 and P5/P6) using a growth curve analysis statistical procedures; and 3) analyzing the group differences among the estimated slopes between the partnership and comparison schools using a statistical procedure referred to as *hierarchical linear modeling* or HLM. This Cohort 3 (and subsequent cohorts) study design is different from that used to analyze the impact of QUIPS on pupil achievement in Cohort 1 and Cohort 2 (discussed in the previous section). In Cohort 1 and 2 the focus of the study was on investigating shifts in class means that could be attributed to the QUIPS interventions. In the Cohort 1 and 2 design pupils are not tested repeatedly. Instead, pupils from target classes (i.e., P3, P4, and P6) are randomly selected to participate in the testing program at the same time each year and these scores are used to estimate the annual class performance over the 2-year intervention period and in the years following direct interventions (see results presented in the section above).

Cohort 3 Data Collection and Test Instruments

Pupils from a sample of 81 schools, 45 QUIPS partnership schools and 36 comparison schools, were administered achievement tests on three test occasions with approximately 6 ½ school months (i.e., school breaks were not including) between each test occasion. Of the 45 partnership schools sampled, 27 or 50% of the 54 partnership schools were randomly sampled from the southern regions. The remaining 18 partnership schools were from the northern regions of Ghana and these 18 schools made up the complete set of partnership schools from the North. Considering this relatively disproportionate number of schools from the North and South (i.e., 50% of the partnership schools from the South and 100% of the partnership schools from the North) a preliminary analysis was conducted to ascertain if this difference did present a noticeable bias. Overall group results for the set of QUIPS partnership schools (i.e., group averages for class performances at each test occasion and learning slope estimates across the subject areas) were not noticeably different when the number of schools in the North were systematically reduced by 50%. Therefore, all of the participating partnership schools were entered into the statistical analyses when possible (see page 9, “Analyses Procedures and Results” for further discussion).

P3 and P5 pupils were tested in mathematics, English literacy and spoken English narrative. Pupils are trained in the test taking activity prior to administering each of the four tests. The tests are not timed and therefore pupils are given as much time as they need to complete the tests. The mathematics and English literacy tests are conducted in classroom settings. Eight pupils (e.g., 4 girls and 4 boys) are randomly selected out of each classroom to participate in the English narrative test. The English narrative test is conducted in small groups and individually.

The test instruments include items that span across the primary school classes, P1 through P6, thereby capturing a baseline for all children regardless of their entry level abilities and allowing sufficient *room to grow* even for the highest performing pupils. The mathematics tests focus on basic mathematics operations, including early basic concepts and story problems. All mathematics test instructions and story problems are presented in both the local language and in English. The English literacy test items are in a multiple choice format and utilize a *cloze procedure*. The English literacy test is designed to test pupils’ ability to read with meaning. In the spoken English narrative tests, pupils retell English stories while

looking at a picture storybook. Prior to telling the story, pupils have the opportunity to hear the English story on three different occasions, both in a small group session and individually.

The mathematics and English literacy tests are scored in a conventional fashion with each item scored as “correct” or “incorrect”, but the English narrative tests are scored using a series of response categories that are measured on an ordinal scale. Pupil responses on the storytelling task are scored according to the following categories: 1) Score of “5”- self-initiated, exceptional verbal response; 2) Score of “4” -self-initiated grammatically accurate verbal response; 3) Score of “3”- self-initiated, incomplete related or grammatically inaccurate verbal response; 4) Score of “2” –correct verbal response to a direct question given by the test administrator; 5) Score of “1” –pointing response to a direct pointing request given by the test administrator; 6- Score of “0” – no response even after test administrator cues given by direct questions and pointing responses.

It should be noted that the baseline data collection was somewhat delayed in Cohort 3. This delay reduced the period of time given to track pupil achievement by approximately four months; that is, four months of school activity in the first academic year of the 2-year intervention cycle. Secondly, in some schools the interventions had been initiated prior to the baseline data collection and therefore early gains attributed to the intervention may have been missed. In consideration of learning growth curves, these latter situations are noteworthy and may have weakened the sensitivity of the analysis slightly, for the following two reasons: 1) the opportunity to observe a full two years of learning growth is constrained by the delay in the baseline assessment during the first academic year; and 2) very early affects of treatment associated with interventions begun prior to the baseline would have been missed.

Analysis Procedures and Results

Growth curve analysis using the hierarchical linear modeling procedure was used to investigate group differences in the estimates of the school learning rates between the QUIPS partnership and control schools for the each of the three subject areas, mathematics, English literacy, and spoken English narrative, for each of the school periods tested (i.e., P3 to P4 and P5 to P6). This analysis procedure is not robust to violations of normality in the data and in some situations there were schools that skewed the distribution substantially and therefore in these situations certain schools were deleted to assure strong statistical conclusion validity. This latter situation occurred for analyses of English literacy, particularly in P3/P4. Five schools, two partnership schools and three comparison schools were deleted from the P3/P4 English literacy data set to insure that the assumptions of the HLM statistical procedure tenable. These were: Kentindrono M/A, QUIPS; Armon Memorial, QUIPS; Sunyani St. Ansem’s Anglican, Control; Koforidua Sarkodee, Control; and Nana Kwaku Boateng, Control. Two control schools were omitted from the P5 English literacy analyses: Koforidua Sarkodee and Nana Kwaku Boateng.

The achievement results for each test occasion and the average estimated rates of learning or *slopes* for the set of QUIPS partnership and Control schools are presented in Table 4 on page 10. The group average achievement results were generated from actual data collected on each of the three test occasions. However, the group averages for learning rate or the *mean slopes* are averages of the slope *estimates* for the set of schools within each group (i.e., partnership versus control). The *slopes* can be described as rates of learning for schools estimated from the growth curves of individual pupils in those schools, compiled across the set of schools in a particular group (i.e., partnership or control). The slope estimates provide a way of predicting achievement in the future. That is, the estimated school slope represents

the number of test points that a pupil from that school would be expected to gain within one inter-test interval (e.g., in this situation, in 6 ½ school months). This is demonstrated in general by taking the group achievement level at baseline, or *Feb '00* and adding the estimated slope value. By doing this you should approximate the achievement level given for the *Nov'00* test occasion. Repeating this once more using the *Nov'00* test occasion as a base should result in an approximation of the *July'01* achievement level.

It can be seen in Table 4 that the estimates of learning rate or *slope estimates* for all subject areas tested were higher on the average for pupils attending the QUIPS partnership schools than for pupils attending the Control schools, although not all results were statistically significant. Group differences in rate of learning were statistically significant for P3 English literacy, P5 English literacy, and P5 spoken English narrative. Figure 5 through Figure 10 show plots of learning curves estimated for P3/P4 and P5/P6 English literacy, Mathematics, and Spoken Narrative for the QUIPS and Control Schools.

Table 4. Summary of Achievement Results and Learning Rates for QUIPS and Control Schools

	Group Mean Achievement Performance Levels						Group Means of Estimated School Learning Rates Average Slopes
	Feb '00		Nov '00		July '01		
	Total	% of Total Possible	Total	% of Total Possible	Total	% of Total Possible	
P3 Mathematics							
QUIPS	18.92	47.3	24.49	61.2	27.44	68.6	4.29
Control	18.82	47.1	23.76	59.4	27.00	67.5	4.08
P3 English Literacy*							
QUIPS	19.19	33.1	22.69	39.1	26.90	46.4	3.80
Control	19.03	32.8	21.16	36.5	24.81	42.8	2.86
P3 English Narrative							
QUIPS	23.39	48.7	28.48	59.3	33.96	70.8	5.43
Control	23.55	49.1	27.07	56.4	32.29	67.3	4.58
P5 Mathematics							
QUIPS	18.68	46.7	21.72	54.3	25.02	62.3	3.12
Control	18.62	46.6	21.40	53.5	24.08	60.2	2.73
P5 English Literacy*							
QUIPS	22.68	40.5	26.45	47.2	30.61	54.7	3.87
Control	22.33	39.9	25.81	46.1	29.52	52.7	3.43
P5 English Narrative*							
QUIPS	36.65	53.9	45.27	66.6	47.95	70.5	5.13
Control	39.77	58.5	45.82	67.4	47.11	69.3	3.08

*Significant group differences for learning rates (i.e., slopes) between QUIPS and Control Schools

Figure 5 through Figure 10 on pages 10-11 present the results above in graphic form. These line graphs demonstrate that pupils in the QUIPS schools are beginning to learn at a faster pace than those in the set of comparisons schools. If these rates of learning were to be sustained, the gap between achievement favoring schools that have participated in the QUIPS program would be expected to increase even further over time. These results underscore the importance of investing in activities that will assist the schools in maintaining the interventions learned and practiced as a part of the QUIPS program.

