Proficiency

Research Report

**Large-scale
Quantitative Study**

**English Proficiency Assessments of
Primary and Secondary Teachers and
Students Participating in English in
Action: Second Cohort (2013)**

ENGLISH in ACTION

English in Action (EIA) (2014) *English Proficiency Assessments of Primary and Secondary Teachers and Students Participating in English in Action: Second Cohort (2013)*. Dhaka, Bangladesh: English in Action.

Authors

Ian Eyres, Tom Power, Robert McCormick and Sonia Burton

For more information about English in Action, visit: http://www.eiabd.com/

Published by English in Action (EIA) in Dhaka, Bangladesh.

English Proficiency Assessments of Primary and Secondary Teachers and Students Participating in English in Action: Second Cohort (2013)

Executive summary

a) Background

The purpose of the study was to assess the student learning outcomes of English in Action's (EIA's) School Based Teacher Development programme, in terms of improved English language (EL) competence, against recognised international frameworks (specifically, the Graded Examinations in Spoken English (GESE) (Trinity College London [TCL] 2013), which map onto the Common European Framework of Reference for Languages (Trinity College London 2007)). Measurably improved student learning outcomes are the ultimate test of success of a teacher development programme. A secondary purpose of the study was to explore whether there was any related increase in teachers' EL competence.

English Proficiency Assessments 2013 is a repeat of the study on the pilot EIA programme (Cohort 1) (EIA 2012).

The students and teachers of Cohort 2 are sixfold greater in number (4,368 teachers, compared with 751 teachers, in schools). To enable this increase in scale, the programme has been delivered through a more decentralised model, with much less direct contact with English language teaching (ELT) experts, a greater embedding of expertise within teacher development materials (especially video), and a greater dependence upon localised peer support.

This report addresses two research questions:

1. To what extent do the teachers and students of Cohort 2 show improved post-intervention EL proficiencies, in speaking and listening, compared with the Cohort 1 2010 pre-intervention baseline?
2. To what extent has the programme been successful in repeating the 2011 post-intervention improvements in EL proficiencies seen in Cohort 1, at the much larger scale of Cohort 2?

b) Research methodology

The design of this study is similar to the study on the pilot intervention, in that a pre- and post-assessment was carried out using samples from the populations of those students and teachers participating in the EIA interventions; however, in this study, EIA's Cohort 1 pre-intervention was also used as the baseline for Cohort 2 (as indicated in Table 1, p.7). Cohorts 1 and 2 are not substantially different in terms of composition by urban-rural location, gender or phase, although they were carried out in different upazilas. There is also evidence that the EL competence of students across Bangladesh has not changed over the years (Hamid 2011). However, selection of Cohort 2 intervention upazilas (within the division) took account of the UNICEF composite deprivation index (CDI), to ensure that EIA specifically targeted areas of greater social deprivation. The sample reported in this study, intended to be representative of Cohort 2, was associated with a greater degree of deprivation, than that of Cohort 1 (see Appendix 6).

Assessments took place through one-to-one, face-to-face interviews, carried out by independent assessors from Trinity College London, which continued until the candidate was judged to have reached the peak of their EL competence, at which point a grade was assigned on the GESE scale. A total of 605 teachers (246 primary, 286 secondary and 73 primary head teachers) and 884 students (463 primary and 421 secondary) were assessed in the sample.

c) Key findings: Students

i) Primary students

Over two-thirds (69.8%) of the primary students tested achieved Grade 1 or above. The proportion of primary students failing the assessment outright was below a third (30.2%). The bulk of primary students in 2013 passed at Grade 1 (43.8%) and Grade 2 (18.4%). A small proportion (6.5%) of primary students attained Grade 3.

More female primary students (74.4%) achieved a pass grade than male primary students (65.1%), but 8.3% of boys achieved Grade 3 compared with 4.7% of girls, and 1.7% of boys achieved Grade 4, compared with just 0.4% of girls. All this indicates that girls are seen to be outperforming boys across the pass grades, which account for more than two-thirds of all students, while boys predominate at the top grades, which account for less than 10% of the student grade distribution (statistically significant; $p<0.05$). Rural primary students performed significantly better than semi-urban and urban primary students ($p<0.01$), with 80.9% achieving Grade 1 or above (as compared with 59.1% and 60%, respectively). There were striking differences in the performance of primary students by division ($p<0.01$): almost all students passed (i.e. scored a Grade 1 or above) in Rajshahi and Chittagong (98.8% and 93.3%, respectively), while in Rangpur and Khulna the proportion attaining a pass was below 30% (29.3% and 23.5%, respectively).

Compared with the 2010 baseline, 2013 post-intervention results show substantial improvement ($p<0.01$). 34.1% more primary students (almost twice as many) achieved Grade 1 or above (69.8% compared with 35.7%), whilst just over 20% more students achieved Grade 2 or above (26% compared with 5.7%). Students achieving Grade 2 increased fourfold from 4.6% (2010) to 18.4% (2013); and the percentage of students achieving the highest grades (3 and above) increased almost sevenfold, from 1.1% (2010), to 7.6% (2013).

Compared with Cohort 1 post-intervention results, 2013 post-intervention results were not just as good, but significantly better ($p<0.01$), with two-thirds (69.8%) of primary students passing the post-intervention assessment, compared with half (50.1%) in 2011. 4.4% more students achieved Grade 1 (39.4% in 2011), while 8.6% more achieved Grade 2 (9.8% in 2011), and whereas only 0.9% of primary students achieved Grade 3 or above in 2011, in 2013 these grades accounted for 7.6% of primary students.

ii) Secondary students

In 2013, over four-fifths (86.2%) of secondary students attained a pass grade (Grade 1 or higher), three-fifths (60.3%) attained Grade 2 or higher, while three-quarters (75.1%) achieved in the grade range 1–3. The highest grade, attained by 2.1% of students, was Grade 5.

There were no statistically significant gender differences in secondary student performance. There were no urban secondary schools in the sample, but students from semi-urban schools performed better ($p<0.05$) than rural students. Although pass rates were similar (85% rural, 89.7% semi-urban), more semi-urban students achieved the higher grades (76.6% at Grades 2–5, compared with 54.8%), with almost half the difference accounted for at Grade 3 (32.7% semi-urban; 18.5% rural).

Compared with the 2010 baseline, secondary students performed considerably better in 2013 ($p<0.01$). Almost 12% more passed (86.2% compared with 74.5%), while just over 14% more achieved Grade 2–5 (rising to 60.3% from 45.8%), and over 11% more achieved Grade 3–5 (rising from 21.7% to 33.2%).

Compared with Cohort 1 post-intervention results, despite the larger scale and more decentralised model, pass rates were similar (89.6% in 2011, 86.2% in 2013), but 10% more students achieved Grade 2 or above in 2011 (69.4%, compared with 60.3% in 2013; $p<0.01$).

d) Key findings: Teachers

i) Primary teachers

All primary teachers attained a pass grade (i.e. none were graded '0') and the great bulk of them (92.7%) attained Grade 2 or above. Grade 2 is the level of English required to teach the primary textbook (*English for Today*) for Class 3 (EIA 2010). Just over three-quarters (76.4%) of primary teachers attained within the grade range 2–4, and 16.2% achieved a Grade 5 or above.

There were no statistically significant gender differences for primary teachers (p=0.160). School location was significant (p<0.05), with rural teachers performing slightly less well, 43.1% achieving Grades 3–4, compared with 52.8% of urban teachers, and 62.9% of semi-urban teachers.

Compared with the 2010 baseline, there is a statistically significant difference (p<0.01), with those in 2013 performing better: 6.6% more primary teachers achieved Grade 2 and above (rising to 92.7 % from 86.1%), and 2.8% more achieving Grade 3 and above (rising from 62.6% to 65.4%).

Compared with the 2011 post-intervention results, the scores achieved in 2013 were not quite as high (in 2011, 4.3% more teachers achieved Grade 2 or above [97% compared with 92.7%] and 9.2% more teachers achieved Grade 4 and above [44.9% compared with 35.7%]). These modest differences were statistically significant (p<0.01).

ii) Secondary teachers

Almost all (99.7%) of the secondary teachers passed the assessment, with half (51.8%) achieving Grades 3–4, and 80.9% attaining Grade 3 or above. Grade 3 competence is necessary to teach the first secondary *English for Today* textbook (Class 6), while Grade 4 gives access to a significant proportion of the Class 7 textbook (EIA 2010). A substantial proportion (15.3%) attained at the highest grades (6–9). There were no statistically significant gender differences (p=0.253), and no statistically significant differences by school location (semi-urban vs. rural, p=0.56), for secondary teachers. (There were no urban secondary schools in the sample.)

Compared with the 2010 baseline, there were no statistically significant differences for the 2013 post-intervention results (p=0.475).

Compared with the 2011 post-intervention results, there were no statistically significant differences with the 2013 post-intervention results (p=0.238).

e) Conclusions

Despite a sixfold increase in scale between Cohort 1 and Cohort 2, and delivery through a more decentralised and peer-supported teacher development programme to teachers and students in upazilas selected for a high CDI, EIA has delivered improvements in student learning outcomes over the baseline study.

For primary students, these improvements are substantially greater than those achieved in the pilot. For secondary students, pass rates were similar to the pilot outcomes, but with slightly fewer students achieving the highest grades. Evidence indicates that girls have benefitted equally (secondary) or similarly (primary) to boys. At the secondary level, rural students performed less well than non-rural students, while at the primary level, they performed better.

Primary teachers improved over the baseline, though with slightly fewer achieving the higher grades, compared with the pilot intervention. There was no statistically significant difference for secondary teachers, with either the pre- or post- intervention scores of pilot teachers.

Table of contents

List of tables and figures

Acronyms

AT   assistant teacher
CEFR   Common European Framework of Reference [for Languages]
CM   cluster meeting
EIA   English in Action
EL   English language
ELT   English language teaching
GESE   Graded Examinations in Spoken English
HT   head teacher
CDI   composite deprivation index
TCL   Trinity College London

1. Introduction

The ultimate test of the success of the English in Action (EIA) schools intervention lies in the proficiency in English of those students participating in the primary and secondary programmes: their ability to communicate in English is expected to improve through the use of EIA methods and materials in the classroom. Further to this, it may be that teachers' English language (EL) proficiency improves as they become more effective in teaching English.

Previous studies of EL proficiency were conducted with the cohort of students and teachers taking part in the pilot EIA intervention (Cohort 1: 2010–11). These studies were designed to investigate the ability of students and teachers to speak and understand English. The 2010 pre-intervention assessment took place during the launch of the pilot programme (February–March 2010), while the 2011 post-intervention assessment was carried out on samples of the same student and teacher populations after taking part in the programme for 12 months (March and April 2011). Pre- and post-intervention assessment findings were published together (EIA 2012).

In keeping with the findings of EIA's earlier Baseline Study (EIA 2009a)[1], attainments of teachers and students in the 2010 pre-intervention study were low: many students failed to achieve any score against the Trinity Graded Examinations in Spoken English (GESE) scale employed, while the EL proficiency of many teachers was found to be at a lower level than the English those teachers were expected to teach. Student progress in the levels of English from one class (school year/grade) to the next (e.g. Class 1 to 2) was minimal. But the 2011 post-intervention assessment showed an improvement in EL proficiency by students and teachers, in both phases of schooling (primary and secondary), compared with the 2010 study – an improvement that was statistically significant in the case of primary and secondary students, and primary teachers (EIA 2012).

Following the pilot intervention, in 2012 the EIA programme up-scaled its implementation with a cohort of 4,368 teachers and an estimated 887,000 students (Cohort 2: 2012–13). To gauge the extent of EL proficiency improvements for this larger cohort, post-intervention assessments of EL proficiency were carried out after a year of participation in the programme (September–October 2013). Whereas the previous study (Cohort 1) had used pre- and post-intervention assessment from the same cohort, this study used a post-intervention assessment, comparing this to the 'pre-intervention' baseline (2010) established by the earlier study (reported in EIA 2012), after ensuring comparability in terms of the sampling of phase, urban-rural locations and administrative divisions. Table 1 shows the relationships between the three studies and the comparisons that will be made in this report.

Table 1: Pre- and post-intervention assessments on Cohort 1 and 2 EIA students and teachers

Pre-intervention baseline	Post-intervention	Comments
Cohort 1: Feb–March 2010	Cohort 1: March–April 2011	Samples from same cohort
Cohort 1: Feb–March 2010	Cohort 2: Sept–Oct 2013	Post-intervention assessment after 12 months' intervention, using previous cohort pre-intervention assessments as baseline

The justification for this comparison is discussed in Section 2, *Methodology.*

---

[1] This 'baseline' was not used as the pre-assessment for Cohort 1 as the sample was skewed as a result of social and political unrest at the time, restricting the sampling of teachers and students as representative of Bangladesh more generally.

As well as presenting the new (2013) findings, this current report makes comparisons between these findings and those of the two earlier studies (2010 and 2011). This study addresses two main questions:

1. To what extent do the teachers and students of Cohort 2 show improved post-intervention EL proficiencies, in speaking and listening, compared with the 2010 pre-intervention baseline?
2. To what extent has the programme been successful in repeating the post-intervention improvements in EL proficiencies seen in Cohort 1, at the much larger scale of Cohort 2?

While the students and teachers assessed in this current study (Cohort 2) have participated in essentially the same programme as those in 2011 (Cohort 1), they are much greater in number. To enable this increase in scale for Cohort 2, the programme has been delivered through a more decentralised model, with much less direct contact with national or international English language teaching (ELT) specialists, a greater embedding of expertise within teacher development materials (especially video), and a greater dependence upon localised peer support.

The first research question seeks to establish if EIA can improve the EL competence of students (and of teachers) over the 'baseline', with this increased scale and more indirect implementation. This is an essential step in moving from the pilot phase (Phase II) to the fully institutionalised phase (Phase IV) in 2014–17.

The second research question seeks to establish if the EL competence of students and teachers is comparable to that obtained in the pilot intervention, as this gives a guide to the likely success of a full scaling up of EIA in the final phase of the programme: in particular, whether the improvements in materials (for students, teachers and those who support them) and training of support staff, including the use of those staff in the Government system, can reproduce the same improvements. There is no reason to expect the 2013 improvement over the baseline to be better than that in 2011 (for Cohort 1), and there may well be reason to anticipate less substantial improvements through a larger scale, decentralised and essentially peer-supported model. Indeed, recent research has shown that success in a pilot phase where robust, randomised, control-trial evidence indicated an effective programme does not guarantee success when this is transferred to a government-implemented (i.e. institutionalised) programme (Bold et al. 2013).

