Education Sector Support Programme in Nigeria (ESSPIN)

Overall findings and technical report of

ESSPIN composite survey 1 (2012)

March 2013

Report number: ESSPIN 060

esspin
Education Sector Support Programme in Nigeria

UKaid from the Department for International Development

www.esspin.org

Report Distribution and Revision Sheet

Project Name: Education Sector Support Programme in Nigeria

Code: **244333TA05**

Report No.: **ESSPIN 060**

Report Title: **Overall findings and technical report of ESSPIN composite survey 1 (2012)**

Rev No	Date of issue	Originator	Checker	Approver	Scope of checking
2	Mar 2013	ESSPIN M&E, Output 3 and Output 4 Teams	Jake Ross	Kayode Sanni	Accuracy, completeness, formatting

Scope of Checking

This report has been discussed with the originator and checked in the light of the requirements of the terms of reference. In addition the report has been checked to ensure editorial consistency, accuracy of data, completeness of scope and responsiveness to client's requirements.

Distribution List

Name	Position
DFID	
Richard Montgomery	Head, DFID Nigeria
Sarah White	Human Development Team Leader, DFID
Barbara Payne	Senior Education Adviser, DFID
Emily Oldmeadow	Education Adviser, DFID
Esther Forgan	Results Adviser, DFID
Tony Gardener	Commercial Adviser, DFID
Olatunji Ogunbanwo	Deputy Programme Manager, DFID
Ben Le Roith	Head of DFID Northern Nigeria Office (Kano)
Adeshina Fagbenro-Byron	Head of DFID Lagos Office (South West)
Olachi Chuks-Ronnie	Head of DFID Enugu Office (South East/South South)
Zainab Lukat	DFID State Representative, Jigawa & Yobe
Ben Fisher	DFID State Representative, Kano & Zamfara
Siaka Alhassan	DFID State Representative, Kaduna & Katsina
Justice Ogoroh	HD Programme Officer, DFID
Roseline Onyemachi	Education Project Officer, DFID
IMEP	
David Wand	National Programme Manager, IMEP
Yvonne Nwokedi	Finance & Admin Manager, IMEP
ESSPIN	
John Martin	Project Director
Kayode Sanni	National Programme Manager

Name	Position
Jake Ross	Deputy Programme Manager
Andy Campbell	Operations Manager
Nguyan Feese	Lead Specialist, Institutional Development
Gboyega Ilusanya	State Team Leader, Lagos
Emma Williams	State Team Leader, Kwara
Olalekan Saidi	State Team Leader, Kano
Christiana Omotayo Odekunle	State Team Leader, Kaduna
Pius Elumeze	State Team Leader, Enugu
Mustapha Ahmad	State Team Leader, Jigawa
John Kay	Lead Specialist, Education Quality
Lilian Breakell	Task Team Leader, School Improvement
Kabiru Abass	Task Specialist, Access, Gender & Equity
Fatima Aboki	Lead Specialist, Community Engagement and Learner Participation
Sandra Graham	Task Team Leader, Civil Society
Bankole Ebisemiju	Communications & Knowledge Management Coordinator
Consortium partners	
David Theobald	Senior Education Adviser, British Council
Georgina Rawle	Education and Social Policy Economist, Oxford Policy Management
Susan Grant	Nigeria Country Director, Save the Children
Sue Phillips	Director, Social Development Direct
Federal partners	
Prof. Ruqayyatu Ahmed Rufa'i, OON	Honourable Minister of Education
Dr. MacJohn Nwaobiala	Permanent Secretary, Federal Ministry of Education
State partners	Honourable Commissioners of Education and SUBEB Chairs

Disclaimer

This document is issued for the party which commissioned it and for specific purposes connected with the captioned project only. It should not be relied upon by any other party or used for any other purpose.
We accept no responsibility for the consequences of this document being relied upon by any other party, or being used for any other purpose, or containing any error or omission which is due to an error or omission in data supplied to us by other parties

Note on Documentary Series

A series of documents has been produced by Cambridge Education as leader of the ESSPIN consortium in support of their contract with the Department for International Development for the Education Sector Support Programme in Nigeria. All ESSPIN reports are accessible from the ESSPIN website. http://www.esspin.org/resources/reports

The documentary series is arranged as follows:

ESSPIN 0-- Programme Reports and Documents

ESSPIN 1-- Support for Federal Level Governance (Reports and Documents for Output 1)

