USAID
FROM THE AMERICAN PEOPLE

EVALUATION of the MALAWI TEACHER PROFESSIONAL DEVELOPMENT SUPPORT (MTPDS) PROGRAM

Final Evaluation Report

May 2013

This publication was made possible by the support of the American people through the United States Agency for International Development (USAID). It was reviewed by the USAID/Malawi Education team and prepared by Social Impact, Inc. (SI) through Dr. Thomas Tilson, Dr. Augustine Kamlongera, Mr. Mateusz Pucilowski, and Dr. Dorothy Nampota.

EVALUATION of the MALAWI TEACHER PROFESSIONAL DEVELOPMENT SUPPORT (MTPDS) PROGRAM

Final Evaluation Report

May 15, 2013

Contract No: AID-612-TO-13-00002

Requisition No: REQ-612-12-000029

TABLE OF CONTENTS

ACRONYMS

ABE	Assistance to Basic Education
CA	Continuous Assessment
CEED	Central East Education Division
COP	Chief of Party
CPD	Continuous Professional Development
CWED	Central West Education Division
DBE	Directorate of Basic Education
DEM	District Education Manager
DEO	District Education Office
DEP	Department of Educational Planning
DLC	District Literary Coordinators
DIAS	Directorate of Inspection and Advisory Services
DTED	Department of Teacher Education and Development
DQA	Data Quality Assessment
EDSA	Education Decentralization Support Activity
EGRA	Early Grade Reading Assessment
EMIS	Education Management Information System
ESIP	Education Sector Implementation Plan
ET	Evaluation Team
FGD	Focus Group Discussion
GoM	Government of Malawi
HT	Head Teachers
IE	Impact Evaluation
KII	Key Informant Interview
M&E	Monitoring and Evaluation
MANEB	Malawi National Examination Board
MIE	Malawi Institute of Education
MoEST	Ministry of Education, Science, and Technology
MTPDS	Malawi Teacher Professional Development Support
NED	Northern Education Division
PCAR	Primary Curriculum and Assessment Reforms
PE	Performance (Project) Evaluation
PEA	Primary Education Advisor
PSIP	Primary School Improvement Plans
PTA	Parent Teacher Association
RCT	Randomized Control Trial
SI	Social Impact, Inc.
SIP	School Improvement Plan
SMC	School Management Committee
SWED	Southwest Education Division
TTC	Teacher Training Coordinator
USAID	United States Agency of International Development

EXECUTIVE SUMMARY

Background

The Government of Malawi has made admirable gains in increasing access to education to the nation's students, even in the most marginalized regions, and doubling enrollment to 95% since 1994. However, this achievement has been coupled with significant challenges to the delivery of high-quality education as evidenced in students' weak literacy and numeracy scores. The 2010 EGRA (Early Grade Reading Assessment) baseline study with students at the beginning of standards 2 and 4 confirms that acquisition of reading skills in the early standards is woefully inadequate.[1] In response to this challenge, USAID/Malawi has shifted the focus of its educational programming from access to quality. USAID/Malawi's investment in the Malawi Teacher Professional Development Support (MTPDS) program shows Malawi's emphasis on increasing the quality of education for primary schools. MTPDS represents an integrated approach designed to improve educational outcomes by building the capacity of the Ministry of Education, Science, and Technology (MoEST), improving teacher efficacy, and improving school management and leadership.

Evaluation Purpose and Methodology

MTPDS was a three-year, $19.9 million USAID-funded program designed to improve learning outcomes for Malawi's primary school children. The program was implemented through three discrete "packages" of support, implemented at the school level:[2]

a) **Level 1 – National Continuous Professional Development:** CPD in-service teacher training includes modules on literacy and leadership that reach all 36,080 Standard 1–4 teachers in all 34 education districts nationwide through in-service training conducted over weekends and school holidays.

b) **Level 2 – Intensive Literacy Intervention:** This includes CPD plus additional days of training, scripted lesson plans, and the *Nditha Kuwerenga* Reader for every Standard 1 student in the seven target districts[3],

c) **Level 3 – Coaching:** This includes all previous interventions plus coaching by MTPDS and Primary Education Advisors (PEA) (select schools in Salima and Ntchisi).

USAID/Malawi commissioned an evaluation of the MTPDS activity comprised of two complementary components. An impact evaluation (IE) approach was to assess the differential impact and cost-effectiveness of the three levels of treatment intensity on student learning in reading. A performance evaluation approach was to document lessons learned and to assess progress towards sustainability (specifically the handing over of the national CPD program to the MoEST). The four-person evaluation team utilized a multi-level, mixed-methods approach to gain both a breadth and depth of information. These methods included: document review, data quality assessment (DQA), direct observation, key informant interviews (KII), focus group discussions (FGD), mini-surveys, and a statistical analysis of quantitative data to derive impact and associated cost-effectiveness estimates. Each technique facilitated triangulation of data during the analysis and reporting phases.

[1] Data from the EGRA baseline study: 73% of G2 and 42% of G3 could not read a single word of the story and 97% of G2 and 69% of G3 could not answer a single comprehension question correctly.
[2] MTPDS defined the interventions in reverse directionality (Level 3 = minimum and Level 1 = maximum). All MTPDS data was recoded to reflect definitional differences..
[3] Salima, Ntchisi, Mzimba North, Thyolo, Blantyre Rural, Zomba Rural, and Ntcheu, representing six educational divisions.

Summary of Conclusions

EVALUATION QUESTION 1 (PROJECT DESIGN):

What is the impact and associated cost-effectiveness of the three training modalities on reading outcomes and teacher performance?

Data Quality Assessment

- MTPDS included an integrated IE designed to attribute changes in literacy competencies to the program. However, given the program's prioritization of development assistance over that of research, MTPDS made a number of methodological concessions. Most importantly, the MTPDS IE design did not include a control group, making it impossible to definitively ascribe impact to the intervention. Regardless, rigorous selection of schools and students facilitated a comparison of Level 1 and Level 3 support.
- The EGRA instrument used to measure changes in literacy was designed and piloted in accordance with international best practices. There were a number of issues with regard to data collection and entry that present cause for concern. On balance, the data were of sufficient quality to justify statistical analysis.
- The classroom observation instrument was found to be deeply flawed, particularly with regard to instrumentation. While the instrument may have been useful for program monitoring purposes, multiple shortcomings inhibit its utility for analyses of impact or for informing policy dialogue.

Literacy

- Level 3 students consistently outperformed Level 1 students in literacy measures (ability to read and answer comprehension questions). This finding held true for zero scores, mean scores, regression analysis, and gender-disaggregated analyses. Students in the high-intensity group made dramatic improvements; the number of students who could not read a single letter dropped by more than 50%, while the average child's ability to read increased by 20 letters per minute. Regression analysis found participation in the high-intensity group to be associated with the ability to read an additional five letters and more than two additional words per minute as compared to students in Level 1.
- Changes in literacy scores for the Level 1 group were nominal. Some measures recorded slight improvements while others registered small decreases. On balance, the Level 1 intervention did not seem to have significant impact on literacy scores. However, without a control group to compare against, it is not possible to determine what would have happened without any MTPDS support.
- Gender disaggregation supported the macro analysis. Both male and female students in the high-intensity group recorded substantially higher increases in literacy outcomes across all three measures. Whereas male students had consistently higher literacy measures at baseline, over the course of MTPDS the two genders became more equitable with regard to basic literacy skills.
- While largely positive for the Level 3 group, these findings have to be viewed within a broader context. First, substantial gains in Level 3 literacy scores notwithstanding, the high-intensity group's endline literacy scores were still poor: nearly half of the group could not read a single word in a minute. Additionally, as documented in the DQA section, there were a number of methodological limitations and potential sources of bias in the MTPDS data. While

methodological and data quality issues limit the internal validity of the preceding conclusions, the general trend (the Level 3 group outperforming the Level 1 group) was triangulated through other data collection and analytical techniques.

Teacher Performance

- Analysis of MTPDS classroom observation data suggests that teachers across the three treatment levels increased their pedagogical abilities. The most pronounced improvements, both in terms of reductions in floor scores and increases in ceiling scores, were registered by teachers in the Level 3 group.
- While teachers improved across all key measures, the most striking finding was a complete absence of endline floor scores for phonics, comprehension, and writing. On the basis of this data, it appears that teachers had the most difficulty operationalizing best practices in the assessment and text-rich environments measures. Barriers in implementing these activities are likely independent of teacher knowledge and motivation, deriving primarily from an overabundance of learners and a dearth of teaching/learning materials, respectively.
- The data paint a very positive picture, with teachers making striking improvements across a variety of measurement areas in the span of a single year. As discussed in the DQA section, however, any analysis of changes in teacher performance using MTPDS classroom observation data should be tempered with regard to instrumentation and data quality concerns.

Cost-Effectiveness Analysis

- Due to the lack of control group data, cost-effectiveness estimates for the Level 3 group were conducted by using Level 1 as a basis of comparison. The *per teacher* cost for providing Level 1 support was calculated to be $112.12, while the *per learner* cost was $1.38. The corresponding costs for Level 3 support were $431.22 and $7.89, respectively.
- According to regression-driven cost-effectiveness analysis, a dollar of Level 3 intervention would be expected to increase reading ability of an average child by an additional 3.26 letters/minute, as compared to Level 1. The corresponding values for words/minute and reading comprehension were 0.66 and 0.002, respectively.
- The analysis presents the cost-effectiveness ration for MTPDS. If MoEST were to continue providing this type of support, the ratio would be more favorable for two reasons. First, the development costs are largely a one-time expense. Revising materials would be less labor-intensive than developing them from scratch. Secondly, MoEST implementation costs would likely be much lower than that of the international firms implementing MTPDS.

Independent Analysis of Package Contents

- The evaluation team administered questionnaires to 104 Standard 1–4 teachers in the course of field work. Responses demonstrated strong support for and utilization of MTPDS components amongst primary school teachers, regardless of treatment intensity. The vast majority of teachers who received trainings found them useful and practical to their teaching, although significantly more teachers reported learning useful information than reported using the information in their classrooms.
- Lesson plans and coaching were reported to be the most important programmatic elements. While training was not ranked as highly, survey data and qualitative findings suggest that it provides teachers with a better understanding of how to use other inputs more successfully.

Overall, MTPDS assumptions of programmatic synergies and strong support for the program amongst beneficiaries were validated by teachers' responses.

EVALUATION QUESTION 2 (LESSONS LEARNED):

What have been the strengths and weaknesses in the implementation of the three levels of training intensity, and how can these findings be used to inform future programming?

- The CPD training program was effective in improving teacher practices, though the IE section makes clear that these gains were not sufficient in and of themselves to translate to increased literacy scores. Some of the techniques are difficult to implement in a resource-constrained environment. Teachers valued the new knowledge and skills they learned across treatment levels and schools.
- Arguably the single-most important contribution of the program was to quicken a return to using the syllabic approach to teach reading. The MoEST curriculum in use during the MTPDS program period utilized a whole-word approach and was generally regarded to be ineffectual. Working with MoEST stakeholders MTPDS promoted a return to the traditional way of teaching reading using letters and syllables. This approach resonated with teachers (particularly those that had experience using the SOSA book) and was widely praised for having improved literacy amongst primary school students.
- Scripted lesson plans and, to a somewhat lesser extent, the *Nditha Kuwerenga* reader are powerful tools to enhance reading skills of learners. Teachers' use of the extra hour was instrumental in achieving literacy by higher intensity schools and was a key contributor towards increased reading competencies.
- Coaching was found to be useful by beneficiary teachers and, if used regularly, would likely lead to greater improvements in teacher efficacy, and hence literacy. However, unless additional resources are obligated, institutional constraints will likely limit the extent to which PEAs are able to support teachers through school visits.
- The leadership training program is valuable and head teachers appreciated the new knowledge and skills that they learned in all of the three treatment level districts. School Report Cards can be useful in assessing teacher actions and new interventions.

EVALUATION QUESTION 3 (OWNERSHIP):

What is the current state of transfer of responsibility and ownership for CPD to the MoEST, and how can the handover be strengthened to bolster program sustainability?

- MTPDS activities have garnered strong support owing in part to opportune timing: The whole-word approach under old Primary Curriculum and Assessment Reforms (PCAR) was rejected in favor of a return to the syllabic approach as the program was being implemented. Given its embrace of the latter, MTPDS was supported and accepted at the grassroots (community and schools) and Ministry levels (district and national). As evidenced by the new primary curriculum and associated materials, especially for Standard 1, which utilize the syllabic approach, the most important contribution of MTPDS will be sustained through the foreseeable future. While this transition would have happened in the absence of MTPDS, the program quickened the return to syllabic instruction and improved the lives of many children that would have not been reached by the revised MoEST materials.
- Responsibility of MTPDS is being accepted by MoEST departments (DIAS, DBE, DTED, DEP, and MIE). However, this process was begun late and could have been more effectively

implemented. There are discussions about adapting MTPDS approaches to complement the revised MoEST curriculum.

EVALUATION QUESTION 4 (SUSTAINABILITY):

To what extent have teachers, school administrators, and other stakeholders embraced and/or bought into the MTPDS approach, including classroom practices and instructional tools, delivered through CPD and the literacy intervention?

- Sustainability of MTPDS components among key stakeholders is significantly associated with access and maintenance of relevant resources as well as the decrease of class sizes.
- Teachers' use of pedagogical skills/techniques shared through Literacy CPDs is likely to continue as will the expansion to additional teachers and classrooms. The teaching of literacy using a *syllabic approach* is likely to continue as it does not require additional resources and the teachers have expressed strong support for it.
- Some pedagogical techniques are difficult to implement as they currently are implemented due to large class sizes and the need for additional resources. The following techniques were found to be the most challenging given these obstacles: continuous assessment, grouping, and print-rich classroom environments.
- Concern for the sustained use of readers may not be urgent given that the MTPDS was designed to provide teachers with methods/skills for teaching using the syllabic approach. With new primary curriculum materials on the way, there should be less need for supplemental materials (including the reader) to the national curriculum.
- Scripted lesson plans were a powerful tool to enhance reading skills and they provided a bridge that allowed students to learn to read while revised textbooks were developed. Lesson plans that rely on the reader will most likely become phased out as the readers degrade in use or lose relevance with respect to the new curriculum.
- Given support for additional time-on-task, it is likely that "extra hour" will continue in the short term in intervention districts but may be reduced/eliminated as MTPDS is phased out.
- PEAs will not be able to increase the number of visits beyond the current average of two per term because of their current list of responsibilities and insufficient resources. The lack of a focused cascade training model severely limits the possibility of school-level support.
- School-based CPDs will likely continue because there is strong support from administrators.
- Community members and especially parents will continue to engage with the schools as they note the improved reading ability of their children. SMCs will probably remain marginally involved in school affairs if their current scopes remain as they are.

Summary of Recommendations

FUTURE PROGRAMMING

- Given planned incorporation of key MTPDS inputs in the revised primary school curriculum it is not recommended that USAID or MoEST implement any of the three packages, as they currently exist, in the future.
- To the extent that the new primary curriculum materials and other Ministry initiatives omit important program components (particularly syllabic approach, scripted lesson plans and extra hour), it is recommended that MoEST discuss with key stakeholders how best to integrate

lessons learned into the formal education system. It is critical to include these changes as part of curricula in the nation's Teacher Training Colleges (TTCs).

- MoEST should integrate the multiple reading programs (MTPDS, Read Malawi, new primary curriculum, etc.) into one coherent curriculum, drawing on the lesson learned and benefits of each approach.
- MoEST, USAID and other stakeholders should discuss how to best fund and implement literacy and leadership CPDs at the district, zonal and school levels. Realizing the resource constraints and the substantial challenges faced by Malawi primary education stakeholders, it is recommended that CPD be prioritized and funded as a low-cost way to increase teacher competencies. Where possible, District Education Managers (DEMs) should be encouraged to obligate more funding for CPD training at the district and zonal levels. In addition, schools should be encouraged to request funds in their SIPs for school-based CPDs.
- The Malawi Institute for Education (MIE), in collaboration with DTED and DIAS, should review and update key MTPDS materials, as well as create new refresher modules to help sustain and extend educational gains.
- The Secretary for Education should identify a point of leadership and responsibility in the Ministry to support the literacy program. In addition, there is a need to define the roles and responsibilities for all of the key Ministry stakeholders.

MTPDS PROGRAM COMPONENTS

Literacy CPD

- Head teachers and, resources permitting, all teachers S1–S8 should be invited to the Literacy CPD training so that they are in a position to support the program at their school.
- The Ministry should develop easy-to-implement continuous assessments, especially for large classes. It is suggested that training on such assessments teach teachers how to assess collectively (alongside a head teacher, classroom aid, fellow teacher, etc.)
- Teachers and school administrators should be provided with further guidance on how to actively engage students during grouping exercises and when grouping is/isn't effective.
- Teachers could be given further guidance on how to use more limited but revolving print materials in a non-secure environment.

Nditha Kuwerenga Readers

- With the new primary and national curriculum, the Ministry should incorporate lessons learned from the use of the readers into new lesson plans for teachers' and student's books.
- Should the new primary curriculum student book be insufficient, the Ministry could consider developing a new reader. The Ministry can provide training on using the readers through CPDs and TTCs.

Lesson Plans

- MIE should adapt scripted lessons plans to reflect the new primary curriculum, involving relevant personnel from MoEST. Lessons should be shortened with optional activities included for more able classes, and guidance provided on how teachers can be weaned off of simply reading the scripted materials.

- CPDs and TTCs should include training on the use of revised lesson plans such that teachers learn how to personalize lessons during implementation.

Extra Hour

- The Ministry should institutionalize the extra hour and provide guidance on how to implement and best utilize this additional time.

Coaching

- Head teachers and deputies of all districts should be trained on how to conduct coaching and incorporating this into their official scopes of work. Coaching could then be provided at a higher frequency and lower cost than the MTPDS approach.
- The Ministry should provide adequate resources needed by PEAs to continue coaching including coverage of transportation costs and vehicles, and revision of workloads.

Leadership CPD

- School-based leadership CPDs should be funded at the local level through PSIPs and better guidance about implementation provided by the Ministry
- The scopes of work for head teachers (HT) should be reduced and/or re-organized in order to better sustain the responsibilities for coaching and organizing CPD training.
- Community participation, especially among parents, in monitoring reading should be re-emphasized through community outreach and through PTAs

FUTURE RESEARCH

- If similar, large-scale data collection activities are planned in the future, it is highly recommended that ample time be dedicated to instrument and data collection protocol design in the early stages of a program. Once finalized, it is imperative that substantive revisions not be made to either during the period of implementation.
- The utilization of MoEST personnel in high-stakes data collection presents both benefits and costs. For future impact evaluations it is recommended that professional and impartial enumerators be used in lieu of MoEST personnel. In the event that this option is not feasible due to budgetary pressures, it is strongly recommended that MoEST personnel engage in data collection not enumerate in areas under their direct responsibility.
- The ability to make inferences from evaluation findings depends on the validity of the underlying data. It is recommended that any party responsible for high-stakes data collection utilize double data entry (if using paper surveys) or electronic data collection methods to ensure that information is of sufficient quality.
- Last, if ascription of educational outcomes to USAID funding is desired in the future, it is imperative that a true control group be integrated into research designs. If withholding treatment is not feasible due to ethical, political, or logistical reasons, the evaluation team recommends phased implementation.

INTRODUCTION

Since the abolition of school fees in 1994, primary school enrollment in Malawi has doubled to 95 percent[4]. While this result is laudable, measures of learning outcomes indicate significant underperformance in schools. The 2010 EGRA (Early Grade Reading Assessment) baseline study conducted in primary schools with children at the beginning of standards 2 and 4 confirms that the learning of children in reading is significantly inadequate.[5] In response to this challenge, USAID/Malawi has shifted the focus of its educational programming from access to quality by concentrating on *teachers*' performance as a direct influence on *students*' performance. The Mission has thus invested in the three-year Malawi Teacher Professional Development Support (MTPDS) project. MTPDS was designed to improve the quality of primary education through a suite of interventions including continuous professional development (CPD) of teachers and school leaders/managers.

Project Background

USAID/Malawi has collaborated with the Government of Malawi (GoM) and its Ministry of Education, Science, and Technology (MoEST) to focus on the MTPDS strategy that emphasizes teacher professional development as a conduit for enhancing literacy among early-grade learners. By achieving targeted results, MTPDS would build the capacity of the Ministry of Education, Science, and Technology (MoEST); improve teacher efficacy in teaching reading; and improve school management and leadership by targeting school governance and management structures.

MTPDS is implemented by the Assistance to Basic Education (ABE) team composed of Creative Associates International Inc., RTI International, and Seward Inc. The project's period of performance is February 2010 – March 2013, and has an estimated total cost of $19,990,000. Originally, MTPDS also complemented national Primary Curriculum and Assessment Reforms (PCAR) and other Programs of Work (POW), including mathematics. The Task Order was modified effective December 5, 2011, to reduce the scope of the activity and to give a larger focus on literacy.

The MTPDS target groups consist of approximately: 34,000 teachers who teach in Standards 1–4; 10,300 head teachers and deputies in public schools in Malawi; 2.7 million students in the first four standards of primary school; and MoEST officials at the central, divisional, and district levels. The activities of MTPDS are organized into five result areas, each of which reflects a current MoEST policy priority:

Result I – Strengthened Teacher Policy, Support, and Management Systems;
Result II – Enhanced Teacher Performance;
Result III – Improved Early Grade Literacy;
Result IV – Enhanced Quality of Primary Teaching and Learning Materials;

[4] http://home.hiroshima-u.ac.jp/cice/12-1MikiKeiDaniJosecDemiJosegAlbNobuSho.pdf
[5] Data from the EGRA baseline study: 73% of G2 and 42% of G3 could not read a single word of the story and 97% of G2 and 69% of G3 could not answer a single comprehension question correctly.

Result V – Improved monitoring and evaluation (M&E) systems focused on teacher competencies and learner outcomes.

MTPDS Project Logic

MTPDS activities follow a multi-pronged approach to achieve these five results, though the focus on student literacy primarily calls for emphasis on those activities that will especially lead to results II and III. The mechanism by which MTPDS strives to achieve these two results is the provision of a suite of teacher training and other interventions. In an effort to gauge the types of inputs having the most profound effects on increasing teacher efficacy and student literacy skills, the project developed three modalities of support (treatment levels) with Continuous Professional Development (CPD) at the core. These treatment levels are defined here such that treatment intensity increases by level[6]:

a) **Level 1 – National Continuous Professional Development:** MTPDS provided two types of CPD: literacy training for Standard 1–4 teachers and leadership training for head teachers and deputy head teachers. Literacy training was composed of two-day modules, and was conducted over weekends and school holidays. Leadership training was implemented similarly, but only covered two two-day modules. CPD in-service teacher training reached all 36,080 Standard 1–4 teachers across all 34 education districts nationwide.

b) **Level 2 – Intensive Literacy Intervention:** Intensive Literacy Intervention includes CPD plus scripted lesson plans, a *Nditha Kuwerenga* Reader for every Standard 1 student, additional days of training, and an optional extension of the school day by one hour to accommodate literacy activities. Level 2 was implemented in seven target districts.[7]

c) **Level 3 – Coaching:** The coaching intervention includes all previous interventions plus coaching by MTPDS officers. Primary Education Advisors (PEAs) have been trained on coaching, and by October 2012 they began to carry out coaching in the seven high-intensity intervention districts. Level 3 was implemented in select schools in all seven target districts.

EVALUATION FRAMEWORK

Purpose of the Evaluation

USAID/Malawi commissioned this summative evaluation of MTPDS for three complementary purposes. *First*, the evaluation was to assess the impact and cost-effectiveness of the three

[6] The Evaluation Team classified and analyzed data at each treatment level in order of intensity beginning with low intensity at Level 1 and high intensity at Level 3. However, the MTPDS team defined the levels in the opposite direction such that Level 1 schools received the full intervention; Level 2 schools received the literacy intervention but no coaching; and Level 3 schools only received the CPD training and essentially no intervention. This difference was captured during the data collection phase of the evaluation and accounted for during data analysis.

[7] Salima, Ntchisi, Mzimba North, Thyolo, Blantyre Rural, Zomba Rural, and Ntcheu, representing six educational divisions.