It is important to see this particular study as part of a set, with two other studies where samples of teachers and students of Cohort 2 are investigated to determine both the classroom practices of teachers and the perceptions of learning English (EIA 2014a & b). In addition to separate reports on these two studies, there will be additional elements:

• An overview study of the three reports will be undertaken to see whether there are apparent connections among the findings based on what is known from the implementation.
• Within the samples of the three studies are overlapping subjects that will enable data on teachers' perceptions, classroom behaviour and EL competence to be related to students' perceptions and EL competence.

These two elements will be the subject of a further report (EIA 2014c). The aim of these elements is to provide evidence on the relationships between the important variables that are likely to lead to improvements in student EL competence.

2. Methodology

2.1 Study design

As indicated in the *Introduction,* the design of this study is similar to that of the pilot intervention (Cohort 1), in that a pre- and post-assessment was carried out using samples from the populations of those students and teachers participating in the EIA intervention. Cohort 1 is not substantially different in nature from Cohort 2 in terms of general key variables (location, gender, phase), although it was carried out in different upazilas within divisions[2]. Thus, the Cohort 1 pre-intervention was used as the baseline for Cohort 2 (as indicated in Table 1). A Comparison of the 2009 and 2010 pre-intervention situation (see *Introduction*), suggests it is unlikely that in the subsequent year the levels of EL competence have improved, and there is evidence that the quality of EL competence of students has not changed over the years (Hamid 2011)[3]; indeed, there is evidence that even after conventional interventions with teacher training in Bangladesh, there are no improvements in the classroom results (Rahman et al. 2006)[4].

The second research question implies a comparison of treatments, in that the results of one year of intervention for Cohort 1 (pilot) are compared with one year of intervention for Cohort 2 (an up-scaled and more indirect implementation of the pilot approach). The hypothesis in this case is that EIA can attain a similar level of student EL proficiency in Cohort 2 as it did in Cohort 1, but in more demanding conditions of implementation. Note that the imperative is for improvement in EL proficiency levels of students, as this is what EIA ultimately aims to achieve; any improvements in EL proficiency of teachers are seen as an additional benefit.

As noted in the *Introduction*, embedded in this pre- and post-intervention study design is a sub-study that enables the variables of student and teacher perceptions and teacher classroom behaviour to be linked to student and teacher EL competence by use of a common set of samples. This design is not described here (see EIA 2014c), though the effect on sampling will be described below.

2.2 Method: the English language proficiency assessment

In this study, both teachers and students underwent assessment by means of a test based on the Trinity College London (TCL) Graded Examinations in Spoken English (GESE). Assessments took the form of one-to-one, face-to-face oral interviews, carried out by an independent assessor. The assessment "replicates real-life exchanges in which the candidate and the examiner pass on information, share ideas and opinions and debate topical issues" (Trinity College London 2009: 6).

The assessment is conducted through an interview, the core of which is a conversation element. This is described as "a meaningful and authentic exchange of information, ideas and opinions, rather than a formal 'question and answer' interview." (Trinity College London 2009: 7). Discussion topics are selected for their potential to elicit the candidate's highest level of EL competence and offer a progression from the familiar to the less familiar and from the 'concrete' to the 'abstract'. Candidates are expected to take increasing responsibility for initiating and maintaining the conversation at each grade, and asking the examiner questions as they arise naturally out of the conversation[5].

---

[2] See the discussion in Section 2.6 *Limitations*.

[3] The the overall framework for ELT is not always supportive to effective classroom practice (Education Watch 2011; EIA 2009b; Hamid and Balfour 2008; Kraft et al. 2009; World Bank 2008).

[4] There is evidence that less than 50% of secondary teachers receive any kind of training (UNESCO 2012: 138), and what training is available to both primary and secondary teachers is weak and has had little effect in the past (Kraft et al. 2008: 8 & 14).

[5] The assessments differed from Trinity's standard procedure in that candidates were not asked to prepare a discussion topic (usual for assessments above Grade 3), but the procedure used nevertheless reproduces the same kind of assessment of communicative English.

The assessor seeks to elicit and facilitate communicative skills, language functions and language items relating to progressively higher grades, ending the interview when the candidate is judged to have reached the peak of his/her capacity. At this point the candidate is assigned a Trinity grade (1–12) and a letter grade (A, B or C) which subdivides each of the numerical grades. In this study, candidates assessed at Grade 0 (a 'fail' grade) were also subdivided into three categories – D1, D2 and D3 – representing levels of failure from 'borderline' (D1) to 'comprehensive' (D3). For the purposes of this report, these finer divisions are not used: only the numerical grades are given.

For teachers, the assessment usually lasts between 10 and 15 minutes; the assessment for students, in the main, takes less time, as they often achieve their peak communicative capacity earlier in the process. Six assessors (all native English speakers) carried out the assessments. They were selected and trained by TCL to ensure consistency and quality of assessments. Five had conducted EL assessments in Bangladesh previously. The assessors received an in-country briefing before embarking on the fieldwork to orient them to the task and the geographical areas in which they would assess.

These assessments were identical to those administered in 2010 and 2011 for Cohort 1. It is a valid and internationally recognised assessment of both EL competence (through its benchmarking to the Common European Framework of Reference (CEFR) for Languages, see Trinity College London 2007) and of the specific communicative approach to ELT promoted by EIA (through the use of the 'conversational' approach indicated above). The reliability of the assessment is ensured by the international experience and high levels of staff training and moderation (Trinity College London 2013).

2.3 Sample

2.3.1 Sample design

A total of 4,368 teachers and approximately 887,000 students participated in EIA's 2012 cohort (Cohort 2), made up of: 1,802 primary teachers (ATs), 900 primary head teachers (HTs) and 1,666 secondary teachers (ATs); and approximately 195,000 primary students and 692,000 secondary students.

A minimum sample size was determined through a power analysis, conducted to ensure the sample was sufficiently large to enable statistically valid comparisons between the 2010, 2011 and 2013 studies (see Appendix 1). The analysis established a minimum sample size of 577 teachers (268 primary; 309 secondary) and 543 students (265 primary; 278 secondary).

A 'core' sample was designed to cover 30 schools (15 primary; 15 secondary). The intention of this sample was to enable the links between the variables of student and teacher perceptions, teacher classroom behaviour and student and teacher competence to be discerned (as noted above). From each school, the plan was to assess:

- 30 students from one class (taught English by an EIA teacher), totalling 900 students;
- 2 EIA ATs, totalling 60 teachers;

Additionally in primary schools the HT (who participated in EIA, as teachers of English) were also tested; a total of 15 HTs.

This core sample of 30 schools was common across the parallel 2013 studies on classroom practice (EIA 2014a), and teacher and student perceptions (EIA 2014b), to enable a more complex analysis across studies to examine the interrelationships among these variables (EIA 2014c).

In addition to the sampling through field visits to schools, it was planned to include teachers attending EIA cluster meetings (CMs) in the same upazilas.

The total planned sample was therefore:

→   900 students (450 primary, 450 secondary);
→   645 teachers (234 primary ATs; 105 primary HTs; 306 secondary ATs).

To achieve this, a multi-layer stratified random sampling strategy was used. All upazilas in which EIA worked with for Cohort 2 were categorised as predominately rural or urban, by reviewing demographic information about where the teachers live – i.e. rural, urban or municipal area[6]. Upazilas, and subsequently schools within selected upazilas, were randomly selected from each division and reviewed to check that the sample reflected the rural/urban balance of the cohort as a whole. Similarly, the English students (of EIA teachers) to be assessed were also chosen at random from within a class, although wherever possible a whole class was assessed.

EL assessments were carried out by six assessors from TCL in ten upazilas over a period of two weeks during September and October 2013. Owing to difficulties in the field (including strikes by primary teachers, national hartals [political strikes] and blockades, and travel difficulties in remote areas), the actual selection was amended as necessary during fieldwork. The final sample achieved was:

→   884 students (463 primary, 421 secondary);
→   605 teachers (246 primary ATs; 73 primary HTs; 286 secondary ATs).

(Note, there were no urban secondary schools in the sample; secondary teachers were chosen from cluster meetings and some of these were from urban schools.)

2.3.2   Statistical comparisons of samples

The sample numbers of students and teachers in each of the EL competence studies vary (see Table 2). As noted above, a power analysis was used to ensure statistically significant comparisons between studies, overall and according to gender, school location (urban, semi-urban and rural), division, and phase (primary or secondary).

Table 2: Comparison of the sample sizes for the three studies (Cohort 1, 2010 & 2011, & Cohort 2, 2013)

Study	Cohort 1, 2010	Cohort 1, 2011	Cohort 2, 2013
Primary teachers	367	230	246
Head teachers	0	0	73
Secondary teachers	176	87	286
Primary students	4,630	785	463
Secondary students	2,609	317	421

---

[6] The 'municipal' and 'urban' categories were merged and classed as 'urban'.

2.4 Ethics

As part of normal ethical procedures adhered to by EIA, prior permission was obtained from the head teachers, the teachers and the students to undertake the research and include them in the sample. Each teacher was asked for his/her verbal consent to be involved in the study at the time of the assessment. All information within the EIA project is held under strict confidentiality and all teachers and students assessed (and their schools) are anonymous in any reporting.

2.5 Data entry, storage management and analysis

The data were entered by data entry operators into a Microsoft Excel spreadsheet from the paper instruments after the fieldwork. Two separate datasets – one of teacher data and one of student data – were then compiled. Random checks were carried out on the data to identify any miscoding and other errors.

Before the analysis was carried out, the data were cleaned to prepare them for analysis (see Appendix 2). Statistical comparisons were conducted through statistical methods such as cross tabulation and statistical significance tests. Results are reported with degrees of freedom and sample size in parentheses, the *p* value and the significance. (All tests of significance along with full data that support the figures used in this report are given in Appendices 3 & 4.)

2.6 Limitations

As noted above, fieldwork plans were disrupted by an uncertain political context in September and October in the run up to the general election, with a series of nationwide hartals and blockades. Further to this, a strike was held by primary teachers which meant primary schools were closed in some areas. These events had the effect of changing almost daily which schools were available for field visits. Flexible and responsive field management and coordination largely overcame these challenges. While the actual sample achieved was a little smaller than planned, it was sufficient to enable comparability; and school and cluster meeting selection remained both random and representative, with regard to the wider cohort under study.

Cohort 1 EIA schools, teachers and students were selected in order to be representative of Government schools across Bangladesh, in terms of national divisions and rural and urban locations. But to further strengthen the programme's social inclusion perspective, selection for the Cohort 2 upazilas (within the division) additionally took account of the UNICEF composite deprivation index (CDI) (UNICEF 2010), to ensure that EIA specifically targeted areas of greater social deprivation. The sample design for this study sought to obtain a representative sample of the Cohort 2 teachers, but the selection of the cohort was inevitably biased towards areas of greater deprivation, compared with the previous cohort and studies (see appendix 6). Therefore the sample reported in this study, intended to be representative of Cohort 2, is associated with a greater deprivation, based on CDI index, than the previous cohort and hence samples. While the use of CDI in the selection for Cohort 2 has meant that the distribution of EIA upazilas by CDI quintile, is now well aligned with the national pattern (appendix 6, figure 1), it also means that Cohort 2 had greater degree of deprivation by CDI than previous studies, which is reflected in the sample (appendix 6, figure 2).

Whilst we suspect it is possible, perhaps probable, that there may be a negative correlation between the CDI and EL proficiency, we have not yet established this quantitatively. Therefore, further investigation would be required, to explore the relationship between the composite deprivation index and EL proficiency. This change in composition, towards upazilas with greater deprivation, is likely to depress the EL achievement of EIA students and teachers in Cohort 2 rather than enhance it.

3. Findings

This section first examines the results for students and then those for teachers. Primary and secondary student and teacher data were analysed separately and in relation to basic demographic data of gender, school location (urban, semi-urban or rural) and division (administrative region).

3.1 Students

For primary, equal proportions of male and female students were assessed (n=463), while in secondary there were more female students (63.5%; n=419). Students were tested across six divisions: Chittagong (12.9%), Dhaka (16.5%), Rajshahi (17.2%), Khulna (14.6%), Sylhet (17.4%) and Rangpur (21.4%). The large majority of secondary students were from rural areas (74.6%), while primary students were mostly from both semi-urban areas (50.8%) and rural areas (42.3%). Only a minority of primary students (6.9%), and no secondary students, were from urban areas (see Table 3). (A full description of the student sample is provided in Appendix 5).

Table 3: Student distribution by school location

Phase	Student distribution		
	Rural	Semi-urban	Urban
Primary	42.3%	50.8%	6.9%
Secondary	74.6%	25.4%	0%

3.1.1 Primary students

2013 overall results

The overall performance of primary students in 2013 is shown in Figure 1.

Figure 1: Primary students' EL competence in 2013, by EL Trinity (GESE) grade
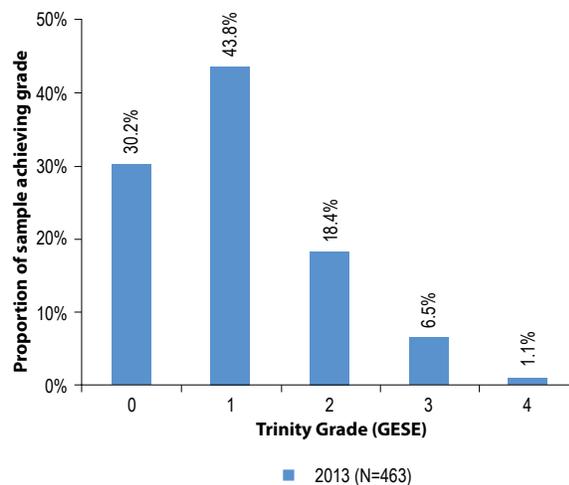
2013 (N=463)

Over two-thirds (69.8%) of the primary students tested achieved Grade 1 or above. The proportion of primary students failing the assessment outright was below a third (30.2%). The bulk of primary students in 2013 passed at Grade 1 (43.8%) and Grade 2 (18.4%). A small proportion (6.5%) of primary students attained Grade 3.

Gender

Analysis by gender (Figure 2) shows that more female primary students (74.4%) achieved a pass grade than male primary students (65.1%). Overall, the results show that girls performed better than boys (statistically significant; $p<0.05$), with them outnumbering boys at Grades 1 and 2: 46.6% of girls achieved a Grade 1, compared with 41% of boys; 22.6% of girls achieved a Grade 2, compared with 14% of boys. The reverse was the case at Grades 3 and 4, where boys are more strongly represented than girls: 8.3% of boys achieved Grade 3 compared with 4.7% of girls and the pattern is repeated at Grade 4 (1.7% compared with 0.4%). All this indicates that girls are seen to be outperforming boys across the pass grades, which account for more than two-thirds of all students, while boys predominate at the top grades, which account for less than 10% of the student grade distribution. This is different from the 2010 and 2011 data (pre- and post-intervention, Cohort 1), where there was no statistical difference between boys' and girls' performance.