ESSPIN 2-- Support for State Level Governance (Reports and Documents for Output 2)

ESSPIN 3-- Support for Schools and Education Quality Improvement (Reports and Documents for Output 3)

ESSPIN 4-- Support for Communities (Reports and Documents for Output 4)

ESSPIN 5-- Information Management Reports and Documents

Reports and Documents produced for individual ESSPIN focal states follow the same number sequence but are prefixed:

JG Jigawa
KD Kaduna
KN Kano
KW Kwara
LG Lagos
EN Enugu

Readers of this report may also be interested *in ESSPIN 061 Gender analysis of key results, ESSPIN composite survey 1 (2012)* and the six State Reports of ESSPIN Composite Survey 1 (2012):
EN201, JG201, KD201, KN201, KW205, LG202

Acknowledgments

ESSPIN gratefully acknowledges the contributions of the following to the Composite Survey 2012 report data collection, processing, analysis, writing, editing and review (alphabetical order):

Fatima Aboki
Lilian Breakell
Edgar Cooke
Domenec Devesa
Allan Findlay
Sani Gar
Sandra Graham
Abbe Katerega
John Kay
Manjola Kola
Juan Munoz
Henry Onimole
Barbara Payne
Gunilla Pettersson
Georgina Rawle
Jake Ross
Kayode Sanni
Sebastian Silvaleander
Simon Thomson
Zoe van der Hoven
Sue Williamson

Thanks are also due to all the children, teachers, head teachers, SBMC members, local and state government officers, SSIT members, NPC field workers, DFID personnel and others, without whose participation and dedication these findings could not have been reported.

Contents

Figures

Tables

Acronyms and Abbreviations

CI	Confidence Interval
CS	Composite Survey
CSO	Civil Society Organisation
DEFF	Design Effect (used in its statistical sense)
DFID	Department for International Development
DPRS	Director(ate) of Planning, Research and Statistics
EDORE	Education Operational Research and Evaluation (project of DFID Nigeria)
EMIS	Education Management Information System
ESSPIN	Education Sector Support Programme in Nigeria
HT	Head Teacher
JSS	Junior Secondary School
LGA/LGEA	Local Government (Education) Area
MLA	Measurement of Learning Achievement
N	Number (of cases observed, interviewed, assessed, etc)
NPC	National Population Commission
SBMC	School Based Management Committee
SE	Standard Error (used in its statistical sense)
SIP	School Improvement Programme (ESSPIN model)
SMO	Social Mobilisation Officer
SQS	State Quality Specialist (ESSPIN team member)
SSIT	State School Improvement Team
SSO	School Support Officer
SUBEB	State Universal Basic Education Board
TPD	Teacher Professional Development (fund of UBEC)
UBEC	Universal Basic Education Commission

Technical terms

The following terms are used in a specific technical, statistical sense in the Composite Survey:

Confidence interval

Confidence intervals of 95% are shown, for example when stating the proportion of primary school teachers who reach the specified teacher competency standard. The Composite Survey assesses the competence of only a sample of teachers out of the whole 'population' of primary school teachers in the state (and skills/knowledge of a sample of pupils, etc). Each time a sample is drawn from a population, it will inevitably vary to some extent from the 'true' level of competence of the wider population from which it has been drawn, and which it represents. The 95% confidence intervals express the range of values (between upper and lower limits) within which there is a 95% probability the true value for teacher competence of the population lies. Another way of expressing it is that if 20 different samples of teachers were drawn from the population and assessed in the same way, in 19 of those samples the level of teacher competence measured would fall between the upper and lower confidence limits reported.

Estimate

An 'estimate' as used in the Composite Survey is a value found by assessing a characteristic of the sample and using that to describe the population from which the sample has been drawn. It is not a guess, in the commonsense meaning of the word. It is an 'estimate' because any sample figure such as head teacher effectiveness is likely to differ slightly from that if all head teachers in the population in question were assessed.