MTPDS training modalities on two key outcomes: (1) student learning, as measured by the Early Grade Reading Assessment (EGRA), and (2) teacher performance, as measured by classroom observation data from MTPDS. After calculating impact estimates, the team was to compare programmatic benefits with their cost to provide policymakers with information to compare the modalities against one another. As discussed in detail in the body of the report, MTPDS operationalized a robust M&E system. This included an integrated impact evaluation (IE) with randomized assignment of schools into the various training modalities and large-scale data collection. The SI Evaluation Team was commissioned to perform an external DQA to assess the validity of MTPDS outcome data and to conduct an impartial analysis of programmatic impacts so as to adhere to the USAID Evaluation Policy mandate for objective evaluation.

Second, the evaluation was to examine MTPDS management, specifically exploring the strengths and weaknesses of implementing the three levels of intensity. The team was to identify key strengths and challenges, and suggested ways to improve the potential for sustainability. *Third*, the evaluation was to assess progress towards sustainability, especially the prospects for handing over the national CPD program and associated activities to the MoEST.

Evaluation Questions

The purposes of the evaluation were met by collecting and analyzing data relevant to the following evaluation questions:

1. What is the impact and associated cost-effectiveness of the three training modalities on reading outcomes and teacher performance? (**Project Design)**
2. What have been the strengths and weaknesses in the implementation of the three levels of training intensity, and how can these findings be used to inform future programming? **(Lessons Learned)**
3. What is the current state of transfer of responsibility and ownership for CPD to the MoEST, and how can the handover be strengthened to bolster program sustainability? **(Ownership)**
4. To what extent have teachers, school administrators, and other stakeholders embraced and/or bought into the MTPDS approach, including classroom practices and instructional tools delivered through CPD and the literacy intervention? **(Sustainability)**

Evaluation Methodology

The evaluation utilized a multi-level, mixed-methods approach to gain both a breadth and depth of information sufficient to answer all four evaluation questions. In this instance, multi-level refers to the incorporation of stakeholders from all levels of the Malawian education system, including the MoEST, district education offices (DEOs), head teachers, Standard 1–4 teachers, members of parent-teacher associations (PTAs), and school management committees (SMCs). *Mixed-methods* refers to the inclusion of various data collection and analytical methods relevant to a quantitative impact evaluation (IE) as well as a qualitative performance evaluation (PE). The specific methodologies used for each of these evaluation approaches are detailed below.

IMPACT EVALUATION: QUESTION 1 (PROJECT DESIGN)

Impact estimates were calculated solely on the basis of outcome data collected by MTPDS: EGRA and classroom observations datasets. However, before assessing the impact of the project's three treatment modalities, the evaluation conducted a data quality assessment (DQA) to determine the reliability of the underlying datasets. The DQA consisted of six phases designed to review the MTPDS data collection approach, from enumeration to data entry and analysis:

1. Review of sampling frames used to select schools and students in the program design;
2. Review the selection of control districts—Dedza and Mwanza;
3. Review of data collection instruments relevant to classroom observations;
4. Review of data collection protocols included in enumerator manuals for the EGRA and classroom observation activities;
5. Review quality of data entry, storage, and cleaning protocols for EGRA and classroom observations by spot-checking relevant databases; and
6. Spot checks to verify MTPDS data through key interviews with school administrators and teachers and verification of scripted lesson plans, readers, and coaching logs.

Lastly, financial data was used to calculate the cost-effectiveness estimates. The analytical methodology for the IE and cost-effectiveness components are described in detail in the body of Chapter I.

PERFORMANCE EVALUATION: QUESTIONS 2–4 (LESSONS LEARNED, OWNERSHIP, AND SUSTAINABILITY)

The performance evaluation approach was comprised of three primary components: (1) Document Review, (2) Key Informant Interviews, and (3) field work. After completing a thorough document review, the team held an in-brief with USAID/Malawi and conducted Lilongwe-based KIIs with stakeholders from the Mission, MTPDS, and MoEST. Subsequently, the team pilot-tested all school-based instruments (classroom observation, teacher questionnaire, FGD, and KII scripts with head teachers and SMC/PTA members). Following a final revision of instruments the Team split into two groups and conducted visits to 24 schools in six districts. In each district, the Teams did the following:

1. Visited district (and where appropriate, divisional) education offices to interview Ministry and MTPDS staff
2. Visited schools to perform the following activities:
    a. Collect basic school information
    b. Conduct KIIs with head teachers

c. Conduct structured classroom observations with 1–2 Chichewa classes (in a few cases the teams observed English to explore the possibility of spillover effects). As part of the DQA, the Team used the MTPDS classroom observation form for a week. Once sufficient data on instrument performance was collected, the Team began using a custom tool better suited for evaluation purposes.

d. Perform direct observations of MTPDS Implementation. The Teams observed whether MTPDS materials were available and/or being used as intended. This included spot checks of readers, lesson plans, and whether the extra hour was evident in the school timetable.

e. Administer a custom teacher questionnaire with Standard 1–4 teachers

f. Conduct FGDs with Standard 1–4 teachers

g. Interview SMC/PTA chair(s) and, in a few instances, additional members

SAMPLING

Selection of districts: On the basis of evaluation purpose and budget, the Evaluation Team visited six districts, stratified on four key variables: treatment intensity, baseline performance, geography, and language. Two districts were selected for each treatment level, at least one district represented each of the three regions, and two districts had schools where Chichewa was not the dominant language. Lastly, the Team included schools from the highest-performing (Central East) and lowest-performing (Central West) divisions.[8] The sampled districts and their divisions were:

- **Level 1**
  - *Dedza*—Central West Education Division (CWED)
  - *Nkhata Bay (non-Chichewa)*—Northern Education Division (NED)
- **Level 2**
  - *Blantyre Rural*—Southwest Education Division (SWED)
  - *Mzimba North* (non-Chichewa)—Northern Education Division (NED)
- **Level 3**
  - *Salima* and *Ntchisi*—both of the Central East Education Division (CEED)

Selection of schools: Each sub-team sampled four schools within each district, for a total of 24 schools. The schools were selected on the basis of a stratified, purposive sampling approach that utilized the 2011 Education Management and Information System (EMIS) database. Within each district, four schools were selected to fit three strata: one school represented resource-rich schools, one represented resource-constrained schools, and two were selected such that they represented an average resource base and were proximate to one another. This approach

[8] Data was derived from a review of multiple sources of MTPDS baseline data, including EGRA, Classroom Observation, and teacher questionnaires.

allowed the team to survey schools covering a broad range of affluence while also ensuring that four schools could be visited in each three-day site visit period. The final list of schools can be found in Annex 2.

EVALUATION TEAM

The Evaluation Team consisted of a select group of education and evaluation specialists with balanced expertise across subject matter, evaluation methodology, and local research context. The team composition included two local Malawians and two expatriates, with local personnel accounting for over 50% of total level of effort. Team members were:

- **Dr. Thomas Tilson** (Evaluation Team Leader) has over 40 years of educational programming, team leadership, and evaluation experience throughout Africa. Dr. Tilson is deeply knowledgeable about the educational, economic and political context in Malawi.
- **Dr. Augustine Kamlongera** (Deputy Team Leader) is an experienced team leader with 29 years of experience in educational development, monitoring and evaluation, and implementation of primary education projects in Malawi.
- **Dr. Dorothy Nampota** (Educational Research Specialist) has over 20 years of leadership and research experience in the Malawian education sector.
- **Mr. Mateusz Pucilowski** (Education Evaluation Specialist) is an evaluation methodologist with experience designing and implementing impact and performance evaluations of education programming.

In addition to the four core team members, two USAID/Malawi staff shadowed the Evaluation Team during fieldwork and intermittently participated in the data collection. The staff members were Mr. Chikondi Maleta and Mr. Chimwemwe Chitsulo.

FINDINGS, CONCLUSIONS, & RECOMMENDATIONS

This section of the report presents key findings, conclusions, and recommendations separated by evaluation question: (1) Project Design; (2) Sustainability; (3) Ownership; and (4) Lessons Learned. Given substantive methodological differences, this report analyzes literacy and teacher performance outcomes separately under the section for Project Design. Both sections contain a brief introduction, a data quality assessment (DQA), and a quantitative impact analysis that presents findings and conclusions. These analyses are then triangulated and further explained through primary teacher questionnaire data collected by the Evaluation Team.

Project Design is followed by performance evaluations of the three remaining evaluation questions. These sections begin as a second volume to the report.

I. IMPACT EVALUATION

This chapter of the report presents key findings, conclusions, and recommendations for Evaluation Question 1: Project Design. Before presenting analysis, the section presents relevant background information as well as a condensed discussion of IE limitations to this evaluation. Literacy and teacher performance outcomes are then presented separately, with the respective sections containing a brief introduction, data quality assessment (DQA), and quantitative impact analysis. These analyses are then triangulated and further explained through primary teacher questionnaire data collected by the Evaluation Team.

Background

In addition to implementing a complex, multi-component program on a national scale, MTPDS incorporated a robust monitoring and evaluation system. The centerpiece of this system was multiple large-scale data collection efforts.[9] MTPDS intended for this data to meet twin goals: to track performance against targets and to estimate programmatic impact. To facilitate the latter goal, the program tracked changes in literacy scores for a subset of randomly selected schools in each of the three treatment modalities. MTPDS was, however, first and foremost a development intervention, not a research project. As such, programmatic design required the IE to make two methodological concessions: (1) due to MTPDS mandate to provide support to all schools in Malawi, the IE could not compare treatment units against a true, non-recipient control group, and (2) due to the fact that not a single district received all three treatment modalities, the IE could not compare the full range of treatment intensities within the same context. In light of these limitations MTPDS designed a hybrid IE that compared Level 3 and Level 2 schools from one set of districts with Level 1 schools in another set of districts.

Salima and Ntchisi districts were purposively selected to participate in the high-intensity intervention.[10] Within each of these districts, MTPDS used the 2009 EMIS database as a sampling frame from which schools were randomly assigned into two levels of intensity: eight schools each were selected for Level 3 and Level 2 treatments.[11] As there were no Level 1 schools within these districts, the program selected the neighboring districts of Dedza and Mwanza as a basis of comparison. These districts were selected by MTPDS on the basis of demographic comparability as well as the relative absence of other literacy interventions. MTPDS randomly selected eight schools in each of these "control" districts, for a total of 16 control schools. Within each of these schools program staff surveyed a random sample of Standard 2 students using the Early Grade Reading Assessment (EGRA). The baseline was conducted before program implementation (November 2010), while the endline occurred two years later (November 2012).

While outside the scope of the MTPDS IE, the evaluation SOW mandated an impact evaluation for a second key outcome: teacher performance. This measure was captured by MTPDS

[9] Three rounds of classroom observation, three rounds of national EGRA, and two rounds of literacy intervention-specific EGRA.

[10] The decision to target Salima and Ntchisi with the intensive reading intervention largely stems from a previous program activity that involved support for complementary basic education. This result area was removed in a contract modification, but work had already begun in the two districts.

[11] Due to the loss of the MTPDS M&E specialist, the specific randomization methodology is unknown.

through a classroom observation tool customized for the project and implemented in 2010, 2011, and 2012, with teachers in all three treatment intensities.

Limitations[12]

Lack of Control Group: The ability of an IE to attribute changes to a program is directly dependent on its ability to estimate the program counterfactual, or what would have happened had the program not been implemented. Given the lack of a true control group in the MTPDS IE (i.e. Level 1 schools in Salima and Ntchisi), the evaluation does not allow for definitive measurement of the true impact of the three treatment intensities. However, given the limited nature of Level 1 support, CDP-only schools were used as a basis of comparison against which to gauge the impact of the literacy intervention.[13]

Selection Bias: Comparisons between the three treatment modalities are valid to the extent that the underlying characteristics of the sampled groups were the same at the time the program was introduced. While the IE design followed methodological best practices in randomly assigning schools in literacy intervention districts into either Level 2 or 3, comparison with schools in Level 1 districts introduces potential selection bias. Given implementation constraints, MTPDS selected the Level 1 schools in as rigorous a manner as was possible. Regardless, baseline differences between the high-intensity (Level 2 and 3) and low-intensity (Level 1) districts call into question the validity of the comparison, as compared to standards of IE research.

Coaching Spillovers: Whereas MTPDS originally intended to compare three discrete treatment levels, the practical distinction between Level 2 and Level 3 disappeared in the course of program implementation. The sole aspect differentiating these two levels, coaching, became integrated within both levels as PEAs and MTPDS staff provided coaching support to Level 2 schools. This finding was corroborated by primary data collection in the course of site visits and correspondence with the lead MTPDS EGRA data analyst who wrote,

> *…the original design of the RCT intended to differentiate between two levels of intensity of coaching visits; however, in practice, coaches—both MTPDS coaches and government PEAs—visited schools both within and outside of their group.*

The result of this spillover of coaching into Level 2 schools limits the utility of comparison between the two high-intensity treatments (Levels 2 and 3). In fact, MTPDS merged EGRA

[12] The following limitations are relevant to the fact that much of the quality of the IE is dependent on the quality of the MTPDS design and availability. While MTPDS was never originally designed to be a "research project" that follows strict statistical and project design techniques (as with an RCT), the Evaluation Team did follow standard guidelines in implementing an impact evaluation (i.e. USAID Evaluation Policy).

[13] According both to the Mission and MTPDS, Level 1 beneficiaries only received three of the four trainings before endline data collection was undertaken. Additionally, the first two modules focused primarily on establishing a national CPD structure that could be transferred to MoEST for continued usage after program completion. As such, Level 1 teachers only received one true literacy module, or two days of training over the three year program. While even this modest amount of support violates a pure estimate of the counterfactual, the corresponding lack of impact supports usage of this group as a control.

observations from the two high-intensity groups, obviating the ability to quantitatively distinguish between the two levels. As a result, the only comparison of EGRA data is for Levels 1 and 3.

Literacy

MTPDS used a customized version of the Early Grade Reading Assessment (EGRA) to measure changes in literacy competencies. For the express purpose of the integrated IE, MTPDS collected two rounds of data (2010 baseline and 2012 endline) consisting of a total of 3,000 randomly selected learners from 150 schools.[14] Enumerator teams randomly selected 20 students at each school ensuring that the following criteria were met: ten boys, ten girls, ten Standard 2, ten Standard 4. The Evaluation Team received from MTPDS cleaned EGRA datasets with a built-in, multi-stage weighting framework.

Data Quality Assessment

Instrumentation

MTPDS spent considerable time and resources contextualizing the internationally field-tested EGRA instrument to the Malawian context.[15] EGRA is a 15-minute oral test intended to measure foundation literacy skills, including phonemic awareness, reading and listening comprehension, and letter/syllable identification. The program conducted a comprehensive adaptation workshop that included participation from an international EGRA expert, MTPDS staff, the Malawi Institute for Education (MIE), the Centre for Language Study, the Malawi National Exam Board, and MoEST. Following this workshop, the instrument was piloted in two rounds of increasing size. The format and content of the EGRA tool used at each data collection phase remained consistent and user-friendly: instructions for enumerators were clear and consistent with experiences in other countries. Literacy skills measured in the EGRA were directly aligned with MTPDS training strategies.

Enumeration

While the instrument was developed in a rigorous manner, MTPDS introduced potential instrumentation bias in the course of data collection by modifying the composition of data collection teams between the two rounds. Baseline data was collected by teams of respondents to job ads posted by the program, while endline data was collected by teams of PEAs. During the latter round, MTPDS intended for PEAs to enumerate exclusively in zones for which they were not personally responsible. However, the Evaluation Team found evidence of PEAs collecting data from their own zones during KIIs with Lilongwe-based program staff, a fact that was later confirmed in discussions with PEAs in the field. Having MoEST personnel simultaneously responsible for collecting data from and overseeing schools is a concern from the perspective of data validity. This bias is all the more acute when Ministry personnel enumerate in their own zones, where they may have an incentive to misrepresent information.

[14] MTPDS also implemented a parallel EGRA data collection effort designed to provide nationally representative measures of literacy performance. This evaluation utilized the literacy intervention dataset in analysis, as (in keeping with its stated purpose) it utilized a much more robust sampling strategy,

[15] http://www.rti.org/page.cfm?nav=528&objectid=E60C72B1-6190-49EF-918317C0BB7E464D

Data entry

In addition to the way enumeration was structured, the Evaluation Team found a number of data integrity issues from a limited comparison of paper surveys and the corresponding electronic databases. The team randomly selected ten paper surveys from two binders storing EGRA data at the MTPDS office (five from the 2010 round, five from the 2011 round). One of the five 2010 surveys recorded a significant error in data entry (incorrectly scoring a Letter Naming category 2 instead of 4), one of the five 2011 surveys consistently over-reported reading sections (marking as "correct" items that were either "incorrect" or left blank), and there were four instances of over-reporting in the remaining 2011 surveys. As validated by MTPDS staff, the total error rate for the 1,041 data elements was 4.8% (3.3% for those items MTPDS consider having a "major" influence on impact evaluation results).[16] Given the small sample size, it is not possible to gauge the extent to which these data entry issues were representative of the EGRA datasets in the aggregate. However, after the findings were corroborated by an independent review by MTPDS staff, it is reasonable to assert that low-level data entry errors were pervasive. In order to strengthen the enumeration and data entry process, MTPDS introduced the use of electronic data collection via tablet computers and "Tangerine" assessment software during the second round of data collection. It is safe to say that the utilization of electronic data collection and entry, particularly using this field-tested approach, greatly increased the accuracy of data.

Conclusions

Whereas the EGRA instrument was designed and piloted in accordance with international best practices, data collection protocols may have introduced measurement bias. Specifically, the decision to utilize teams of PEA enumerators, while beneficial from the vantage point of capacity building, created the possibility of misreporting due to conflicting incentives on the part of MoEST personnel. [17] Regardless, the benefits of PEA enumeration significantly outweigh possible biases and continued engagement of MoEST personnel (e.g. PEAs, teachers, etc.) is essential for long-term sustainability of EGRA data collection. The only reason to dispossess MoEST of EGRA data collection responsibilities would be to adhere to the higher requirements for data integrity in future impact evaluations. In these instances the USAID Evaluation Policy (2011) requires that evaluation be conducted by objective and external parties.

DQAs performed by both the Evaluation Team and MTPDS identified a 5% data entry error rate, a figure 10 times higher than the commonly accepted limit for IEs.[18] An error rate of this magnitude means that, on average, one out of every 20 items was entered incorrectly. While this level of error presents significant data quality concerns, misreporting in the sample was found to be bidirectional (both over and underreporting performance). As such, there is some

[16] In response to the Evaluation Team's DQA, MTPDS conducted an independent review of data with 18 randomly selected surveys. Error rates from this sample were very similar to those identified above: 5.1% total error rate, 3.7% "major" error rate.

[17] Due to the fact that all EGRA data collection occurred before the evaluation team was contracted, it was not possible to gauge the extent to which PEA enumeration resulted in reporting bias.

[18] Organizations such as Innovations for Poverty Action (IPA), the Abdul Latif Jameel Poverty Action Lab (J-PAL) and Social Impact (SI) use 0.5% as the threshold for acceptable data entry error rates.

evidence that data entry errors were random and not systematically biased. In subsequent rounds of data collection, it is highly likely that MTPDS significantly reduced these error rates due to utilization of electronic data collection methods. If EGRA data are to be utilized for substantive policy purposes, the aforementioned data quality concerns may warrant a more robust DQA.

Analysis of Impact

Whereas the EGRA collects data on a large number of variables, this analysis prioritizes three measures that are indicative of literacy performance: number of letters read in one minute, number of words read in one minute, and number of reading comprehension questions answered correctly (maximum possible score = 5/5) Additionally, to reflect the challenges of substantively improving literacy scores in a context where the majority of learners cannot read, analysis was performed on two complementary measures: zero scores and mean scores. Zero scores, or the proportion of respondents that did not register a single correct answer on a given survey question, are important to gauge literacy gains amongst the lowest-performing students. Mean scores, on the other hand, indicate total reading proficiency for an average learner. By focusing on changes at the threshold of literacy and on competencies of average students, the analysis can provide a more complete picture of programmatic effects. In reading the following analysis, it is important to note two facts. First, literacy gains affect the two measures in opposite directions: increases in the ability of students to read causes zero scores to drop and mean scores to rise. Second, as tasks become more difficult (e.g., reading words instead of letters), it is normal for scores to drop. This is to be expected, as the skills are cumulative.

BALANCE CHECKING

Before exploring changes in outcomes, Level 1 and Level 3 schools were compared along key baseline characteristics to check for balance (Table 1). [19] Both groups exhibited similarity across basic literacy measures and a range of contextual variables ($p > .05$). The only exceptions were the proportion of children who speak the same language at home and school, and the availability of printed materials at home.[20] Since access to print materials and familiarity with the language of instruction are important covariates of literacy outcomes, this imbalance raises the possibility that other factors besides the MTPDS interventions may account for differences in observed changes.

Table 1: Balance Checking of Key Variables at Baseline[21]

Level of	Variable	Level 1	Level 3	p

[19] The validity of any IE, and any subsequent causal assertions, rests on the comparability of evaluation groups: in this instance, Level 1 and 3. Specifically, for the purpose of difference-in-differences analysis, baseline comparability matters to the extent these characteristics affect differential changes over time.

[20] Children in Level 1 schools were, on average, 10% more likely than Level 3 children to speak the same language at home as in school. Similarly, Level 1 children were 10% more likely to have access to print materials outside of school than Level 3 children.

[21] Continuous variables were analyzed using two-tailed *t*-tests. Nominal variables were tested using chi^2.

Measurement				
Continuous	Letters per Minute	1.63	1.47	0.382
	Syllables per Minute	0.56	0.51	0.799
	Words per Minute	0.38	0.21	0.302
	Reading Comprehension	0.01	0.01	0.321
	Age	9.09	9.02	0.630
	Grade Size	253.34	183.80	0.267
Nominal	Gender (Female)	49.01	48.95	0.950
	Teacher (Female)	75.12	57.35	0.335
	Teacher (Trained)	98.63	98.25	0.891
	Shift (Full Day)	86.07	90.19	0.809
	Language (Same)	97.07	87.30	0.020*
	Textbooks (Yes)	24.17	17.17	0.174
	Printed Material (Yes)	32.00	22.53	0.022*
	TV (Yes)	9.70	15.96	0.145
	Bicycle (Yes)	71.26	68.90	0.695

ZERO SCORES

At the time the program began (baseline), both groups recorded very low levels of literacy: across all 1,452 assessments, three quarters of respondents could not read a single letter, 95% could not read a single word, and only seven children could answer a single reading comprehension question (.005%). While the two groups had similarly high proportions of zero scores in the words per minute and reading comprehension measures, the Level 3 group had a significantly higher (11%) proportion of respondents who could not read a single letter.

From this low starting point, the two groups charted very different trajectories over the course of the two-year study period. As illustrated in Table 2, the high-intensity group improved substantially more than the low-intensity group across all three outcome measures. This result was especially pronounced with regard to the most fundamental aspect of reading: the proportion of children not able to read a single letter in a minute dropped by more than half (58%) in the Level 3 group. During the same time period, the proportion of students who could not read one letter increased by 5% in the low-intensity group. This puzzling finding is most

likely an artifact of data collection and not an indication that literacy skills worsened in any substantive manner.

The story was similar for the words per minute and reading comprehension outcomes, on which the high-intensity group achieved zero-scores reductions of 45 and 28 percent, respectively. Level 1 schools marked negligible improvements for both of these measures, resulting in very low endline scores: among this group, almost no children could read a single word or answer a single reading comprehension question.

Table 2: Percent Zero Scores, by Treatment and Time

Variable	Treatment	Baseline	Endline	Change	Difference in Differences
Letters per Minute	Level 1	67.73	72.31	4.58	-62.59
	Level 3	78.71	20.70	-58.01	
Words per Minute	Level 1	95.80	94.64	-1.16	-43.69
	Level 3	93.24	48.39	-44.85	
Reading Comprehension	Level 1	99.48	98.34	-1.14	-26.67
	Level 3	99.16	71.35	-27.81	

MEAN SCORES

Findings were quite similar when focus was shifting to an analysis of the average student: baseline scores were very low for both groups and the Level 3 group significantly outperformed the Level 1 group across each of the three measures. Similar to the zero-score analysis, Level 3 improvements were the most dramatic in the ability to read letters (22% increase) and words (9% increase). Level 1 performance was substantially worse. While Table 3 shows slight improvements across all three outcomes, they were all within the margin of error.[22] The farthest-right column ("difference in differences") demonstrates that children in the high-intensity group outperformed their peers by being able to read, on average, an additional 20 letters and nine words per minute.