Figure 2: Primary students' EL competence in 2013 by gender and EL Trinity (GESE) grade
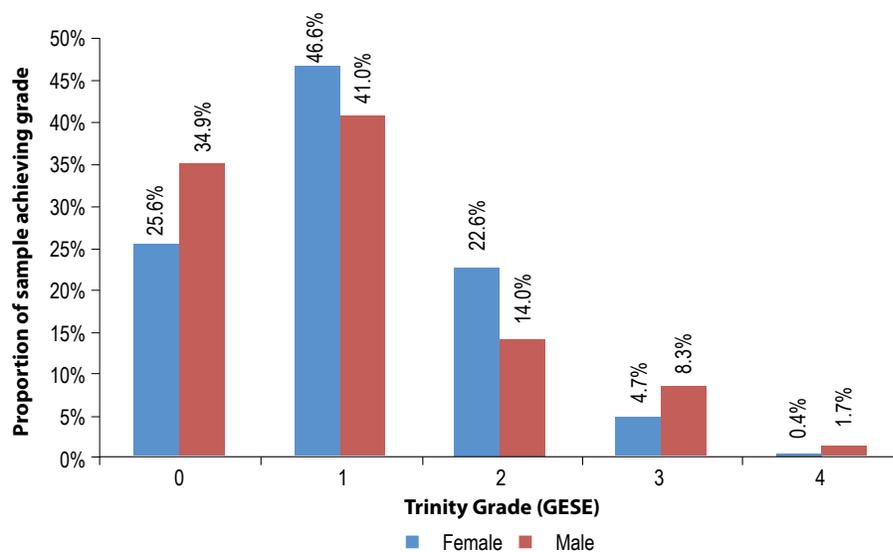
School location

Comparison according to school location (Figure 3) shows that, perhaps surprisingly, rural primary students performed statistically significantly better than semi-urban and urban primary students ($p<0.01$), with 80.9% achieving Grade 1 and above (as compared with 59.1% and 60%. respectively). This contrasts with 2011, where the statistical significance was in the reverse direction (i.e. rural students performed worse). The majority of rural primary students (57.9%) achieved Grade 1, a slightly smaller proportion achieved Grade 2, compared with urban and semi-urban counterparts. At higher grades, a slight advantage can be shown for students in semi-urban schools, where 30.9% attained Grade 2 or above and 10% attained Grade 3 or above. For rural students, 23% achieved a Grade 2 and above and 6.6% achieved a Grade 3 and above; for urban students the proportions are 23.3% and 3.3% respectively.

These shortcomings at higher levels notwithstanding, rural students, who might be expected to do less well than their urban and semi-urban counterparts, performed strongly in most grades and achieved the highest proportion of pass grades. (This will be returned to in Section 4, *Discussion*.)

Figure 3: Primary students' EL competence in 2013, by school location and EL Trinity (GESE) grade
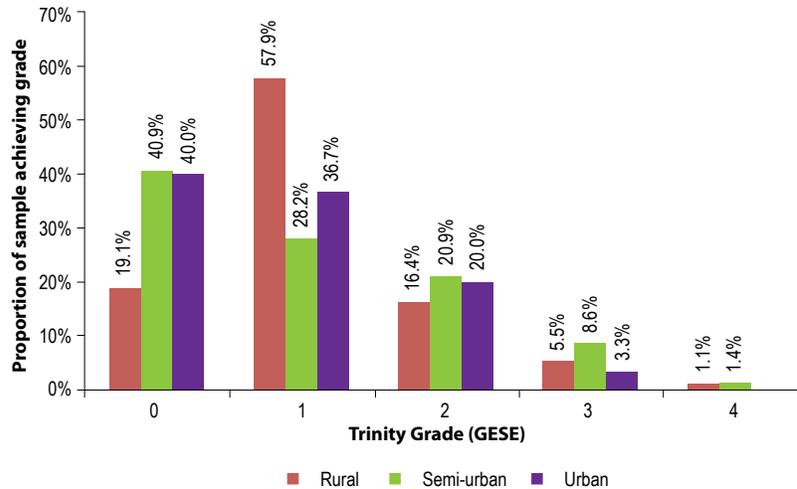
Division

There were striking and statistically significant ($p<0.01$) differences in the performance of primary students between the six divisions of the country in which data were collected (Figure 4), with Rangpur and Khulna students performing least well. This reflected the situation in 2011, which showed there was a statistically significant difference among districts.

Figure 4: Primary students' EL competence in 2013, by division and EL Trinity (GESE) grade
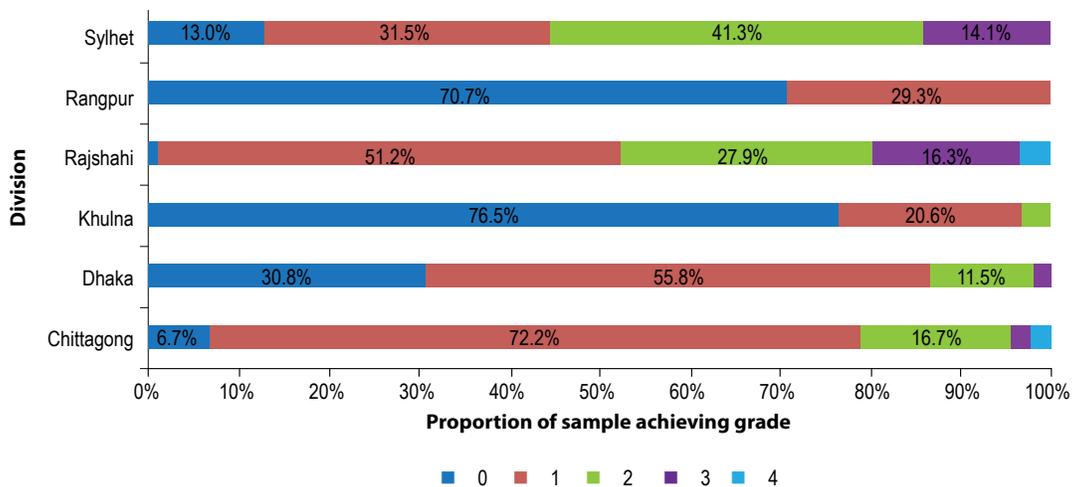
In Rajshahi and Chittagong, almost all students passed, i.e. scored a Grade 1 or above (98.8% and 93.3% respectively), while in Rangpur and Khulna the proportion of students attaining a pass was below 30% (29.3% and 23.5% respectively). In Rajshahi and Sylhet, around half the students (47.7% and 55.4% respectively) attained Grade 2 and above. In four divisions – Rangpur, Khulna, Dhaka and Chittagong – more than 90% attained no better than Grade 2; in fact, in Khulna, only 2.9% of students attained higher than Grade 1 and in Rangpur no students at all attained higher than that. At the other end of the spectrum, 55.4% of primary students in Sylhet and 47.7% in Rajshahi attained Grade 2 or higher. Some students in Rajshahi and Chittagong managed to attain Grade 4.

Addressing the research questions: Primary students

As explained in Section 1, in earlier reports (EIA 2012), pre-intervention assessment data collected from Cohort 1 (2010) was taken as a baseline against which post-intervention data (2011) for the same cohort was compared. This current study compares post-intervention data from Cohort 2 (2013), with:

a)   the pre-intervention baseline from Cohort 1 (2010), to examine the extent to which teachers and students show improvements over baseline;

b)   Cohort 1 post-intervention data (2011), where appropriate, to evaluate to what extent the large-scale programme has been successful in repeating the improvements seen in the pilot study.

Figure 5: Primary students' EL competence in 2010, 2011 and 2013, by EL Trinity (GESE) grade
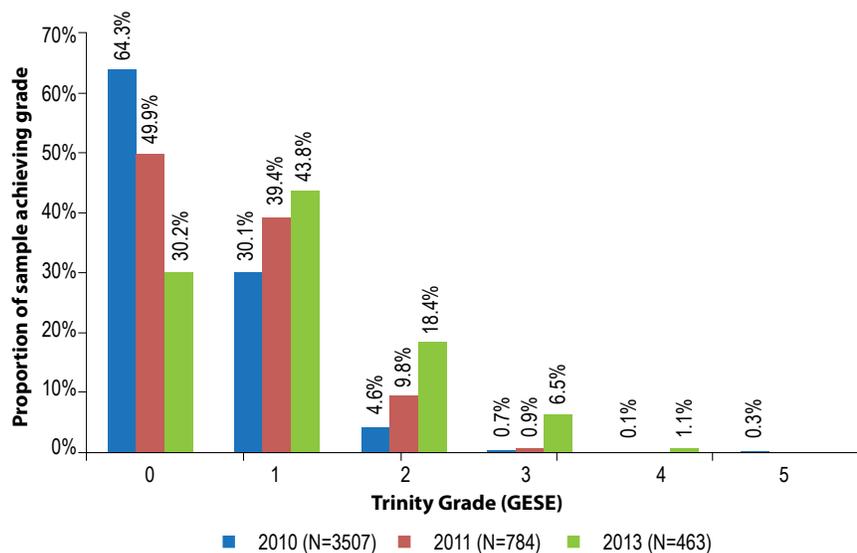
a) Comparison with pre-intervention (2010) baseline

As anticipated, 2013 post-intervention results show substantial improvement over those of the 2010 baseline, and the difference is statistically significant ($p<0.01$). Just over two-thirds (69.8%) of primary students passed the assessment in 2013, whereas only a little over one-third passed in 2010 (35.7%). The bulk of students in 2013 passed at Grade 1 (43.8%) and Grade 2 (18.4%). 13.7% more students achieved Grade 1 in 2013 than in 2010, and 13.8% more primary students achieved a Grade 2 than in 2010, increasing greatly, from 4.6%, to 18.4% (2013). Another interesting development is the increase in the proportion attaining the higher grades of 3 and above. In 2010 this amounted to 1.1%, while in 2013 these grades accounted for 7.6% of primary students.

b) Comparison with post-intervention (2011) assessment from pilot

Two-thirds (69.8%) of primary students passing the post-intervention assessment in 2013 represents a statistically significant increase ($p<0.01$) over post-intervention assessments in 2011 (50.1%). 4.4% more students achieved Grade 1 than in 2011 (39.4%), while 8.6% more achieved Grade 2 compared with 2011 (9.8%). Whereas, in 2011, only 0.9% of primary students achieved Grade 3 or above, in 2013 these grades accounted for 7.6% of primary students.

3.1.2 Secondary students

2013 overall results

The overall results for secondary students in 2013 are shown in Figure 6.

Figure 6: Secondary students' EL competence in 2013, by EL Trinity (GESE) grade
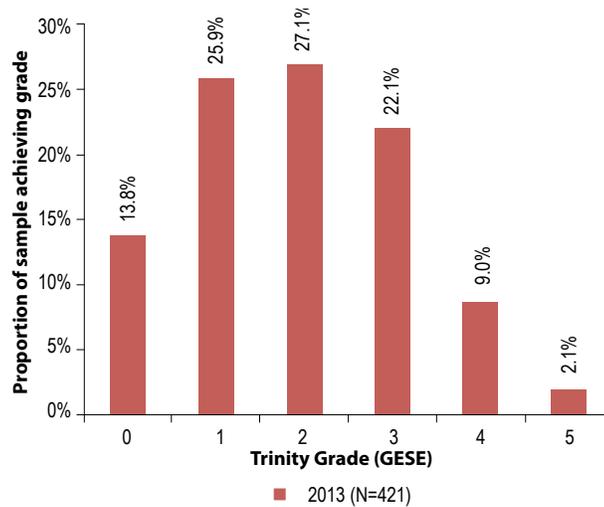
In 2013, over four-fifths (86.2%) of secondary students attained a pass grade (Grade 1 or higher), three-fifths (60.3%) attained Grade 2 or higher, while three quarters (75.1%) achieved in the grade range 1–3. The highest grade attained, by 2.1% of students, was Grade 5.

Gender

Analysis by gender (Figure 7) shows that, in many respects, male and female secondary students scored similarly: differences at each specific grade were no greater than 2.5% at Grades 0, 1, 4 and 5 (there is not statistical difference between the two groups – a similar result to that in 2011). Taking all those achieving a pass grade together, there is again a close similarity between female and male secondary students: 85.3% female, compared with 87.6% male.

Figure 7: Secondary students' EL competence in 2013, by gender and EL Trinity (GESE) grade
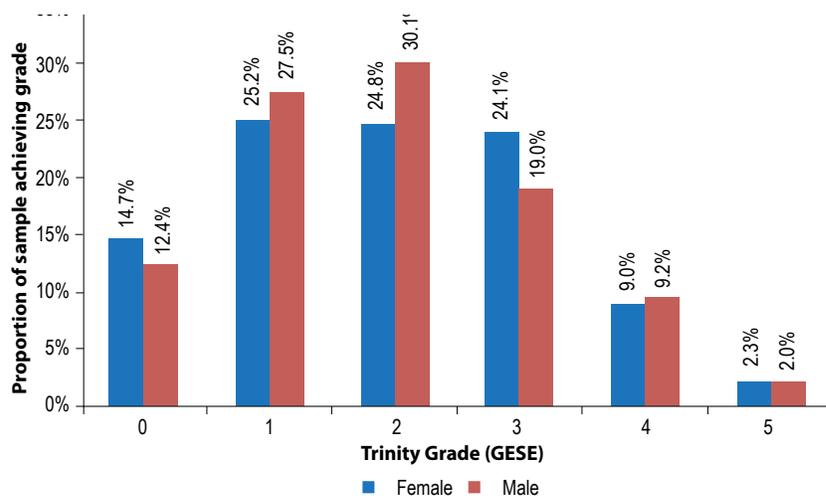
School location

There were no urban schools in the secondary student sample. The proportion of students passing the assessment (Grade 1 and above) were quite similar for rural and semi-urban areas: 85% of rural secondary students attained Grade 1 or above, compared with 89.7% of semi-urban secondary students. However, semi-urban students outperformed their rural counterparts in the higher grades (2–5, see Figure 8), and overall there is a statistically significant difference in performance according to location, with semi-urban students doing better ($p<0.05$). Again, this was a finding in 2011 (though with higher confidence level, $p<0.01$). This was most marked at Grade 3, which was achieved by 32.7% of semi-urban students, but only 18.5% of rural students (a difference of 14.2%). Correspondingly, rural students were much more prevalent at the lowest pass grade, with 30.3% of rural students assessed at Grade 1, as opposed to just 13.1% of semi-urban students.