Level of statistical significance

The report describes the different levels of performance for distinct groups within the study. Eg, in Jigawa teacher competence for teachers who have been exposed to the ESSPIN interventions is 87.1%* whereas it is 62.2% for those who have not. In such cases, an accompanying statement is shown along the lines that 'The asterisk * signifies that the mean estimate is significantly different from the mean estimate for the control school group… at the 0.05 level of statistical significance'. This means that we can be 95% sure that the difference observed between the samples of ESSPIN and control group teachers (87.1% - 62.2% = 24.9%) reflects an actual difference in measurable competence between the *populations* they represent (ie, the populations of all ESSPIN-supported and all control school teachers in Jigawa state respectively). It does ***not*** mean that the figures quoted are more than 5% apart so therefore they are statistically significantly different (see Annex A). Being statistically significantly different at a given level confers a high degree of assurance that a real effect has been picked up, although it does not in itself say what *caused* any difference observed. Conversely, failure to demonstrate a statistically significant difference does not necessarily mean that no difference exists—only that the observations did not satisfy the requirements for statistically significant difference to be reported. Typically, the size of the sample in relation to the population will influence whether a finding is significant or not, even if by eye there appears to be a relatively large or small difference. 95% is conventionally regarded as an exacting but realistic standard to use in social science; 90% or 99% are alternatives also used on occasion, in which case 0.10 and 0.01 would be the levels of statistical significance.

Abstract

In July 2012, representative stratified samples of public primary schools, head teachers, teachers and pupils were surveyed in the six Nigerian states where the DFID/UKaid-funded Education Sector Support Programme in Nigeria works. This report presents the findings with respect to ESSPIN outputs, outcomes and impact, in pilot schools, roll-out schools and control schools. Early indications are that ESSPIN-supported schools are associated with significantly more competent teachers, more effective school development planning, better functioning School Based Management Committees which reflect women and children's concerns, and (to a degree, in accordance with the ESSPIN theory of change) pupil learning outcomes. The origin of these differences requires analysis of a follow-up Composite Survey in 2014. Head teacher effectiveness and school inclusiveness may require more intensive or adjusted interventions, going forwards. This Technical Report describes the intended and achieved samples, the statistical tests and instruments used, the key data and findings, and next steps for the ESSPIN programme in view of the results and lessons learnt.

The report in one sentence

In schools supported by ESSPIN, levels of teacher competence, school development planning, governance and (to a lesser extent) inclusiveness and leadership are higher than in control schools and are associated with higher learning outcomes by pupils on several measures; however, standards of pupils' English literacy and numeracy remain very low state-wide and will require a longer intervention covering all schools and teachers before significant impacts on children's learning outcomes can be seen on a large scale.

Executive Summary

1. The ESSPIN Composite Survey (CS) process serves two main functions: periodically assessing the effects of ESSPIN's integrated School Improvement Programme (SIP), and reporting on selected indicators of the quality of education in the six ESSPIN-supported states. The CS addresses five Output pillars of the SIP, namely teacher competence, head teacher effectiveness, school development planning, school based management committee functionality and inclusive practices in schools. It also provides estimates of one Outcome indicator—school quality; and one Impact indicator—pupil learning achievement. The CS is wide-ranging but not exhaustive: it complements other ESSPIN/state monitoring and evaluation processes in areas such as institutional development, school enrolments and infrastructure. It brings together into a single exercise baseline surveys that were conducted by ESSPIN in 2010, hence 'composite' survey.

2. A two-stage sampling strategy was used in each state: public primary schools were sampled, and then teachers and p2/p4 pupils from within the sampled schools. Results were analysed according to 'programme phase', ie, Phase 1 schools which had participated in the ESSPIN pilot stage intervention supported by UKaid; Phase 2 schools scheduled to do so at state expense from mid-2012 onwards; and Control Schools which can be considered as a baseline. The ESSPIN intervention varies from state to state, but head teacher development, teacher development, school development planning/direct

school funding, school based management committee development and inclusive practices are at the core. Expectations of impacts observed state-wide at this point should be modest, in line with the limited number of schools covered, limited personnel participating within those school communities, and the short duration of the intervention at the time of CS1 (July 2012). In many places Kwara data are not included with the other states where programme-wide results are reported, as Kwara did not adopt the same two-phase strategy as other states.