Variable	Treatment	Baseline	Endline	Change	Difference in Differences
Correct Letters	Level 1	1.63	1.54	-0.09	20.14

[22] The full tables, with floor and ceiling scores (upper and lower bounds), are presented in Annexes 3 and 4.

per Minute	Level 3	1.47	21.52	20.05	
Correct Words per Minute	Level 1	0.38	0.24	-0.14	8.88
	Level 3	0.21	8.95	8.74	
Reading Comprehension	Level 1	0.01	0.02	0.01	0.55
	Level 3	0.01	0.57	0.56	

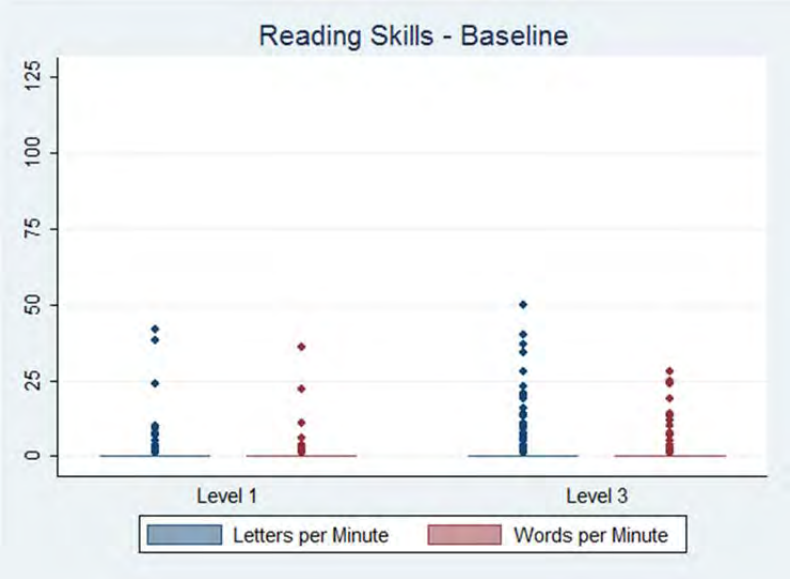
Table 3: Difference in Differences of Key Literacy Mean Scores, by Treatment

The differences in literacy outcomes between the two groups are illustrated in two ways below. Figures 1 and 2 demonstrate the distribution of mean scores for the two groups at baseline and endline, respectively. Given low scores and modest gains on the reading comprehension measure, only letters/minute and words/minute outcomes are presented. Figure 1 demonstrates that the two groups started with similarly low scores, while Figure 2 illustrates notable progress by students in Level 3 and a lack of progress by those students in Level 1.[23]

Figures 3 through 5 demonstrate changes for the two groups between baseline (0) and endline (1). Each figure illustrates both the percent change for zero scores (decrease is positive) and the change in mean value of correct responses (increase is positive). These graphs further reiterate the finding that Level 3 students consistently outperformed Level 1 students in both zero-score and mean-score measures. We see, once again, how the rate of improvement decreased as the tasks became more difficult.

Figures 1 and 2: Reading Skills at Baseline and Endline

[23] The solid boxes represent the middle half of scores (25-75%), the black line within the box denotes the median score (value for which half of respondents scored higher and half scored lower), the line extending from the box represents the upper quartile (highest score), while the dots are outliers (1.5 times the upper quartile).


Reading Skills - Baseline
125
100
75
50
25
0
Level 1
Level 3
Letters per Minute
Words per Minute

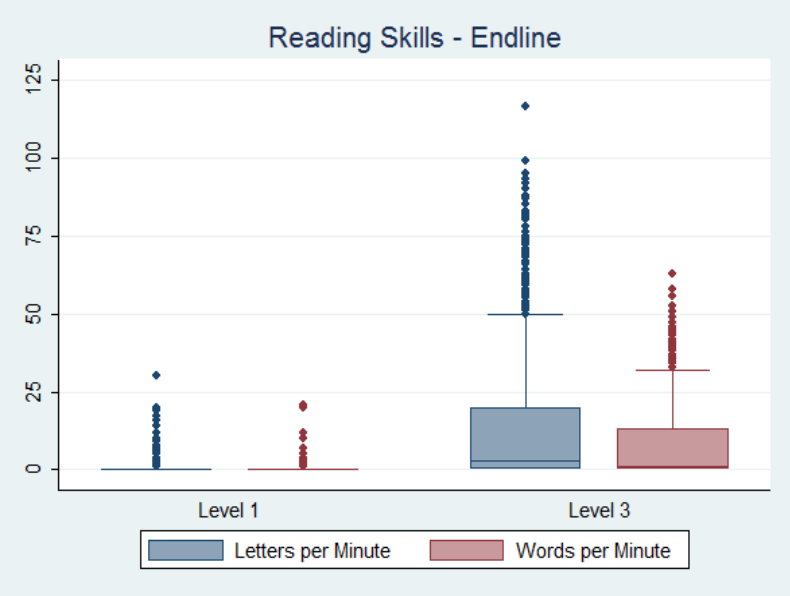
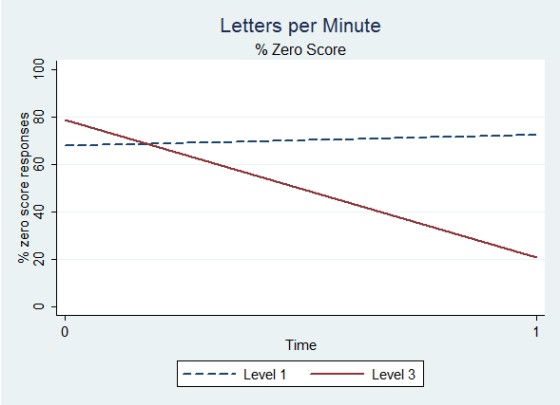

Reading Skills - Endline
125
100
75
50
25
0
Level 1
Level 3
Letters per Minute
Words per Minute

Figure 3: Letters per Minute (Change in Zero and Mean Scores)



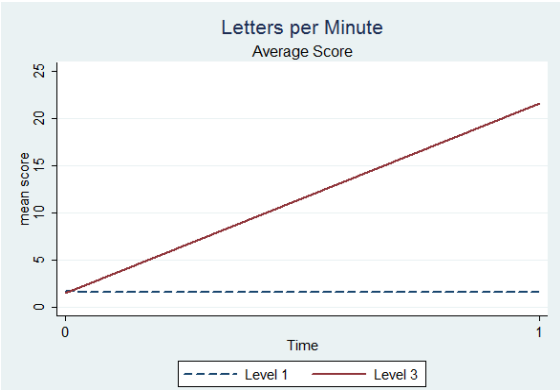
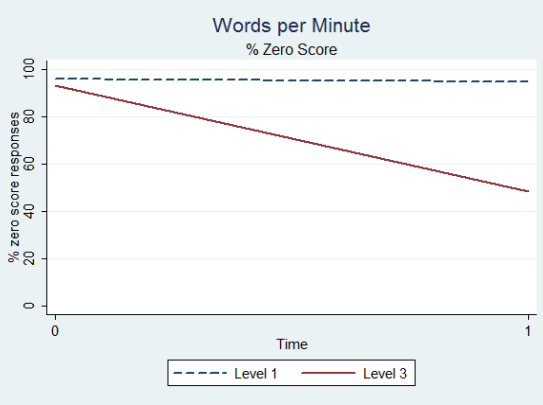


Figure 4: Words per Minute (Change in Zero and Mean Scores)



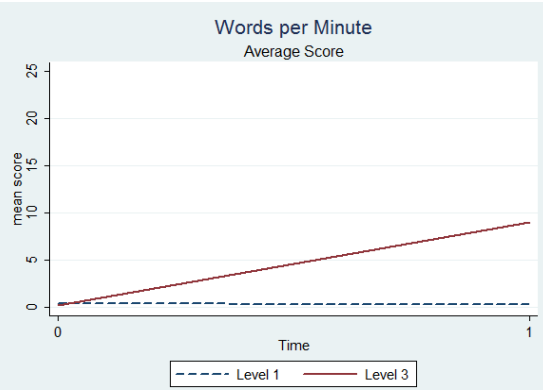
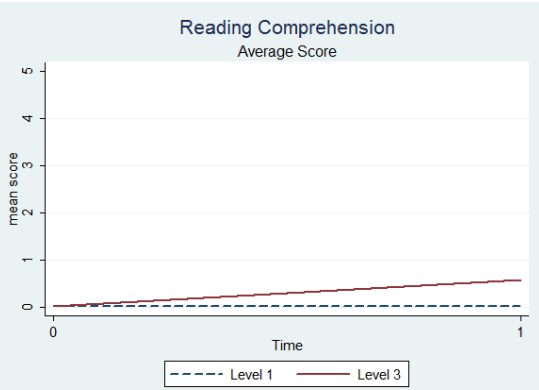


Figure 5: Reading Comprehension (Change in Zero and Mean Scores)



GENDER DISAGGREGATION

Whereas the preceding analysis established general trends between the two groups, this section presents disaggregated findings differentiating the experiences of male and female children (see Table 4 & 5). Across all three literacy outcomes, girls were more likely to have zero scores at baseline than boys. This was particularly acute with regard to the letters per minute measure, on which boys outperformed girls by 10% in Level 1 and 4% in Level 3.

Changes in literacy outcomes in the high-intensity group were similar for both genders, with girl students slightly outperforming boy students in the letter/minute zero-score measure. The picture for Level 1 children was very different. Whereas the proportion of girls unable to read any letters did not change substantively, boys recorded a 9% increase in the proportion of zero scores, suggesting a sizable drop in the literacy outcome. While it is tempting to attribute this decline in basic literacy skills to participation in the low-intensity program (i.e. a negative impact), it is important to keep in mind the lack of a true control group. There are compelling reasons supporting the notion that national literacy scores would have declined over this time period,[24] and it could very well be the case that participation in the Level 1 treatment reduced the rate of decline. Unfortunately, without a clean comparison with this counterfactual scenario it is not possible to definitively attribute causality. Regardless of the reason, basic literacy among Level 1 boys declined over the course of the study.

Looking at the far-right column in Table 3 (difference in differences), it is clear that Level 3 children consistently and substantially outperformed their peers in Level 1 schools. The largest recorded difference between the two groups was for male students in the letters/minute measure, on which the high-intensity group improved by an average of 65% more than the low-intensity group. The endline results for the two genders were, on balance, much more similar than at baseline. This was especially apparent for the letters/minute measure, although this result was partly driven by the adverse changes experienced by Level 1 boys.

[24] School enrollment has been increasing at an annualized rate of 3% without a corresponding increase in the availability of school block or trained teachers. Additionally, it has been argued that the detrimental effects of the whole-word PCAR materials were cumulative. In other words, the longer the material was implemented, the worse students would do.

Table 4: Sex-Disaggregated Percent Zero Scores, by Treatment and Time

Variable	Gender	Treatment	Baseline	Endline	Change	Difference in Differences
Letters per Minute	Female	Level 1	72.68	72.56	-0.12	-60.37
		Level 3	80.75	20.26	-60.49	
	Male	Level 1	62.96	72.06	9.10	-64.70
		Level 3	76.75	21.15	-55.60	
Words per Minute	Female	Level 1	94.78	96.93	2.15	-50.75
		Level 3	95.83	47.23	-48.60	
	Male	Level 1	91.77	92.25	0.48	-46.65
		Level 3	95.77	49.60	-46.17	
Reading Comprehension	Female	Level 1	100.00	99.47	-0.53	-26.39
		Level 3	99.86	72.94	-26.92	
	Male	Level 1	98.35	97.17	-1.18	-28.25
		Level 3	99.11	69.68	-29.43	

With regard to baseline mean scores, boy students outperformed girl students across all three outcome areas in both Level 1 and Level 3 groups. However, unlike the zero-score analysis, the greatest baseline mean-score discrepancy between the genders was the word/minute measure: male students were, on average, able to read an additional 3.5 words per minute, as compared to female students.

Consistent with the zero-score analysis, Level 3 students improved substantially more than did their peers in Level 1 schools. Furthermore, improvements in the high-intensity treatment group were similar for both genders, triangulating findings from the zero-score analysis above. Interestingly, the puzzling finding of negative impact for Level 1 male students disappeared in the mean-score analysis (the slightly negative result below is within the margin of measurement error). This finding indicates that, whereas the endline round of data collection picked up many more Level 1 male students who could not read any letters as compared to baseline, on average male students had similar scores across both years.

Table 5: Gender-Disaggregated Mean Scores, by Treatment and Time

Variable	Gender	Treatment	Baseline	Endline	Change	Difference in Differences
Letters per Minute	Female	Level 1	1.10	1.49	0.38	19.94
		Level 3	1.39	21.71	20.32	
	Male	Level 1	2.14	1.59	-0.55	20.32
		Level 3	1.55	21.32	19.77	
Words per Minute	Female	Level 1	0.10	0.13	0.03	8.51
		Level 3	0.20	8.74	8.54	
	Male	Level 1	0.65	0.35	-0.30	9.24
		Level 3	0.22	9.17	8.95	
Reading Comprehension	Female	Level 1	0.00	0.01	0.01	0.53
		Level 3	0.00	0.54	0.54	
	Male	Level 1	0.02	0.04	0.01	0.57
		Level 3	0.01	0.59	0.58	

REGRESSION ANALYSIS

The analysis of zero and mean scores suggests that participation in MTPDS was associated with improved literacy among Level 3 primary school learners. However, as previously discussed, evidence of change does not automatically connote causation. Many factors may have contributed to the differential changes outlined above, which could have occurred even in the absence of the treatment (influencing outcomes in either positive or negative directions). Because the two groups varied with respect to several important variables at baseline (language of instruction, availability of print materials, and letters/minute zero scores), data were analyzed using a multivariate difference-in-differences regression model.[25] Regression analysis is a statistical technique that enables exploration of relationships between variables. In this case, the focus is on the relationship between participation in the MTPDS program and scores on the three outcome variables. The primary benefit of regression analysis, as compared to the direct comparison approach detailed in preceding sections, is that it allows for holding other important variables (i.e. those that would

[25] $Y_i = \beta_0 + \beta_1 T + \beta_2 P_i + \beta_3 T*P + \beta_4 C1_i + \beta_5 C2_i + \beta_6 C3_i + \beta_7 C4_i + \beta_8 C5_i + \beta_9 C6_i + \mu_i$
Where *Yi* represents the literacy outcome, β_0 is the constant (*y*-intercept), β_1 is a coefficient capturing the initial difference between the two treatment levels, *T* is a dummy variable for time (baseline/endline), β_2 captures the impact of the program on teacher efficacy, as measured by the intent to treat, *P* is a dummy variable for treatment intensity (level 1/3), β_3 captures the effect of the interaction of treatment and time (*T***P*), β_4–β_9 capture the effect of covariates C1–C6 (respondent age, grade size, native language, availability of non-textbook print items, gender of teacher, bicycle ownership), and μ_i is the error term.

plausibly influence the outcome measure) constant. Including these covariates in the regression model is important for two reasons: first, it allows for the estimation of each variable's influence on the outcome measure and, second, it controls for this influence, yielding a more precise measure of programmatic impact.[26] Key results of the regression analysis are presented below, and the summary of each regression is detailed in Table 6.

Letters per Minute: Participation in Level 3 was correlated with the ability to read five additional letters per minute more than a comparable child in Level 1 (p = 0.00, adjusted R^2 = .389).

Words per Minute: Participation in Level 3 was correlated with the ability to read 2.3 words per minute more than a comparable child in Level 1 (p = 0.00, adjusted R^2 = .277).

Reading Comprehension: Participation in Level 3 was correlated with a .14 increase in reading comprehension scores as compared to Level 1 students (p = 0.00, adjusted R^2 = .206).

Table 6: Summary of Regression Analysis, by Outcome

	Letters per Minute		Words per Minute		Reading Comprehension	
	Coefficient	Std. Error	Coefficient	Std. Error	Coefficient	Std. Error
Outcome	5.07***	0.25	2.28***	0.24	0.14***	0.01
Age	-0.07	0.06	-0.04	0.03	0.00	0.00
Grade Size	0.00	0.00	0.00	0.00	0.00	0.00
Language	2.30	1.56	1.10	0.69	0.05*	0.02
Books	2.78	1.49	2.06*	1.02	0.13*	0.06
Bicycle	0.49	0.62	0.38	0.40	0.04*	0.02
Teacher Gender	0.51	0.73	1.07	0.65	0.06	0.05

[26] Regression analysis was performed using multi-stage sample weights built into the dataset by MTPDS. Robustness checks were conducted by re-specifying the weighting framework (e.g. the manner in which missing values were treated), with tests yielding consistent results.

Constant	-2.53	2.75	-2.59	1.93	-0.16	0.09
R^2	0.39		0.28		0.21	
Observations	2691		2678		2688	

* $p < .1$
** $p < .05$
*** $p < .01$

CONCLUSIONS

On the basis of the MTPDS EGRA data, children attending Level 3 schools appear to have been significantly more successful than their Level 1 peers, across all three key literacy outcomes. This finding held true for analysis of zero scores, mean scores, gender-disaggregated data, and multivariate regression. Over the course of the two-year study, data suggest that respondents in the high-intensity group made dramatic improvements: the number of children who could not read a single letter dropped by more than 50%, while the average child's ability to read increased by 20 letters per minute. While these numbers are impressive, pre-existing differences between the groups at baseline may have mediated these gains. Regression analysis was conducted to control for these differences, with results painting a very similar picture. Participation in the high-intensity treatment was associated with the ability to read an additional five letters and 2.3 additional words per minute, as compared to students in Level 1. In addition to improving the most basic of literacy skills, the high-intensity group improved in their ability to answer reading comprehension questions. This was particularly evident with regard to zero scores, where reductions of over 25% were registered for both girls and boys.

The story of children in low-intensity treatment schools was quite different. Changes in literacy scores were inconsistent. For the most part, participation in Level 1 did not lead to any substantive changes in literacy measures. The one exception was an adverse trend in boys' letters/minute scores. These apparent reductions in literacy could have been caused by a number of factors. However, without a control group it is not possible to determine whether these losses are representative of what happened in Malawian schools or are the result of idiosyncratic data collection and entry processes of MTPDS. By and large, participation in Level 1 did not lead to any substantive changes in literacy measures. This result is not surprising, given the limited intensity of the Level 1 treatment. Particularly in the face of such substantive obstacles (large and growing class sizes, dearth of teaching/learning materials, etc.), it is not surprising that such a limited intervention did not significantly improve learning outcomes—of the four modules, the first focused on institutionalizing the national CPD, while the fourth was not delivered in time to influence results.

When disaggregated by gender, the macro analysis held true. Both male and female students in the high-intensity group recorded substantially higher increases in literacy outcomes for all three measures. Whereas male students had consistently higher literacy measures at baseline, over the course of the MTPDS program the two genders became more equitable with regard to basic literacy skills.

Across all methodological approaches, Level 3 students consistently outperformed their peers in the low-intensity treatment schools. These findings, however, have to be viewed

within the broader context of evaluation validity. As documented above in the DQA section, there were a number of methodological limitations and potential sources of bias. First, without a true control group it is impossible to make definitive claims of program impact.[27] However, given the limited nature of the Level 1 intervention and the lack of changes in its literacy measures, it is plausible to substitute the low-intervention group for a true control group. While not in keeping with research standards, this substitution is warranted if one believes that four day-long trainings do not constitute sufficient grounds for significantly improving literacy scores amongst students.

Secondly, the way in which Level 1 schools were selected (i.e., from a different district and not subject to the same randomized assignment process) introduced the potential for selection bias. This is borne out in data of the two groups at baseline, where significant differences existed across a number of important variables. Third, while the EGRA instrument was representative of global best practices in education research, having Ministry personnel (even if only in isolated instances) enumerate within their own zones of responsibility may have introduced bias into the data collection. Fourth, as demonstrated by two independent data quality assessments, there were not-insignificant issues with data entry errors.

On balance, the Evaluation Team concludes that the evaluation results are of a sufficient quality to warrant consideration in policy discussions. While the exact extent of impact is likely not captured by this analysis, the general trend (i.e., the Level 3 group outperforming the Level 1 group) is empirically sound, being triangulated through other data collection and analytical techniques (discussed in the Independent Analysis section of this Impact Evaluation chapter as well as in the Performance Evaluation chapter).

Teacher Performance

The primary mechanism through which MTPDS aimed to improve primary school literacy was improved teacher performance. According to the MTPDS theory of change, if teacher performance increased, so would the reading competence of their students. MTPDS developed a classroom observation instrument to measure changes in key teacher competencies. The Evaluation Team received three classroom observation datasets from MTPDS, one each for the three rounds (2010, 2011, 2012).

DATA QUALITY ASSESSMENT

MTPDS developed, piloted, and implemented for the 2010 baseline a classroom observation tool comprised of 76 items. However, program staff decided to revise the tool after the initial round of data collection to better capture teacher performance. The result was a more concise, user-friendly tool containing 23[28] questions clearly delineated into thematic sections. This updated tool was used in both the 2011 and 2012 rounds of data collection. While the revision certainly increased the utility of the instrument, the instrument was changed so fundamentally (both with regard to question selection and scoring) that data derived from the two tools does not allow for comparison of changes for a single measure.

[27] Since comparisons can only be made between treatment modalities, not against the counterfactual, it is not possible to control for general time trends.

[28] The instrument originally had four additional math-related questions. These were dropped when MTPDS ceased to focus on numeracy. Due to the change in program focus, MTPDS dropped the math component of the classroom observation tool for the 2012 round.

For purposes of estimating program impact, the inability to compare data between baseline and endline severely restricts measurement, let alone ascription, of change associated with the intervention. Accordingly, this analysis focuses exclusively on changes between the 2011 and 2012 rounds.
A related impediment to calculating impact estimates of teacher performance stems from inconsistent use of the revised instruments. Whereas observational data from the 2011 round are presented alongside important context variables (for example gender, years of experience, and extent of teacher training), none of these variables was captured for the 2012 round. This data gap prevents the inclusion of control variables in regression analysis, severely limiting the explanatory power of statistical models.

The Evaluation Team tested the efficacy of the revised MTPDS classroom observation instrument through independent assessment. Before using the tool in the field, the team requested training in the use of the instrument from MTPDS. During this session, and in the course of 63 classroom observations in 24 schools, the team identified a number of issues with the instrument:

- Question 5: Potential ambiguity between response options 3 *("learners practice literacy independently or in groups—mostly with teacher assistance")* and 4 *("learners actively involved in guided and independent learner tasks")*.
- Question 18: Ambiguity in how to differentiate between *"limited"* (3) and *"large amount"* (4) of print rich material on display.
- Question 19: Ambiguity in how to calculate a composite observational score. The enumerator is asked to count how many sub-items are observed. However, some of these sub-items are themselves composed of multiple components. There is a lack of clarity regarding whether one or all conditions must be met in order for an item to be marked. Examples include: *"Teacher marks exercises, solicits and provides feedback in class, and makes use of learner feedback during class"* and *"learners encouraged to show parents their homework and teachers monitor homework, checking for parent involvement."*
- Question 23: Ambiguity in how many methods constitute "a variety" in response to option 3: *"provides a variety of methods for students to establish good writing mechanics such as…."*. In addition, there is a perceived mistake with scoring *"writing letters in the air"* (3) at a higher level than *"students copy words from the board and teacher corrects"* (2).

While lack of clarity on the proper way to use the classroom observation tool may be a result of the Evaluation Team receiving only a one-hour training session, the ambiguity in the instrument raises the potential for data entry errors.

FINDINGS

In the analysis of quantitative data, the ability to make causal inference about observed changes is dependent on the comparability of evaluation groups. In addition to the absence of a pure control group, the lack of useful baseline data severely limits the ability to compare

the three MTPDS treatment intensities.[29] Specifically, as of midline (2011), the three groups appear to be quite different from one another on various teacher performance metrics. Table 7 presents this variability, including both floor (lowest possible) and ceiling (highest possible) scores across a number of key variables. This imbalance is likely explained by the fact that, by the time of the 2011 round of data collection, schools had been receiving varying levels of MTPDS support for a year, and program effects had begun to emerge. If impact estimates were derived from changes between 2011 and 2012, they would present a false picture of program performance. Given the aforementioned limitations in comparing treatments, the following analysis focuses primarily on changes within treatment groups, not between them.