Figure 5.

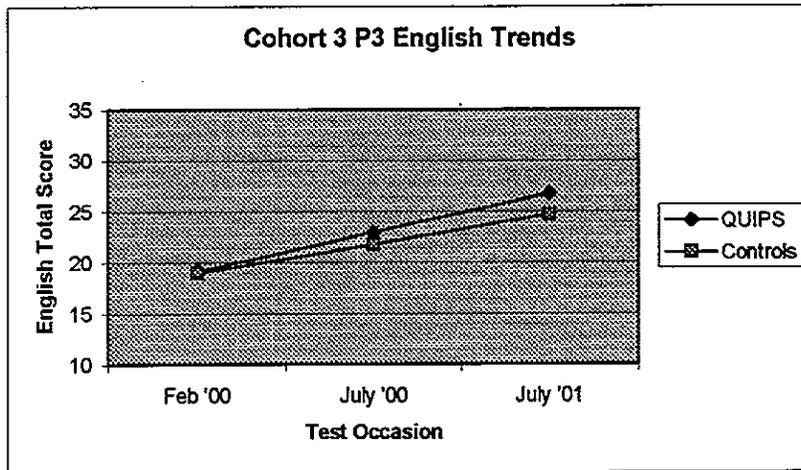


Figure 6.

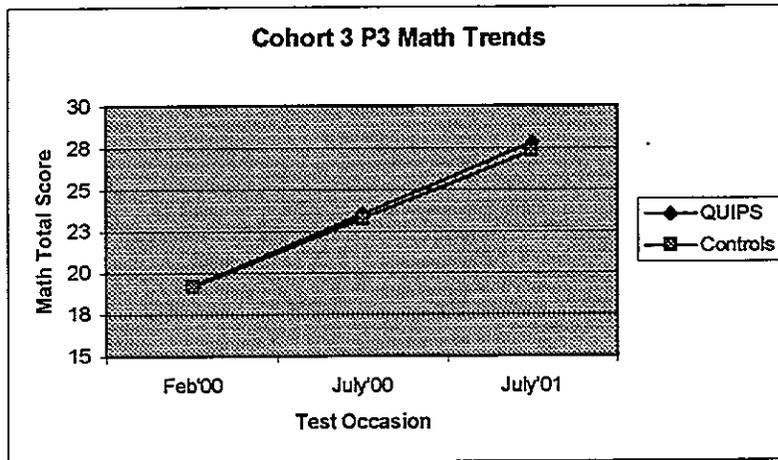


Figure 7.

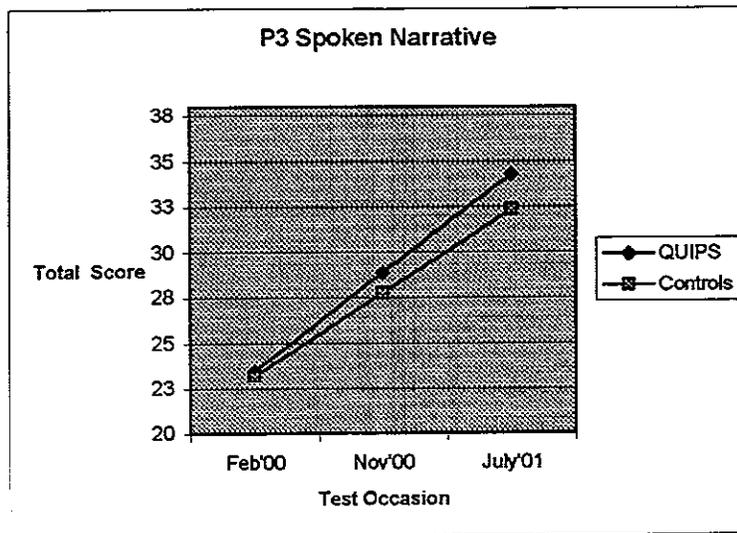


Figure 8.

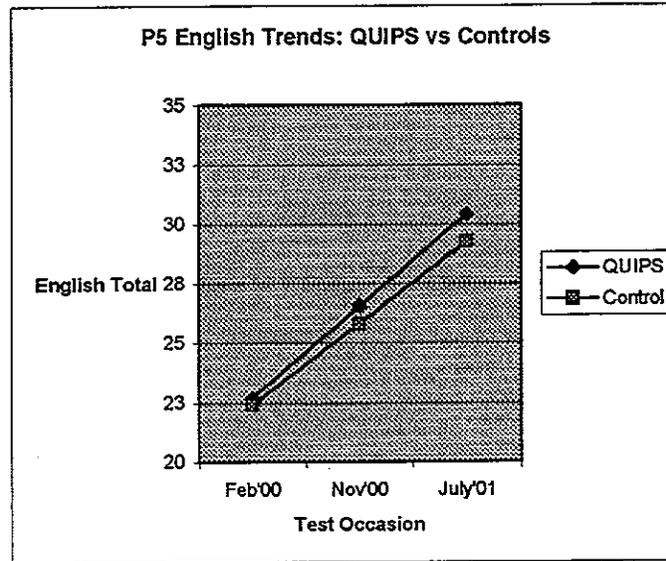


Figure 9.

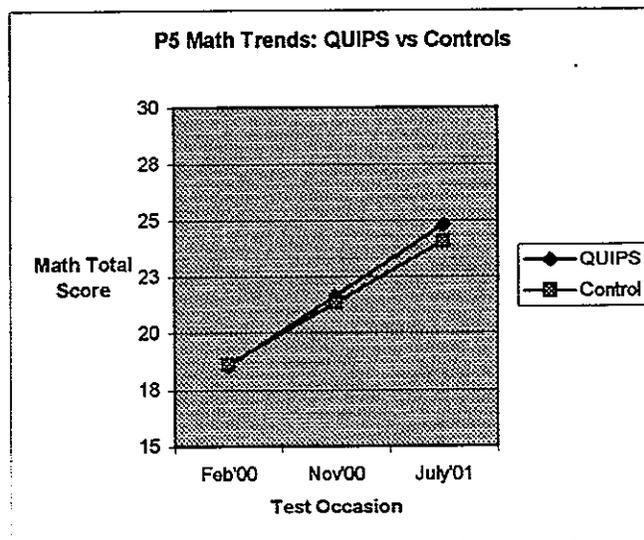
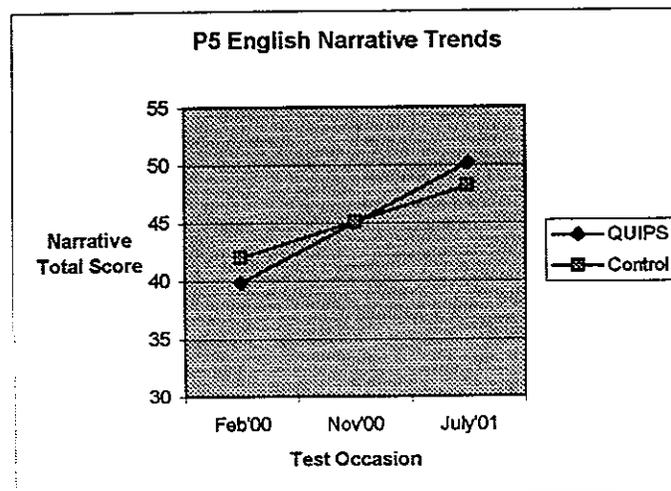


Figure 10.