Figure 8: Secondary students' EL competence in 2013, by school location and EL Trinity (GESE) grade
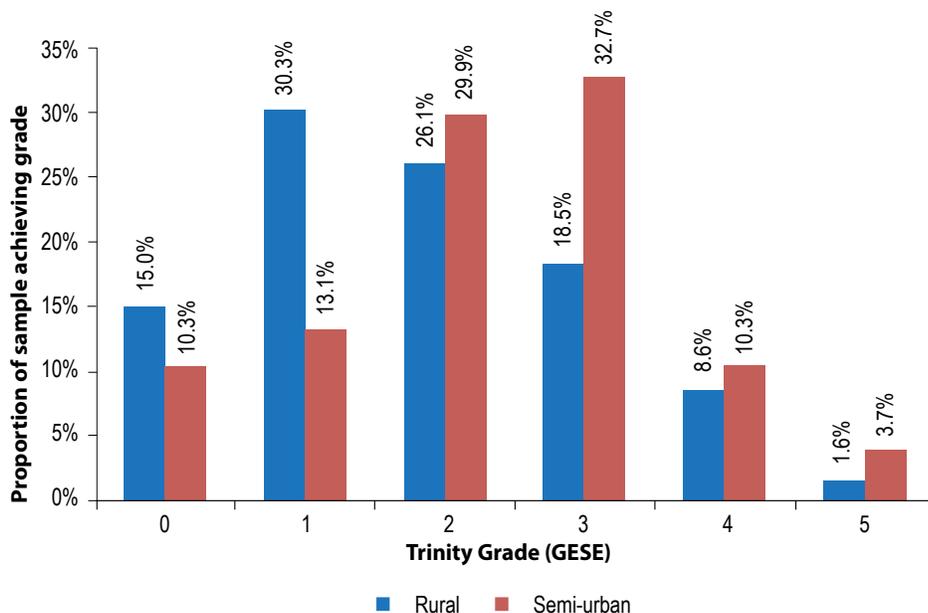
Division

As with primary students, distribution of grades varies considerably between divisions (Figure 9), and there is a statistically significant difference ($p<0.01$; as was the case in 2011). However, the picture is not the same as for primary students. In Sylhet, just over half (51.6%) passed, and the highest grade achieved was Grade 3. In Rajshahi, all students passed and a good proportion (6.1%) attained Grade 5. Rajshahi also had the highest proportion of students attaining Grade 2 and above (92.5%), compared with 20.9% in Sylhet. Comparing the primary and secondary student differences, it is evident that Rajshahi and Chittagong do relatively well in both sectors.

Figure 9: Secondary students' EL competence in 2013, by division and EL Trinity (GESE) grade
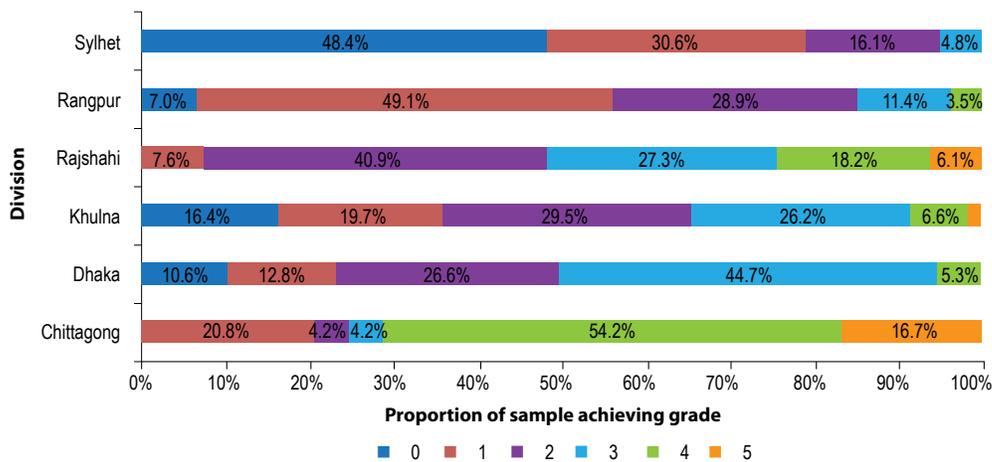
Addressing the research questions: Secondary students

Figure 10 shows secondary students' EL competence for the three studies.

Figure 10: Secondary students' EL competence in 2010, 2011 and 2013, by EL Trinity (GESE) grade
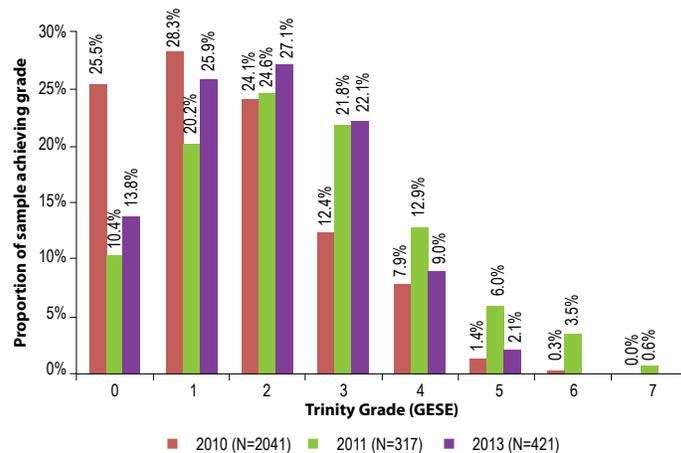
a) Comparison with pre-intervention (2010) baseline from pilot (Cohort 1)

Compared with the 2010 baseline, secondary students performed considerably better in 2013, with the difference being statistically significant ($p<0.01$). In 2013, 86.2% attained Grade 1 or above, which is considerably better than the 2010 figure (74.5%).

Compared with the 2010 baseline, a higher proportion of 2013 secondary students attained Grades 2–5, and these accounted for a substantial majority (60.3%, as opposed to just 45.8%, in 2010). Whereas just over a quarter (25.5%) of the baseline students failed to achieve any grade in 2010, only 13.8% failed to achieve Grade 1 or above, post-intervention, in 2013.

b) Comparison with post-intervention (2011) assessment from pilot

Results for 2011 and 2013 were broadly similar, with a modest but statistically significant difference in the distribution ($p<0.01$), with students in 2011 performing overall better. In 2013, 86.2% attained Grade 1 or above, which is comparable to the figure for 2011 (89.6%). However, approximately 10% more 2013 students attained the lower grades (1–3) than their 2011 counterparts (75.1%, as opposed to 66.6%).

3.2 Teachers

Of the 605 teachers assessed in 2013, 246 were primary teachers, 73 were primary HTs and 286 were secondary teachers. Overall, the proportion of male teachers tested was higher than that of female teachers (58.3%; n=605). Of the primary teachers, the majority of the sample – three-fifths (59.3%) – was female; while for secondary, the large majority (77.6%) were male. Proportionally this is a similar to the 2010 sample, and this reflects the gender proportions found in the Bangladeshi primary and secondary teacher population more generally[7]. For primary HTs, just under three-fifths (57.5%) were male, while 42.5% were female.

In terms of school location, most teachers tested were located in rural areas (58.6%), and smaller proportions were located in urban (9.5%) and semi-urban (31.9%) areas (n=570). For primary teachers, three-fifths (61.9%) were from rural areas, 15.2% were from urban areas and 22.9% were from semi-urban areas (n=236). The secondary teacher sample was almost entirely made up of rural (55.7%) and semi-urban (43.6%) teachers (n=264)[8].

As in the case of students, teachers were tested across six divisions: Chittagong (14.5%), Dhaka (18.8%), Rajshahi (18.3%), Khulna (12.9%), Sylhet (17.5%) and Rangpur (17.9%).

For Cohort 2, the primary and secondary interventions included specific programmes for head teachers (HTs). The role of HTs differs between sectors. Unlike their secondary counterparts, some primary HTs teach English regularly to classes in the same way as assistant teachers (ATs) do, and for this reason a sample of primary HTs was assessed. This is the first time HTs have been included in the teacher assessments, so comparison with earlier years is not possible. The relatively small size of the HT sample also means that analysis according to demographic factors was not appropriate.

3.2.1 Primary teachers

This section deals with the attainment of primary teachers excluding data on HTs.

2013 overall results

All primary teachers attained a pass grade (Figure 11) (i.e. none were graded '0') and the great bulk of them (92.7%) attained Grade 2 or above. Grade 2 is the level of English required to teach the primary school textbook (*English for Today*) for Class 3 (EIA 2010). Just over three-quarters (76.4%) of primary teachers attained within the grade range 2–4, and 16.2% achieved a Grade 5 or above.

---

[7] Official statistics for 2012 show 63% of primary teachers and 24.4% of secondary teachers were female (BANBEIS 2014a & b)

[8] Note that although there were no urban secondary schools visited for EL assessments, teachers were also tested at cluster meetings and this is where urban teachers were chosen as part of the sample.

Figure 11: Primary teachers' EL competence in 2013, by EL Trinity (GESE) grade

Figure 11 bar chart values:
Grade 1: 7.3%, Grade 2: 27.2%, Grade 3: 29.7%, Grade 4: 19.5%, Grade 5: 4.9%, Grade 6: 8.9%, Grade 7: 2.0%, Grade 8: 0.4%

Y-axis: Proportion of sample achieving grade
X-axis: Trinity Grade (GESE)
Legend: 2013 (N=246)
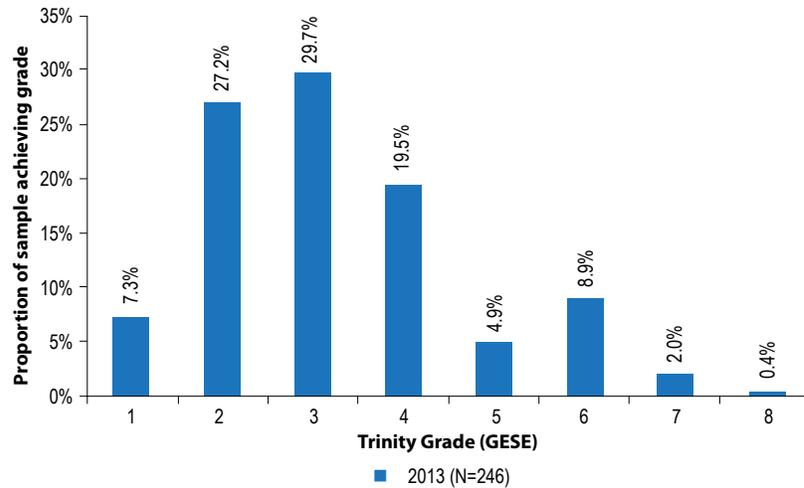
Gender

In terms of gender (Figure 12), there is no statistically significant difference for male or female primary teachers (*p*=0.160). Female primary teachers were better represented in Grades 2 and 3 (30.8% and 32.2% respectively, compared with 22% and 26% for males), and more men attained the lower Grade 1 (11% as opposed to 4.8%) and the higher grades, 4–7 (37% as opposed to 30.8%); the differences at Grades 5 and 6 were extremely small. This comparative advantage of men was also found in 2011, but the difference was statistically significant.

Figure 12: Primary teachers' EL competence in 2013, by gender and EL Trinity (GESE) grade

Figure 12 bar chart values (Female / Male):
Grade 1: 4.8% / 11.0%, Grade 2: 30.8% / 22.0%, Grade 3: 32.2% / 26.0%, Grade 4: 17.1% / 23.0%, Grade 5: 4.8% / 5.0%, Grade 6: 8.9% / 9.0%, Grade 7: 0.7% / 4.0%, Grade 8: 0.7%

Y-axis: Proportion of sample achieving grade
X-axis: Trinity Grade (GESE)
Legend: Female (N=146), Male (N=100)
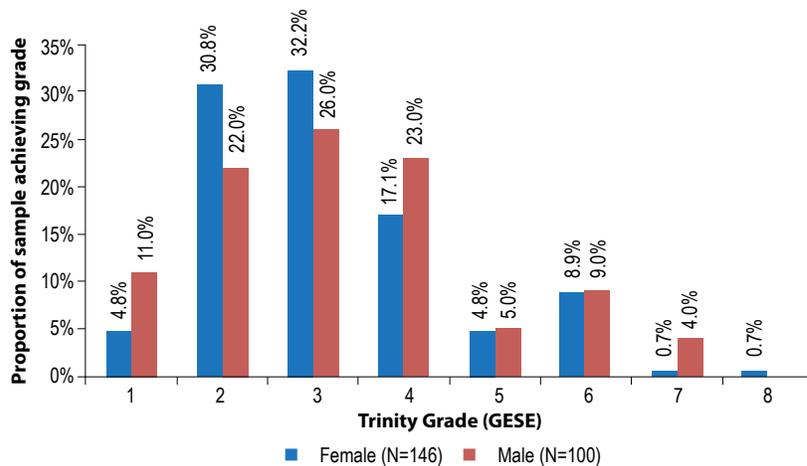
School location

The difference according to school location (Figure 13) shows primary teachers in rural schools fared slightly less well than those in semi-urban and urban schools; the difference is statistically significant (*p*<0.05; in 2011 there was no statistically significant difference). Rural schools were most strongly represented at the lowest grades, 1 and 2, while in Grades 3 and 4 the picture is reversed. Above Grade 4 the picture is a mixed one.

Figure 13: Primary teachers' EL competence in 2103, by school location and EL Trinity (GESE) grade
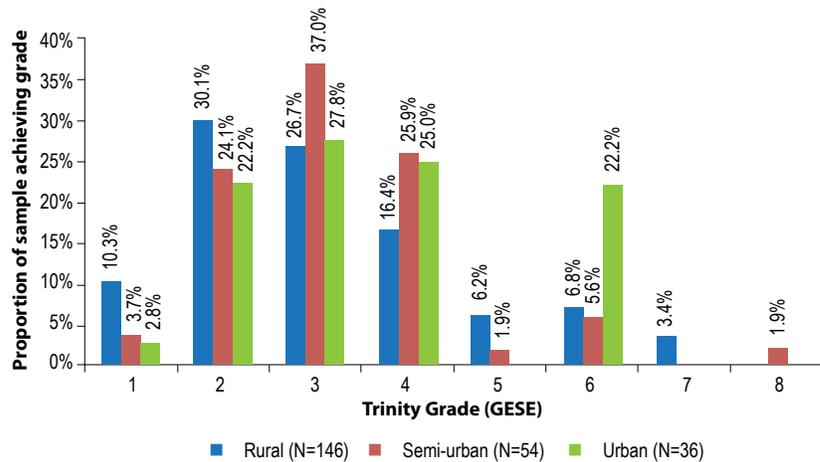
Division

As with students, primary teacher performance varied from division to division (Figure 14), though all but two divisions included teachers attaining as high as Grade 7 or 8. The two divisions where performance is poorest were Rangpur, where nearly four-fifths (78.4%) of teachers attained within the lowest three grades (1–3), and Chittagong, where just over two-thirds obtain these lowest grades. (For primary students, it was Rangpur and Khulna that performed least well.) In contrast, in Khulna, almost two-thirds (64.3%) attained above that level. Looking at the highest grades, 42.8% of teachers in Khulna attained Grade 5 or higher, while in Chittagong only 6.3% attained within this range.