3. Four data collection methods were used to complete ten questionnaires: interviews, record schedules, observation and oral/written tests. The total sample covered 595 schools/head teachers/SBMCs, 2,975 teachers and 9,520 pupils. Enumerators drawn from State School Improvement Teams and education officials were trained and then mobilised to collect the data over a six week period, with field supervision by NPC and ESSPIN. Data entry, cleaning and checking took longer than intended due to several technical problems. Each indicator of education quality was underpinned by a variety of objectively observable criteria. Estimates (values drawn from a sample to describe the population as a whole) are shown within 95% confidence intervals. In the case of Kano (and to a lesser extent Kaduna) some values are insufficiently precise to include in programme-wide aggregates. Mean estimates for ESSPIN-supported schools and non-ESSPIN supported schools are compared, and said to be significantly different at the 0.05 level (ie, where there is at least a 95% probability that the values for Phase 1 and Control Schools are actually different from one another). For certain numeracy measures, a comparison of the difference between 2010 and 2012 values for Phase 1 and Control Schools is possible. In most cases, such 'difference in differences' calculations will have to wait until the CS is repeated in 2014 and beyond. Although those CS 2012 results which show a significant difference between Phase 1 and Control Schools cannot necessarily be ascribed to 'the ESSPIN effect' (since other characteristics of schools in those categories could actually determine the difference), in the absence of evidence for an alternative cause it is reasonable to suppose that ESSPIN interventions are having the intended effect. This is particularly true of the Output and Outcome indicators but less likely with respect to Impact (children's learning outcomes) at this stage in the programme. The basis of allocation of schools to Phase 1 in each state is reported, to aid critical consideration of any selection bias.

4. Findings are presented by defining the ESSPIN logframe standard used to monitor each indicator; reporting the estimate of performance in all public primary schools (not only Phase 1 ESSPIN intervention schools); highlighting any significant differences between Phases; illustrating the results graphically; explaining how to read the graph; tabulating results; and providing a narrative commentary where relevant. More detailed findings and explanations are available in the six individual State Reports of the Composite Survey. The findings are listed in line with the ESSPIN intervention logic and results chain: teacher competence, head teacher effectiveness, school development planning, school inclusiveness, SBMC functionality, school quality and learning outcomes.

5. Results are summarised in the table below (Source: Composite Survey 2012, see Annex D for details):

Indicator	Phase 1 schools	Phase 2 schools	Control schools	All schools
Output indicators: (5 states)				
% competent teachers	80%*	72%	63%	67%
% schools with competent teachers	74%*	58%	39%	44%
% schools with effective head teacher	24%	14%	11%	13%
% schools with effective school development planning	24%*	9%	0%	3%
% schools that meet needs of all children (inclusive)	19%	16%	17%	17%
% schools with functioning SBMC	47%*	13%	19%	21%
% schools where SBMC reflects women's concerns	39%*	10%	7%	10%
% schools where SBMC reflects children's concerns	23%*	6%	4%	5%
Outcome indicator: (5 states)				
School quality	15%*	7%*	0%	2%
Impact indicators: (4 states)				
% p2 pupils with skills for reading comprehension	8%	9%	5%	9%
% p4 pupils with skills for reading comprehension	8%*	9%	2%	4%
% p2 pupils able to perform p2 arithmetic	19%*	16%	10%	12%
% p4 pupils able to perform p4 arithmetic	8%	7%	8%	7%

Estimates marked * are significantly different between Phase 1 (or 2) and Control Schools at the 0.05 level, ie, there is a high degree of certainty that ESSPIN intervention schools are significantly different from non-intervention schools.

6. **Six out of eight Output indicators plus the single Outcome school quality indicator are significantly better in ESSPIN-supported schools than in Control Schools** across the five states where a phased programme operates. **Two out of four Impact indicators are also significantly higher in ESSPIN supported schools**, although this is more surprising at this early stage in the programme lifetime.

7. At the more detailed level of criteria by which the standards are measured, significant differences were observed between ESSPIN-supported and Control Schools in 4/4 teacher competence criteria, 4/5 school development planning criteria, 10/10 SBMC criteria, 4/4 women inclusiveness criteria and 4/4 child inclusiveness criteria. At the state level (where sample sizes are smaller so significant results are harder to detect) 19/40 Output indicators showed a significant difference between ESSPIN-supported and Control Schools; a further 16/40 indicators were positive but not statistically significantly so; whilst the remaining 5/40 showed either little difference or a marginally negative result. Head teacher effectiveness and school inclusivity appear to be the least responsive to the ESSPIN intervention at this stage: findings which will be explored further and responded to operationally where needed.

8. The raw percentages for school quality and learning outcomes remain a cause for serious concern in all six states, especially in northern Nigeria, as seen in earlier surveys. But that finding does justify DFID Nigeria's growing education programme. And it is encouraging that there is now evidence from the Composite Survey that standards are higher in schools which have ESSPIN School Improvement Programme support. This comes after decades of ineffective projects and programmes. It should help states to advocate increased fund release for more schools to access SIP roll out, in order to reap the quality of *outcomes* benefits that states' massive annual investments in educational infrastructure, access, demand and school running costs deserve.