Two other descriptive analyses were used to study the impact of the QUIPS Program on pupil learning. The first of these was to investigate the distribution of QUIPS and Control schools that fell in the upper and lower 25% performance groups with respect to learning growth rates. This was accomplished by identifying for each subject and grade the schools whose estimated learning growth rates or *slopes* fell above the 75th percentile and by identifying the schools whose estimated learning growth rates fell below the 25th percentile. The relative number of QUIPS and Control schools that fell in the upper and lower performance groups are presented in Table 5 below. For example, for P3/P4 mathematics 15 QUIPS partnership schools had learning rates that fell above the 25th percentile rank. These 15 schools represented 75% of the total number of schools (19 schools) whose slopes for P3/P4 mathematics fell above the 75thile.

It can be seen in Table 5 that for every subject area tested and for every grade proportionately more QUIPS schools than Control schools had learning rates that fell above the 75th Percentile (i.e., in the upper quarter) and in most subject areas a larger number of Control schools than QUIPS schools fell below the 25thile (i.e., in the lower quarter). This differential pattern for the QUIPS and Control schools was especially noticeable for P3 English, P3 Mathematics, P5 English and P5 Spoken Narrative.

Table 5. Number of QUIPS and Control Schools whose Learning Growth Rates Fall above the 75th Percentile and the Relative Number that fall Below the 25th Percentile

		P3 to P4		P5 to P6	
		QUIPS	Control	QUIPS	Control
Mathematics					
Above 75 th ile	Number schools in upper group.	15	5	13	7
	Percent of total in upper group.	75	25	65	35
Below 25 th ile	Number schools in lower group.	9	11	10	11
	Percent of total in lower group.	45	55	48	52
English Literacy					
Above 75 th ile	Number schools in upper group.	13	6	14	6
	Percent of total in upper group.	68	32	70	30
Below 25 th ile	Number schools in lower group.	7	12	9	11
	Percent of total in lower group.	37	63	45	55
Spoken English					
Above 75 th ile	Number schools in upper group.	13	7	15	5
	Percent of total in upper group.	65	35	75	25
Below 25 th ile	Number schools in lower group.	10	10	8	12
	Percent of total in lower group.	50	50	40	60

Finally, individual QUIPS schools/communities were ranked in order to allow the intervention teams to investigate the relative performance of the QUIPS partnership schools with respect to the school estimates of pupil performances. Table 6 presents the school learning growth rates for P3/P4 mathematics, English literacy, and spoken English narrative ranked within each subject area. The 90th, 75th, 50th, 25th and 10th percentile cutoffs are provided to support the qualitative analysis of these data.

%ILES	P3 MATH		P3 ENGLISH		P3 NARRATIVE	
		Learning Rate		Learning Rate		Learning Rate
	Nangodi L/A	6.58	Busunu L/A	7.34	Adaklu Kodzobi	13.28
	Bogoso Methodist	5.47	Nangodi L/A	7.07	Toh Kpalime L/A	11.54
	Ntankoful M/A	5.42	Tsavanya L/A	6.13	Trom Nyeredede L/A	9.52
	Abrem Dwabor D/C	5.35	Sunyani Sacred He	5.84	Peki Blengo E.P	9.13
90%ILE	Nkanfo Catholic	5.34	Kentinkrono M/A	5.70	Nator	8.76
	Kentinkrono M/A	5.27	Bogoso Methodist	5.59	Bulugu	8.65
	Nator	5.26	Kobedi A.M.E. Zio	5.53	Kami L/A	8.64
	Amomaso D/C	5.21	Tizza R/C	5.41	Nangodi L/A	8.25
	Kedadwen D/C	5.14	Armon Memorial	5.19	Ampunyasi R/C	8.10
	Baazu L/A	5.00	Ntankoful M/A	5.14	Tubong	7.83
	Zabugu	4.98	Nkanfo Catholic	4.91	Kokodei Methodist	7.76
75%ILE	Sunyani Sacred He	4.96	Toh Kpalime L/A	4.81	Abu Bonsra L/A	7.45
	Tubong	4.89	Amomaso D/C	4.80	Gbefi St. Peters	7.42
	Kami L/A	4.83	Peki Blengo E.P	4.71	Yapei Presby	7.39
	Kokodei Methodist	4.81	Ve Wudome L/A	4.28	Nkanfo Catholic	7.26
	Amede Presby	4.72	Kedadwen D/C	4.24	Tsavanya L/A	6.33
	Toh Kpalime L/A	4.53	Abu Bonsra L/A	4.12	Baazu L/A	6.32
	Avenopeme R/C	4.50	Ampunyasi R/C	3.96	Kedadwen D/C	6.29
	Tsavanya L/A	4.44	Nator	3.91	Ayensuano D/C	6.10
	Watania E/A	4.41	Yapei Presby	3.78	Isadiah E/A	5.74
	New Asuoayaa R/C	4.38	Gbefi St. Peters	3.76	Dapouh R/C	5.70
	Trom Nyeredede L/A	4.29	Tubong	3.70	Kpafinig E/A	5.67
50%ILE	Abu Bonsra L/A	4.22	Trom Nyeredede L/A	3.62	Watania E/A	5.55
	Abutia Agorve EP	4.21	Fufulso Presby	3.61	Ve Wudome L/A	5.32
	Ampunyasi R/C	4.18	Isadiah E/A	3.59	Tizza R/C	5.31
	Obretema L/A	4.16	Abutia Agorve EP	3.57	Amomaso D/C	5.21
	Busunu L/A	4.13	Abrem Dwabor D/C	3.44	Avenopeme R/C	5.11
	Kalbeo/ Tindonsol	4.13	Adaklu Kodzobi	3.42	Kobedi A.M.E. Zio	5.06
	Isadiah E/A	4.09	Kalbeo/ Tindonsol	3.33	Amede Presby	4.69
	Fufulso Presby	4.08	Baazu L/A	3.16	Zabugu	3.75
	Kobedi A.M.E. Zio	4.06	Kokodei Methodist	3.10	Abutia Agorve EP	3.61
	Ayensuano D/C	4.05	Avenopeme R/C	3.09	Kalbeo/ Tindonsol	3.20
	Bulugu	4.02	Amede Presby	2.99	Bogoso Methodist	3.14
25%ILE	Yapei Presby	3.91	Daffiama R/C	2.92	Daffiama R/C	2.91
	Dapouh R/C	3.82	New Asuoayaa R/C	2.83	Fufulso Presby	2.83
	Ve Wudome L/A	3.76	Tongo L/A	2.77	Abrem Dwabor D/C	2.82
	Tizza R/C	3.47	Ayensuano D/C	2.70	Obretema L/A	2.53
	Armon Memorial	3.45	Santrokofi L/A Gb	2.62	Santrokofi L/A Gb	2.40
	Gbefi St. Peters	3.36	Watania E/A	2.45	New Asuoayaa R/C	1.97
	Peki Blengo E.P	3.24	Obretema L/A	2.41	Kentinkrono M/A	1.96
10%ILE	Adaklu Kodzobi	3.22	Zabugu	2.32	Sunyani Sacred He	1.61
	Tongo L/A	2.85	Dapouh R/C	2.04	Ntankoful M/A	1.60
	Daffiama R/C	2.70	Bulugu	2.01	Tongo L/A	0.62
	Santrokofi L/A Gb	2.16	Kami L/A	1.49	Armon Memorial	0.18
	Kpafinig E/A	1.84	Kpafinig E/A	0.99	Busunu L/A	-0.15

**Table 7. Ranked Estimates of Learning Growth Rates for QUIPS Partnership Schools
P5/P6 Mathematics, English Literacy, and Spoken English Narrative**