Figure 14: Primary teachers' EL competence in 2013, by school location and EL Trinity (GESE) grade
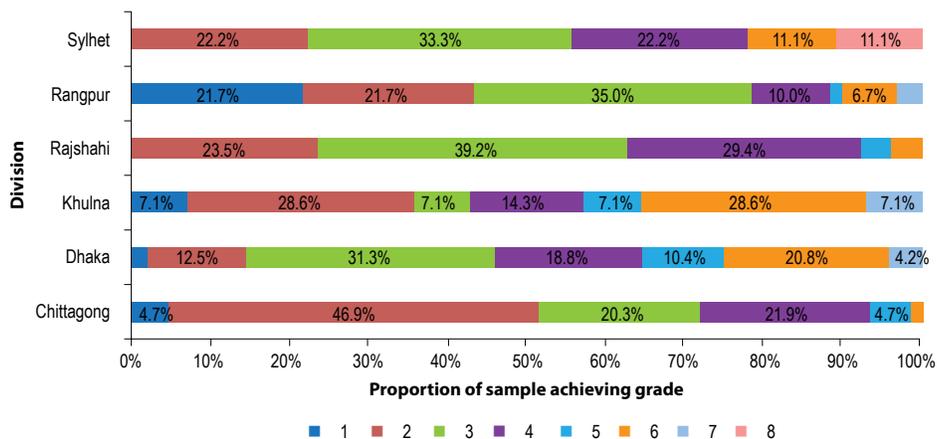
Addressing the research questions: Primary teachers

Figure 15 compares primary teachers' performance over the three years in question.

a) Comparison with pre-intervention (2010) baseline from pilot (Cohort 1)

Figures for 2013 show improvements over 2010, and these differences are statistically significant ($p<0.01$). In contrast to 2010, no teachers failed in 2013; there was a decrease in the percentage of teachers scoring the lowest grades (6.6% less teachers scored Grades 0 or 1); and a similar increase

in the percentage of teachers scoring the higher grades (7.9% more primary teachers scored Grade 4 or above). In both 2010 and 2013, over half of all primary teachers scored Grade 2 or 3 (58.3% in 2010, 56.9% in 2013). In 2013, 6.6% more primary teachers achieved Grade 2 and above (the criterion level for teaching Class 3).

b) Comparison with post-intervention (2011) assessment from pilot

Improvement above the baseline was similar for the larger-scale Cohort 2 to that achieved in the smaller-scale Cohort 1, which also saw no teachers failing the post-intervention assessment (Grade 0), and an increase in teachers scoring Grades 2 and above, and 4 and above. However, the scores achieved in 2013 were not quite as high as those achieved in 2011, when 4.3% more teachers achieved Grade 2 or above (97% in 2011, 92.7% in 2013), and 9.2% more teachers achieved Grade 4 and above (44.9% in 2011, 35.7% in 2013). These modest differences were statistically significant ($p<0.01$).

Figure 15: Primary teachers' EL competence in 2010, 2011 and 2013, by EL Trinity (GESE) grade
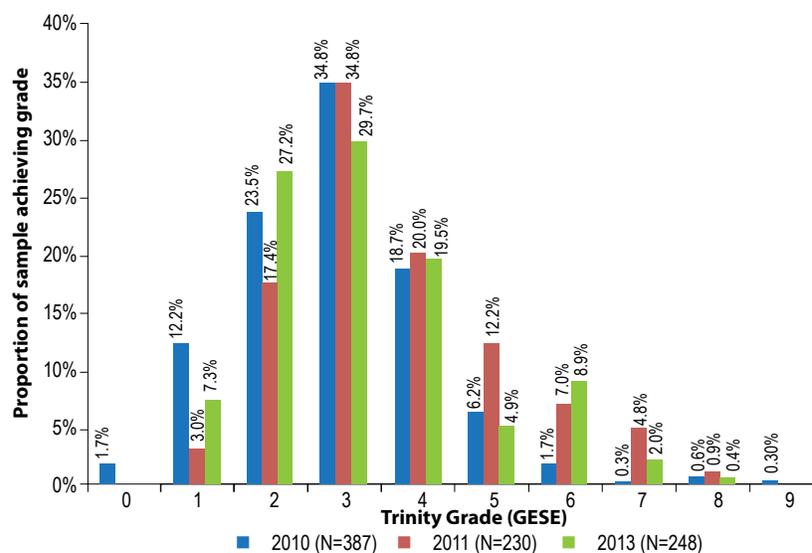
3.2.2 Primary head teachers

The results for primary HTs (Figure 16) are distributed fairly evenly across the grades. Over four-fifths (82.2%) passed the assessment (achieved a Grade 1 or higher). The large majority (78.1%) attained in the grade range 0–3, with the remainder mostly achieving Grades 4 and 5 (19.2%). Notably the only fail grades among primary ATs and HTs combined were among primary HTs (17.8%); one possible explanation for this is that not all the HTs assessed are in fact regular teachers of English. As noted earlier, this is the first time that HTs have been assessed, so no comparison with earlier years is possible. In addition, the small size of the sample makes it impossible to carry out any valid analysis according to gender, school location or division.

Figure 16: Primary head teachers' EL competence in 2013, by EL Trinity (GESE) grade
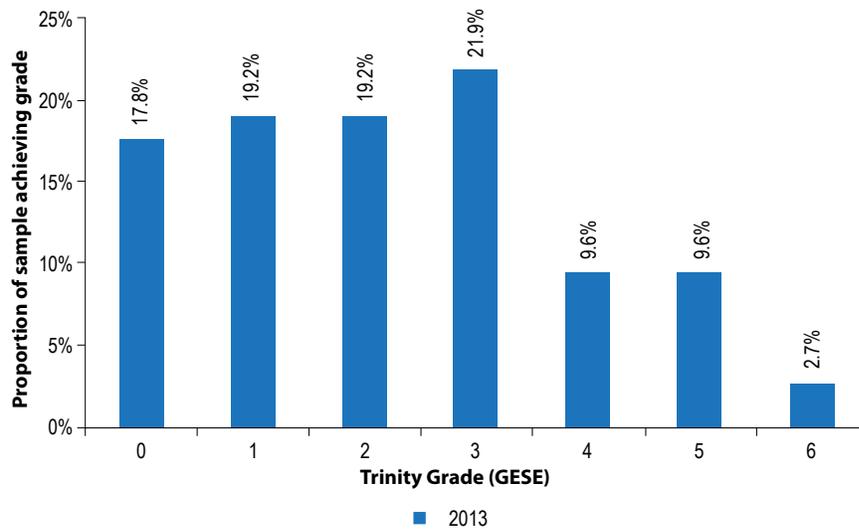
3.2.3 Secondary teachers

2013 overall results

Almost all (99.7%) of the secondary teachers (Figure 17) passed the assessment. They performed strongly in the mid grades – Grades 3 and 4 – with 51.8% achieving these grades, while 80.9% attained Grade 3 and above. Grade 3 competence is necessary to teach the first secondary *English for Today* textbook (Class 6), while Grade 4 gives access to a significant proportion of the Class 7 textbook (EIA 2010). A substantial proportion (15.3%) attained at the highest grades, 6–9, while 19.1% attained Grade 2 or lower, a figure which includes a single teacher who failed the assessment.

Figure 17: Secondary teachers' EL competence in 2013, by EL Trinity (GESE) grade
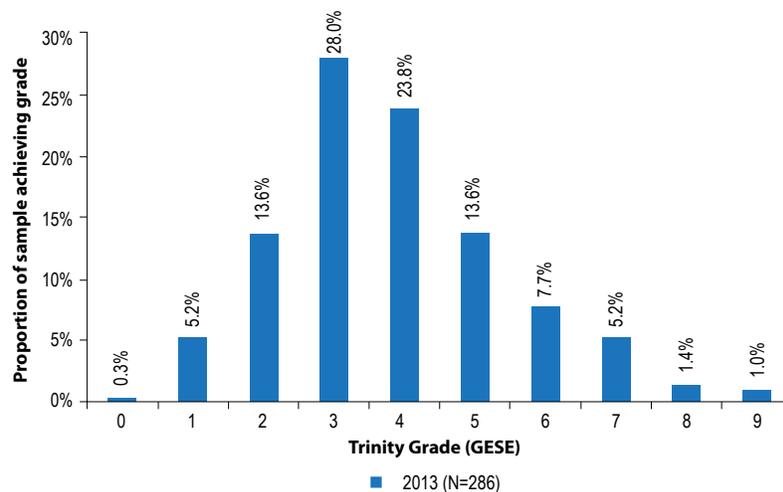
Gender

It is worth noting here that female secondary teachers made up less than a quarter of the secondary teacher sample. However, overall there is no statistical difference between the performance of male and female teachers (Figure 18; p=0.253), a similar result to 2011. That said, it is possible to discern some advantage for men, who are better represented in the grade range 2–5 (81.2% as opposed to 71.8%).

Figure 18: Secondary teachers' EL competence in 2013, by gender and EL Trinity (GESE) grade
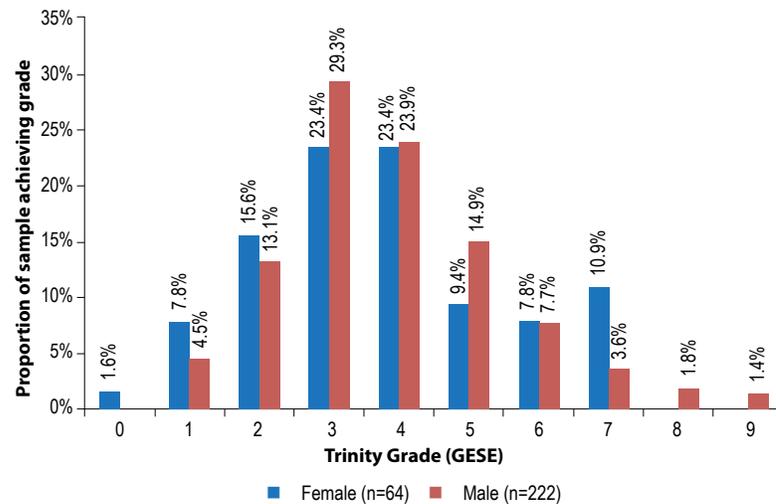
School location

Differences in terms of school location (Figure 19) are again complex and no conclusions can be drawn in respect of urban schools, since only two teachers in such establishments were assessed (one attaining Grade 2, the other Grade 5)[9].

Figure 19: Secondary teachers' EL competence in 2013, by school location and EL Trinity (GESE) grade
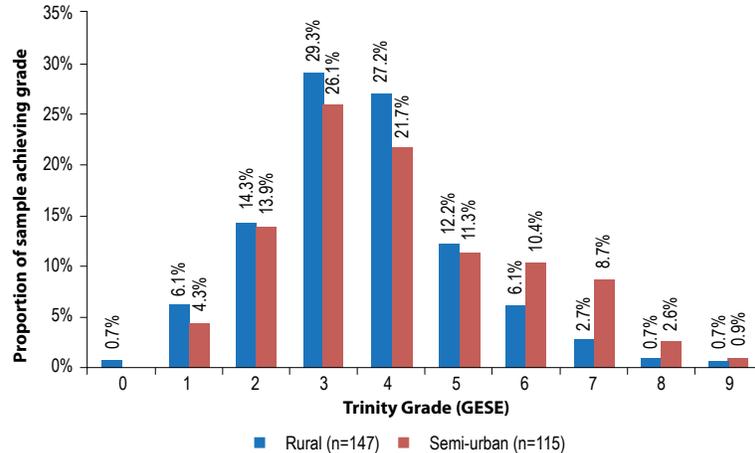
There is no statistically significant difference between rural and semi-urban secondary teacher Trinity grades (p=0.56). The sample size for urban teachers was too small to test.

The distributions of rural and semi-urban teachers were very similar; the grade range 3–7 accounts for over three-quarters of each category (77.5% of rural teachers and 78.2% of semi-urban). There are 12.5% more rural teachers represented in the lower grades up to Grade 5 (89.8% rural teachers as opposed to 77.3% semi-urban) and correspondingly 12.4% more semi-urban teachers in the higher grades (6–9), but these are not significant differences: statistically, the performance of rural and semi-urban teachers is equivalent.

---

[9] No statistical test could be applied across the three categories because of the low number of urban schools in the sample for secondary teachers.

Division

In terms of variation by division (Figure 20), another complex picture emerges, though overall there is a highly significant statistical difference for this variable (*p*<0.01). In some divisions, relatively few teachers attained at the lower grades (1–3): in Rajshahi the figure is 28.1%; in Chittagong, 33.3%. In others, the proportion is higher, with around half of secondary teachers in Sylhet and Khulna (48% and 50.8% respectively) attaining in this range (a parallel case with secondary student data, see Figure 9), while in Rangpur the proportion is almost four-fifths (78.7%). In Chittagong, no teacher attained above Grade 5. Sylhet and Khulna had the highest proportion of secondary teachers attaining at Grade 6 or above (21% and 28.6% respectively).

Figure 20: Secondary teachers' EL competence in 2013, by division and EL Trinity (GESE) grade
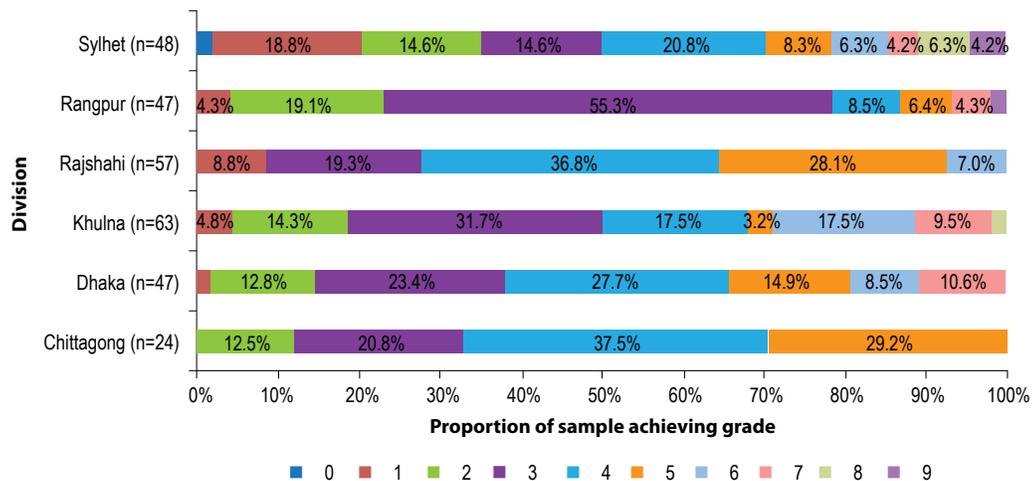
Addressing the research questions: Secondary teachers

Figure 21 compares the performance of secondary teachers over the three years in question.

a) Comparison with pre-intervention (2010) baseline from pilot (Cohort 1)

The difference between 2013 and 2010 levels of attainment is not statistically significant (*p*=0.475). However, there was a slight improvement in the proportion of teachers who achieve Grade 3 or above, the criterion level for secondary school teachers to be able to teach Class 6 students (the first year of secondary school).

b) Comparison with post-intervention (2011) assessment from pilot

While the statistically significant improvement across all the pass grades up to Grade 6 observed in 2011 (*p*<0.05) was not repeated in 2013, there was substantial progress in Grades 3 and 4. However, in terms of improvement in those who are at or above these grades (related to their needs to teach the textbook), the results in 2013 are approximately 10% below that of 2011. Overall there is no statistically significant difference between 2011 and 2013 results (*p*=0.238).