Table 7: Midline Scores, by Treatment

Treatment	Floor Scores			Ceiling Scores		
	Assessment	Phonics	Writing	Assessment	Phonics	Writing
Level 1	22.02	26.61	36.70	3.67	1.83	0.92
Level 2	25.00	28.12	37.50	3.12	6.25	9.38
Level 3	39.13	39.13	52.17	8.70	13.04	4.35

To focus analysis on the most important aspects captured by the observation tool, the Evaluation Team focused on the six items identified by MTPDS to be most representative of teacher performance in literacy. The observation items, their abbreviations for this report, and their scale are presented in Table 8.

Table 8: MTPDS Key Literacy Components

Classroom Observation Item	Abbreviation	Scale
"*Practice with guided/independent reading*"	Reading	1–4
"*Provision of a stimulating learning environment*"	Environment	1–4
"*Use of assessment*"	Assessment	1–4
"*Use of phonics*"	Phonics	1–5
"*Teaching of comprehension strategies*"	Comprehension	1–5
"*Teaching mechanics of writing*"	Writing	1–5

As classroom observation data were collected at the ordinal level (rank-ordering scales) it is only possible to count, not measure, the data. For instance, this means that a calculation of

[29] Unlike the EGRA dataset, for which MTPDS collapsed Level 2 and Level 3 treatments into a single high-intensity group, the classroom observation dataset presented observations for three discrete treatment groups.

the average score is not justified.[30] Accordingly, the six variables have been recoded into two variants for the purpose of this analysis: floor (minimum values), and ceiling scores (maximum values).[31] Highlights from this analysis are reported for each measure below. Table 9 displays percent floor and ceiling scores for all six priority outcomes across the three treatment groups.

Floor Scores: Across all treatments and variables, the proportion of observations registering minimal scores dropped. The trend was most evident for the phonics, comprehension, and writing variables, which did not register a single minimum score at endline. In fact, with the exception of the environment and assessment variables, all three MTPDS groups almost eliminated floor scores over by 2012.

Ceiling Scores: From relatively low proportions of classrooms registering perfect marks in 2011, MTPDS beneficiaries, and particularly teachers in the Level 3 group, showed impressive gains. By 2012, at least two-thirds of teachers in Level 3 were logging perfect scores in Reading, Phonics, and Writing.

Table 9: Percent Change in Classroom Observation Floor Scores, by Treatment

Variable	Treatment	Floor Scores			Ceiling Scores		
		Midline	Endline	Reduction	Midline	Endline	Increase
Reading	Level 1	13.76	3.55	10.21	8.26	11.24	2.98
	Level 2	12.50	6.00	6.50	9.38	16.00	6.62
	Level 3	13.04	0.00	13.04	13.04	62.50	49.46
Environment	Level 1	53.21	44.38	8.83	2.75	2.96	0.21
	Level 2	59.38	28.00	31.38	3.12	12.00	8.88
	Level 3	26.09	18.75	7.34	13.04	12.50	-0.54
Assessment	Level 1	22.02	7.10	14.92	3.67	13.02	9.35
	Level 2	25.00	2.00	23.00	3.12	16.00	12.88
	Level 3	39.13	18.75	20.38	8.70	25.00	16.30
Phonics	Level 1	26.61	0.00	26.61	1.83	1.78	-0.05
	Level 2	28.12	0.00	28.12	6.25	20.00	13.75
	Level 3	39.13	0.00	39.13	13.04	56.25	43.21
Comprehension	Level 1	58.72	0.00	58.72	6.42	2.96	-3.46

[30] Assumption of equal distances between scores is not justified with ordinal data (e.g. one cannot assume that the difference between a score of 1 and 2 is the same as the difference between a score of 2 and 3).
[31] Observations of math classes and standards outside the program target range of 1–4 were excluded from analysis.

	Level 2	65.62	0.00	65.62	0.00	6.00	6.00
	Level 3	60.87	0.00	60.87	8.70	25.00	16.30
Writing	Level 1	36.70	0.00	36.70	0.92	1.78	0.86
	Level 2	37.50	0.00	37.50	9.38	14.00	4.62
	Level 3	52.17	0.00	52.17	4.35	37.50	33.15

Midline to endline trends for the two most important variables (reading and writing) are illustrated in Annex 6. While the three groups recorded similar breakdowns of scores at midline for both measures, the results at endline were varied. While Levels 1 and 2 registered modest improvements, Level 3 significantly outperformed the other groups. This was particularly evident for the reading measure, where Level 3 finished with more than half of observations registering the maximum score. Similar illustrations for the other four variables, as well as graphs demonstrating changes in floor and ceiling scores can be found in Annexes 3 and 4. Regression analysis was not performed on classroom observation data, as the lack of information about the teachers and their context (control variables) undermines the utility of the statistical technique.

CONCLUSIONS

This analysis of MTPDS classroom observation data suggests that teachers across the three treatment levels improved in their pedagogical abilities between the midline (2011) and endline (2012) data collection events. The most pronounced improvements, both in terms of reductions in floor scores and increases in ceiling scores, was registered by teachers in the Level 3 group. While teachers improved across all six measures, the most striking finding was a complete absence of floor scores for Phonics, Comprehension, and Writing. On the basis of this data, it appears that teachers had the most difficulty operationalizing best practices in the continuous assessment and text-rich environments measures. As discussed in detail in the following sections, barriers in implementing these activities are primarily independent of teacher knowledge and motivation, and are largely the result of an overabundance of learners and a dearth of teaching/learning materials, respectively.

The data paints a positive picture, with teachers making striking improvements across a variety of measurement areas in the span of a single year. As discussed in the DQA section, however, any analysis of changes in teacher performance using MTPDS classroom observation data should be tempered with regard to aforementioned issues with instrumentation and data quality. The largest detriment for analysis of impact is the lack of comparable baseline data. While trends were positive between the latter two data points, without information about what occurred between 2010 and 2011, it is impossible to pass judgment on the efficacy of the MTPDS intervention in improving teacher effectiveness. Additionally, given concerns with instrumentation raised in the DQA section, the MTPDS classroom observation data have to be approached cautiously. For example, whereas seven percent of all observations at endline recorded a perfect score on the Environment variable, this finding is at odds with experiences of the Evaluation Team, who did not observe a single classroom that scored a 5 on the same assessment form. Last, without a true control group,

these positive changes cannot be definitively attributed to the MTPDS program, as they could result from any number of external factors.

While it is not possible to calculate sound impact estimates on the basis of this data the Evaluation Team concludes that the general trends, Level 3 outperforming the lower-intensity treatments, are valid. This conclusion is drawn primarily from the EGRA data and qualitative data collected by the team in the course of school visits. Given serious issues with the MTPDS classroom observation form and associated processes, it is not recommended that the data be used for policy purposes.

Cost-Effectiveness Analysis

FINDINGS

Given serious constraints in estimating MTPDS impact on teacher performance outlined above, cost-effectiveness analysis was conducted solely on the literacy outcome. Additionally, given the composition of the MTPDS EGRA datasets, comparisons could only be conducted between Levels 1 and 3.

The process began when the Evaluation Team requested detailed budgetary information from MTPDS. Program staff supplied aggregated, fully-burdened labor costs and disaggregated, direct costs for Levels 1 and 3. While in the field, team members met with the MTPDS COP and the CPD finance lead to discuss the budgets as well as assumptions underpinning direct cost calculations: minor revisions were made to the MTPDS document as a result of this meeting. Subsequently, *per teacher* and *per student* unit costs were calculated for both treatments.

Level 1: The *per teacher* cost for providing Level 1 support was calculated to be $112.12. Of this number, $63.47 went to MTPDS labor costs[32] and $48.65 was spent on total direct costs.[33] This value was computed by dividing total labor and direct costs by the number of teacher participants (34,500). The Level 1 cost *per learner* was calculated to be $1.38 ($0.78 for labor and $0.59 for direct costs), using 2,803,335 standard 1–4 learners in target districts as the denominator.

Level 3: The estimated *per teacher* cost for providing all Level 3 inputs was $431.22 ($287.33 for labor and $143.90 for direct costs),[34] while the *per learner* cost was $7.89 ($5.26 from labor and $2.63 from direct costs). These calculations used 3,750 teachers and 205,000 learners as assumptions for Level 3 beneficiaries.

These unit costs were paired with estimates of program impact (calculated in the Literacy section above) to derive cost-effectiveness ratios. Cost-effectiveness analysis requires a basis of comparison against which to judge performance. As such, the lack of a control group precludes an analysis of the efficacy of the low-intensity treatment. Given the limited scope of the Level 1 intervention and the fact that literacy scores within the group did not

[32] All labor costs, for both Level 1 and 3, represent field and HQ support staff from Creative Associates, RTI, and Seward. These include base salaries, fringe, overhead, G&A, and fees.
[33] Constituent components included materials review workshop with MoEST, training of trainers, training of facilitators, CPD delivery at cluster level, and printing costs.
[34] These included training of trainers, workshop delivery, printing, and recurrent coaching costs

substantively change over the course of the project, the group can reasonably be substituted for a true control group for purposes of cost-effectiveness comparison.[35]

Cost-effectiveness estimates comparing Level 3 to Level 1 were computed using two separate estimates of impact: direct measurement and regression analysis. The first methodology simply compares changes in mean scores between the two groups. As previously mentioned, these direct measurements do not account for general time trends that may have occurred in the absence of the program. As such, some of the differences in outcomes may be caused by external factors, not the program. The second approach, regression analysis, is a more robust way to measure programmatic effects, as it allows controlling for many of these key factors.

According to the direct measurement approach (using Level 1 schools as a substitute for a control group), it cost $0.41 of Level 3 intervention to improve an average child's ability to read one additional letter per minute. Increasing the ability to read an additional word per minute cost one dollar, while increasing reading comprehension scores by one correct answer cost $15.

Comparing cost-effectiveness estimates between the two impact estimate approaches allows for a more nuanced understanding of true program impact. Unsurprisingly, every one of the direct measurement estimates is higher than those calculated using regression. Interestingly, each direct estimate is about four times less expensive than the corresponding regression-derived calculation. This uniformity seems to suggest that direct measures consistently over-represent program impact.

Table 10: Cost-Effectiveness Estimate, by Treatment and Impact Estimate Methodology

Treatment	Unit Cost (Learner)	Measure	Regression Analysis			Direct Measurement Change in Mean Scores		
			Letter/ Minute	Word/ Minute	Read. Comp.	Letter/ Minute	Word/ Minute	Read. Comp.
Level 3	$ 7.89	Program Effect*	5.07	2.28	0.14	20.05	8.74	0.56
		Cost-Effectiveness	$1.72	$3.77	$60.68	$0.41	$0.97	$15.17

CONCLUSIONS

While the lack of control group data does not allow for an exploration of the counterfactual, given the lack of literacy gains among Level 1 students, it can be inferred that the low-intensity intervention is not a cost-effective way to increase literacy scores. In and of themselves, the positive Level 3 impact estimates clearly indicate the supremacy of the high-intensity intervention.

[35] The danger of using Level 1 as a substitute for the counterfactual is that schools without the CPD intervention may have done even worse without MTPDS support. In this case, we would be underestimating program impact. Conversely, if schools would have done better without MTPDS support, we would be overestimating program impact.

As compared to Level 1 (in this case used as substitute for a control group), the high-intensity treatment was clearly the more cost-effective option. Of the two methodologies presented in the preceding analysis, regression analysis is much more likely to better approximate the true cost-effectiveness ratio. Using this approach, a dollar of Level 3 intervention would be expected to increase of reading ability of an average child by an additional 3.26 letters per minute, as compared to Level 1. The corresponding values for words per minute and reading comprehension are 0.66 and 0.002, respectively.

It should be highlighted that these estimates represent cost-effectiveness of the MTPDS project, as implemented by Creative Associates, RTI, and Seward. If MoEST were to continue CPD in the future, cost estimate would be different for two reasons. First, MTPDS contractors and the Government of Malawi have different cost structures. Without substantial expatriate time both in the field and in remote support positions, labor costs would doubtless reduce. Further cost reductions would be realized from reductions in fringe rates, overhead, and fees. Secondly, if the materials produced by MTPDS were used for future CPD trainings, development costs would not reoccur. As such, the direct cost portion of the estimate would also reduce. While it is not possible for the Evaluation Team to definitively calculate these hypothetical situations, it is plausible that the cost-effectiveness of these interventions would increase if they were implemented directly by MoEST.

RECOMMENDATIONS

Future Programming:

- On the basis of the preceding analyses, as well as findings collected through the course of primary data collection, it is clear that Level 3 schools outperformed Level 1 schools. However, given the planned incorporation of key MTPDS inputs in the revised primary school curriculum it is not recommended that USAID or MoEST implement any of the three packages, as they currently exist, in the future.
- To the extent that the new primary curriculum materials and other Ministry initiatives omit important program components (particularly syllabic approach, scripted lesson plans and extra hour), it is recommended that MoEST discuss with key stakeholders how best to integrate lessons learned into the formal education system. It is critical to include these changes as part of curricula in the nation's Teacher Training Colleges (TTCs).
- MoEST should integrate the multiple reading programs (MTPDS, Read Malawi, new primary curriculum, etc.) into one coherent curriculum, drawing on the lesson learned and benefits of each approach.

Future Research:

- If similar, large-scale data collection activities are planned in the future, it is highly recommended that ample time be dedicated to instrument and data collection protocol design in the early stages of a program. Once finalized, it is imperative that substantive revisions not be made to either during the period of implementation.
- The utilization of MoEST personnel in high-stakes data collection presents both benefits and costs. For future impact evaluations it is recommended that professional and impartial enumerators be used in lieu of MoEST personnel. In the event that this option is not feasible due to budgetary pressures, it is strongly recommended that MoEST personnel engage in data collection not enumerate in areas under their direct responsibility.

- The ability to make inferences from evaluation findings depends on the validity of the underlying data. It is recommended that any party responsible for high-stakes data collection utilize double data entry (if using paper surveys) or electronic data collection methods to ensure that information is of sufficient quality.
- Last, if ascription of educational outcomes to USAID funding is desired in the future, it is imperative that a true control group be integrated into research designs. If withholding treatment is not feasible due to ethical, political, or logistical reasons, the evaluation team recommends phased implementation.

II. PERFORMANCE EVALUATION

Background

This chapter of the evaluation report answers evaluation questions 2–4 through empirically derived findings, conclusions, and recommendations. As with the previous chapter, we first provide an overview of evaluation limitations. Each evaluation question is presented independently with a brief introduction and a presentation of empirical findings, synthesized conclusions and actionable recommendations.

Limitations

Hawthorne effect:[36] In situations where participants are aware that they are involved in a study, it is not uncommon for them to change behavior in response to being evaluated. This concern is particularly acute with classroom observations, where the very presence of Evaluation Team members invariably causes teachers to adjust the manner of instruction to present a more favorable impression. Without the use of candid recording equipment, it is not possible to completely control for this bias. However, to minimize distortions stemming from schools changing normal behavior to prepare for observations, visits by the Evaluation Team were all unannounced.

Sample size: Given the allocated resources and time, the Evaluation Team was not able to visit all treatment districts, much less all schools. As a result the generalizability (*external validity*) of Performance Evaluation findings is limited in comparison with evaluation question 1 (which utilized robust secondary data). To maximize the representativeness of site visits, the Evaluation Team utilized a stratified sampling process capturing key dimensions of school performance. Specifically, a bracketed sample of schools and districts yielded a balance of units across critical variables: high/low intensity training, high/low performing schools, and geographic spread.

School accessibility: Given that the evaluation was carried out during the rainy season, several of the sampled schools were inaccessible due to the condition of roads. To the extent possible, the Evaluation Team selected replacement schools using the same criteria.

Package components: A major strength of MTPDS is a design that includes three distinct modalities of intervention and, thus, the opportunity to better understand what components affect the impact on students and teachers. While it is possible to compare the efficacy of the three intervention packages, it was difficult to disaggregate and assess each of the

[36] The term Hawthorne Effect stems from industrial studies between 1924 and 1932 in which basic changes in the environment of factory workers impacted their productivity.

intervention components separately. The Evaluation Team utilized qualitative data collection approaches as well as a custom teacher questionnaire to explore the relative merits of package components. See Annex 7 for the evaluation team's complete analysis of package components.

Limited implementation period: Although the two initial districts of Ntchisi and Salima implemented MTPDS in 50 schools for a full year prior to the data collection in November 2012, this was not true of the five additional target districts and even the subsequent schools added in Ntchisi and Salima. In the five more recent target districts, the children had two terms[37] of the MTPDS intervention by the time of the November 2012 EGRA data collection. Thus, the test results in these districts cannot reflect the potential impact of the reading intervention since these students were exposed to the program for two terms instead of the full three terms.

Evaluation Question 2: Lessons Learned

As described in the introduction, the MTPDS program had different components and implemented with various levels of intensity. This section traces teacher experiences with the various components at each level of intensity and draws lessons learned in the process. This evaluation question asks the extent to which MPTDS components were successfully implemented or met with challenges, and then draws lessons learned from this data. Findings, conclusions and recommendations in this section are presented under each component for each level of intensity: Level 1: Literacy CPD and Leadership CPD; Level 2, adding scripted lesson plans and *Nditha Kuwerenga* readers; Level 3: adding on coaching.

FINDINGS

Literacy CPD

Strengths. The Analysis of Package Components (see Annex 7 for complete analysis) findings from a teacher questionnaire used by the Evaluation Team (Annex 9) show that among all teachers in the six districts visited during the evaluation who received training[38], most responded positively that they learned useful information from the literacy CPD training (69% strongly agreed and 30% agreed). In addition, the percentage of teachers who said they used knowledge and skills from the training in their classrooms increased across treatment intensities, either strongly agreeing or agreeing (38% in Level 1; 50% in level 2; 61% in level 3). When asked whether or not the trainings have made them more effective, the majority of teachers responded positively to the statement, irrespective of treatment level and almost all said that they would like to attend more training, if offered (93%).

Overall, the syllabic approach included in the literacy CPD was found to have broad-based support and utilization. The Evaluation Team's classroom observation data, representing 61 classrooms visited, shows that nearly three-quarters of the teachers observed embraced

[37] Note that because the intervention for the five new target districts began so late in the previous academic year, most schools have continued to use the S1 scripted lessons and readers in S2 begin with the new school year in September 2012.

[38] 107 teachers across three treatment levels completed the Evaluation Team's teacher questionnaire. 96% of all teachers who completed the teacher questionnaire received CPD training across all levels, with an average of three trainings.

skills that supported the syllabic approach to promoting literacy[39]: 78% used vocabulary, 70% used phonological awareness, and 65% used alphabetical principles. These were more pronounced in Level 2 and 3 districts. However, fewer than half the teachers used the remaining two skills of fluency (43%) and comprehension (37%), and these were used only in Level 2 and 3 schools. Also touted as helpful from CPD trainings by the teachers was the concept of the learning cycle. The learning cycle necessarily involved four steps that the teachers had to go through in a lesson: *advance organizer, modeling, guided practice, and independent practice*. There was considerable evidence of use of some of these techniques, especially the first three among the teachers of Level 2 and 3 schools; although almost no teacher in Level 1 schools demonstrated their use.

Use of the Chichewa alphabet[40], as opposed to the English alphabet, is another technique that was found to be new and useful by the teachers. A Teacher at Chorwe School (Level 3) stated that "Using the Chichewa alphabet has helped me a lot in teaching my children how to read and write. I did not know this before."

The literacy CPDs also embraced a number of teaching aids such as word cards and a print-rich environment (e.g., text on walls), plus interactive techniques including songs and praise. Classroom observations showed that use of these aids in classrooms was not dependent on treatment level of the school. For example the 83% of the teachers that used various teaching aids, the most common of which were word cards, were almost evenly spread in the three treatment schools. Songs and praise were also widely used especially in S1 classes in all schools. However, other techniques met with significant difficulty. This is further explained as a weakness below.

Weaknesses. While teachers reacted positively to the literacy CPD trainings, weaknesses were noted across all treatment levels. Concerning the CPD training process, a number of issues were raised by teachers during the interviews. The vast majority complained about the limited allowances received and the limited number of days for the training, which were eight (two days for each session). In addition, one facilitator was used for the entire training period (two-day sessions), which led to fatigue and hence a less dynamic training environment. Also, since only S1–4 teachers were trained, other teachers who might be rotated[41] into these classes were not prepared. A related shortcoming was the decision not to invite head teachers to the trainings unless they taught in S1–4. Since these head teachers did not have information about the Literacy CPD, they had a difficult time supporting their teachers who had been trained.

Of the various teaching aids and pedagogical techniques included in the Literacy CPDs, those that were most challenging to implement were as such due mainly to large class sizes and poor resources: continuous assessment, grouping of students, and providing a print-rich environment. *Continuous Assessment* (CA) was defined by the Evaluation Team as a more formal process by which teachers kept a list of each student and indicated specific measures of achievement. We observed that only a few of the teachers conducted CA in the way we

[39] Five reading skills enforced through CPD and MPTDS were: phonological awareness, alphabetic principles, fluency, vocabulary, comprehension, and teaching of writing.

[40] ANIKUMETO are the commonly used letters in Chichewa language, both written and spoken.

[41] The Evaluation Team noted that there is a rather high rate of rotation (re-assignment) of teachers in Malawian primary schools (both within and between).

defined it. The vast majority of teachers, however, used the more informal measures mentioned such as thumbs up and down, clapping of hands, etc. In fact, in the 39% of the classes (20 of 51 observed) where CA was used, there was limited success due to large class sizes as often when the teacher was assessing some learners, others were sitting idle and making noise. According to the Evaluation Team's teacher questionnaire, the average class size of the sampled lessons was 113, ranging from 23 to 452 students. CA in large classes was challenging as it takes a long time and parents often do not provide folders for portfolios where learner assessment reports are kept.

Use of *group work* ("grouping") to enhance learner interaction during lessons was a technique appreciated by many teachers, but difficult to implement in large classes, even though large classes was a rationale for using groups. Across all three levels, group work was used with limited success. There was no interaction of learners in the groups, or some groups were left without much to do while the teacher worked with one of the groups.

The provision of *print-rich environments* was limited by security concerns as many classrooms did not have lockable doors and other lessons were held under a tree. For all of the three treatment level schools, inadequate classrooms and lack of head teacher offices led to minimal use of "print-rich" materials as teachers often complained of lack of storage space and materials to post in classrooms. As such, print-rich environments were observed only in 36 classrooms observed (59%).

Nditha Kuwerenga Readers

Nditha Kuwerenga readers were developed by MTPDS and distributed only to S1 classes in the L2 and 3 districts, although in practice these were also used by S2 teachers in some schools, especially L2 districts.[42] 77% and 68% of all teachers received the readers in Levels 2 and 3, respectively.

Strengths. Virtually all teachers in both Level 2 and 3 schools claimed that they use the readers. According to the teacher questionnaire, a majority of teachers in S1–2 reported using them every day. Those teachers said the readers were useful, especially in Level 3 districts that had used them for a longer period of time. Almost all teachers agreed that the readers helped the children acquire literacy skills.

Weaknesses. Based on classroom observations of 31 Standard 1 classes in Level 2 and 3 schools, many of the teachers did not use the readers, in spite of the positive report on usage in the teacher questionnaire and during interviews. In fact, readers were seen in only 42% (13 of the 31 observed) of the S1 and S2 MTPDS classes observed and some of the L1 schools did not have the reader at all. However, it is possible that some teachers felt pressured by the observation and decided not to complicate the lesson by using the readers. The evaluation team encountered differential usage of the readers: in some schools the readers were visibly worn to the point of falling apart (in Senga Bay, the head teacher had set up a stapling station to reconstruct the books), while in others it was evident that the

[42] Only in early 2013 did MTPDS distribute some scripted lessons and readers to L1 teachers.

readers were new and one teacher at Wantaya school (Level 2) was using them for the first time during the time of the classroom observation. Of those teachers who tried to use the readers while being observed, there was limited success. In many classes, learners struggled to find the referenced page due to limited numeracy vocabulary. For example, two adjacent pages use similar symbols and often have similar letters so learners cannot easily identify the page being referred. On a less technical note, the book is too big for the learners' desk or lap in crowded classes.