%ILES	P5 MATH	Learning	P5 ENGLISH	Learning	P5 NARRATIVE	Learning
		Rate		Rate		Rate
	Abrem Dwabor D/C	8.08	Tubong	5.51	Baazu L/A	11.53
	Tubong	6.38	Busunu L/A	5.12	Tizza R/C	9.99
	Nator	5.04	Nkanfo Catholic	5.06	Zabugu	9.82
	New Asuoyaa R/C	4.86	Tongo L/A	5.06	Kpafinig E/A	9.38
90%ILE	Kedadwen D/C	4.74	Abutia Agorve EP	5.05	Isadiah E/A	8.99
	Armon Memorial	4.63	Armon Memorial	4.93	Kedadwen D/C	8.46
	Kobedi A.M.E. Zio	4.56	Dapouh R/C	4.85	Tubong	8.34
	Ampunyasi R/C	4.23	Nator	4.83	Watania E/A	8.10
	Kalbeo/ Tindonsol	4.16	Daffiama R/C	4.78	Amomaso D/C	7.97
	Kokodei Methodist	4.12	Nangodi L/A	4.63	Kokodei Methodist	7.53
	Nkanfo Catholic	4.01	Avenopeme R/C	4.59	Gbefi St. Peters	7.51
75%ILE	Kpafinig E/A	3.97	Bogoso Methodist	4.47	Abutia Agorve EP	7.49
	Kentinkrono M/A	3.88	Abrem Dwabor D/C	4.39	Kami L/A	7.38
	Isadiah E/A	3.66	Adaku Kodzobi	4.38	Bulugu	6.90
	Dapouh R/C	3.63	Tizza R/C	4.35	Abrem Dwabor D/C	6.81
	Nangodi L/A	3.55	Kobedi A.M.E. Zio	4.32	Kalbeo/ Tindonsol	6.26
	Baazu L/A	3.50	Peki Blengo E.P	4.22	Toh Kpalime L/A	6.15
	Ntankofui M/A	3.50	Ampunyasi R/C	4.14	Adaku Kodzobi	6.05
	Abu Bonsra L/A	3.44	Gbefi St. Peters	4.11	Dapouh R/C	6.01
	Watania E/A	3.23	Kentinkrono M/A	4.08	Nator	5.47
	Yapei Presby	2.96	Tsavanya L/A	4.07	Kobedi A.M.E. Zio	5.42
	Abutia Agorve EP	2.78	Sunyani Sacred He	4.02	Ayansuano D/C	5.27
50%ILE	Adaku Kodzobi	2.78	Toh Kpalime L/A	4.01	Peki Blengo E.P	5.14
	Zabugu	2.75	Ve Wudome L/A	4.00	Tongo L/A	4.78
	Daffiama R/C	2.75	New Asuoyaa R/C	3.99	Avenopeme R/C	4.68
	Trom Nyeredede L/A	2.53	Ntankofui M/A	3.89	Daffiama R/C	4.63
	Peki Blengo E.P	2.49	Kedadwen D/C	3.77	Trom Nyeredede L/A	4.58
	Toh Kpalime L/A	2.44	Abu Bonsra L/A	3.74	Busunu L/A	4.54
	Avenopeme R/C	2.40	Trom Nyeredede L/A	3.70	Nangodi L/A	4.53
	Sunyani Sacred He	2.36	Ffulso Presby	3.54	Ampunyasi R/C	3.75
	Ayansuano D/C	2.26	Santrokofi L/A Gb	3.54	New Asuoyaa R/C	3.52
	Amede Presby	2.21	Kalbeo/ Tindonsol	3.51	Ve Wudome L/A	3.36
	Obretema L/A	2.20	Zabugu	3.43	Ntankofui M/A	3.05
25%ILE	Busunu L/A	2.19	Baazu L/A	3.29	Kentinkrono M/A	2.70
	Kami L/A	2.19	Yapei Presby	3.26	Santrokofi L/A Gb	2.59
	Tsavanya L/A	2.10	Obretema L/A	3.09	Tsavanya L/A	2.41
	Amomaso D/C	2.01	Amede Presby	2.94	Ffulso Presby	2.17
	Bulugu	2.01	Kami L/A	2.93	Obretema L/A	2.15
	Tongo L/A	2.00	Ayansuano D/C	2.92	Yapei Presby	2.08
	Bogoso Methodist	2.00	Amomaso D/C	2.89	Abu Bonsra L/A	2.00
10%ILE	Santrokofi L/A Gb	1.97	Kokodei Methodist	2.69	Nkanfo Catholic	1.66
	Tizza R/C	1.79	Kpafinig E/A	2.67	Amede Presby	1.60
	Gbefi St. Peters	1.51	Isadiah E/A	2.20	Bogoso Methodist	1.46
	Ve Wudome L/A	0.86	Bulugu	1.78	Armon Memorial	0.82
	Ffulso Presby	-0.28	Watania E/A	1.41	Sunyani Sacred He	-1.81

Cohort 3 Gender Analysis

Table 6 summarizes the Cohort 3 achievement results disaggregated by gender. Overall the relationships between male and female results were similar in the QUIPS partnership and comparison schools. Therefore, only the summaries for the QUIPS Partnership Schools are reported.

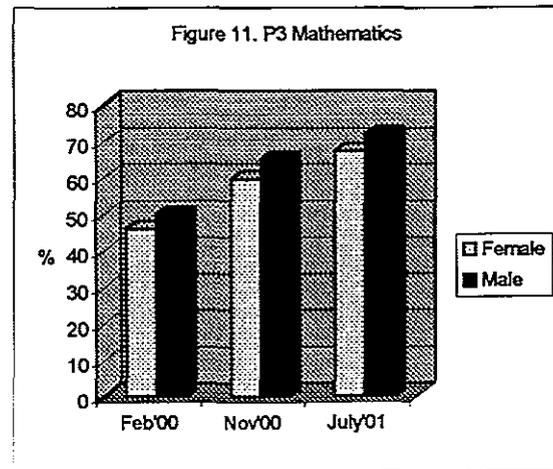
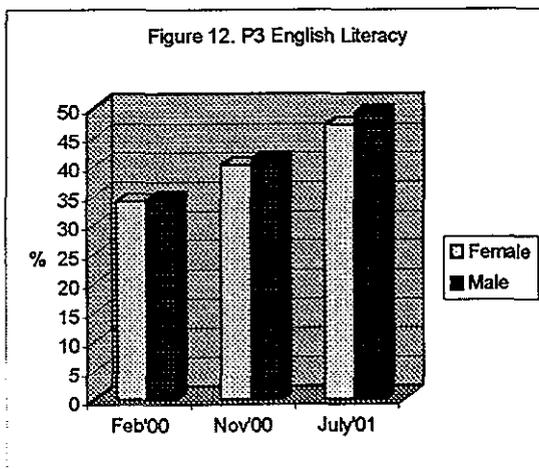
It can be seen in Table 6 and that for every subject area and every test occasion the males, on the average, obtained higher scores than did the females, although these differences were sometimes very slight and not always statistically significant. Compiled results across the three test occasions revealed significant group differences between males and females in mathematics and spoken English narrative for both P3/P4 and P5/P6. Analyses of the rates of learning growth for males and females were not different even though the overall performance levels were.