Figure 21: Secondary teachers' EL competence in 2010, 2011 and 2013, by EL Trinity (GESE) grade
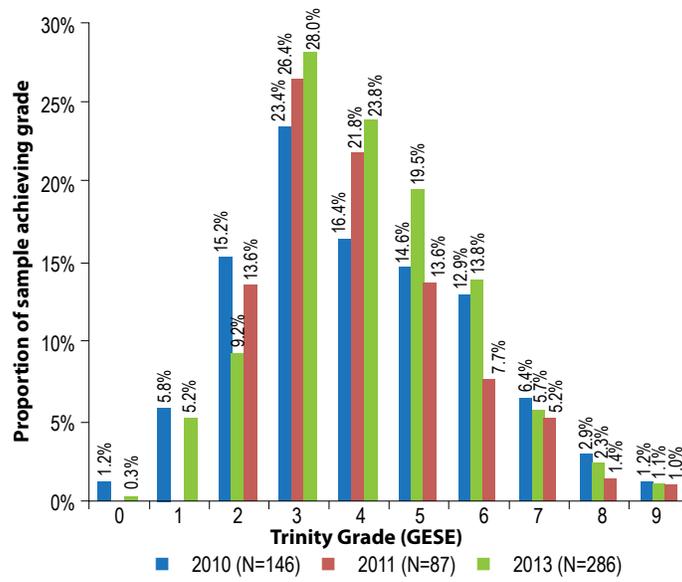
4. Discussion

4.1 Research questions

This study addresses two main questions:

1. To what extent do the teachers and students of Cohort 2 show improved post-intervention EL proficiencies, in speaking and listening, compared with the 2010 pre-intervention baseline?
2. To what extent has the programme been successful in repeating the post-intervention improvements in EL proficiencies seen in Cohort 1, at the much larger scale of Cohort 2?

4.1.1   Research question 1: Cohort 2 (2013) improvement over pre-intervention

The present study shows that the second cohort of students participating in EIA show highly statistically significant improvements over the 2010 baseline. Just over two-thirds (69.8%) of primary students passed the assessment in 2013, whereas only a little over one-third passed in 2010 (35.7%). Similarly, the secondary student pass rate increased from 74.5% (2010) to 86.2% (2013).

Among primary and secondary students, the study shows not just an increase in the pass rate, but increases in proportions of students scoring at the higher grades. In the primary sector, 13.7% more students achieved Grade 1 (over the baseline), 13.8% more students achieved Grade 2, and 6.5% more students achieved Grade 3 or higher. In the secondary sector, 14.5% more students achieved the higher grades (2–5) than in 2010 (60.3% compared with 45.8%).

Again, there is a highly statistically significant difference between the performance of primary teachers at the baseline and in 2013, with those in 2013 performing better across the board. In both 2010 and 2013, over half of all primary teachers scored Grades 2 or 3 (58.3% in 2010, 56.9% in 2013), but in 2013 there was a decrease in the percentage of primary teachers scoring the lowest grades (6.6% fewer teachers scored Grades 0 or 1) and a similar increase in the percentage in those scoring the higher grades (7.9% more scored Grade 4 or above). Of educational significance is the fact that 6.5% more primary teachers now have a level of English that enables them to teach at Class 3. For secondary teachers, there was no statistically significant difference between 2013 and the 2010 baseline. But equally, there was no statistically significant difference between 2013 and the 2011 post-intervention assessment.

In general, then, there were statistically significant improvements above the baseline for all students, and for primary teachers. Differences were not statistically significant for secondary teachers. Primary students showed very substantial improvements in EL competence, which were greater than anticipated, and indeed are remarkable given the increased scale and indirectness of the implementation for Cohort 2.

For secondary teachers there is no statistically significant difference between the 2013 results and either those for 2011 (p=0.238) or the 2010 baseline (p=0.475). Although this quantitative study cannot identify reasons for its results (something that will be more possible in the composite study [EIA 2014c]), two possible explanations are: i) the sample of secondary teachers was too small and subject to bias; and ii) the focus on upazilas of greater deprivation (CDI, see Appendix 6) may have resulted in teachers having a lower starting point. The fact that results for secondary students showed significant improvement despite relatively low teacher attainment, perhaps suggests improvements in classroom pedagogy outweighed EL competence of teachers (see Appendix 7 for a fuller account).

4.1.2 Research question 2: Comparison with pilot intervention

For primary students, the figures for 2013 show improvements over and above those achieved in 2011 (highly statistically significantly so), despite the increase of scale and programme implementation (as noted above). Compared with the 2011 outcomes, in 2013: 4.4% more primary students achieved Grade

1; 8.6% more achieved Grade 2; and 6.7% more achieved Grade 3 or higher. This is a remarkable result, showing that EIA is helping almost all students learn more effectively than they were in the pilot phase. For secondary students, the improvements at the greater scale were broadly similar to those in 2011 (there is no statistically significant difference), with an 82.6% pass rate (as opposed to 89.6% in 2011). This result means that the up-scaled and more indirect implementation of Cohort 2 is still able to reproduce broadly the same improvements as the pilot phase. This was exactly what EIA hoped for, as the basis for the complete up-scaling and institutionalisation in the next phase.

For primary teachers, the improvements in EL competence were not quite as good as those achieved in the pilot study (but a difference with modest statistical significance), with approximately 9% less primary teachers achieving the highest grades (4 or above) than in 2011. For secondary teachers, there were no statistically significant differences between the 2013 results and the 2011 results. As with secondary students, the performance of teachers is similar to that in 2011, post-intervention, indicating that the up-scaled and indirect implementation is largely still able to reproduce the pilot results.

In general, then, for primary students, the improvements were greater than those achieved in the pilot; for secondary students and primary teachers, the improvements were similar, but marginally lower at the highest grades. For secondary teachers, the results were statistically equivalent to those achieved in 2011.

4.2 Reporting on social inclusion

Social inclusion (SI) has always been an important consideration in the design and implementation of EIA and there is some evidence of the effectiveness of this policy in the findings of this study.

4.2.1   SI: Gender

For three of the four populations studied (secondary students, and both primary and secondary teachers), there was no statistically significant difference in EL competence attributable to gender.

The only statistically significant gender difference was found in primary students. Here, significantly more girls passed (74.4%, compared with 65.1% of boys). However, 4.9% more boys achieved the higher grades (3–4).

These results are indicative of EIA's capacity to improve EL learning outcomes for boys and girls alike. As these improved learning outcomes are attributed to improved classroom practice, and, in particular, an increasingly communicative approach to EL learning (EIA 2014a), the implication is that girls are benefitting equally from the improved language learning opportunities provided through the programme. Similarly, female teachers are achieving the same level of EL competence as their male counterparts.

4.2.2   SI: Rurality

Although rural primary students performed less well than semi-urban students (there were no urban students) in 2011, this rural student disadvantage was not evidenced as strongly in 2013. Rural primary students performed robustly and although they did slightly less well at the highest grades, in terms of the proportion achieving pass grades (80.9%), they surpassed semi-urban (59.1%) and urban (60%) students by a margin of 20%. For secondary students, the percentages passing the assessment were similar from rural and semi-urban locations, but again a greater proportion (18 more) of semi-urban students achieved the higher grades (3–5).

For teachers, there was no significant difference by location for secondary (and in any case too few urban teachers to conclude anything on location), although 12.4% more semi-urban secondary teachers achieved the top four grades (6–9) than their rural counterparts. Again, for primary teachers, a greater proportion of urban (10%) or semi-urban teachers (20%) achieved the Grades 3–4 than their rural counterparts.

Differences in EL proficiency between the various administrative divisions of the country are intriguing. While teachers and students in one division (Rajshahi) appear to perform consistently relatively well in both primary and secondary schools, elsewhere the picture is more varied. In Khulna, for example, primary teachers are among the best performers, while their students are among the weakest.

As already noted, the purpose of large-scale quantitative studies is to gauge and report on performance, rather than explain it. While it is clear even from this brief discussion that there are a number of matters which would benefit from further, detailed research, it is equally clear from the results of this study that EIA is continuing to make a significant impact on learning outcomes.

References

BANBEIS (2014a) 'Chapter 2: Primary Education', available online at:
http://www.banbeis.gov.bd/webnew/images/2pe.pdf

BANBEIS (2014b) 'Chapter 3: School education', available online at:
http://www.banbeis.gov.bd/webnew/images/3se.pdf

Bold, T., Kimenyi, M., Mwabu, G., Ng'ang'a, A. and Sandefur, J. (2013) *Scaling-up what works: experimental evidence on external validity in Kenyan education*. Working Paper No. 321. Washington, DC: Centre for Global Development.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Education Watch (2011) 'Exploring low performance in Education: The Case Study of Sylhet division', available at: www.campebd.org/download/EW2009-10FullReportEnglish.pdf

EIA (2009a) *An Assessment of Spoken English Competence Among School Students, Teachers and Adults in Bangladesh, (Baseline Study 1)*. Dhaka: EIA.

EIA (2009b) *An audit of current materials for teaching English in Bangladesh (Baseline Study 4). English In Action*, Dhaka, Bangladesh

EIA (2010) *Textbook analysis*. (Internal EIA document.) Dhaka: EIA.

EIA (2011a) *The Classroom Practices of Primary and Secondary School Teachers Participating in English in Action (Study 2a1)*. Dhaka: EIA.

EIA (2011b) *Perceptions of English Language Learning and Teaching Among Primary and Secondary School Teachers and Students Participating in English in Action (Study 2b3b).* Dhaka: EIA.

EIA (2012) *English Proficiency Assessments of Primary and Secondary Teachers and Students Participating in English in Action (Study 3a2).* Dhaka: EIA.

EIA (2014a) *The Classroom Practices of Primary and Secondary Teachers Participating in English in Action: Second Cohort (2013).* Dhaka: EIA.

EIA (2014b) *Perceptions of English Language Learning and Teaching Among Primary and Secondary Teachers and Students Participating in English in Action: Cohort 2 (2013).* Dhaka: EIA.

EIA (2014c) *English in Action: A Consideration of Primary and Secondary Students' and Teachers' English Language Competence and Perceptions and Teachers' Classroom Practice: Cohort 2 (2013). A Composite Report.* Dhaka: EIA. (forthcoming)

Hamid, M.O. & Baldauf, R.B.J. (2008) 'Will CLT bail out the bogged down ELT in Bangladesh?', *English Today,* 24:16–24

Hamid, O.M. (2011) 'Planning for failure: English and language policy and planning in Bangladesh'. In J.A. Fishman and O. García (eds.) *Handbook of Language and Ethnic Identify: the success-failure continuum in language and ethnic identify efforts* (pp. 192–203). (Volume 2). Oxford: Oxford University Press.

Kraft, R.J., Ehsan, A. and Khanam, R. (2009) *Comprehensive Primary Teacher Education and Continuing Professional Development: A Framework for Reform.* Dhaka: UNICEF.

Rahman, A., Kobir, Md. M. and Afroze, R. (2006) *Effect of BRAC-PACE Training on English Language Learning of Teachers in Rural Non-Government Secondary Schools.* Dhaka: BRAC Centre.

Trinity College London (2007) *Relating the Trinity College London GESE and ISE examinations to the Common European Framework of Reference – project summary.* London: Trinity College London.

Trinity College London (2009) *Graded Examinations in Spoken English (GESE). Syllabus from 2010.* London: Trinity College London.

Trinity College London (2013) Exam Information: Graded Examinations in Spoken English (GESE). London: Trinity College London

UNESCO (2012) 2012 Education for all Global Monitoring Report – Youth Skills and Work: Statistical Tables – long version.

UNICEF (2010) A Case for Geographic Targeting of Basic Social Services to Mitigate Inequalities in Bangladesh, Dhaka, available online at:
http://www.unicef.org/socialpolicy/files/Geo_targeting_web.pdf

Verspoor, A. (2005) The Challenge of Learning: Improving the Quality of Basic Education in Sub-Saharan Africa. Paris: ADEA.

World Bank (2008) 'Education for all in Bangladesh: Where does Bangladesh stand in achieving EFA Goals by 2015?', Bangladesh development Series, Paper no. 24 Dhaka: World Bank.

The sample sizes were determined using R statistical program with 'pwr' package developed for power analysis. Power analysis provides probability of correctly rejecting the null hypothesis when it is false. In the R package 'pwr' sample sizes (n2) can be determined providing the known parameters (effect size [d], sample size [n1] to which the next study will be compared, significant level and the power). The effect sizes (d) used here are taken from Cohen (1988) as parameters (means and standard deviation) for determining actual effect sizes from the existing studies could not be retrieved. Cohen's suggested effect sizes ranged from small to large (0.2 to 0.8). Smaller effect size needs higher sample size to prove statistically significant difference. R 'pwr' package was set for two group independent sample (N1 and N2) t-test.

Teachers	N1 (2010)	Effect size	Power1	Suggested N2 (2013) for Power1	Power2	Suggested N2 (2013) for Power2	Power3	Suggested N2 (2013) for Power3	Significant level	SAMPLE SIZE
Primary	367	0.2	70%	174	75%	213	80%	268	0.05	645 teachers (720-75) • 720 Ts at CMs (testing at 30 CMs with 24 Ts at each) (-75Ts in schools). 11 Pri CM; 5 PHT CM; 14 Sec CM. 264 P Ts; 120 P HTs; 336 S Ts • 75 Ts in schools (30 schools, 2 Ts in each, + 1HT) 15 primary schools; 15 secondary schools. 30 P Ts; 30 S Ts; 15 PHTs
Secondary	176	0.3	70%	188	75%	235	80%	309	0.01	
TOTAL								577		306 Sec; 234 Pri; 105 Pri HT

Students	N1 (2010)	Effect size	Power1	Suggested N2 (2013) for Power1	Power2	Suggested N2 (2013) for Power2	Power3	Suggested N2 (2013) for Power3	Significant level	SAMPLE SIZE
Primary	4630	0.2	70%	213	75%	237	80%	265	0.01	900 students (30 Ss per class, in 30 schools)
Secondary	2609	0.2	70%	221	75%	247	80%	278	0.01	
TOTAL								543		450 P Ss; 450 S Ss

Appendix 2: Data cleaning steps

The following actions were taken to clean the dataset:

- a data screening exercise by range checking, and checking variable values against predefined maximum and minimum bounds to catch spurious values or data entry;
- contingency tables constructed to carry out consistency checks;
- missing data, non-responses, data imputation for missing values dealt with, and outlier detection to ensure the data is in right shape and format for analysis;
- data transformation, involving re-categorising and altering variables (e.g. from original string to numerical variable);
- derived/newly created variables from existing variables.