Scripted Lesson Plans

Scripted lesson plans were received only by S1 teachers of L2 and L3 schools by design. However, while L2 districts had received them about five to six months before the evaluation, L3 schools had used them for more than a year since they were the first intervention districts. All L1 teachers reported not to have received any and so the findings in this section only refer to L2 and L3 schools.

Strengths. Observational data collected by the Evaluation Team show widespread use of scripted lesson plans in Standard 1 in all intervention districts and in S2 in the five new districts.[43] In the two L2 districts of Mzimba North and Blantyre Rural, the lesson plans were found to be used by S2 teachers as well although they were meant for SI. When interviewed, teachers said that they were doing so because they had received the lessons plans when the S1 learners were in term 3, so they extended their use to the next academic year and therefore S2.

In general, teachers appreciated the lesson plans for "making our jobs easier" during the interviews with the Evaluation Team. Among Level 2 and 3 teachers who received the scripted lessons and completed the questionnaires, 83% and 95% respectively, reported using them every day. The lessons are valued for helping to integrate the five principles of teaching reading, "something that teachers wouldn't be able to do themselves" (CEED M&E officer). Teachers at the Ngumbe School (Level 2) reported that "with the introduction of the scripted lesson plans for teachers, we have grown in confidence. We are able to face the class confidently and deliver the lesson as required." Of teachers who completed the teacher questionnaire, a great majority said that the scripted lessons made them more effective (57% agreed/ 40% strongly agreed across all levels). Overall, the main strengths of the scripted lesson plans were that they made the teachers' job easier; they helped integrate the five principles of teaching reading and increased teacher confidence.

Weaknesses. A number of weaknesses about the scripted lesson plans were noted in the course of the evaluation. In every L2 and L3 school visited, teachers complained that the scripted lessons are too long. There is the generic challenge of teaching effectively in large classes, which affects MTPDS as well as other subjects. The end result was that learners became bored or got distracted as their attention limit was often reached in the middle of the lesson. Some teachers interpreted use of scripted lessons to mean that they did not have to prepare for their classes, but could just read through the scripts as they teach. The Evaluation Team saw both verbatim and non-verbatim use of the lesson plans during other

[43] The scripted lessons are designed for S1, but since the new five target districts came on board late during the last school year, the S1 classes did not complete the scripted lessons for the whole year. Thus, for the current year, many of the S2 classes have continued with the S1lessons. This should only be a one-year phenomenon.

classroom observations. For instance, during a classroom observation one teacher at Wantaya School (Level 2) in Mzimba North read the scripted lesson plans verbatim, making limited progress in the lesson and failing to adequately engage all learners. In sharp contrast to this, a teacher at Madise School (Level 3) in the same district taught a lesson integrating all of the five literacy principles in a very successful 30-minute lesson without making verbatim reference to the scripted lesson plans. In general, teachers who used the scripted lessons as a guide tended to be more engaging than teachers who just read the scripted lessons.[44]

Some respondents complained that both the CPD modules and the scripted lessons have some mistakes and typographical errors.

Extra Hour

An extra hour for teaching literacy was recommended by MTPDS to all literacy intervention schools in formed L2 and L3 districts. The expectation was that the teachers in such schools would use the extra hour to teach using the scripted lesson plans that they had received.

Strengths. Almost all Level 2 and 3 schools, except for the Mgaga School (Level 3)[45], extended their school days by one hour, especially for S1 and S2. In most schools, the extra hour, especially for S1, has become school policy, as evidenced by its inclusion on the timetables/school schedules observed.

The extra hour is used for reading (mostly S1) or remediation (S2–4), and is usually placed in the middle of the morning, although some schools used the scripted lessons or held remedial classes at the end of the morning. Most S1 teachers said that the extra hour was not only valuable, but essential for teaching the scripted lessons, which is an add-on to the school subjects. This prioritization of the MTPDS (or EGRA lesson as some schools referred to it) is quite significant as it is presented during a time of the day when students are most attentive.

With respect to the value of the extra hour, most S1 teachers said that the extra hour was not only valuable, but essential for teaching the scripted lessons, which is an add-on to the school subjects. Of the 38 teachers L2/3 teachers who responded to the question of whether the extra hour made them a more effective teacher, 45% agreed and 55% respondents strongly agreed with the statement.

Weaknesses. Most Level 2 and 3 teachers identified that one main challenge was that an hour is too long for learners to attend to a single subject. Students' attention faded over the hour and they became tired or hungry. Moving other classes to the later hours gave the impression that the other subject areas are a lower priority.

[44] Based on the teacher questionnaire, three-quarters of the S1 teachers and two-thirds of the S2 teachers received the scripted lessons in Level 1 and 2 districts, and about 90% of these teachers reported using them every day. Very few teachers in the Level 1 districts had the scripted lessons, and those that did have them did not use the lessons.

[45] The Maganga School is a double-shift school, which limits possibilities of having the extra hour as the classes have to be used by the next shift. It was learned that for this school, teachers have tried to integrate the MTPDS approach inherent in the lesson plans into their normal Chichewa lessons.

"When you put the extra hour in between other lessons, learners may manage to concentrate for the one hour. However, when they do concentrate it becomes difficult to teach them the next lessons as they become very tired. On the other hand, when you have the extra hour at the end of the school day, learners are not able to concentrate the whole time." (Teachers, Engocongolweni School – Level 2)

A second challenge is that the extra hour is still voluntary, although there seemed to be wide compliance by teachers. Yet, since the extra hour is not yet MoEST policy, it is not strictly adhered to in all schools.

Coaching

As mentioned above, Level 3 was intensified through the addition of coaching by MTPDS staff in two initial districts—Ntchisi and Salima—and later in five more target districts. While the MTPDS staff coached teachers in selected schools, the coaching function quickly evolved into a responsibility of PEAs in all target districts.

Strengths. The Evaluation Team found that coaching provides useful support to the teachers implementing the reading program and is valued by teachers. Teachers appreciated the support they received from the coaching and, for those who received several coaching sessions, they found them to help their professional growth. Teachers at the Ngumbe School (Level 2) stated that "it is very useful and allows us to improve the teaching of literacy especially in areas that we are not doing great; but would love to increase the frequency that coaching is undertaken." Indeed, responses from the teacher questionnaire showed that 63% of Level 2 and 54% of Level 3 teachers strongly agreed that coaching made them more effective teachers because they said it helped correct the mistakes that they were making in the course of their lessons. Based on the teacher questionnaire, 79% of the teachers in the L2 and 85% of teachers in L3 districts reported to having received coaching from either a PEA or an MTPDS staff member and the vast majority reported to have been coached at least once in the previous term.

Weaknesses. A major challenge for the coaching process is that the PEAs already have too many responsibilities and too many schools to support. Thus, they are unlikely to be able to provide a substantial amount of support to teachers. In addition, there was wide variability in the coaching support; the number of visits in one term ranged from 0 to 7 and the average number was 2.2 visits. In general, however, it was unusual for a teacher to have coaching more than twice a term. In some cases, PEAs opted to coach teachers at the cluster or zone level because the schools were inaccessible or just to reach more teachers during a limited time. There was some additional variability in the type of coaching support provided. While some teachers received observation, individualized coaching, demonstration of teaching techniques, and follow-up, others only received the first two as individuals and received the latter in groups.

PEAs also reported difficulty in travel due to lack of transportation resources that would provide reliable motor bikes/vehicles and fuel for travel.

There was limited evidence of head teachers providing coaching support to teachers, mostly because they did not attend the literacy CPDs unless they taught in S1–4.

Leadership CPD

Leadership CPD modules were taught to all head teachers and many deputy head teachers throughout the county. The emphasized the organization of school-based CPDs and the monitoring reading by SMC/PTA and community members (especially parents). These trainings were intended to bolster systems and understanding of the literacy intervention components in order to promote a reading and literacy enabling environment.

Strengths. The first module of leadership CPD centered on how schools could organize school-based CPDs. Visits to the schools in all three level districts revealed minimal evidence of school-based training. The presence of CPD training on the school timetable was more prevalent in L2 and L3 schools. Some zones in Level 1 districts also showed evidence of school-based CPDs, one example of which is Chihame II School (Level 1) in the Nkhata Bay district, where the two schools visited had school-based CPD timetables and a stockbook in which notes on what was discussed were recorded by the school secretary.

The leadership training modules did not only encourage participation of SMC in reading but also that of the whole community. The findings show that in many L3 schools, the community is more engaged in education in general, particularly reading, than they were before the program. The causal pathway seems to be a combination of the Parent Teacher Association (PTA) and children disseminating program benefits to their peers and parents, respectively. As parents attend the PTA meetings, which are held at least once a month as part of normal school management, parents are informed about how the learners are improving their reading and this gets disseminated to the rest of the parents that may have otherwise not attended the meeting. The incorporation of literacy fairs as part of the school open days was encouraged by MTPDS, especially in L3 schools. One learner in Ntchisi was teaching other children in the community and was shown on a display presented during the literacy fair. Through the fair, such practices were disseminated to the wider public and parents were encouraged to take an interest in what their children are doing at school.

Weaknesses. Although there was wide support/demand for school-based CPDs, there is little implementation to date. Institutionalization of school-based CPDs in some districts was negatively affected by a number of factors, including a lack of teacher allowances.[46] Similarly, while some Level 1 districts such as Nkhata Bay had widespread school-based CPDs, Dedza did not have any, raising questions about the success of the cascade model for CPD implementation.

The evaluation team encountered widespread confusion about the role and implementation of the *school report cards,* which contained information about the school, the teachers, recording absenteeism rates, as none of the forms had been distributed and even intended champions of the process appeared confused[47]. The two most significant observations regarding school report cards were that (1) there was a significant lack of coordination with the former EDSA program, which also created a "school report card" for school assessment[48]. Although MTPDS says that the purpose of their version is a bit different from EDSA, it seems like there is unnecessary duplication. There was also confusion with the name, as most teachers assumed that the school report card is what gets sent home to

[46] At Ntonda school in Blantyre Rural (Level 2), there was a complaint that teachers refused to have the school-based CPD for free while their colleagues received some allowances for it.
[47] Except for pilot zones which were managed by MTPDS staff in the 5 level 2 districts.
[48] EDSA was a decentralization activity working at the community level. Whereas, MTPDS was a teacher professional development activity that worked between the school and district levels.

parents; (2) the report card has no guidance on timing (# of times per term, or what part of the term to conduct activities). The SMCs have limited capacity to implement the school report card since it requires statistical information that can best be collected by head teachers.

Conducting "open days" and literacy fairs was not compulsory for L2 and 3 districts and not encouraged for Level 1 districts, which could be regarded as a missed opportunity.

Lessons Learned

From the foregoing findings, a number of lessons can be learned with regards to the implementation of a program and intervention to improve literacy. The lessons learned can be distinguished for program design and implementation.

Program design. The program design included levels of treatment intensity, provision of inputs, and a varied target population. The findings of the study show that L3 schools showed higher achievement both in reading ability of the learners and the degree of buy-in of the ideas taught through the CPDs by the teachers, while L1 schools performed consistently poorly. One lesson learned is that CPD alone may not be sufficient to change teacher practice; rather, it has to be associated with other inputs to support the teachers.

Secondly, MTPDS literacy CPDs targeted only S1–4 teachers. While this had the advantage of focusing attention to teachers who teach early grade reading, two challenges were noted. The first is that there is a tendency for schools to swap teachers around in different classes due to prevailing conditions. This resulted in some S5-S8 teachers, who have not been trained, being transferred to the lower standards and, thus, not prepared to teach the new reading program. The second is that head teachers who were not teaching S1–4 classes were not invited to attend the CPDs. Thus, they were not prepared to monitor and support their teachers who were trying to implement the MTPDS program. The lesson learned is that, where possible, all teachers should be involved in the CPDs and head teachers have to be targeted for any interventions on teaching and learning whether they are directly involved or not.

Program implementation. The study showed that scripted lesson plans and, to a somewhat lesser extent, the *Nditha Kuwerenga* reader are powerful tools to enhance reading skills of learners if used properly. However, their long-term usefulness of these two inputs will be limited by their lack of alignment to existing Ministry guidelines and curricula. One lesson learned is that it is important to align any inputs given to schools to existing guidelines, norms, and curricula if a program is to have a lasting impact. Such alignment will also help to ensure teacher acceptance of the intervention. The same can be said about the extra hour, which though useful, was not enforced by Ministry policy. The MTPDS coaching component, which was quickly picked up by PEAs as part of their job, is but one good example that shows how alignment of inputs to existing structures would lead to better results.

CONCLUSIONS

Literacy CPD and Intervention Inputs

- The literacy training program was useful in improving teacher practices, and teachers valued the new knowledge and skills that they learned in all three treatment levels.

- Scripted lesson plans and, to a somewhat lesser extent, the *Nditha Kuwerenga* reader are powerful tools to enhance reading skills of learners.
- Coaching was found to be useful by beneficiary teachers and, if used regularly, will lead to greater improvements in teaching reading and promoting literacy. However, the PEAs are unlikely to be able to increase their number of visits to each teacher, which is currently about two per term, and this limited contact may not make a big difference in teacher performance or student learning.
- Use of the extra hour was found to be instrumental in achieving literacy by Level 2 and 3 teachers.
- Continuous assessment was used with mixed success in Level 2 and 3 districts.

Leadership CPD

- The leadership training program is valuable and head teachers appreciated the new knowledge and skills that they learned in all of the three treatment level districts.
- As they are currently being implemented, use of school report cards is not sustainable due to lack of distribution and sufficient training and confusion with the EDSA report cards and over which one should be used (school report cards versus student report cards).

RECOMMENDATIONS

The following are recommendations based on the findings and conclusions above:

- There should be continued literacy and leadership CPDs at zonal, cluster, and school levels with funding from district CPD funds and School Improvement Plans (SIPs) in all districts. These CPDs can introduce new teachers to the skills promoted by MTPDS, and new refresher modules should be created to help sustain and extend the gains made by teachers and head teachers. Also, all head teachers should be invited to the literacy CPDs, and these CPDs should be open to all teachers through S8.
- Scripted lessons should be continued but in a revised form that builds on and supports the new primary curriculum. Lessons should be shortened and include additional activities for more able classes. While decreasing the size of classes is a challenge to be addressed in the long term, revised lesson plans should include activities that aim to address the challenges of teaching in large classrooms such as the scripting of roles for each student placed in a group (for group work) or strengthening scripts for 'guided instruction' components. In addition, there should be guidance on how teachers can be weaned off of just reading the scripted lessons. MIE could take the lead in revising the lesson plans.
- The Ministry should phase out the *Nditha Kuwerenga* readers and realign the reading activities in the scripted lessons to the new primary curriculum readers as another activity of MIE.
- One of the highest-impact ways to increase literacy capabilities is to increase time on task. Thus, there is need for MoEST to formalize into policy (1) extending the

school day for S1–2, (2) formalizing integration of the additional reading curriculum into daily instruction, (3) extending the MTPDS intervention components into S2–4, and (3) integrating the MTPDS approach into other subjects.

- Coaching would be greatly improved by training head teachers and deputies of all schools in all districts on how to conduct coaching and incorporating this into their SOWs. Then, coaching could be provided at higher frequency and lower cost. Additionally, skills presented through coaching should be mindful of the challenge of large classes and include remedies for school leaders, HTs, and teachers to utilize.
- In order for PEAs to continue the coaching, the Ministry will need to provide adequate funding for the motorbikes, fuel, and other travel costs. MTPDS provided travel subsidies for PEAs, which won't be continued when the program ends. However, the Ministry has a process for providing PEAs with money for fuel, but the level of funding needs to be increased. Also, the Ministry needs to provide additional funding for motorcycle maintenance and replacement.
- The Ministry should think of alternative ways of doing continuous assessment for all schools, as the current process does not appear to work well with large classes. For example, there might be more oral assessment combined with shortened tests as part of normal lessons as well as the inclusion of co-teachers and head teachers.
- Since there is already a report card introduced through the PSIP program, there is no need to reinvent the wheel. The MTPDS report card could be integrated into PSIP and be rolled out to all districts as part of the same PSIP. Use of this integrated version of the school report card would be strengthened by encouraging schools to include this activity in their SIPs.
- The Ministry should consider encouraging the coordination of literacy fairs through school open days for all schools.
- Community participation, especially among parents, in monitoring reading needs to be emphasized and followed up in all schools.

In conclusion, MTPDS is a valuable program. It has introduced several new and innovative ideas in Malawi that have resulted in significant gains in reading scores, especially in Level 2 and 3 districts. Teachers and community members have valued the program because they see that their children are learning how to read in a much earlier grade. The key remaining challenge is to develop and implement a transition plan that will ensure that the key components of MTPDS be integrated into the education system and the new National Reading Strategy.

Evaluation Question 3: Ownership of MTPDS

As MTPDS comes to a close, it is expected that the MoEST will become responsible for and take ownership of the literacy and leadership CPDs, as well as incorporating some of key activities such as scripted lessons (whether modified or not), the use of readers, and the extra hour per day. The Evaluation Team defined *ownership* as the extent to which relevant senior officers and departments of the MoEST (i.e. (DIAS, DBE, DTED, and DEP) buy into and embrace responsibility for key MTPDS components. The transfer of ownership should enable sustainability of key MTPDS components and a program to revise the components to better fit the Malawi education context. The Ministry at all levels would be committed to

incorporating the revised program in training teachers, head teachers, PEAs, and SMC/PTA chairs, and the use of the revised materials in classrooms.

The Evaluation Team discussed the ownership of MTPDS products and activities by MoEST and community based on two pillars: leadership and responsibility. These pillars describe the various roles and responsibilities of MTPDS and the Ministry and its departments (DIAS, DBE, DEP, and DTED), DEMs, primary school teachers in S1 and S2, head-teachers, MIE and SMCs/PTAs. The key areas of inquiry regarding the question of ownership were: how the MTPDS activities were designed; who participated in the activities; and the relationship among the various key players in designing, planning, and implementing the program. This section focuses on the current state of responsibility and ownership of MTPDS from its inception to the present and the transfer of MTPDS components to the MoEST.

FINDINGS

Leadership

Leadership refers to actions of Ministry personnel to shape the development and implementation of MTPDS and, currently, actions taken to incorporate MTPDS fully into the Ministry.

Strengths. The Ministry was involved in the initial design of MTPDS and recognized that it had an important role to strengthen the development of the MTPDS, the CPD programs, and the related reading intervention. The Ministry recognized that MTPDS had a role to play in improving the quality and relevance of primary education as stated in the National Education Sector Plan 2008-2017 (MoEST 2008) and its related Education Sector Implementation Plan 2009-2013 (MoEST 2009). Since the Ministry already had a CPD program and system in place, it was deemed best to build on this model for carrying out the training related to MTPDS.

The evaluation shows that MTPDS designed the project in conjunction with the Directorate of Inspection and Advisory Services (DIAS), Directorate of Basic Education (DBE), Department of Teacher Education and Development (DTED) and, to a limited extent, Directorate of Education Planning (DEP). A sense of ownership was also enabled by MTPDS involvement of DIAS, DTED, DBE and MIE personnel in the initial stages of developing CPD, Nditha Kuwerenga readers, and scripted lesson plans.

Such a participatory design was done on the advice and guidance of USAID. By working with Ministry personnel, MTPDS was able to benefit from their advice and experience, and to take into account their concerns. The project provided the Ministry an opportunity to champion the activities as leaders in their respective departments and duties. At different stages of the implementation there were different leaders. For instance, DTED appeared to lead the CPD design, whereas DIAS was on the forefront in developing teaching and learning materials.

Other current evidence of leadership regarding MTPDS and the CPD activity is shown by DBE, which remarked that after the first priority (teaching and learning materials), the second priority in the coming financial year ought to be CPD. Also, DIAS was looking forward to working with DTED and MIE in making CPD a nationwide exercise with special attention given to reading and leadership. DIAS intends to continue the Literacy CPD

training along with other CPD trainings as one way of improving reading in Malawi. For example, the PEAs took it as their duty to monitor progress, whereas central personnel under DIAS were used as resource persons at national and divisional trainings.

Weaknesses. Although there was initial support from MoEST, the leadership role decreased over time. Such leadership on MTPDS activities was largely on an individual basis due to the MoEST's failure to commit its staff for long-term activities or meet MTPDS' expected deliverables/outputs, all as a consequence of other relevant demands such as the department's tight schedule. Second, over time, the collaboration among MTPDS, the Ministry, and MIE decreased, perhaps because of the pressure on MTPDS to meet their contractual deliverables within a given time. In fact, MTPDS had to play catch-up on lost time since the implementation of the project itself was delayed (largely through no fault of the project). Overall, participation in MTPDS activities by individuals at the expense of working on behalf of their directorate and /or institution did not advance and promote teamwork and institutional leadership. Worse still, the Department of Education Planning, as the key policy and planning mover, was hardly involved.

A final finding on the weakness of the leadership can be summed up by the following words from the Director of DIAS, "There was [a] need to construct the 'roads' and 'bridges' if the project was to be successful and minimize conflict and contradictions between and among key players from the three entities" (government, funding agency, and implementing agency). There was a need for written roles and responsibilities (roadmap) that could make the three organizations (MoEST, MIE, and MTPDS) work together and minimize the unwarranted delays on both policy direction and implementation issues. In essence, we can argue that MTPDS activities did not have a definitive roadmap that could make all players commit, align, adhere/follow, and execute as part of their normal duty in time and without impediment. Thus, MTPDS should have ensured that leadership was deliberately defined and not assumed under the project.

Over time, the role of the Ministry shifted from providing leadership to just participating, which is covered below. For example, although MTPDS always tried to work through appropriate institutional channels, in reality, the project tended to work.

Examples of working with individuals at the expense of the institution and tight schedule were stated by key informants within both MIE and DIAS. For example, these key informants highlighted that MTPDS largely developed the scripted lessons and readers on their own or they worked with individuals from MoEST institutions such as MIE, but not through the institutions themselves. Although MTPDS states that they communicated regularly through official channels of the participating institutions, the perceptions of the individuals interviewed were that the formal links were weak and that assistance was provided more on an individual basis than through institutions. In fact, the Director of MIE had requested early on in the life of MTPDS that the project be placed at this institute, since MIE is the place where new curricula are developed. He was disappointed that the project remained in Lilongwe.

The Evaluation Team learned of the late planning for the sustainability of MTPDS activities which was scheduled in early 2013 (after the Evaluation Team left) with the Ministry to try to work out a way forward. As commented by the Director of MIE, "Why wasn't this done at the beginning of the project?" Another major issue throughout the life of the project, and

of particular importance now as the project is ending, is the lack of a lead department in the Ministry to oversee and provide leadership and a home for MTPDS activities. At different stages of the implementation there were different leaders. Interestingly, DBE also felt that they had a say given that its primary teachers and head teachers participated in the implementation of CPD. The Evaluation Team found that the DEP–Budget Section was not aware of MTPDS activities.

MTPDS dealt directly with individuals at DIAS, MIE, and Malawi National Examination Board (MANEB) and not through the official institutional channels. As one official said, "MIE was left out of the loop on planning various activities other than dealing with specific individuals from MIE mainly the language section." Another interviewee said, "The procedures for developing the materials were dictated by MTPDS officials and not based on a consensus."

Another example of apparent unilateral decision-making by MTPDS, was the design of the CPD that did not follow the existing MIE's "Technical Manual for the Organization of CPD" (2008). Instead of a school-based CPD as per MoEST/MIE guidelines, MTPDS chose to use a cluster model. Although MTPDS did promote school-based CPD in Leadership Module 1, it never formalized or supported this model on the ground.

Lack of engagement and collaboration with the ministry and MIE was systemic but it was further exacerbated by factors beyond the project's control, such as the death of the teacher training expert at MTPDS and the modification of the project that resulted in the departure of two senior MTPDS staff. After the program modification and the death of the teacher training expert, there was reduced engagement between the three institutions (MTPDS, MoEST, and MIE) on developing CPD materials.

MTPDS set up a parallel management structure (District Literacy Coordinators, Division M&E Officers) instead of using existing DIAS and DBE officers at division and district levels. Early on, MTPDS had a "light" structure using District Literacy Coordinators and Division M&E Officers, but this structure was substantially strengthened in December 2011 as a result of the contract modification, including an expansion to five more districts and a stronger M&E system. MoEST suggested that the limited time to meet deliverables after the modification resulted in MTPDS establishing a parallel structure. However, it was expected that the MTPDS staff would be counterparts to the division and district officers, thereby building capacity among the Ministry staff. However, capacity building through this counterpart modality was not fully realized because the DLC and divisional/district officers were visiting schools at different intervals and frequencies. It was also noted that the project staff and MoEST officials had no established times to share experiences and each group had its own reporting structure.