9. As children progress through school, an ever-increasing proportion falls behind grade-appropriate standards of numeracy and especially English literacy. The Composite Survey reveals that pupils commonly attain basic knowledge and skills but struggle with meaningful application of those capabilities. This is unsurprising given the limitations of the teaching force with respect to core curriculum-related knowledge and skills, as well as pedagogical understanding and weak school leadership for academic success. However, a smaller proportion of children in ESSPIN-supported schools remain in the lower performance bands, and a higher proportion is found in the upper bands, than in Control Schools (Annex E). This is strongly suggestive of an 'ESSPIN effect' on learning outcomes, although proof of causality will not be available until 2014 due to the need to eliminate possible selection bias as a rival explanation. ESSPIN will work with state partners to focus on the concepts of *progression* through the curriculum and *differentiation* of pupils' learning needs. Further investigations will be undertaken in action research mode to deepen state partners' professional engagement with these challenges, and develop appropriate ways to work effectively with the children most in need of help.

Introduction

Background and purpose of composite survey

10. ESSPIN's intended Impacts are 'More children achieve basic literacy and numeracy; and more children, especially girls, enter and complete basic education'. In Nigeria, basic education is delivered principally in public primary schools and junior secondary schools (JSS). The programme aims for Nigeria's own resources to be used more efficiently and effectively to improve participation and learning achievement of pupils across six focus states, through better teaching in schools of improving quality: Enugu, Jigawa, Kaduna, Kano, Kwara and Lagos. The composite survey (CS) is central to ESSPIN's internal monitoring and evaluation strategy and accountability for results. It is designed to provide robust evidence about the effects ESSPIN is having at key milestones and the end of programme (currently scheduled for July 2014).

11. The Composite Survey is so called because it is comprised of the essential elements of several hitherto separate baseline studies conducted by ESSPIN in 2009/10. Based on this first round of the CS, an interim note on pupil learning results was produced in November 2012, and individual State Briefs summarise the state-level results. This technical report complements these documents by providing an overview of the key findings across all six states combined, as well as technical information on the sampling strategy and analytical approach. Equally, the Composite Survey is intended to make a substantive contribution to knowledge about standards of education in Nigeria and what interventions work to improve the quality of schooling and children's learning.

12. The core of ESSPIN's intervention is a comprehensive school improvement programme (SIP) which combines various school and community level interventions with support for education systems reform at local, state and federal level. The SIP is derived from ESSPIN's theory of change, which proposes that an **integrated** set of interventions at various levels are needed to deliver sustained improvement in school quality, participation and learning achievement. Details of the SIP and ESSPIN's theory of change, as captured in a results chain detailing anticipated outputs, outcomes and impact, are in Annex B.

13. The design of the CS, as a key monitoring and evaluation tool, is driven by ESSPIN's theory of change and logframe. A survey concept paper is available which provides full details on the purpose, research questions, sampling strategy and risks (ESSPIN, July 2012). The remainder of this background section summarises the most relevant information.

14. There will be at least two rounds of the CS: the first took place in June/July 2012 (covered in this report) and the second will be in 2014. The CS has three main objectives:

- To **validate the school and community level estimates of ESSPIN's outputs**, which are being assessed by states as an integral part of ESSPIN-supported development processes. The outputs are the expected results of the school and community level interventions of the SIP (see Annex B), namely teacher development; head teacher development; school development planning; School Based Management Committee (SBMC) development; and inclusive practices in schools[1].

[1] The sixth school-level element of the SIP is infrastructure improvement. This will be addressed separately.

- To **provide estimates of one outcome and one impact indicator** which ESSPIN interventions are expected to contribute to, and which are central to the success of the programme, namely overall school quality (outcome) and pupil learning achievement (impact).
- To **demonstrate that ESSPIN has (or has not) contributed to improvements in school quality (outcome) and learning achievement (impact)**. Data from multiple rounds of the survey will be used to assess the impact of ESSPIN at various levels of the results chain[2].