Appendix 3: Statistical tables for the figures used in the report

A) PRIMARY STUDENT DATA

Year data: 2010, 2011, 2013

	Grade						Total
	0	1	2	3	4	5	
2010 (N=3507)	64.3%	30.1%	4.6%	0.7%	0.1%	0.3%	100%
2011 (N=784)	49.9%	39.4%	9.8%	0.9%			100%
2013 (N=463)	30.2%	43.8%	18.4%	6.5%	1.1%		100%
Total	60.1%	32.4%	6.1%	1.1%	0.2%	0.2%	100%

Gender

	Grade					Total
	0	1	2	3	4	
Female	25.6%	46.6%	22.6%	4.7%	0.4%	100.00%
Male	34.9%	41.0%	14.0%	8.3%	1.7%	100.00%
Total	30.2%	43.8%	18.4%	6.5%	1.1%	100.00%

School location

	Grade					Total
	0	1	2	3	4	
Rural	19.1%	57.9%	16.4%	5.5%	1.1%	100.00%
Semi-urban	40.9%	28.2%	20.9%	8.6%	1.4%	
Urban	40.0%	36.7%	20.0%	3.3%		100.00%
Total	31.6%	41.3%	18.9%	6.9%	1.2%	100.00%

Division

	Grade					Total
	0	1	2	3	4	
Chittagong	6.7%	72.2%	16.7%	2.2%	2.2%	100.0%
Dhaka	30.8%	55.8%	11.5%	1.9%		100.0%
Khulna	76.5%	20.6%	2.9%			100.0%
Rajshahi	1.2%	51.2%	27.9%	16.3%	3.5%	100.0%
Rangpur	70.7%	29.3%				100.0%
Sylhet	13.0%	31.5%	41.3%	14.1%		100.00%
Total	30.2%	43.8%	18.4%	6.5%	1.1%	100.00%

B) SECONDARY STUDENT DATA

Year data: 2010, 2011, 2013	Grade								Total
	0	1	2	3	4	5	6	7	
2010 (N=2041)	25.5%	28.3%	24.1%	12.4%	7.9%	1.4%	0.3%	0.0%	99.9%
2011 (N=317)	10.4%	20.2%	24.6%	21.8%	12.9%	6.0%	3.5%	0.6%	100.0%
2013 (N=421)	13.8%	25.9%	27.1%	22.1%	9.0%	2.1%			100.0%
Total	21.2%	30.9%	23.8%	13.6%	7.8%	1.9%	0.7%	0.1%	100.0%

Gender	Grade						Total
	0	1	2	3	4	5	
Female	14.7%	25.2%	24.8%	24.1%	9.0%	2.3%	100.0%
Male	12.4%	27.5%	30.1%	19.0%	9.2%	2.0%	100.0%
Total	13.8%	26.0%	26.7%	22.2%	9.1%	2.1%	100.0%

School location	Grade						Total
	0	1	2	3	4	5	
Rural	15.0%	30.3%	26.1%	18.5%	8.6%	1.6%	100.0%
Semi-urban	10.3%	13.1%	29.9%	32.7%	10.3%	3.7%	100.0%
Total	13.8%	25.9%	27.1%	22.1%	9.0%	2.1%	100.0%

Division	Grade						Total
	0	1	2	3	4	5	
Chittagong		20.8%	4.2%	4.2%	54.2%	16.7%	100.0%
Dhaka	10.6%	12.8%	26.6%	44.7%	5.3%		100.0%
Khulna	16.4%	19.7%	29.5%	26.2%	6.6%	1.6%	100.0%
Rajshahi		7.6%	40.9%	27.3%	18.2%	6.1%	100.0%
Rangpur	7.0%	49.1%	28.9%	11.4%	3.5%		100.0%
Sylhet	48.4%	30.6%	16.1%	4.8%			100.00%
Total	13.80%	25.90%	27.10%	22.10%	9.00%	2.10%	100.00%

C) PRIMARY TEACHER DATA

Year data: 2010, 2011, 2013	Grade										Total
	0	1	2	3	4	5	6	7	8	9	
2010 (N=367)	1.7%	12.2%	23.5%	34.8%	18.7%	6.2%	1.7%	0.3%	0.6%	0.30%	100.0%
2011 (N=230)		3.0%	17.4%	34.8%	20.0%	12.2%	7.0%	4.8%	0.9%		100.0%
2013 (N=246)		7.3%	27.2%	29.7%	19.5%	4.9%	8.9%	2.0%	0.4%		100.0%
Total	0.9%	8.2%	22.9%	33.6%	19.0%	7.5%	5.2%	2.0%	0.6%	0.10%	100.0%

Gender	Grade								Total
	1	2	3	4	5	6	7	8	
Female (N=146)	4.8%	30.8%	32.2%	17.1%	4.8%	8.9%	0.7%	0.7%	100.0%
Male (N=100)	11.0%	22.0%	26.0%	23.0%	5.0%	9.0%	4.0%		100.0%
Total	7.3%	27.2%	29.7%	19.5%	4.9%	8.9%	2.0%	0.4%	100.0%

School location	Grade								Total
	1	2	3	4	5	6	7	8	
Rural (N=146)	10.3%	30.1%	26.7%	16.4%	6.2%	6.8%	3.4%		100.0%
Semi-urban (N=54)	3.7%	24.1%	37.0%	25.9%	1.9%	5.6%		1.9%	100.0%
Urban (N=36)	2.8%	22.2%	27.8%	25.0%		22.2%			100.0%
Total Urban (N=36)	7.6%	27.5%	29.2%	19.9%	4.2%	8.9%	2.1%	0.4%	100.0%

Division	Grade								Total
	1	2	3	4	5	6	7	8	
Chittagong	4.7%	46.9%	20.3%	21.9%	4.7%	1.6%		1	100.0%
Dhaka	2.1%	12.5%	31.3%	18.8%	10.4%	20.8%	4.2%		100.0%
Khulna	7.1%	28.6%	7.1%	14.3%	7.1%	28.6%	7.1%		100.0%
Rajshahi		23.5%	39.2%	29.4%	3.9%	3.9%			100.0%
Rangpur	21.7%	21.7%	35.0%	10.0%	1.7%	6.7%	3.3%		100.0%
Sylhet		22.2%	33.3%	22.2%		11.1%		11.1%	100.00%
Total	7.3%	27.2%	29.7%	19.5%	4.9%	8.9%	2.0%	0.4%	100.00%

D) PRIMARY HEAD TEACHER DATA

Year data: 2010, 2011, 2013	Grade							Total
	0	1	2	3	4	5	6	
2013	17.8%	19.2%	19.2%	21.9%	9.6%	9.6%	2.7%	100.0%
Total	17.80%	19.20%	19.20%	21.90%	9.60%	9.60%	2.70%	100.0%

E) SECONDARY TEACHER DATA

Year data: 2010, 2011, 2013	Grade										Total
	0	1	2	3	4	5	6	7	8	9	
2010 (N=176)	1.2%	5.8%	15.2%	23.4%	16.4%	14.6%	12.9%	6.4%	2.9%	1.2%	100.0%
2011 (N=87)			9.2%	26.4%	21.8%	19.5%	13.8%	5.7%	2.3%	1.1%	100.0%
2013 (N=286)	0.3%	5.2%	13.6%	28.0%	23.8%	13.6%	7.7%	5.2%	1.4%	1.0%	100.0%
Total	0.5%	4.6%	13.3%	26.4%	21.3%	14.9%	10.2%	5.6%	2.0%	1.1%	100.0%

Gender	Grade										Total
	0	1	2	3	4	5	6	7	8	9	
Female(n=64)	1.6%	7.8%	15.6%	23.4%	23.4%	9.4%	7.8%	10.9%			100.0%
Male(n=222)		4.5%	13.1%	29.3%	23.9%	14.9%	7.7%	3.6%	1.8%	1.4%	100.0%
Total	0.3%	5.2%	13.6%	28.0%	23.8%	13.6%	7.7%	5.2%	1.4%	1.0%	100.0%

School location	Grade										Total
	0	1	2	3	4	5	6	7	8	9	
Rural(n=147)	0.7%	6.1%	14.3%	29.3%	27.2%	12.2%	6.1%	2.7%	0.7%	0.7%	100.0%
Semi-urban(n=115)		4.3%	13.9%	26.1%	21.7%	11.3%	10.4%	8.7%	2.6%	0.9%	100.0%
Urban(n=2)			50.0%			50.0%					100.0%
Total	0.4%	5.3%	14.4%	27.7%	24.6%	12.1%	8.0%	5.3%	1.5%	0.8%	100.0%

Division	Grade										Total
	0	1	2	3	4	5	6	7	8	9	
Chittagong(n=24)			12.5%	20.8%	37.5%	29.2%					100.0%
Dhaka(n=47)		2.1%	12.8%	23.4%	27.7%	14.9%	8.5%	10.6%			100.0%
Khulna(n=63)		4.8%	14.3%	31.7%	17.5%	3.2%	17.5%	9.5%	1.6%		100.0%
Rajshahi(n=57)			8.8%	19.3%	36.8%	28.1%	7.0%				100.0%
Rangpur (n=47)		4.3%	19.1%	55.3%	8.5%	6.4%		4.3%		2.1%	100.0%
Sylhet (n=48)	2.1%	18.8%	14.6%	14.6%	20.8%	8.3%	6.3%	4.2%	6.3%	4.2%	100.00%
Total	0.3%	5.2%	13.6%	28.0%	23.8%	13.6%	7.7%	5.2%	1.4%	1.0%	100.00%

Appendix 4: Statistical significance tests used in the report

A) PRIMARY STUDENT STATISTICAL SIGNIFICANCE TESTS

Year data: 2010, 2011, 2013

2010 vs 2013, there is significant different; $p < .001$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	408.724[a]	5	.000
Likelihood Ratio	304.601	5	.000
Linear-by-Linear Association	339.207	1	.000

2011 vs 2013: there is significant different; $p < 0.01$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	83.163[a]	4	.000
Likelihood Ratio	84.777	4	.000
Linear-by-Linear Association	84.777	1	.000

Gender

Female vs male: there is significant different, $p < 0.05$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	13.035[a]	4	.011
Likelihood Ratio	13.251	4	.010
Linear-by-Linear Association	.639	1	.424

School location

Sig different among rural, semi-urban and urban area, $p < 0.01$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	40.824a	8	.000
Likelihood Ratio	42.050	8	.000
Linear-by-Linear Association	2.282	1	.131

Division

Differences among divisions are significant, $p < 0.01$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	207.926a	20	.000
Likelihood Ratio	208.354	20	.000
Linear-by-Linear Association	7.859	1	.005

B) SECONDARY STUDENT STATISTICAL SIGNIFICANCE TESTS

Year data: 2010, 2011, 2013

2010 vs 2013: There is a significant difference; $p<0.01$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	63.234[a]	7	.000
Likelihood Ratio	62.365	7	.000
Linear-by-Linear Association	37.628	1	.000

2011 vs 2013 There is a significant difference: $p<0.01$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	31.535[a]	7	.000
Likelihood Ratio	36.240	7	.000
Linear-by-Linear Association	21.227	1	.000

Gender

There is no gender difference among secondary school students.

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2.729[a]	5	.742
Likelihood Ratio	2.745	5	.739
Linear-by-Linear Association	.075	1	.784

School location

There is a significant difference between rural and semi-urban: $p<0.05$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	20.078[a]	5	.001
Likelihood Ratio	20.750	5	.001
Linear-by-Linear Association	12.898	1	.000

Division

There are significant differences among divisions. $p<0.001$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	216.797a	20	.000
Likelihood Ratio	173.371	20	.000
Linear-by-Linear Association	71.487	1	.000

C) PRIMARY TEACHER STATISTICAL SIGNIFICANCE TESTS

Year data: 2010, 2011, 2013

2010 vs 2013: There is a significant difference. $p<0.01$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	34.156[a]	9	.000
Likelihood Ratio	37.611	9	.000
Linear-by-Linear Association	10.672	1	.001

2011 vs 2013: There is a significant difference, $p<0.05$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	21.433[a]	7	.003
Likelihood Ratio	21.902	7	.003
Linear-by-Linear Association	9.038	1	.003

Gender

No gender difference is found, $p=.160$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.536[a]	7	.160
Likelihood Ratio	10.893	7	.143
Linear-by-Linear Association	.457	1	.499

School location

There are significant differences among areas: $p<0.05$

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	26.837[a]	14	.020
Likelihood Ratio	28.102	14	.014
Linear-by-Linear Association	3.236	1	.072
N of Valid Cases	236		

D) SECONDARY TEACHER STATISTICAL SIGNIFICANCE TESTS

Year data: 2010, 2011, 2013

2010 vs 2013: There is no significant difference, *p*=.475

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	17.714[a]	18	.475
Likelihood Ratio	21.958	18	.234
Linear-by-Linear Association	2.348	1	.125
N of Valid Cases	549		

2011 vs 2013: There is no significant difference, *p*=.238

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.397[a]	8	.238
Likelihood Ratio	13.530	8	.095
Linear-by-Linear Association	6.203	1	.013

Gender

There is no gender difference. *p*=.253

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	10.174[a]	8	.253
Likelihood Ratio	10.827	8	.212
Linear-by-Linear Association	.081	1	.776

School location

There is no significant difference between rural and semi-urban; sample size for urban too small for valid test.