Responsibility

Responsibility refers to the extent to which Ministry officials or departments were committed to managing certain aspects of the project, especially the CPD component. That is, to what extent did they feel committed to and acted accordingly to support MTPDS? This is an important component of ownership and eventual sustainability, but at a slightly lower level than leadership as described above.

Strengths. As described above, the leadership of the MTPDS activities by the Ministry was fragmented, especially after the initial project activities. Yet, the sense of responsibility towards MTPDS activities by MoEST departments including DIAS and DTED is promising.

In fact, interviews with policymakers and implementers at the district, school, and community levels showed there is acceptance of the roles and duties bestowed on the PEAs, head teachers, and teachers in making sure that the MTPDS project was implemented according to Ministry expectations. The PEAs have taken on a major training role on behalf of MTPDS throughout the country and, more recently, in the seven target districts, have assumed an active coaching role. The teachers in L2/3 districts reflected that they have accepted the responsibility of using the scripted lessons, the *Nditha Kuwerenga* reader, adding an extra hour per day, and trying to conduct continuous assessment as arising from their training under MTPDS. An example of accepting responsibility and change can be construed when teachers spoke highly of phonological awareness and syllables as important skills for teaching Chichewa amongst S1 learners. Likewise, the SMC/PTA representatives, who had limited leadership training, took it upon themselves to accept the role of "watchdog" by stating that the MTPDS approach was worthwhile because they observed that children were able to read at an early stage.

Increased responsibility is also evident from the CPD program. MTPDS helped to promote and expand CPD training. Positive sentiments came from all of the three levels of MTPDS intervention. Policymakers stated that MTPDS has made CPD a reality because, unlike in the past when the Ministry was failing to implement CPD, it is now documented, budgeted, and implemented. Furthermore, increased responsibility is also shown at the school level in that there was interest in the inclusion of MTPDS CPD (both literacy and leadership) in the Primary School Improvement Plans (PSIPs). In the past, PSIPs mostly focused on construction and procurement of teaching and learning materials. Now, CPDs are also being financed through PSIPs, something that MTPDS has precipitated. As stated by a DEM from a Level 1 district and a Director at MoEST, MTPDS has shown that quality of teachers and other educational personnel was critical if quality and relevance of learning were to be advanced in primary schools. Overall, we conclude that MoEST headquarters, DEMs, PEAs and head teachers have duly accepted the responsibility of including CPD in their annual activities.

In addition to interviewees making positive remarks about CPD, it was observed that the departments (DBE, DTED, and DIAS) and MIE were ready to accept their expected CPD roles and responsibilities. For instance, DBE remarked that after the first priority (teaching and learning materials), the second priority in the coming financial year ought to be CPD; and DIAS was looking forward to working with DTED and MIE in making CPD a nationwide exercise with special attention to reading and leadership. In line with committing to MoEST strategies and accepting responsibility, DIAS intends to continue the Literacy CPD training along with other CPD trainings as one way of improving reading in Malawi. Actually, the essence of the ministry being committed and accepting responsibility was shown when the

Operations and Guidance Plan for CPD, which spells out the MTPDS approach, was designed in consultation with and signed off by Ministry officials.

One notable sign of taking responsibility for an MTPDS activity was noted at the district level. Some districts, such as Salima, had released unbudgeted funds (MK76,000) for inclusion under the Literacy CPD after realizing the importance of such an intervention. Likewise, Dedza, a district that had leadership and Literacy CPD training only, allocated funds to schools for CPD and was making a deliberate effort to include CPD under its 2013–2014 budget.

Officers from MoEST and MIE were part of the MTPDS activities from the initial design activities and they have retained some involvement, even if just in one's personal capacity or assigned by the department. Thus, MTPDS activities cannot be labeled as "foreign." The participation of individuals in MTPDS activities provides a basis for institutional responsibility and continued support upon the closure of the project.

Also, at inception stage, MoEST departments and MIE participated in developing the workplan, and most of the initial meetings to develop the readers and CPD trainings were done with the blessing of respective directorates and MoEST in general.

Finally, it was noted that all those who attended Leadership and Literacy CPD training (head teachers, teachers, and SMC/PTA community representatives) felt empowered and stated that the training led them to accept various responsibilities and advance the reading agenda without fear of contradiction or reprisal from other members of staff and community.

Weaknesses. However, non-acceptance of responsibility was observed in Nkhata Bay and Blantyre Rural, where they were not preparing any literacy or leadership CPDs under their PSIP because the DEMs were not conversant in the Literacy CPD training. The findings in Blantyre and Nkhata Bay showed that where the DEM was not aware of CPD or, more specifically, the MTPDS program, the CPD program as a whole may not be budgeted, or if so, it may not include a Literacy CPD.

Second, the acceptance of responsibility had its own challenges among DIAS, DTED, and MIE in terms of what entails CPD. The CPD, which the MTPDS project was advancing, was not a complete mirror of the "Technical Manual for the Organization of CPD" by MIE. Thus, the MIE manual intended to advance CPD at the school level, whereas the MTPDS CPD was at the cluster level and the focus for MIE training was a variety of school subjects and issues, such as absenteeism, welfare, and discipline, among others. The MTPDS was focused on just reading. So there was some resistance as to what CPD connotes to some MIE officials.

CONCLUSIONS

Although MTPDS was intended to be integrated within the MoEST system from the central level through the districts and down to the school level, in addition to including many stakeholders from across the education landscape, engagement was insufficient. There is evidence to demonstrate some leadership and responsibility by the Ministry institutions and departments, yet the results have been mixed, leading to challenges for sustainability. Such a shortfall limited the prospects for a smooth and timely transfer of MTPDS activities to MoEST (particularly since there is no clear owner). That being said, the MTPDS activities are

on course to being owned by the districts, communities, and schools, and there are promising indicators of leadership and responsibility at the Ministry level.

Second, responsibility is being accepted by MoEST departments (DIAS, DBE, DTED, DEP, and MIE), though late and gradual, and there are even arguments for modifying/changing MTPDS approaches to fit MoEST expectations and curriculum. Overall, there is some sense of ownership of MTPDS activities, although more work has to be done to institutionalize. A key indicator of ownership of MTPDS will be the budget allocation to support CPD training and the revision and distribution of materials.

Finally, it was validated by the Evaluation Team through discussions with directors of the MoEST budget team and DEMs that ownership can only be realized and enhanced if and when the Government of Malawi allocates funding for CPD and related activities in the annual budget at national, district, and school (local) levels. Furthermore, such government commitment will reflect political will and national ownership of intent to improve literacy in general and reading specifically in education.

RECOMMENDATIONS

Ownership of MTPDS activities (CPDs, extra hour, coaching, readers, and lesson plans) will be embraced and accepted by MoEST and other relevant stakeholders, and program sustainability will be bolstered if and when the following is done:

- In developing the new National Reading Strategy, the Ministry should identify appropriate ways to include MTPDS activities, as well as other programs such as Read Malawi. Regarding MTPDS, the Ministry needs to determine which project components should be accepted, modified, or rejected. For example, there seems to be substantial support for CPD training, yet the analyses of the EGRA results show no benefit from just CPD training (Level 1), yet substantial benefits to the more intensive intervention. Which of the components (e.g., scripted lessons, readers, extra hour, and coaching) should be accepted or modified?
- The Secretary for Education should appoint a key leader/directorate to provide leadership and support for the literacy program, of which MTPDS is an important part of the integrated curricula and CPD. For example, might DBE or DIAS take the lead role, or should there be a new coordinating position as with the PSIP? In addition, there is a need to define the roles and responsibilities for all of the key Ministry stakeholders regarding policies and activities surrounding literacy, including such issues as adding an extra hour to the timetable, lesson plans, budgeting, CPD, developing/integrating readers, participation of SMCs and PTAs, etc. One department should lead, but not usurp the roles and responsibilities of the other participating departments. Include a role for the Department of Education Planning to facilitate the transfer of the MTPDS activities into the Ministry (e.g., annual Programs of Work and budget) and to ensure that each department includes its activities and budgets accordingly.
- DBE should continue to support and encourage community participation through the SMCs and PTAs. The MoEST and communities must work together for CPD and related activities to be effective. It is important for the reading program to be supported by villagers/community if it is to be owned.

- MoEST should initiate dialogue on MTPDS activities in order to more fully inform key stakeholders about the most effective results and promise of MTPDS activities as lessons learned, especially the reading program, in order to strengthen the ownership and buy-in from the stakeholders. A key initial objective would be to lay out a strategy for institutionalizing the key elements of MTPDS. An EGRA conference could kick-start the dialogue.
- The Ministry should conduct periodic meetings, perhaps on a quarterly basis, on the reading initiatives with the key stakeholders. These meetings would disseminate the findings and status of the reading programs and seek feedback and suggestions for strengthening the programs. The proposed EGRA Committee should oversee these periodic meetings.
- DTED should include Literacy CPD training in the ESIP II, Programs of Work (PoW), and PSIP, and DEP should ensure that resources are allocated beginning with the 2013–2014 financial year. Likewise, DEMs should ensure that all public schools include CPD on their SIPs.

Evaluation Question 4: Sustainability of MTPDS

The prospects for MTPDS sustainability were determined, in part, by assessing the extent to which Ministry officials at the central and district levels, teachers, school administrators, and beneficiaries embraced and/or bought into the MTPDS approach. This was covered in the previous section on Ownership. In this section, we assess sustainability in terms of the specific components – Literacy CPD, readers, lesson plans, the extra hour, coaching, and Leadership CPD.

FINDINGS

The "findings" of the various MTPDS components – Literacy CPD, scripted lessons, readers, coaching, the extra hour, and the Leadership CPS – are covered in the section Lessons Learned above and, thus, will not be repeated here. This section will focus on sustainability issues of these components based on the "findings" and reflected in the Conclusions and Recommendations that follow.

CONCLUSIONS

Literacy CPD

Generally, the Evaluation Team concludes that sustainability of MTPDS among teachers, school administrators, and school-level stakeholders is significantly associated with access to and maintenance of relevant resources as well as a decrease of class sizes. Related to the findings in the Lesson Learned section, the team concludes the following:

- The teaching of literacy using a *syllabic approach* is likely to continue, as it does not require additional resources and the teachers have expressed strong support for it. The approach is also something that was familiar to teachers in the prior SOSA materials. Sustainability is especially feasible if the syllabic approach is reinforced by the new primary curriculum material.
- Continuous assessment is not likely to continue or increase in use by more teachers, without a reduction of class size or the creation of commonly understood or quickly administered tools.

- Without a reduction of class sizes, grouping is not effective, will continue to be poorly implemented, and will not likely continue being used or expand to other teachers/classrooms.
- Teachers will not be able to implement the print-rich classroom environments without additional resources and materials, and a way to protect them.
- Teachers' use of pedagogical skills/techniques shared through Literacy CPDs is likely to continue, as will the expansion to additional teachers and classrooms. These skills/techniques enable previously trained approaches to interactive learning and use of teaching aids, and they are well liked. Most importantly, these do not require significant resource outlays.

Nditha Kuwerenga Readers

With MTPDS drawing to a close, there is no current plan to print additional readers. Considering this, the Evaluation Team concludes the following:

- In classes where readers are extensively used, they are falling apart, making the prospects for sustainability nil.
- However, the concern for the sustained use of readers may not be urgent given that the MTPDS learning intervention was designed to provide teachers with a way to teach reading/literacy using the syllabic approach.
- With new primary curriculum materials on the way, there should be less need for supplemental materials (including the reader) to the national curriculum.

Lesson Plans

Given significant teacher buy-in, use of the MTPDS scripted lesson plans is likely to continue and is certainly sustainable as long as the scripted lesson plans are available. This is particularly so given the following additional conclusions:

- The scripted lesson plans are a powerful tool to enhance reading skills of learners, especially in a context where teachers are undertrained and there is a dearth of textbooks (and poor quality of the ones that exist). The lesson plans provided a bridge that allowed countless children to learn to read while MIE revised primary curriculum materials.
- As with CPD, the primary utility of this input was the provision of a syllabic approach to teaching Chichewa literacy.
- Lesson plans that rely on the reader will become less useful as the readers degrade in use or lose relevance with respect to the new curriculum.

Extra Hour

Given teacher and administration support for additional reading time-on-task, it is likely that extra hour will continue for the foreseeable future in literacy intervention districts.

However, full embrace or sustainability of the extra reading hour will be challenged when considering the following conclusions:

- If the MTPDS materials degrade, there may not be sufficient materials and resources for schools to justify an additional hour.
- If new head teachers are assigned to intervention schools, they may cut the extra time as they are unfamiliar with MTPDS rationale.
- In time, the extra reading hour may be reduced/ eliminated as MTPDS is phased out.

Coaching

Coaching will most likely remain a responsibility of PEAs, but they may not be able to sustain the current coaching model based on the following conclusions:

- PEAs will not be able to increase the number of visits beyond the current average of two per term because of their current list of responsibilities (SOW) and insufficient resources for fuel and repair of motorbikes.
- The lack of a focused cascade model (particularly the omission of head teachers from CPD trainings) severely limits the possibility of school-level support.

Leadership CPD

Based on the Evaluation Team's findings regarding the sustainability of Leadership CPDs at the three levels indicated above, the following conclusions can be drawn:

- School-based CPDs will likely continue because there is strong support for CPD at national and local levels.
- SMCs will probably remain marginally involved in school affairs. Their participation will continue to depend on the personalities involved in HT and SMC roles.
- The school Report Card will not be sustainable, in part because of its duplication with the similar EDSA Report Card and the likelihood that SMC participation in school academic activities, other than reading, will continue to be marginal.
- Community members and especially parents will continue to engage with the schools as they note the improved reading ability of their children.

RECOMMENDATIONS

Literacy CPD

Overall, the Evaluation Team recommends the following as complimentary to incorporating lessons learned in future MTPDS and curricular programming:

- Head teachers and, resources permitting, all teachers S1-S8 should be invited to the Literacy CPD training so that they are in a position to support the program at their school.

- The Ministry should develop easy-to-implement continuous assessments, especially for large classes. It is suggested that training on such assessments teach teachers how to assess collectively (alongside a head teacher, classroom aid, fellow teacher, etc.)

- Teachers and school administrators should be provided with further guidance on how to actively engage students during grouping exercises and when grouping is/isn't effective.

- Concerning the print-rich environment, teachers could be given further guidance on how to use more limited, but revolving print materials in a non-secure environment.

- Considering significant buy-in and use of supplementary pedagogical skills/techniques the Ministry should incorporate these approaches into the national curricula and the TTCs.

Nditha Kuwerenga Readers

- Given the prospect of the new national curriculum and efforts to implement cost-effective and sustainable strategies, the Evaluation Team recommends that new production of Nditha Kuwerenga readers be phased out.

- With the new primary and national curriculum, the Ministry should incorporate lessons learned from the use of the readers into the latest lesson plans for teachers and student book.

- Should the primary curriculum student book be insufficient, the Ministry could consider developing a new reader. The Ministry can provide training on the use of the readers through CPDs and TTCs.

Lesson Plans

The ultimate utility of the scripted lesson plan is limited because it doesn't harmonize with MoEST curriculum. Thus, the Evaluation Team recommends the following:

- The scripted format is a powerful tool and should be maintained, but current lesson plans should be revised to reflect the new primary curriculum and literacy materials.
- The revised lesson plans should be shortened and provided in a modular format with additional activities that can be used in more able classes (scaffold difficulty level).
- CPDs and TTCs should include training on the use of the revised lesson plans such that teachers learn how to personalize lessons during implementation as opposed to just reading the scripts.

Extra Hour

- In order to sustain and promote the full embrace of the extra reading hour among all schools, the Ministry should institutionalize the extra hour and provide guidance on how to implement and best utilize this additional time.

Coaching

Coaching would be greatly improved if:

- HTs and deputies of all schools in all districts were trained on how to conduct coaching and incorporating this into their official scopes of work. Coaching could then be provided at a higher frequency and lower cost than the MTPDS approach.
- The Ministry provides adequate resources (i.e. lighter work load/ additional personnel and transportation for visiting schools) needed by PEAs to continue coaching.

Leadership CPD

Sustainability would be enhanced if the following recommendations were adopted:

- School-based CPDs are funded at (1) local level through PSIPs and (2) national level.
- The Ministry should provide further guidance on the implementation of school-based CPD.
- The scopes of work for HTs should be reduced and/or reorganized if they are to take on responsibilities for coaching and organizing CPD training with their work/teaching load.
- Community participation, especially among parents, in monitoring reading should be re-emphasized through community outreach and through PTAs.

CONCLUDING STATEMENT

In conclusion, MTPDS is a valuable program. It has introduced several new and innovative ideas in Malawi that have resulted in significant gains in reading scores, especially in Level 2 and 3 districts. Teachers and community members have valued the program because they see that their children are learning how to read in a much earlier grade. The key remaining challenge is to develop and implement a transition plan that will ensure that the key components of MTPDS be integrated into the education system and the new National Reading Strategy.

ANNEX 1: MTPDS EVALUATION SCOPE OF WORK

** *Abridged*

SECTION C DESCRIPTION/SPECIFICATION/STATEMENT OF WORK
C.1 PURPOSE OF THE EVALUATION

The purpose of the evaluation is for the Contractor to evaluate the performance of the Malawi Teacher Professional Development Support (MTPDS) activity.

C.2 OBJECTIVES

The primary objective of the evaluation is to understand the effectiveness of the three levels of intensity in the teacher training model employed by MTPDS. The effectiveness relates to how results achieved are impacting learning outcomes, especially changes in literacy levels of pupils in impact areas as well as at the national level. A secondary objective is to assess progress by MTPDS in handing over the national CPD program to the MoEST and identify barriers inhibiting transition of the CPD models to the MoEST. Findings, conclusions, and recommendations will be used to inform future programming of USAID and form part of USAID/Malawi's learning process to better understand implementation fidelity and understand the feasibility and sustainability of the key components related to: an expanded school day; and additional learning materials, coaching, and in-service training that are anticipated to effectively increase early-grade literacy skills. The evaluation will be part of USAID's efforts to demonstrate value-for-dollar for U.S. investments in the Education sector in Malawi. USAID is keenly interested in understanding the relative cost-effectiveness within the levels of intensity of the literacy intervention at increasing literacy skills and in helping USAID/Malawi reach its targets to improve early-grade reading skills for primary students.

[…]

C.4 TASKS

The Contractor shall conduct an evaluation and analysis of the MTPDS activity to document actual/cumulative results by performing the following tasks:

I. Evaluation Questions

The Contractor must, at a minimum, address the following questions in the final evaluation report:

A. Program/ Activity Design

- How effective are the different levels of intensity of MTPDS' intervention in improving student reading skills?
- How effective are the different levels of intensity of MTPDS' intervention in improving classroom instruction skills for teachers?
- What is the relative cost-effectiveness of the different levels of intensity in relation to each other and student outcomes?
- Provide an external audit of the monitoring and evaluation data collected under the activity, particularly the internal impact evaluation that can verify or substantiate whether the data and results have minimized bias and conserved

reliability and validity. MTPDS will collect the final round of EGRA data and classroom observations in November 2012.

B. Lesson Learning

- What are the most important lessons learned from MTPDS implementation?

  ii. What are the strengths and weaknesses in the current implementation with three levels of intensity in regards to attaining desired outcomes?

C. Sustainability

- What progress has the activity made in developing sustainability and transferring responsibility and ownership for CPD to the MoEST and the respective directorates?

  i. How can the handover of MTPDS to MoEST be improved, such that program sustainability is strengthened?

  ii. To what extent have teachers embraced and or bought into the classroom practices and instructional tools delivered through CPD and the literacy intervention?

Upon completion of the evaluation and analysis the contractor shall submit to USAID/Malawi a media device that includes all instruments and data in formats suitable for reanalysis and submit the USAID accepted Final Report to the USAID Development Experience Clearinghouse (DEC). The primary audiences for the evaluation report shall be USAID, the Ministry of Education, Science and Technology, Development Partners working in the Malawi education sector, and interested stakeholders, including the implementing partners.

II. The Final Evaluation Report

The Contractor shall provide an electronic copy of the final report in both PDF and MS Word format. The contractor shall provide 50 hard copies to USAID/Malawi. Forty-five of these will be for dissemination at the Findings Workshop. The final report format will comply with the requirements set forth in the Agency's 2011 Evaluation Policy, and shall at a minimum include:

a. USAID branded cover page

b. Executive summary

   i. 3–5 pages summarizing key points, including activity purpose and background, key evaluation questions, methods, findings, conclusions, and recommendations.

c. Data Methods and Analysis

d. Findings, Conclusions and Recommendations

e. Appendices as appropriate

In the final Evaluation Report the Contractor must provide a full description of methodology (or methodologies) to answer each evaluation question. Evaluation methodology shall be explained in detail and all tools used in conducting the evaluation such as questionnaires, checklists and discussion guides will be required in an Annex in the final report. The final evaluation report is needed by USAID/Malawi no later than March 1, 2013, to meet internal reporting requirements. The final evaluation report must include the final scope of work and modifications approved by the Contracting Officer as an annex. All modifications to the scope of work, whether in technical requirements, evaluation questions, Evaluation Team composition, methodology or timeline need to be agreed upon in writing by the USAID Contracting Officer. Limitations to the evaluation shall be disclosed in the report, with particular attention to the limitations associated with the evaluation methodology (selection bias, recall bias, unobservable differences between comparison groups, etc.). Sufficient information must be provided so that a reader can make an informed judgment as to the reliability, validity and generalizability of the findings. Disclosure of conflict of interest: All Evaluation Team members will provide a signed statement attesting to a lack of conflict of interest, or describing an existing conflict of interest relative to the MTPDS activity being evaluated. If a difference arises in the interpretation of the results from the various stakeholders, the evaluation report will include a statement identifying any significant unresolved differences of opinion on the part of funders, implementers, and/or Contractor and its members of the Evaluation Team.

[…]

C.6 EVALUATION METHODOLOGY

I. Research Design

The Contractor must adequately answer the evaluation questions identified in section C.4. (I) "Evaluation Questions" with a detailed methodological approach using quantitative, qualitative, or mixed methods. The Contractor must use existing data to the greatest extent possible using impact evaluation methodology where appropriate. The Contractor shall use data from MTPDS-developed baselines and follow-up data, where applicable, to answer evaluation questions. Where existing data is insufficient, sample districts and schools (and their surrounding communities) shall be purposively selected for the study based on sampling methods that can draw conclusions about the relative effectiveness of the three levels of intensity employed in the teacher training models used by MTPDS. As part of the audit of data collected by MTPDS, Contractor must conduct spot checks of MTPDS data. MTPDS will be collecting endline EGRA data via tablet Kindle Fire devices during the November 2012 round of data collection and conducting enumerator training, field work, and classroom observations during the month of November 2012. This data will be provided to the Contractor. The Contractor shall limit the collection of additional primary data on student learning outcomes to the greatest extent possible.

II. Sampling Design and Geographic Coverage

The Contractor shall purposefully sample and spot-check MTPDS interventions. The sampling must include at a minimum two districts that only receive intensity level 1: national CPD; one district that will be entering its full second year of the level 2: intensive Standard 1 literacy intervention (Salima and/or Ntchisi), and one of the five intervention districts that received the intervention beginning in 2012 since the contract modification. At a minimum, the Contractor's sampling will include zones with the Level 2 and 3 districts that enable findings to distinguish differences between schools that receive coaching from MTPDS DLC (or District Teacher Training Coordinators (DTTCs)) in comparison to schools that receive coaching only from PEAS in the seven Standard 1 literacy intervention districts in comparison to those districts and zones that have only received the more limited level 1 national CPD in-service training. Sampling must also enable findings to identify any differences in benefits from the MTPDS activities by sex for students and for school teachers and administrators. It is expected that the Evaluation Team shall be both time- and cost-effective, and able to address the key questions and supporting questions identified in the Evaluation Questions section above. If additional data collection is required, it is anticipated that a maximum of ten schools per sampled district will be included in the sampling framework.