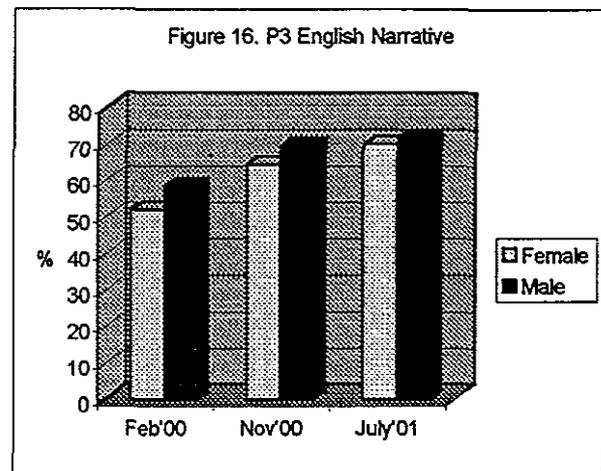
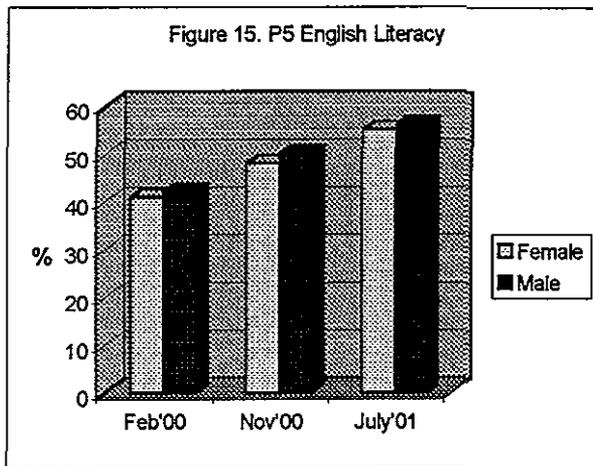
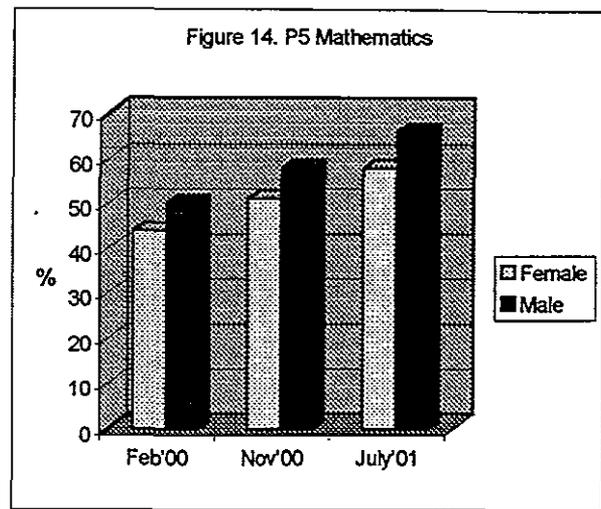
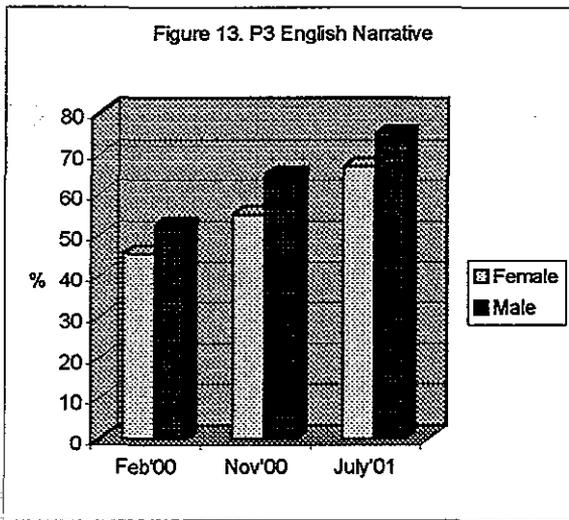
Table 6. Summaries of Performances by Females and Males (Average scores reflect the percentage of total possible).

Grade 3 to Grade 4				Grade 5 to Grade 6			
P3 Mathematics*				P5 Mathematics*			
	Feb'00	Nov'00	July'01		Feb'00	Nov'00	July'01
Female	45.91	59.75	66.90	Female	43.95	51.36	57.60
Male	49.71	64.59	71.87	Male	50.13	57.83	66.22
P3 English Literacy				P5 English Literacy			
Female	34.01	39.93	47.33	Female	40.94	47.57	54.79
	34.25	41.39	49.25	Male	41.86	49.55	56.22
P3 Narrative*				P5 Narrative*			
Female	45.07	54.83	66.90	Female	52.30	63.84	69.52
Male	52.17	64.91	75.31	Male	55.76	69.09	71.94

*Significant group differences between males and females ($p < .015$)

Figures 11-16 present the above results.





Achievement Test Performance Analyses

TMG also conducted a test item analyses to attempt to achieve a better understanding about the types of skills that pupils are able to do in reading and mathematics and spoken English narrative at the end of Grade 3, Grade 4, Grade 5, and Grade 6. The analyses were conducted on data collected from the QUIPS achievement testing program in July 2001. The P3 and P5 pupil data were obtained from Cohort 4 schools and the P4 and P6 pupil data were from Cohort 3 schools.

Mathematics and English Literacy

The study revealed that in mathematics and English literacy the majority of pupils in each of these grade levels (i.e., pupils just completing P3, P4, P5, and P6) were performing approximately two years below grade level. The descriptive study attempts to describe what the majority of pupils can and cannot do in mathematics and English literacy upon completion of P3, P4, P5, and P6 and then considers any shifts in the distribution of skills that

may be attributed to the QUIPS Program. The studies of mathematics and English literacy abilities were conducted using a three-step procedure. First of all, the overall grade level mastery for the majority of pupils is estimated for the group as a whole; that is, where applicable to pupils in QUIPS and the Control schools. If there is a difference in the overall ability level between pupils in the QUIPS and Control schools then these differences are pointed out to the reader. The overall ability levels were determined by studying the grade and content of test items that 50% or more of the pupils in both the QUIPS and Control school were able to do. The second major activity was to describe the kinds of skills that the majority (i.e., 50% or more of the pupils) were and were not able to do. The descriptions provided apply to all pupils tested unless there are special exceptions for pupils who have attended a QUIPS school. The *exceptions* are skills that pupils from the QUIPS schools have mastered that are still problems for pupils in the Control schools. These exceptions are addressed in the third study activity. In this study activity the abilities of pupils in the majority of QUIPS schools are compared to the abilities of pupils from the Control schools to see if there are any shifts in the ability levels of children who participated in the QUIPS program or any specific skills that pupils from the QUIPS schools are able to do that pupils from the Control schools are not able to do. Table 7 and Table 8 on page 19 and page 20, respectively summarize the findings for the descriptive study of mathematics and English literacy achievement.

It should be noted that the abilities and ability levels discussed in Table 7 and Table 8 are general statements about the majority of children in the grades tested. It is not to say that all of the pupils fall in the majority group. There is enormous diversity among pupils in Ghana's primary schools and there are pupils who are performing at grade level. However, this descriptive study tends to focus on the strengths and weakness of the majority and is included to help provide some insight into the nature of pupil learning and to provide some insight into the skill areas where QUIPS has had a special impact or where a shift in the focus of interventions may be needed.

However, it is important to consider the results with caution because the tests in this context are being used in some sense for diagnostic reasons when indeed they were not designed for this purpose. This information is probably most useful as a general guide for practitioners and intervention teams to identify areas where further investigation may be warranted. For example, as you can see in Table 7 and Table 8, Ghanaian pupils, even in the later primary grades have difficulty with multiplication problems that involve carrying and are not able to multiply with double digits. This is not surprising because pupils are very slow in learning addition that involves carrying, particularly where multiple digit numbers are involved. Therefore, it may be useful to try to learn why children aren't readily picking up these processes and what new teaching skills teachers may need to meet this challenge.

As can be seen in Table 7 and Table 8 pupils are late to develop English literacy. Though there is some breakdown in various literacy skill areas, the most notable finding is that the majority of pupils in the Control schools are not reading simple sentences and passages until Grade 5. This is an area where QUIPS had significant impact. Pupils attending QUIPS partnership schools were reading simple sentences and passages at Grade 4. English literacy development continues to present a challenge to Ghanaian primary school pupils. One reason for this may be the challenge presented by the multilingual context. If children haven't developed sufficient mastery of English speaking and listening comprehension skills delays in English literacy development are to be expected. Descriptive summaries of pupils' spoken English follow the results presented for mathematics and English literacy tests.