15. The CS sampling strategy was driven partly by the design, timing and roll-out of the ESSPIN programme. A simplified programme timeline covering the six years of the programme is set out below:

    - inception phase: July 2008-July 2009;
    - implementation phase starts: July 2009;
    - baseline surveys, training and preparation of State School Improvement Teams (SSIT) and partner Civil Society Organisations (CSO): July 2009 to mid-2010;
    - interventions in **phase 1 schools start: mid-2010**;
    - interventions in **phase 2 schools, staggered start from around: mid-2012**;
    - interventions in phase 3 schools expected to start: mid-2013;
    - end of programme: July 2014.

16. ESSPIN operates in six focus states: Enugu, Jigawa, Kaduna, Kano, Kwara and Lagos. Enugu was the last to join the programme in March 2010. During the first three years of ESSPIN (2008 to 2011), the SIP has been introduced in each of the focus states. There has been some variation in the content, timing and coverage of the SIP in the different states (more details in next section) which has implications for the degree of change that can be expected in some key estimates.

17. The key features of the CS sampling strategy (see Annex A for details) are:

- Sufficient sample size to produce **state-level estimates** with sufficient precision to feed into ESSPIN state-level monitoring and evaluation**;**
- Restriction of **sample** to **public primary schools** (and mission schools in Enugu), as there are too few JSS intervention schools (and those few are in only two states);
- **Two stage sampling** of public primary **schools** (first stage), and then within each sample school, **teachers** and **grade 2 and grade 4 pupils** respectively (second stage).
- **Two rounds of the composite survey in 2012 and 2014** using the same sample of schools so that a panel data set can be used for impact analysis in 2014 (applying a difference-in-difference approach).
- **Inclusion of schools in the sample** from the ESSPIN **Monitoring Learning Achievement (MLA) baseline survey** which took place in 2010, to facilitate assessment of trends in learning achievement over time.

[2] Note: (1) 'impact' indicators are at the top of the results chain and refer to final measures of welfare for beneficiaries. In this report, the word 'impact' also refers to changes caused by the programme anywhere along the results chain. (2) Panel data analysis will seek to address the potential bias in single year estimates of programme impact arising from the non-random selection of schools into the programme. Using panel data removes any time-invariant characteristics which are correlated with performance, although there is still the possibility that time-varying school characteristics may be correlated with performance.

This analysis has been possible for one of the key measures of primary numeracy skills (presented later in this report).

- Schools divided into **six strata per state** (except Enugu where there were four strata[3]) according to two variables: (i) programme phase: phase 1, phase 2 or control[4]; and (ii) participation or not in the 2010 MLA survey.
  - This stratification means that separate estimates for each programme phase are available to feed into the impact assessment in 2014.
  - At the time of the 2012 survey, most phase 2 schools had yet to receive ESSPIN interventions and so the 2012 phase 2 estimates serve as a baseline for this group[5].

18. In interpreting the findings from the 2012 CS presented later in this report, it is important to note the following limitations:

- **Estimates are subject to sampling error**, and sampling errors are notably large for some of the pupil learning indicators in particular. This means that some of the estimates are not very precise, and 95% confidence intervals are wide.
- **Some of the observed differences in estimates for different groups** (e.g. states or ESSPIN phases) **are not statistically significant**. This means that any such differences should not be taken to mean that some groups are performing notably better than others. Differences which are statistically significant are marked in the tables using asterisks.
- Even if there is a statistically significant **difference between 2012 estimates for phase 1 and control schools, this should not necessarily be taken as a measure of the impact of ESSPIN** at this stage. The extent to which any statistically significant difference reflects the impact of ESSPIN's interventions rather than pre-existing characteristics of schools depends largely on whether the schools were randomly selected into the programme[6]. A systematic random selection of schools to participate in the ESSPIN pilot phase did not take place in any state—not that random selection would have been expected for a development programme as opposed to a research study. That said, there does not appear to have been any systematic attempt to target better performing schools (more details are in Annex D). So, in the absence of any other intervention found uniquely in the phase 1 schools, or any other distinguishing characteristics, the reader can form their own judgement about how reasonable it is to suppose that any observed differences between phase 1 schools and control schools with respect to aspects of schooling which ESSPIN worked upon, were or were not indeed *caused* by ESSPIN.

ESSPIN programme details

19. The SIP consists of five types of intervention (excluding infrastructure improvement which is not covered by the CS). The exact nature of the interventions delivered varies across the six states, but a

[3] Enugu had not identified its phase 2 schools at the time of the survey.

[4] Control schools have received no ESSPIN interventions and have not been selected for phase 2 of the programme.