III. Research Methodology

The Contractor's methodology and approach for the evaluation for individual or group interviews or focus groups will be conducted as primary data source research as part of the spot check of MTPDS activities. The Contractor must incorporate the following stakeholders to the greatest extent possible:

a) USAID Mission staff,

b) Implementing partner staff,

c) Government of Malawi Officials (such as MoEST and relevant Directorates, such as the Department of Inspectorate and Advisory Services, the Department of Teacher Education and Development, the Department of Basic Education, among others; District Education Managers; Divisional Education Managers; and Primary Education Advisors),

d) Head Teachers and Standard 1–4 Teachers,

e) School staff, parents, and school management committee/parent teacher association (SMC/PTA) officials,

f) Standard 1–5 Students.

IV. Logistics

The academic year is from September until early July, with term breaks in December and April. MTPDS conducts CPD training during weekends and school holidays. The early standards attend school Monday to Friday from 7:30 to 10:30 am. In the seven intervention districts, the school day has been extended by one (1) hour for Standard 1 classes to accommodate the MTPDS literacy intervention. School, classroom, teacher and student observations must be conducted during the morning hours during the academic year. The

Contractor must align evaluation timeline with the academic calendar and must minimize scheduling field work during academic breaks when teachers and students will not be in schools. Teacher interactions will need to be arranged by the Contractor with the head teacher and classroom teachers to occur after classroom instruction to ensure that the evaluation does not impinge upon classroom instruction.

V. Findings: Empirical facts collected during the evaluation

The Contractor must present evaluation findings as analyzed facts, evidence, and data and not based on anecdotes, hearsay, or the compilation of people's opinions. Findings must have sufficient evidence and documentation that a reader of the findings can be confident that the findings are based on actual data. Evaluation findings must highlight any regional variations or discrepancies as well as identify outcomes or impacts that affect male and female teachers and students differently. Findings must be specific, concise, and supported by strong quantitative or qualitative evidence. The Contractor must identify sources of information/data and list in an annex.

VI. Conclusions: Interpretations and judgments based on the findings

The Contractor must present evaluation conclusions for each finding based on the evidence collected or analyzed by the Evaluation Team. Conclusions must logically follow from the gathered data and findings. Because conclusions involve interpretation of collected data, they must be explicitly justified. If and when necessary, the contractor must state his/her assumptions, judgments, and value premises so that readers can better understand and assess them.

VII. Recommendations: Proposed Actions for Management.

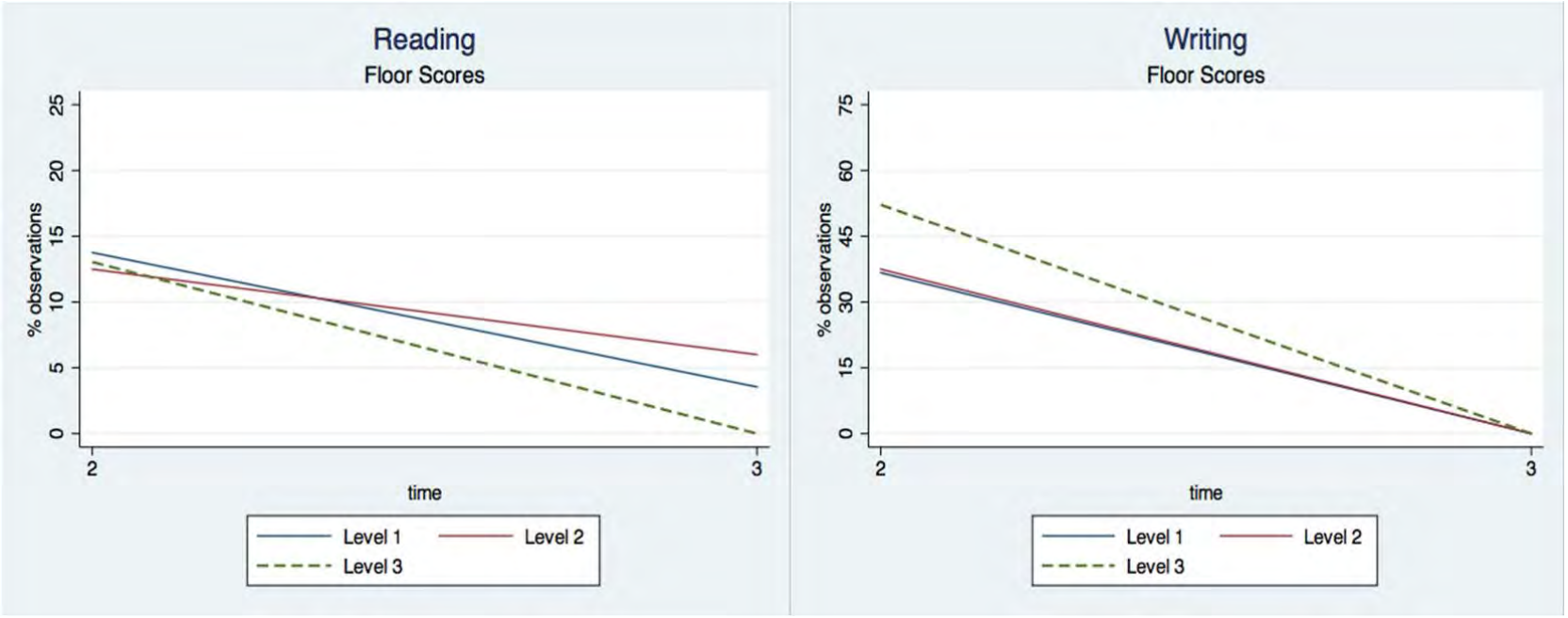
The Contractor must link recommendations to specific findings. Recommendations must be action-oriented, practical, and specific, with defined responsibility for the action. Contractors must take into consideration the economic and political context of the MTPDS activity, the strengths and weaknesses of the implementing partners, the MoEST, available resources, and the feasibility of change and innovation while framing recommendations.

ANNEX 2: SCHOOL SELECTION

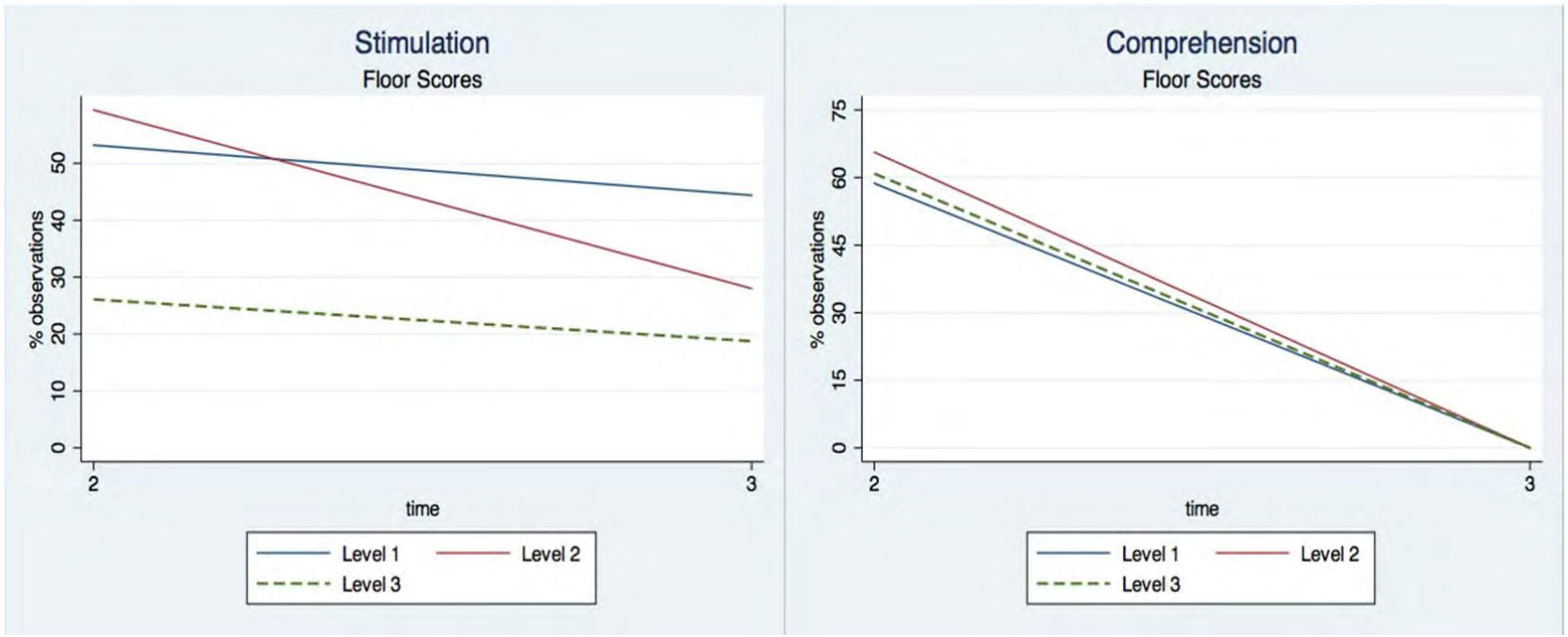
Level	District	School Name	Date of Visit	Enrollment	Number of Teachers	Other Reading Initiatives	Literacy Initiatives in SIP	Readers	Scripted Lessons
				TOTAL	TOTAL				
1	Dedza	Dedza Government Primary School	1/15/2013	2586	38	Yes	Yes	No	No
1	Dedza	Dzeza Primary School	1/16/2013	1337	16	Yes	No	Yes	Yes
1	Dedza	Mkomera	1/16/2013	2147	10	Yes	No	Yes	Yes
1	Dedza	Magunditsa Primary	1/17/2013	732	10	No	No	No	No
1	Nkhota Bay	Chihame II	1/21/2013	1017	9	Yes	Yes	No	No
1	Nkhota Bay	Chigawi	1/21/2013	371	6	No	No	No	No
1	Nkhota Bay	Mandezu	1/22/2013	279	5	No	Yes	No	No
1	Nkhota Bay	Chimbere	1/23/2013	496	5	No	No	Yes	No
2	Blantyre Rural	Ng'onga	1/22/2013	950	18	Yes	No	Yes	Yes
2	Blantyre Rural	Ngumbe	1/23/2013	2460	50	Yes	Yes	Yes	Yes
2	Mzimba	Engcongolweni	1/18/2013	646		No	No	Yes	Yes
2	Mzimba	Chilukwa	1/16/2013	192	5	No	No	Yes	Yes
2	Mzimba	Wantaya	1/17/2013	275	4	No	No	Yes	Yes
2	Ntchisi	Mingu FP	1/11/2013	604	8	Yes	Yes	Yes	Yes
2	Salima	Ngolowindo	1/10/2013	1579	19	Yes	Yes	Yes	Yes
3	Blantyre Rural	Mlambe II	1/22/2013	696	9	Yes	Yes	Yes	Yes
3	Blantyre Rural	Ntonda	1/23/2013	1904	33	Yes	Yes	Yes	No
3	Mzimba	Madise FP	1/16/2013	1292	14	No	No	Yes	Yes
3	Ntchisi	Chipwapwate	1/14/2013	870	16	Yes	No	Yes	Yes
3	Ntchisi	Zaoneka	1/10/2013	424	7	Yes	No	Yes	Yes
3	Ntchisi	Chorwe	1/14/2013		11	No	No	Yes	Yes
3	Salima	Chikombola	1/10/2013	338	9	No	No	Yes	Yes
3	Salima	Mgaga F.P.S	1/10/2013	1932	19	No	Yes	Yes	Yes
3	Salima	Senga Bay F.P. School	1/11/2013	1775	24	No	No	Yes	No

ANNEX 3: FLOOR SCORES

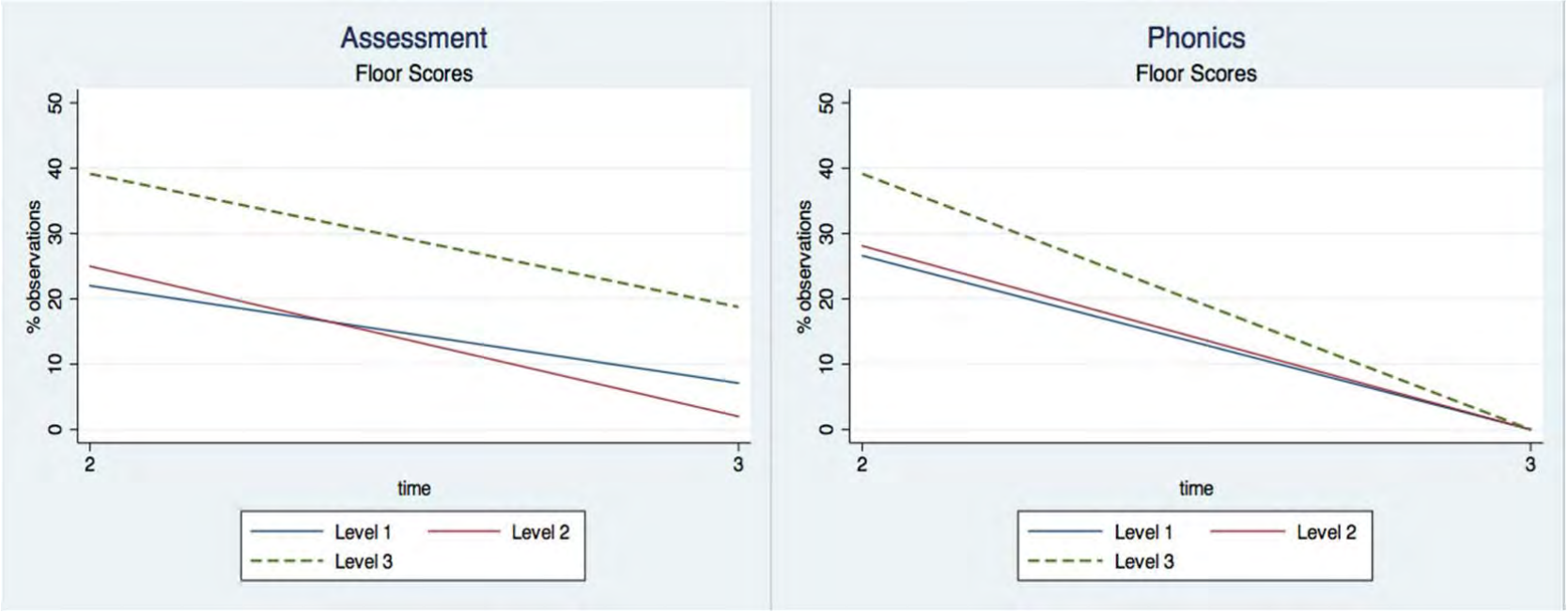
Annex 3.1 Reading and Writing Floor Scores



Annex 3.2 Stimulation and Comprehension Floor Scores

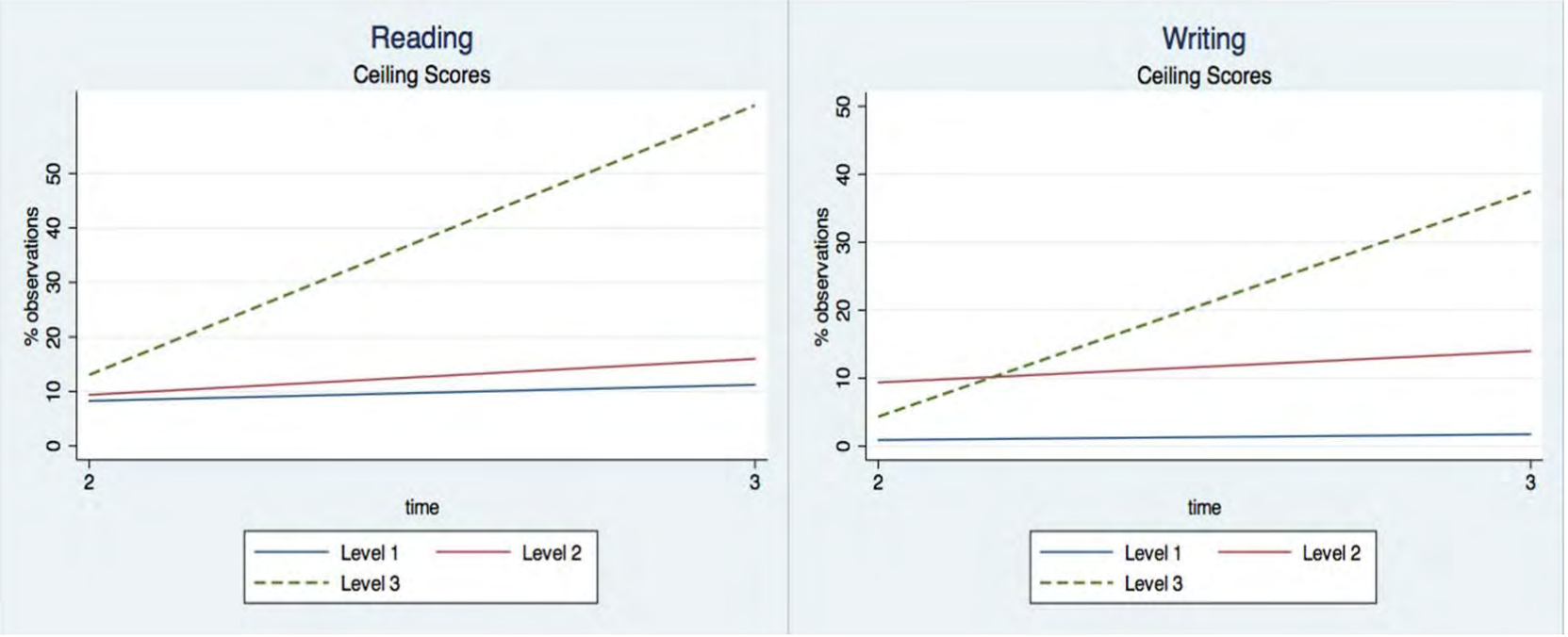


Annex 3.3 Assessment and Phonics Floor Scores

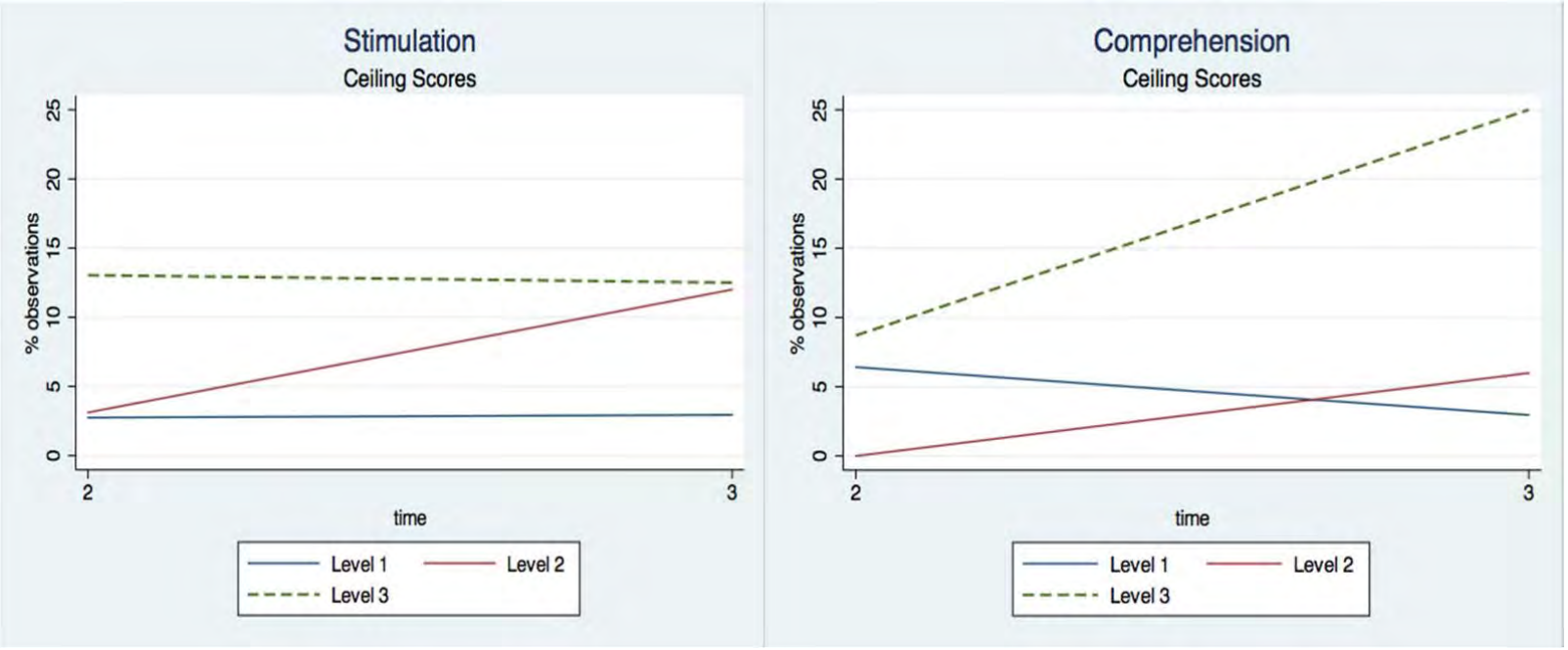


ANNEX 4: CEILING SCORES

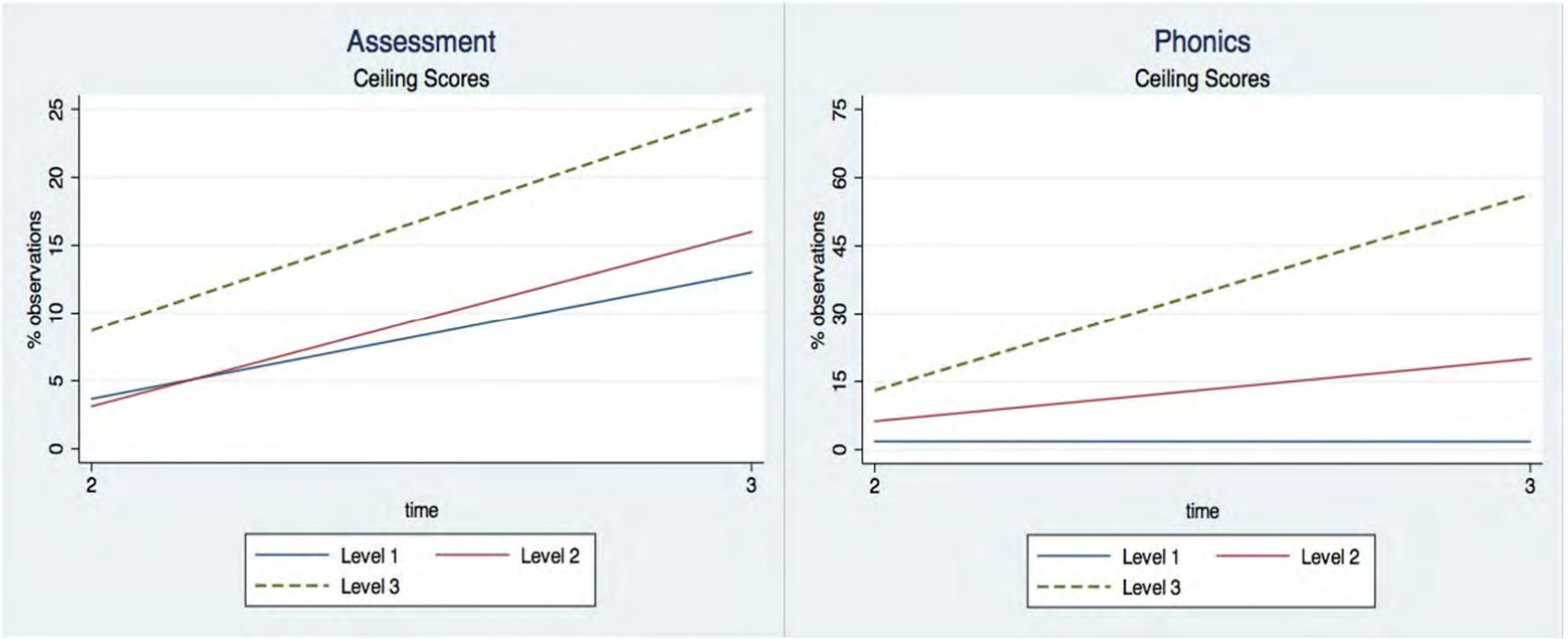
Annex 4.1 Reading and Writing Ceiling Scores