Table 7. Pupil Performance Levels and Abilities in English Literacy

GRADE	Estimated Performance Level for majority.	Skills 50% or more ARE able to do.	Skills fewer than 50% pupils can do.	Shifts in ability level and special skills attributed to QUIPS impact.
Grade 3	Class 1 to Class 2 (QUIPS and Controls)	Recognizes simple words and letters presented verbally.	Match words to picture. Read simple sentences. Less than 40% of pupils are reading any sentences.	Most pupils in QUIPS schools are able to match words to pictures.
Grade 4	Control Schools: Class 2 QUIPS: Class 2- Class 3	Match words to pictures. Complete open-ended sentences. Beginning to read simple 3-4 word sentences with picture support.	Majority of pupils in Control schools cannot read sentences. Less than 50% of pupils in Control schools were able to read any sentences. (QUIPS pupils were reading).	Most P4 pupils in QUIPS schools were reading simple sentences and passages. Notably higher performance overall, estimated to be at the P2 to P3 level.
Grade 5	Class 3 (QUIPS and Controls)	Match words to pictures (P1 TO P4 Vocabulary). Complete open-ended sentences (P1 TO P4 Vocabulary). Reads simple sentences without picture support (P1 to P3 Vocabulary)	Read full passages with meaning.	P5 pupils in QUIPS and Control schools performed similarly.
Grade 6	Class 3 to Class 4 (QUIPS and Controls)	Reads simple sentences without picture support (P1 to P4 Vocabulary) Read simple passages with understanding.	Higher level vocabulary and more advanced passages presented problems for the majority of pupils.	Pupils in QUIPS schools were reading more of the difficult sentences and passages.

Table 8. Pupil Performance Levels and Abilities in Mathematics

GRADE	Estimated Ability Level for Majority	Skills 50% or more ARE able to do.	Skills fewer than 50% pupils can do.	Shifts in ability level and special skills attributed to QUIPS impact.
Grade 3	Class 2	Simple addition without carrying Simple subtraction of single digits and no renaming. Simple multiplication facts. Very simple story problems (presented verbally with picture support).	Addition with carrying. Subtraction with renaming. Subtraction of numbers with double digits. Addition of fractions with common denominator. Measuring lengths.	Pupils in the QUIPS A general positive shift in distribution of difficulty levels (e.g., more children were successful on more of the test items). Majority of pupils from QUIPS schools were able to subtract double digits.
Grade 4	Class 3	Addition with carrying Subtraction with renaming. Simple multiplication without carrying. Simple story problems involving simple numbers to add, subtract, multiply, divide.	Addition or subtraction of fractions with common denominator. Multiplication with carrying. Very simple, single digit division.	Pupils in the QUIPS A general positive shift in distribution of difficulty levels (e.g., more children were successful on more of the test items). Majority of pupils from QUIPS schools were able to complete the measuring tasks.
Grade 5	Class 4	Addition with carrying. Subtraction with renaming (no renaming involving double digit problems). Simple multiplication without carrying.	Addition or subtraction of fractions with common denominator. Multiplication with carrying or double digits. Long division. No understanding of percentage. Measuring tasks.	P5 pupils in QUIPS and Control schools performed similarly.
Grade 6	P4 to P5	Addition with carrying. Subtraction with renaming. Simple multiplication without carrying.	Addition or subtraction of common fractions. Multiplication with carrying or double digits. No long division. No understanding of percentages.	Pupils in the QUIPS A general positive shift in distribution of difficulty levels (e.g., more children were successful on more of the test items). Could do complex story problems & simple fractions.

Spoken English Narrative

The spoken English component of the QUIPS achievement tests provide pupils with an opportunity to retell an English story using a picture storybook presentation. Prior to retelling the story, the pupils hear the story read to them by the test administrator on three occasions, twice in a group setting with other pupils and once in an individual session. As discussed previously, pupil responses on the storytelling task are scored according categorically, as follows:

Score of “5”- self-initiated, exceptional verbal response.

Score of “4” -self-initiated grammatically accurate verbal response;

Score of “3”- self-initiated, incomplete related or grammatically inaccurate verbal response;

Score of “2” –correct verbal response to a direct question given by the test administrator;

Score of “1” –pointing response to a direct pointing request given by the test administrator;

Score of “0” – no response even after test administrator cues are given such as direct questions and pointing requests.

Through this scoring processes it is possible to ascertain a general level of skill development in English speaking and listening comprehension. If children are able to meaningfully retell a story on his/her own, even if the grammatical accuracy is poor should have the needed “English literacy readiness”, at least with respect to a language fluency. On the other hand, if pupils are not able to answer direct questions about a story with picture prompts available to him/her or pupils don’t understand simple English words (i.e., fail to point to pictures in the story upon the verbal request of an examiner) then it is not likely that they are ready to learn to read English words.

In the following we discuss the general levels of spoken English development demonstrated by pupils who have just completed P3, P4, P5, and P6. The focus of our analysis is on the use of self-initiated verbal responses that pupils were able to give while trying to retell the English story. From this perspective, pupils in the QUIPS schools performed similarly to pupils in the Control schools. The one exception was the case of Grade 6. In P6 pupils from the QUIPS schools demonstrated a larger percentage of responses that were spontaneous verbal responses and grammatical accuracy was slightly higher for pupils in the QUIPS schools.

The greatest jump in English speaking was demonstrated between Grade 3 and Grade 4. In Grade 3 the majority of children had some difficulty retelling a story in English. When attempting to retell the story pupils were using self-initiated verbal responses only about 50% of the time. The remaining responses were answers to direct questions or pointing responses to the examiner’s picture identification requests. By the end of Grade 4; however, the majority of pupils were using self-initiated verbal responses around 83% of the time. For the most part the pupil’s verbal responses were not grammatically accurate in either Grade 3 or Grade 4.

The stimulus story in P5 and P6 was more complex than the story used in P3 and P4 and children’s use of self-initiated verbal responses declined slightly from P4 to P5 in the context

of the more advanced story. The majority of pupils just finishing Grade 5 used self-initiated verbal responses approximately 65% of the time and this was increased to 88% of the time in Grade 6.

One important aspect of the English narrative results is the marked variation across schools. In some schools there are no pupils who are able to attempt to retell a story in English in Grade 3 or Grade 4. On the other hand there are schools where most of the children had mastered English speaking and listening and the storytelling activity presented no challenge. Considering the importance of English fluency on English literacy development it is recommended that overall pupil ability in English speaking and listening be considered individually for each school and that pupils' English speaking and listening skills be considered as a key developmental step toward acquiring English literacy.

Figure 17 through Figure 24 show the distribution of pupil responses in the storytelling task for the QUIPS and Control schools. How did the QUIPS interventions impact pupils' spoken English development as assessed through the narrative exercise? Overall, the distribution of responses were among pupils just finishing their P3, P4, P5 and P6 classes was similar for pupils in the QUIPS partnership schools and the Control schools; however, results from the analysis of pupil learning growth rates that overall, rates of change for spoken English from P5 to P6 was significantly higher for pupils attending the QUIPS schools than for pupils from the Control schools. It can be seen in the Figures 19-20 and Figures 23-24, comparisons of the response patterns of pupils in the QUIPS and Control schools in P4 and P6 that though the overall patterns were similar, particularly with respect to the proportion of responses that were self-initiated, the proportion of responses that were grammatically accurate verbal responses (i.e., scored *SI-Complete*) were higher for pupils in the QUIPS schools than for pupils in the Control schools.