[5] In two states (Kaduna and Lagos), some phase 2 schools started receiving interventions in the months before the survey took place, so strictly speaking the 2012 estimates will not provide pure baseline information for these groups.

[6] There is also the possibility of self-selection issues, which occur when specific types of pupils enrol in intervention schools precisely because of the intervention. Since increased participation is an explicit objective of ESSPIN, if observed this must count as a programme success, despite its interference with the CS research process.

brief generic description of the highest level of intervention which schools have received under each component is set out below.

- **Head teacher development**: up to 16 days of training for head teachers on academic leadership, school development planning, management of teachers, and working with the community.
- **Teacher development**: up to 16 days of training for teachers (selected from programme schools) on generic basic teaching skills, basic literacy teaching (initial reading skills), basic numeracy teaching (number concepts, and addition/subtraction), use of teaching aids, classroom organisation and encouraging children's self-esteem.
- **School development planning and school grants**: up to two school grants received in phase 1 schools, in consecutive years to be spent on priorities for school improvement included in a school development plan based on school self-evaluation.
- **SBMC development**: up to 16 days training for community members on establishing an SBMC covering, school planning and management, inclusive schools (focusing on the inclusion of women and children in decision making), resource mobilisation and financial processes, and child protection.
- **Inclusive practices**: woven throughout SBMC, teacher and head teacher development rather than treated as a stand-alone activity.

20. SIP interventions have been delivered by ESSPIN in a selected number of public primary schools in six states and in just a few JSS in two states. Since the CS covers public primary schools (not JSS), **Table 1** summarises the coverage of SIP in ESSPIN-supported public primary schools at the time of the survey in each state, and also the type of SIP intervention delivered. The introductory pages of the individual CS State Briefs provide details of the ESSPIN coverage and content of the SIP in each state.

Table 1: Summary of ESSPIN SIP interventions delivered in public primary schools by state and phase at the time of the 2012 composite survey

	Enugu	Jigawa	Kaduna		Kano	Lagos		Kwara	
Proportion of public primary schools covered by SIP by time of survey (%)	**Phase 1** 10%	**Phase 1** 6%	**Phase 1** 3%	**Phase 2** 8%	**Phase 1** 4%	**Phase 1** 7%	**Phase 2** 51%	**Not applicable** 83%	17%
Type of SIP intervention									
Head teacher development	✓	✓	✓	✓	✓	✓	✓	✓	✓
Teacher development	✓	✓	✓	✓	✓	✓	✓	✓	✓
School development planning & grant	✓	✓	✓		✓	✓			✓
SBMC development	✓	✓	✓		✓	✓			✓
Inclusive practices[1]	✓	✓	✓		✓	✓			✓

Note: (1) Inclusive practices is ticked automatically if SBMC, head teacher and teacher development have been ticked.

21. A number of points stand out from Table 1 which have implications for the interpretation of the results which follow in the key findings section:

- **Coverage of the SIP has been limited to a pilot scale during phase 1** of the programme in five states (except Kwara), ranging from 3-10% of public primary schools included in the programme. Under these circumstances, it is reasonable to have modest expectations about likely gains in overall state level

outcome and impact indicators—although it is state-wide results which are measured in the ESSPIN logframe.

- The **rollout of the SIP in Kwara has been different from the other states.** In Kwara all public primary schools received head teacher and teacher development interventions, and a comparatively high proportion of primary schools (17%) received all SIP interventions. The phase terminology is not applicable to Kwara because there are no control (non-intervention) schools.
- **Phase 2 interventions had already started in Kaduna and Lagos** at the time of the 2012 CS which means that phase 2 estimates for these states are not baseline estimates. In both states, head teachers and teachers had received four days and two days of training respectively during the six months prior to the 2012 CS fieldwork.

Survey instruments, sample and implementation

Survey instruments

22. The 2012 CS used ten different questionnaires. These were of four types:

- **Interview**: oral questions to individual respondents. For example head teachers were asked about their lesson observation practices. Often the questions require the respondent to produce written evidence of action.
- **Record schedule**: for collecting information from written records. This was used to collect information on primary 2 and 4 enrolment, and teacher numbers, which the data collectors used to draw the pupil teacher samples.
- **Observation**: for recording information on activities taking place during lessons.
- **Test**: written/oral questions given to pupils on English literacy and numeracy

A summary of the questionnaires is in Table 2.