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	14.523[a]	16	.560
Likelihood Ratio	14.262	16	.579
Linear-by-Linear Association	4.008	1	.045

Division

There are significant differences among divisions, *p*<0.01

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	115.882[a]	40	.000
Likelihood Ratio	117.261	40	.000
Linear-by-Linear Association	2.294	1	.130

Appendix 5: Sample descriptions & population demographics – Cohort 2 students and teachers

SAMPLE DESCRIPTION; STUDENTS

Pri	463	52%
Sec	421	48%
TOTAL	884	

GENDER	Pri		Sec		Total	
Male	229	49.5%	153	36.5%	382	43%
Female	234	50.5%	266	63.5%	500	57%
TOTAL	463		419		882	
No data	0		2		2	

DIVISION	Pri		Sec		Total	
Chittagong	90	19.4%	24	5.7%	114	12.9%
Dhaka	52	11.2%	94	22.3%	146	16.5%
Khulna	68	14.7%	61	14.5%	129	14.6%
Rajshahi	86	18.6%	66	15.7%	152	17.2%
Rangpur	75	16.2%	114	27.1%	189	21.4%
Syhlet	92	19.9%	62	14.7%	154	17.4%
TOTAL	463		421		884	

LOCATION	Pri		Sec		Total	
Rural	183	42.3%	314	74.6%	497	58.2%
Semi-urban	220	50.8%	107	25.4%	327	38.3%
Urban	30	6.9%	0		30	3.5%
Total	433		421		854	
Blanks	30				35	
TOTAL	256					

EIA STUDENT POPULATION (COHORT 2)*

Pri	195000	22%
Sec	692000	78%
TOTAL	887000	100%

GENDER**	Pri	Sec	Overall
Male	52%	49%	51%
Female	48%	51%	50%
TOTAL	100%	100%	100%

LOCATION	Pri	Sec
Rural	81%	not available
Urban	19%	not available
TOTAL	100%	

*Overall student population figures are derived from those estimated in the EIA logical framework.

** Data derived from returned teacher 'entry' (i.e. registration) questionnaires completed by teachers at EIA orientation workshops. 79% of questionnaires were returned.

SAMPLE DESCRIPTION: TEACHERS

Pri	246	41%
Sec	286	47%
Pri HT	73	12%
TOTAL	605	

GENDER	Pri		Pri HT		Sec		Total	
Male	100	40.7%	31	42.5%	222	77.6%	353	58.3%
Female	146	59.3%	42	57.5%	64	22.4%	252	41.7%
TOTAL	246		73		286		605	

DIVISION	Pri		Pri HT		Sec		Total	
Chittagong	64	26.0%	0	0.0%	24	8.4%	88	14.5%
Dhaka	48	19.5%	19	26.0%	47	16.4%	114	18.8%
Khulna	14	5.7%	1	1.4%	63	22.0%	78	12.9%
Rajshahi	51	20.7%	3	4.1%	57	19.9%	111	18.3%
Rangpur	60	24.4%	1	1.4%	47	16.4%	108	17.9%
Sylhet	9	3.7%	49	67.1%	48	16.8%	106	17.5%
TOTAL	246		73		286		605	

LOCATION	Pri		Pri HT		Sec		Total	
Rural	146	61.9%	41	58.6%	147	55.7%	334	58.6%
Semi-urban	54	22.9%	13	18.6%	115	43.6%	182	31.9%
Urban	36	15.3%	16	22.9%	2	0.8%	54	9.5%
Total	236		70		264		570	
Blanks	10		3		22		35	
TOTAL	246		73		286		605	

EIA TEACHER POPULATION (COHORT 2)*

Pri	1802	41%
Sec	1666	38%
Pri HT	900	21%
TOTAL	4368	100%

GENDER**	Pri	Sec	Overall
Male	45%	79%	62%
Female	55%	21%	38%
TOTAL	100%	100%	100%

DIVISION	Pri	Sec	Overall
Chittagong	15%	16%	15%
Dhaka	28%	24%	26%
Khulna	10%	14%	12%
Rajshahi	16%	19%	17%
Rangpur	15%	18%	16%
Sylhet	17%	10%	13%
TOTAL	100%	100%	100%

LOCATION	Pri	Sec	Overall
Rural	56%	67%	62%
Non-rural	44%	33%	39%
TOTAL	100%	100%	100%

* Data regarding EIA's Cohort 2 teacher population demographic make-up is derived from returned teacher 'entry' (i.e. registration) questionnaires completed by teachers at EIA orientation workshops. 79% of questionnaires were returned.

Appendix 6: CDI and quintile value analysis for upazilas of 2011 (Cohort 1, post-test) & 2013 (Cohort 2, post-test) EL assessment

Notes:

All CDI and quintile values extracted from UNICEF (2010).
Composite Deprivation Index (CDI) value is a number on a scale running between -15 and +15, where -15 expresses maximum deprivation and +15 maximum advantage.

Quintile values are a number on a scale between 1 and 5, where 1 represents a least deprived upazila and 5 a most deprived upazila

2013 assessments	Quintile value	CDI index value	STUDENTS			TEACHERS		
			No assessed	Proportion	Weighted CDI value	No assessed	Proportion	Weighted CDI value
Anowara	3	-2	114	0.13	-0.26	88	0.15	-0.29
Gangni	2	4	129	0.15	0.58	58	0.10	0.38
Kishoregonj sadar	3	-2	94	0.11	-0.21	47	0.08	-0.16
Mohammadpur	3	0	52	0.06	0.00	0		
Patgram	2	1	189	0.21	0.21	108	0.18	0.18
Puthia	2	2	152	0.17	0.34	18	0.03	0.06
Sunamganj sadar	3	-6	154	0.17	-1.05	48	0.08	-0.48
Mirpur	2	4				67	0.11	0.44
Boriagram	1	7				70	0.12	0.81
Jagganuthpur	3	-4				15	0.02	-0.10
Khustia Sadar	2	3				20	0.03	0.10
Natore Sadar	5	-13				23	0.04	-0.49
Syhlet sadar	3	-3				43	0.07	-0.21
			884	1.00		605	1.00	

			STUDENTS			TEACHERS		
TOTAL VALUES 2013 ASSESSMENTS	34	-9	STUDENTS		-0.37	TEACHERS		0.24
AVERAGE VALUES 2013 ASSESSMENTS	.6	0.69	STUDENTS		-0.05	TEACHERS		0.02

2010 assessments	Quintile value	CDI index value	STUDENTS			TEACHERS		
			No assessed	Proportion	Weighted CDI value	No assessed	Proportion	Weighted CDI value
Beanibazar	2	2	420	0.06	0.12	19	0.03	0.07
Dumuria	1	9	1260	0.17	1.57	42	0.08	0.70
Golapganj	2	5	762	0.11	0.53	21	0.04	0.19
Laxmipur	3	-1	1066	0.15	-0.15	46	0.08	-0.08
Lalmonhirat	3	0	1369	0.19	0.00	32	0.06	0.00
Nalchiti	1	6	1237	0.17	1.03	40	0.07	0.44
Tungipara	2	4	1115	0.15	0.62	30	0.06	0.22
Bhandaria	2	3				22	0.04	0.12
Burichang	2	4				17	0.03	0.13
Chandonaish	2	4				4	0.01	0.03
Chattak	4	-9				21	0.04	-0.35
Chowgasa	3	2				19	0.03	0.07
Kapasia	2	5				37	0.07	0.34
Monohordhi	2	3				31	0.06	0.17
Paba	2	3				35	0.06	0.19
Pangsha	2	16				16	0.03	0.47
Ramu	3	-5				21	0.04	-0.19
Rangamati	3	-6				38	0.07	-0.42
Shyamhagar	3	-4				20	0.04	-0.15
Syedpur	2	2				15	0.03	0.06
Ullapara	3	0				17	0.03	0.00
			7229	1.00		543	1.00	

	Quintile value	CDI index value	STUDENTS			TEACHERS		
TOTAL VALUES 2010 ASSESSMENTS	49.00	43.00	STUDENTS		3.71	TEACHERS		2.01
AVERAGE VALUES 2013 ASSESSMENTS	.33	.05	STUDENTS		0.53	TEACHERS		0.10
Difference 2013 vs 2010	-0.28	2.74	STUDENTS		0.58	TEACHERS		0.08

Quintile comparison - actuals	2013 sample	2010 sample	EIA upazilas	National upazilas
1	1	2	13	53
2	5	11	37	167
3	6	7	42	181
4	0	1	7	55
5	1	0	8	25
Totals	**13**	**21**	**107**	**481**

Quintile comparison - proportions	2013 sample	2010 sample	EIA upazilas	National upazilas
1	8%	10%	12%	11%
2	38%	52%	35%	35%
3	46%	33%	39%	38%
4	0%	5%	7%	11%
5	8%	0%	7%	5%
Totals	**100%**	**100%**	**100%**	**100%**

Appendix 6, figure 1: National and EIA distribution of Upazilas, by UNICEF CDI quintiles.
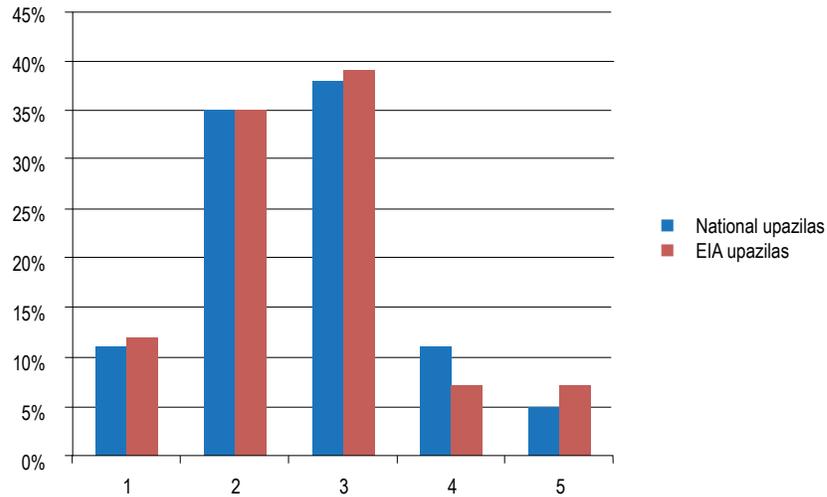
Appendix 6, figure 2: Sample of Upazilas in EIA ELC studies 2010 & 2013, by UNICEF CDI quintiles
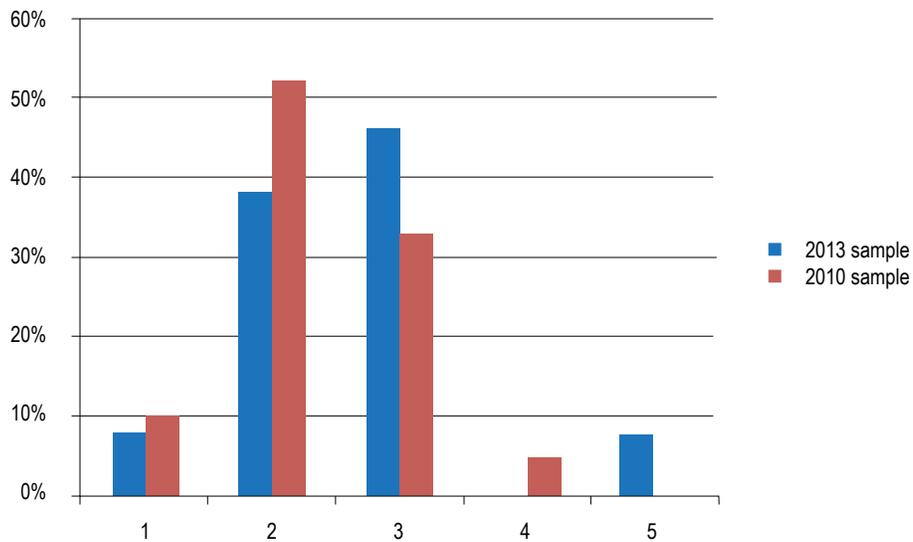
Appendix 7: Discussion regarding secondary teacher performance differences from 2011 to 2013

This appendix considers some possible explanations for the differences in EL competence demonstrated by secondary teachers in 2011 and 2013. It also reflects on the fact that students in 2013 made a degree of progress from the baseline greater (and in the case of secondary students, much greater) than that of their teachers.

First, there may well be sample size issues in relation to the 2011 study, which produced results that were described at the time as 'surprising' (EIA 2012: 21). In that study, the sample sizes were calculated through a power analysis to give a sample size for teachers as a whole (primary and secondary), giving a total figure of 326 (with a minimum of 300). In the event, this number was just about reached (317 primary and secondary teachers assessed). In the light of the results for that cohort, it was decided to calculate the 2013 sample by a different method. When the power analysis was conducted for this sample, primary and secondary teachers (and students) were calculated separately (see Appendix 1), giving a higher figure (309 – of whom 286 were assessed, as opposed to 87 in 2011)[10].

A second possible explanation relates to the assumption in the original design of the evaluation that Cohort 1 and Cohort 2 populations would be of very similar socioeconomic composition. (Naturally, during the early stages of the pilot phase, the Cohort 2 upazilas had not been selected.) In the event, this was a false assumption, since for Cohort 2, in order to fulfil EIA's need to address social inclusion, upazilas with a higher UNICEF composite deprivation index (CDI) were chosen.

As shown in appendix 6, figure 1, the distribution of upazilas EIA is now working in, is very similar to the national distribution of upazilas, by CDI quintile.

This had been less of a priority during the pilot phase, but gained importance in Phase III as a result of the project's Annual Review. EIA's Research, Monitoring and Evaluation resources had been irrevocably committed to Phase III on the basis of a design which used the Cohort 1 pre-intervention as the baseline for Cohort 2. As shown in Appendix 6, figure 2, the sample of upazilas from which teachers were assessed in 2010, was slightly skewed towards quintiles 1 & 2 (less deprived upazilas), and the weighted CDI values for teachers, show the 2010 sample to be relatively advantaged (2.01) compared to the 2013 sample (0.24).

It seems reasonable to speculate that had a baseline study been conducted in these relatively disadvantaged upazilas it would have shown teachers' EL competence to be at a lower level than the 2010 baseline, and therefore a similar degree of value added in 2013 as seen in 2011. However, though plausible, this idea is reflected only to a lesser extent in the results of primary teachers and not borne out at all by the results of primary and secondary students. Of course, there may well be other factors at work to produce different outcomes for the different groups.

Given the apparently minor improvements shown by this cohort of secondary teachers, it is noteworthy that their students still performed strongly. A possible explanation for this apparent anomaly may lie in the multiplicity of factors influencing the English of specialist English teachers. EIA will make a relatively small contribution to their lifetime's experience of English, whereas for students (and to some extent primary teachers), classroom experience is their only encounter with the language.

There was also a difference in the implementation of the two cohorts. Cohort 1 teachers enjoyed a relatively high degree of support from a team of international and national EL specialists, using pilot training materials, while Cohort 2 teachers had much less direct 'expert' support, but their pedagogy was supported by training materials that had been completely rewritten; these were much more focused on authentic classroom video and teachers' own reflective practice. This may have given a disproportionate boost to their pedagogy, resulting in higher student attainment, despite lower teacher proficiency. The composite study, which will relate teachers' EL competence to their classroom practice, perceptions, their students' EL competence and other factors revealed in the qualitative school studies may shed some light on this conjecture.

---

[10]Note that the sample size for primary teachers was much higher in 2011 (230), and less likely to suffer from any distortion (and the results for 2011 and 2013 for primary teachers are more comparable).