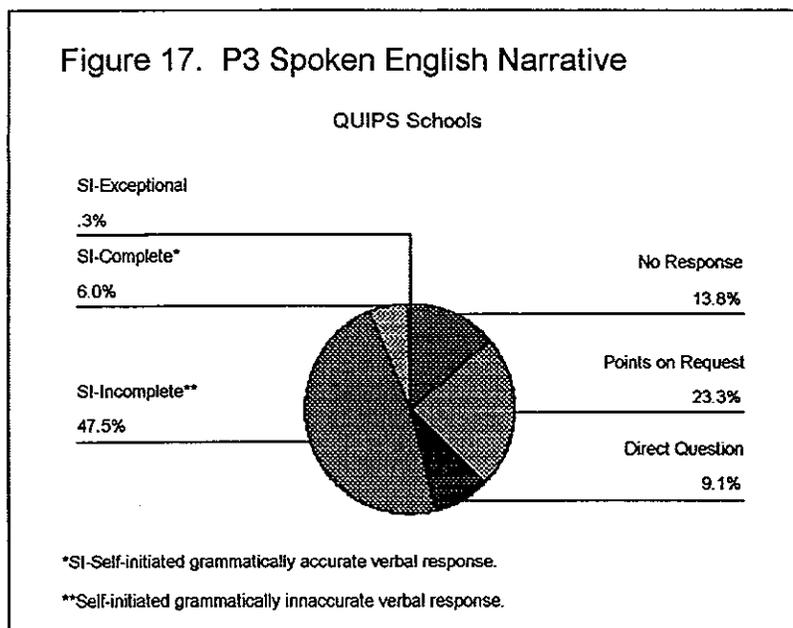


Figure 18. P3 Spoken English Narrative

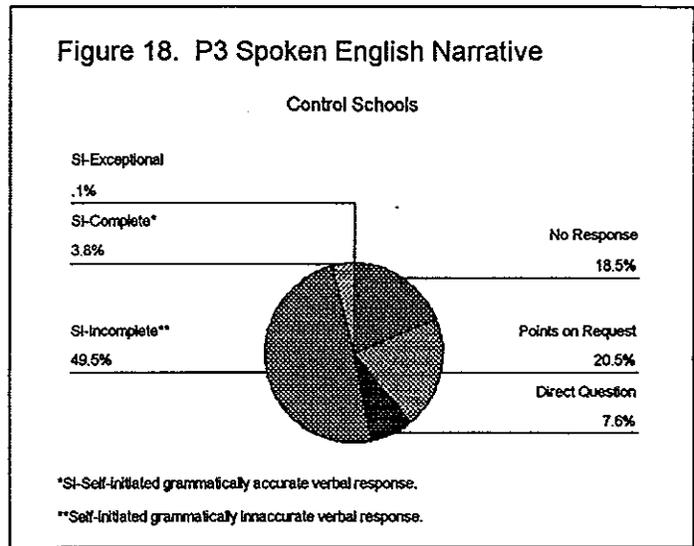


Figure 19. P4 Spoken English Narrative

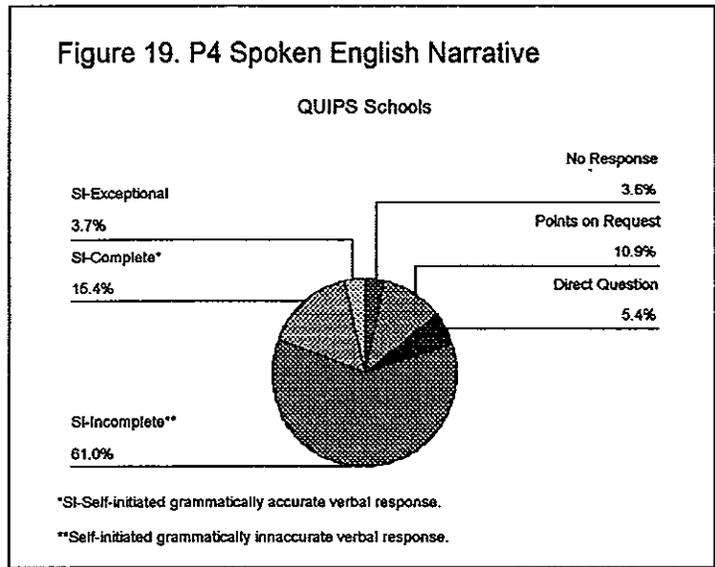


Figure 20. P4 Spoken English Narrative

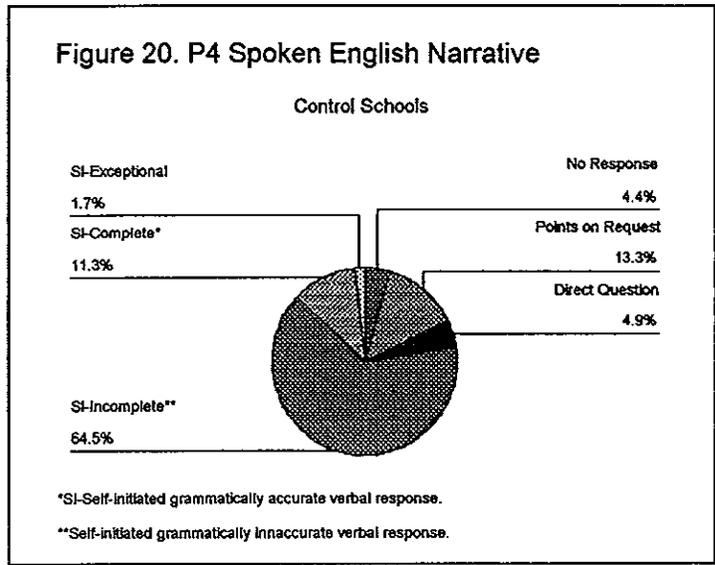


Figure 21. P5 Spoken English Narrative

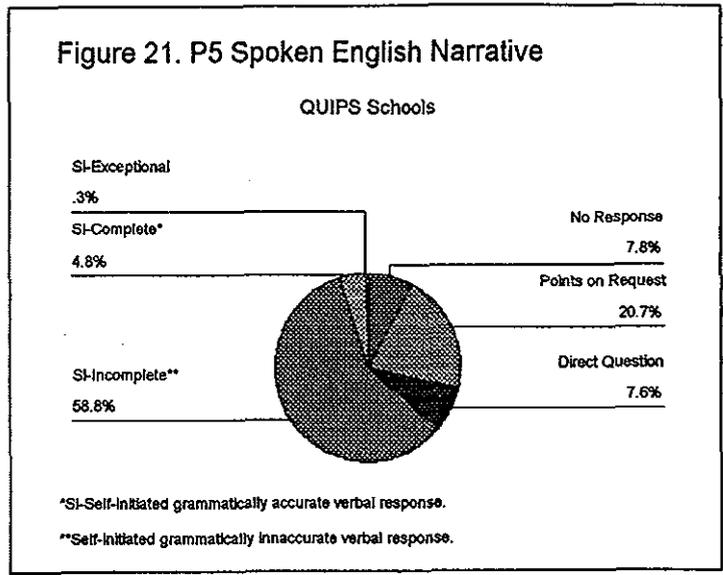


Figure 22. P5 Spoken English Narrative

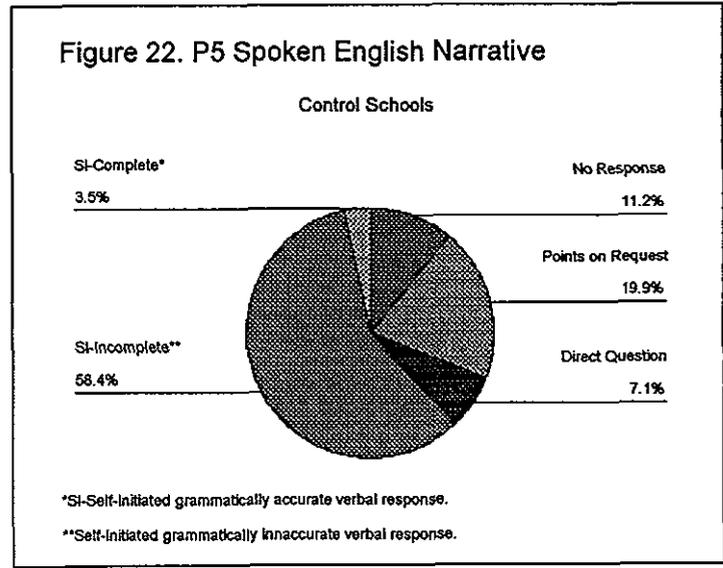
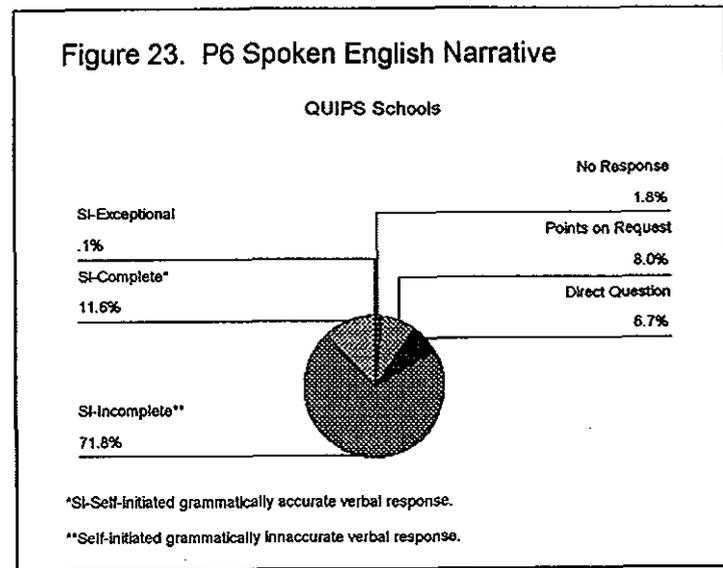
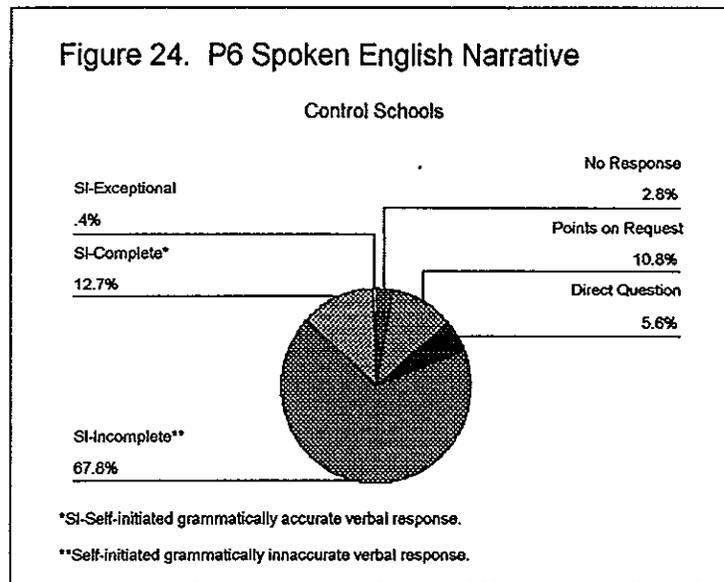