Annex 4.2 Stimulation and Comprehension Ceiling Scores

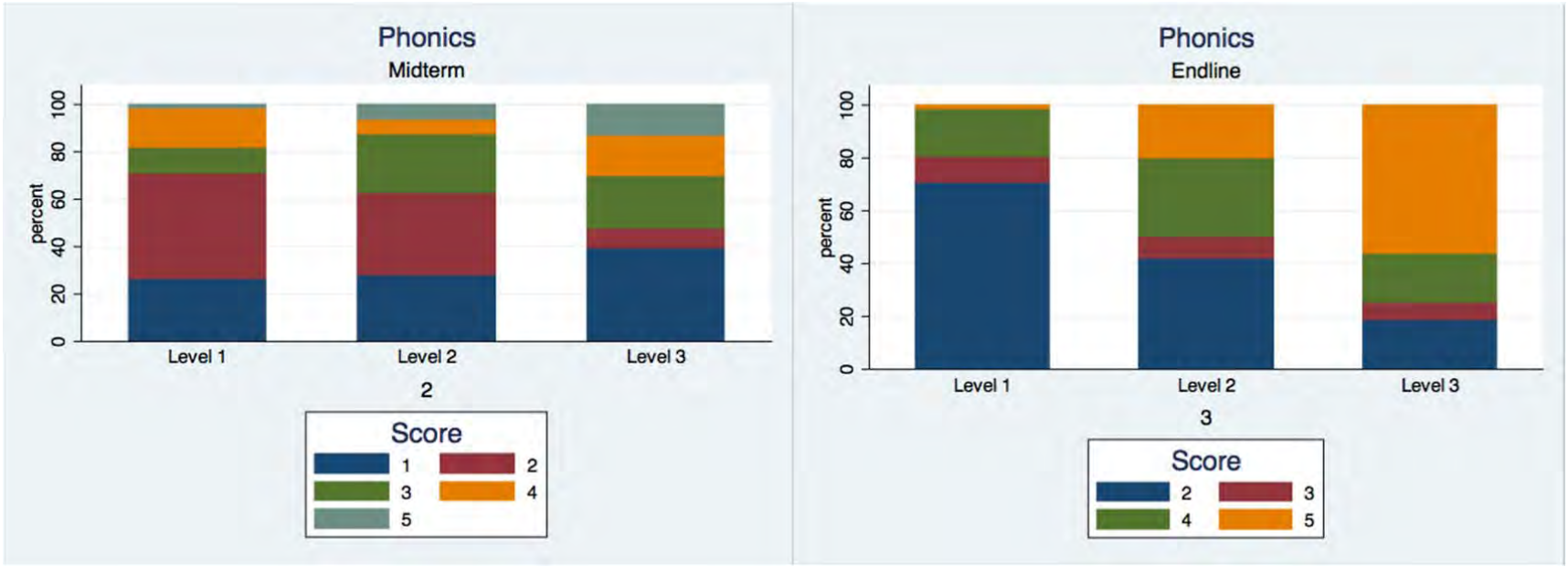


Annex 4.3 Assessment and Phonics Ceiling Scores

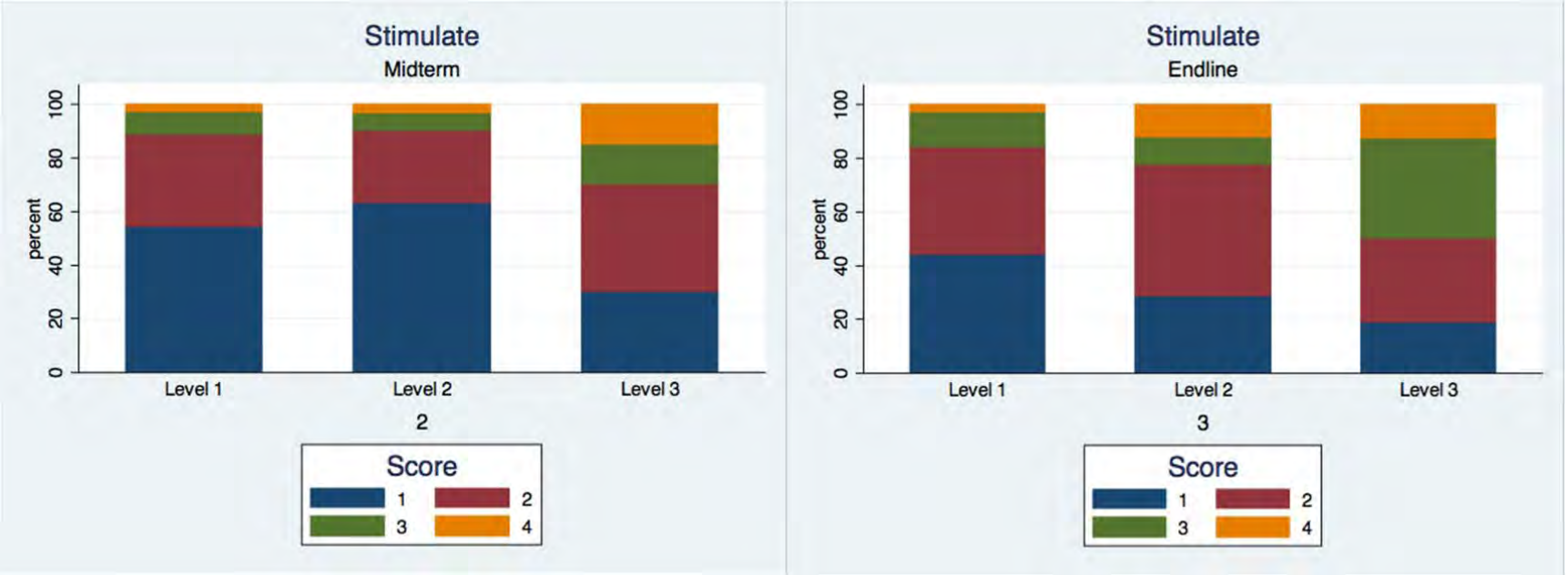


ANNEX 5: MIDTERM AND ENDLINE SCORES

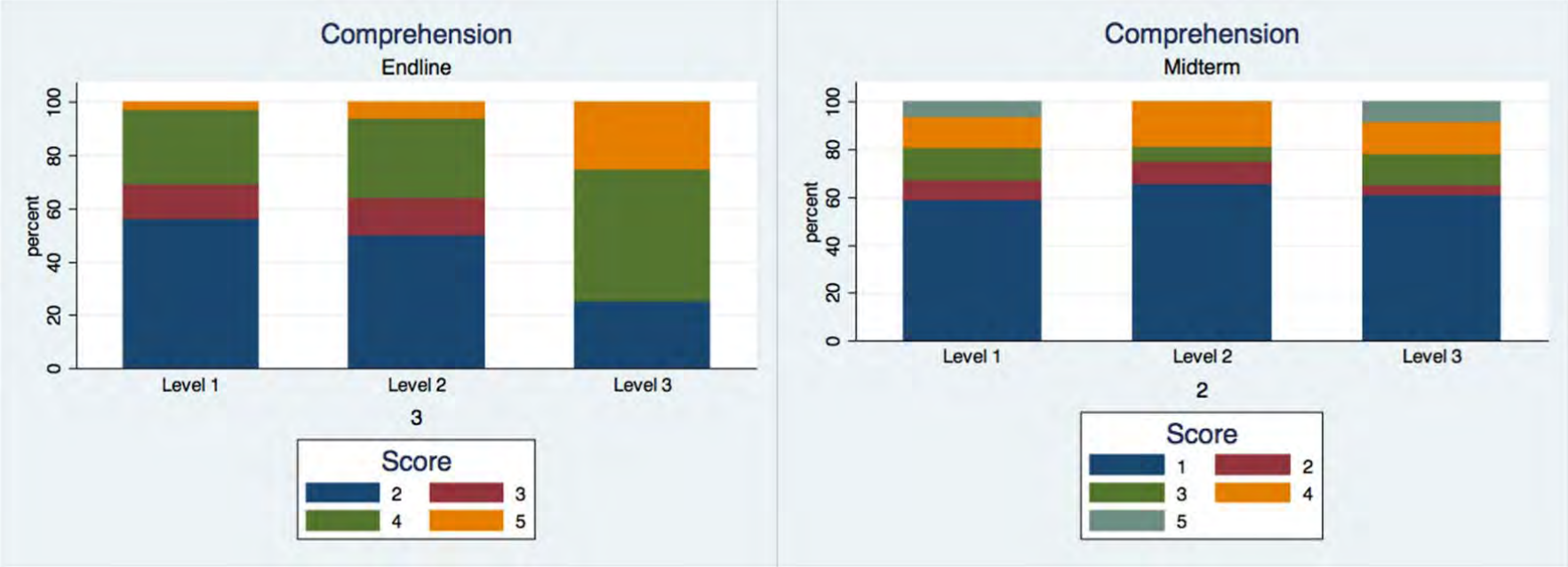
Annex 5.1 Phonics Midterm and Endline Scores



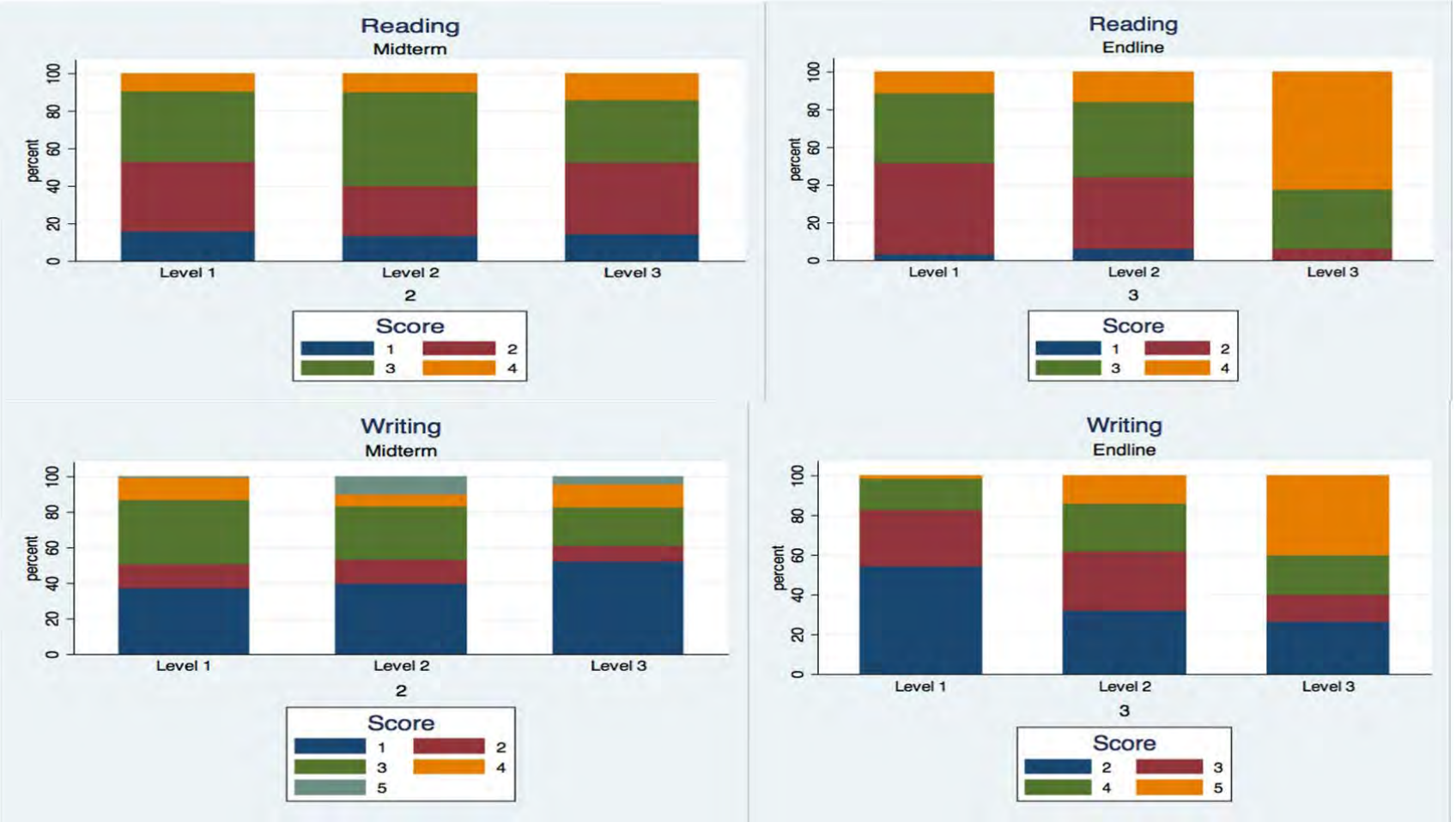
Annex 5.2 Stimulate Midterm and Endline Scores



Annex 5.3 Comprehension Midterm and Endline Scores



ANNEX 6: COMPARISON OF LONGITUDINAL READING AND WRITING SCORES ACROSS TREATMENT LEVELS



ANNEX 7: INDEPENDENT ANALYSIS OF PACKAGE COMPONENTS

OVERVIEW

The EGRA and classroom observation instruments were designed to shed light on package performance across two key outcomes. What they could not do, however, is to assess the relative merits of package components. While project staff spoke of proven synergies between package components, the lack of an analytical lens through which to test this assumption prompted the Evaluation Team to develop an instrument that solicited beneficiaries' opinions about relative merits of package components. Specifically, the teacher questionnaire was designed to measure teachers' use and perceived utility of the five programmatic inputs: CPD training, scripted lesson plans, *Nditha Kuwarenga readers*, the extra hour, and coaching support. In the course of site visits in six districts, the Evaluation Team administered the questionnaire to 104 primary school teachers (Standards 1–4). Two thirds of respondents taught Standards 1 or 2, and classroom sizes varied widely (ranging from 23 to 452, with an average of 113 students per class). Surveys were administered to participants of focus group discussions. To minimize response bias, respondents were asked to complete the forms independently before discussions began.

FINDINGS

Of the 104 surveys, the vast majority of respondents participated in at least one CPD training, and almost all strongly agreed that they would attend more trainings if they were offered. Furthermore, the majority of teachers who received training, regardless of treatment level, reported using what they learned every day. Respondents from the higher-intensity groups, however, were found to be 8 points more likely to use that knowledge on a daily basis.

Among teachers who participated in at least one CPD training, 69% strongly agreed and 29% agreed that they learned useful information: teachers from higher-intensity districts, on average, were more likely to report having learned useful information. All participating teachers agreed, and half strongly agreed, with the statement that they use what they learned in the classroom; however, consistent with the previous trends, teachers from higher-treatment districts tended to apply the knowledge more. Regardless of treatment intensity, all teachers agreed that CPD trainings had made them more effective teachers: two thirds strongly agreed. While almost all teachers thought both the scripted lesson plans and readers made them more effective, there was more support for lesson plans. Almost all teachers agreed that the extra hour and coaching support made them more effective teachers. While support for the notion that CPD training helped students acquire literacy skills was widespread, stronger of support was recorded in the high intensity districts.

Teachers consistently reported utilizing all programmatic components in their classrooms, with lesson plans as the most frequently used input (~90% use every day). Readers and training knowledge were used by more than three-quarters of teachers every day. Almost all respondents agreed that inputs made them more effective teachers, with support strongest for trainings. Coaching was the second highest rated, lesson plans and the extra hour were tied for the third most effective input, while readers rated last. The vast majority of respondents felt that the program inputs they received helped their students acquire literacy skills. The most support was registered for coaching. The other inputs, in order of support,

were training, lesson plans/extra hour (tied), and readers. For all inputs other than readers, the number of teachers who strongly agreed exceeded those who agreed.

CONCLUSIONS

These findings demonstrate strong support for and utilization of MTPDS components amongst primary school teachers, regardless of treatment intensity. The vast majority of teachers who received trainings found them useful and practical to their teaching, although significantly more teachers reported learning useful information than reported using the information in their classrooms. Support for the trainings was very high, with almost all respondents expressing interest in more training sessions.

From this sample of 104 teachers, lesson plans and coaching seem to be the most important self-reported programmatic elements. While training was not ranked as highly, survey data and qualitative findings suggest that it provides teachers with a better understanding of how to use other inputs more successfully. Overall, the teacher questionnaire exercise validated MTPDS' assumption of programmatic synergies and highlighted strong support for the program amongst beneficiaries.

RECOMMENDATIONS

MTPDS invested significant time and resources into the development and implementation of large-scale data collection. These efforts provided the program with powerful monitoring data and allowed for empirical assessments of performance vis-à-vis targets. Additionally, the information was intended to feed into rigorous, quantitative analyses of program impact. Unfortunately, multiple data quality issues call into question the validity of both the classroom observation and EGRA data. The independent DQA concluded that problems with EGRA were not of a sufficiently serious nature to undermine data utility. Classroom observation data, however, are compromised from a number of different factors and should not be used for purposes of estimating causal relationships between the MTPDS modalities. Investment in classroom observations could have been more useful had the program been consistent in their design and application. If similar activities are planned in the future, it is highly recommended that ample time be dedicated to instrument and data collection protocol design in the early stages of a program. Once finalized, it is imperative not to make substantive revisions to either during the period of implementation.

The utilization of MoEST personnel in the course of high-stakes data collection should be deliberated. Whether or not there were actual instances of misreporting in the field, the sheer existence of perverse incentives on the part of enumerators (PEAs reporting on performance of schools for which they are responsible) calls into question the validity of data. If impact evaluations are planned for the future, it is recommended that professional and impartial enumerators be used in lieu of MoEST personnel. This decision would, however, present a tradeoff. While data would likely be more sound, in light of resource constraints MoEST personnel would likely have to assume responsibility for EGRA if it were implemented in the future. This participation also has the benefit of capacitating the Ministry in data collection and statistical analysis. As such, if the primary purpose of future CPD interventions is not to attribute causality, it is recommended that MoEST personnel continue to engage in data collection but not enumerate in areas under their direct responsibility.

Last, if ascription of educational outcomes to USAID funding is desired in the future, it is imperative that a true control group be integrated into research designs. If withholding treatment is not feasible due to ethical, political, or logistical reasons, the Evaluation Team recommends phased implementation where successive waves of implementation facilitate a comparison between recipients and non-recipients while ultimately providing programmatic benefits to everyone.

While the Level 3 intervention consistently and substantively outperformed Level 1, the comparison does not warrant a clear-cut recommendation. The limited nature of the low-intensity treatment[49] should be kept in mind when reviewing the preceding analyses. Specifically, the limited impact of the MPTDS CPD should not be grounds for assuming that all CPD interventions are ineffective without supporting, costly activities. Regardless, given the strong performance, and to the extent resources permit, it is recommended that MoEST integrate the lessons learned from the high-intensity treatment into future curricula and policies. It is recommended that USAID and other donors review the new primary curriculum, and where gaps persist, work with MoEST to integrate lessons learned the MTPDS literacy intervention in the design of future literacy programming.

[49] Only three trainings were delivered in time for the endline, of which two were not focused on literacy.

ANNEX 8: MTPDS EVALUATION CLASSROOM OBSERVATION FORM

MTPDS EVALUATION CLASSROOM OBSERVATION FORM

Section 1 – School and Class information

Name of school: Standard:
Name of teacher: # of coaching sessions:
Number of children: Class arrangement:

Section 2 – Checklist of key indicators and summary descriptions

1. Text on walls: Y / N How many walls: 1 / 2 / 3 / 4 Used during lesson: Y / N
   - Description

2. Use of scripted lessons: Y / N
3. Use of MTPDS readers: Y / N
4. Use of continuous assessment: Y / N
   - Description

5. Use of teaching aids during the lesson: Y / N
   - Description

6. Use of grouping during the lesson: Y / N
   - Description

7. Use of any of the following specific reading skills
   a) Phonological awareness Y / N
      (*Ability to hear, recognize, and produce separate sounds in words*)
   b) Alphabetic Principle or Phonics Y / N
      (*Letters represent sounds/are printed symbols for sounds*)
   c) Fluency Y / N
      (*Ability to read smoothly, accurately, and with varied expression*)
   d) Vocabulary Y / N
      (*Word parts, context clues, and definitions*)
   e) Comprehension Y / N
      (*Strategies to extract and construct meaning from what is read*)
   f) Teaching of writing: Y / N

8. Do students practice reading?
   a. Syllables Y / N As a whole group Y / N; As individuals Y / N
   b. Words Y / N As a whole group Y / N; As individuals Y / N
   c. Sentences Y / N As a whole group Y / N; As individuals Y / N

9. Did the teacher include interactive activities (songs, games, role-play) Y / N
   - Description

Section 3 – Narrative description of the class

- The general structure and pedagogical approach and the extent to which the lesson was focused on clear learning activities.
- Details of special activities
- Overall impression

ANNEX 9: TEACHER QUESTIONNAIRE

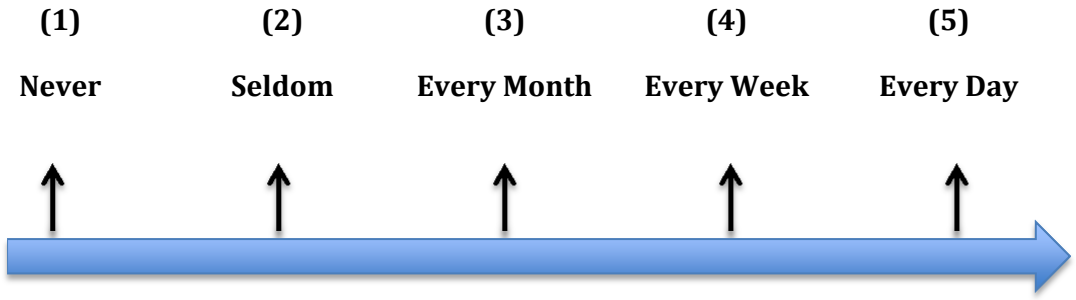
Level:

TEACHER QUESTIONNAIRE

District: _____ Sex: Male / Female
School: _____ Years teaching: _____
Standard: _____ Number of students: _____

1. Have you attended any literacy CPD trainings? YES / NO
2. How many trainings sessions have you attended? _____
3. Have you received scripted lesson plans from the training? YES / NO
4. Have you received a copy of the *Maziko a Kuwerenga Reader*? YES / NO
5. Have you used the optional extra hour during this school year? YES / NO
6. Have you received coaching from a PEA or MTPDS staff member? YES / NO
7. How many times have you received coaching? _____

For the next series of questions, please use the following five-point scale to respond to the questions:



1. How often do you use the lesson plans in your classroom? 1 / 2 / 3 / 4 / 5
2. How often do you use the reader in your classroom? 1 / 2 / 3 / 4 / 5
3. How often do you apply what you learned in your classroom? 1 / 2 / 3 / 4 / 5

For the next series of questions, please use the following five-point scale to respond to the statements:

(1)	(2)	(3)	(4)	(5)
Strongly Disagree	**Disagree**	**Neither Agree nor Disagree**	**Agree**	**Strongly Agree**

4. I learned useful information from the literacy CPD training. 1 / 2 / 3 / 4 / 5
5. I have used what I learned from the trainings in my class. 1 / 2 / 3 / 4 / 5
6. I would attend more trainings if they were offered. 1 / 2 / 3 / 4 / 5

TEACHER EFFECTIVENESS

7. The trainings have made me a more effective teacher. 1 / 2 / 3 / 4 / 5
8. The lesson plans have made me a more effective teacher. 1 / 2 / 3 / 4 / 5
9. The readers have made me a more effective teacher. 1 / 2 / 3 / 4 / 5
10. The extra hour has made me a more effective teacher. 1 / 2 / 3 / 4 / 5
11. The coaching has made me a more effective teacher. 1 / 2 / 3 / 4 / 5

LITERACY SKILLS

12. The trainings have helped my students acquire literacy skills. 1 / 2 / 3 / 4 / 5
13. The lesson plans have helped my students acquire literacy skills. 1 / 2 / 3 / 4 / 5
14. The readers have helped my students acquire literacy skills. 1 / 2 / 3 / 4 / 5
15. The extra hour has helped my students acquire literacy skills. 1 / 2 / 3 / 4 / 5
16. The coaching has helped my students acquire literacy skills. 1 / 2 / 3 / 4 / 5

U.S. Agency for International Development
USAID/Malawi
Education Office
NICO House, City Centre
P.O. Box 30455
Lilongwe 3
Malawi