Figure 23. P6 Spoken English Narrative





Summary of Findings

In summary, the QUIPS interventions have made a consistent and positive impact on pupil achievement. Results from Cohort 1 and 2 findings immediately following the QUIPS two year intervention cycle demonstrated that there was a significant positive shift in the average class achievement from baseline in December 1998 to the end of the interventions for both cohorts in December 2000. In July 2001, one year after the QUIPS interventions had been discontinued in Cohort 1, a follow-on study of pupil achievement was conducted to evaluate the ability of the QUIPS partnership schools to maintain the gains made in the years following direct QUIPS support. Results showed that, with the exception of P5 English class performance levels obtained at the end of the QUIPS interventions were sustained over the first year following the withdrawal of QUIPS support. The follow-on study as originally planned focused on a study of the QUIPS schools only and did not include the Cohort 1 and 2 control schools. Therefore, no comparisons with Control schools could be made in the years following QUIPS support. It is recommended that in the future a random sample of the Control Schools also be included in the follow-on studies of pupil achievement.

The Cohort 3 impact evaluation focused on analyses of group differences in the rates of pupil learning in English literacy, mathematics and spoken English narrative. The result of these analyses indicated that for each subject area tested, the rates of pupil learning growth were higher for the QUIPS partnership schools than for the comparison schools, although group differences were not always statistically significant. Analyses of P3/P4 English literacy, P5/P6 English literacy and P5/P6 spoken English narrative did yield statistically significant results. The fact that the QUIPS partnership schools consistently, though not dramatically, demonstrated higher rates of learning growth than the Control schools is noteworthy. If these shifts in the growth curves are maintained, then the gains in achievement that have resulted from the QUIPS Program should be expected to increase even more in the years to come.

Therefore, it is critical that capacity is built within the system to provide ongoing support to the schools and communities that have been involved in the QUIPS Program to assist them in sustaining the good QUIPS school and community practices and thereby maintaining the positive shifts in the learning growth curves of pupils benefiting from the practices.

It would also be of benefit to continue to follow Cohort 3 (and Cohort 4) pupils who were tested in P3 and P4 through P5 and P6 in order to investigate the learning trends over through the two years following the QUIPS direct interventions. Such longitudinal analyses will provide information about the longer term impact or *sustained impact* of the QUIPS Program in the context of pupil learning growth for at least two years following the period of direct QUIPS interventions. Similarly to the recommendation made for follow-on testing in Cohort 1 and Cohort 2, it would be important to include a set of comparison schools in the follow-on testing program for Cohort 3 and Cohort 4. In order to accomplish this it would be necessary to equate the P3/P4 and P5/P6 test instruments using formal item equating procedures. In order to accomplish this, a somewhat large core of P3-P6 pupils would need to take both the P3/P4 mathematics and English tests and the P5/P6 mathematics and English tests at one time. TMG (and EARC) have the necessary human resources to administer the tests, though some support would be required for short term technical assistance and training on the procedure and to purchase the needed software (e.g., BILOG from Scientific Software, Inc.). This exercise could provide the additional training in Item Response Theory to the TMG team that may very well be needed in the future as the GES Performance Monitoring Test undergoes further developments. For example, some thought has been given to building a substantive database of test items that are properly aligned to the PMT test designs so that new tests can be more readily be generated from year to year. In order to reach this goal, application of the equating procedures being recommended for the QUIPS tests would need to be applied to the PMT test item banks.

Finally, TMG conducted a descriptive analysis of pupil abilities and ability levels in mathematics and English literacy by analyzing item responses on the QUIPS achievement tests. The study provided evidence that pupils were not performing at Grade level in mathematics or English. In general, pupils were performing approximately 1 year behind in mathematics and approximately 2 years behind in English. Pupils failed to master some of the most basic mathematics operations such as addition with carrying and subtraction with renaming until their 4th year of primary school. By the end of primary school the majority of pupils still were unable to do multiplication and division problems that involved carrying or multiplication of double digits. Long division, addition and subtraction of fractions and use of percentages also challenged most pupils finishing Grade 6.

Delays in developing English literacy are likely confounded by the multilingual challenges Ghanaian pupils face, particularly in the rural regions where exposure to the second language, English, is often non-existent outside of the school environment. Results indicated that the majority of pupils in Grade 3 have difficulty retelling a very simple English story designed for a pre-primary aged child. This situation is improved greatly by Grade 4 where the majority of children are able to attempt the storytelling task on their own. This delay in the ability to use spoken English until Grade 4 suggests that in many situations primary school pupils may not have sufficient readiness for learning to read in English until Grade 4. As performance in English speaking and listening varies considerably from school to school, it is recommended that schools consider the levels of English speaking mastery of pupils when considering their focus for English literacy intervention. English literacy interventions may want to consider a component that focuses on enhancing pupils' mastery of English speaking

in the early grades as a way of shifting pupils' readiness for learning to read in English to an earlier stage and thereby shifting the onset of English literacy development to a level that is more consistent with curriculum expectations.

Indeed, this shifting of the onset of meaningful reading appears to have been accomplished by the QUIPS interventions. One of the most positive findings of the abilities analysis was that the QUIPS Program did have a substantial impact on early pupil literacy. This was demonstrated by the fact that the majority of pupils just completing Grade 4 from the Control schools were not reading simple sentences with meaning. The majority of pupils from the QUIPS schools; however, who had just completed Grade 4 were reading sentences and simple passages with meaning. In fact, the data suggested that the majority of pupils from the QUIPS schools who just finished Grade 4 were performing overall at a Grade 2 to Grade 3 level while pupils from the Control schools were performing overall at a Grade 2 level.

Recommendations

Three primary recommendations are made regarding the impact evaluation program.

1. Comparison schools be included in all of the follow-on assessments.
2. Conduct a study that will allow the P3/P4 and P5/P6 achievement tests to be statistically equated and provide the necessary training to the TMG team in item response theory and item equating so that the team can be called upon by the GES to apply these techniques to further develop and improve the Performance Monitoring Test.
3. School intervention teams consider the results of the achievement performance analyses and where indicated further investigate the nature of constraints on specific skill development areas.