Table 2: Questionnaires used in the 2012 Composite Survey

Questionnaire		Respondent	Type of instrument
1	School background	Head teacher	Record schedule (and sampling)
2	School leadership	Head teacher	Interview
3	School governance	SBMC chair & secretary	Interview
4	Teacher	Teacher	Interview
5	Lesson observation	Teacher	Observation
6	Classroom mapping	Teacher	Observation
7	P2 literacy	p2 pupil	Test
8	p4 literacy	p4 pupil	Test
9	p2 numeracy	p2 pupil	Test
10	p4 numeracy	p4 pupil	Test

Sample

23. In each of the six focus states, the intended sample for the 2012 CS was 105 primary schools, except in Enugu where phase 2 schools had not been identified at the time of the survey and the intended

sample was 70 schools. This gives a total sample size of 595 schools. In each school the head teacher (N~595) and five other teachers who had received ESSPIN-sponsored training (N~2,975) and five other teachers who had not received such training (N~2,975) were expected to be interviewed except in cases where a sample school had fewer than five teachers (of either category) in which case all teachers were interviewed. Four primary 2 pupils were to be assessed in literacy and four primary 2 pupils in numeracy in each school, and similarly for primary 4 pupils (N~ 9,520). More details on the intended sample size in each state are in Annex A (Table A.1 to Table A.6).

24. The actual sample interviewed was inevitably lower than intended; this occurs in almost all sample surveys for a variety of reasons. A comparison of the intended and actual number of records obtained for all six states combined is in Table 3. The highlighted column shows that 99% of schools and 96% of pupils were sampled as intended, which is a good response rate. The comparable figure falls to 72% for teachers, but this is largely explained by the fact that many schools had fewer than 10 teachers on the staff and so fewer teachers were interviewed by necessity[7]. The third and fourth columns provide information on one aspect of the quality of the records: missing data. All records could not be used in the analysis because of missing data, and these columns show the highest and lowest number of records that were actually used to generate key estimates. It is clear that missing data is a particular problem for school records (questionnaire 2 and 3). For at least one school-level key estimate, 17% of records had no data. A state breakdown of Table 3 is in Annex A (Table A.8).

Table 3: Sample units selected and interviewed for all 6 states combined

Unit	Intended sample	Total # of records	Highest # of records used[1]	Lowest # of records used[2]	# of records/ intended sample	Highest # of records/ total	Lowest # of records/ total
	[1]	[2]	[3]	[4]	[2]/[1]	[3]/[2]	[4]/[2]
Public primary schools	595	587	583	485	99%	99%	83%
Teachers	5,950	4,297	4,121	3,939	72%	96%	92%
Pupils (p2 & p4)	9,520	9,106	8,923	8,923	96%	98%	98%

Source: Composite Survey 2012. Notes: (1) Not all records could be used because of missing data for some questions; this column shows the highest number of records which were used in estimating the key indicators. (2) This column shows the lowest number of records which were used in estimating the key indicators because of missing data.

Survey implementation

25. Data collectors were trained using a two-step layered approach. Senior data collectors from each state were trained over a two week period in May 2012, using copies of the questionnaires and classroom and field-based practice. This group then trained the remaining data collection team during an intensive eight day period in early June. A total of 25 field teams undertook the survey in each state, each with two members. The data collectors were either members of the State School Improvement Teams (SSITs), School Support Officers (SSOs) or Social Mobilization Officers (SMOs), and so were experienced in observing classroom practice and other aspects of school management and governance. To try to ensure objectivity, the data collectors were not deployed to their current assigned areas and districts. Participation in the survey had significant professional development benefits for the data

[7] Another reason was the absence of the sampled teacher during all five days of the survey.

collectors with respect to their school support roles. This represents a lasting benefit in comparison to the alternative of buying data collection services from an external supplier.

26. Fieldwork took approximately six weeks to complete, from mid-June to late-July 2012. During this period there was serious civil unrest in throughout the north of Nigeria, which disrupted the field work in Kaduna and Kano in particular. The possibility that this traumatic environment affected some of the results in these states cannot be ruled out.

27. The senior member of each field team was responsible for data quality during the collection process. They checked questionnaires and ensured that sampled units were interviewed. In addition, ten roving quality control officers (hired from the National Population Commission (NPC)), checked up on the work of the field teams by verifying sampling procedures and checking the accuracy of data collection. These officers filled in separate quality control questionnaires. Members of ESSPIN's state and Abuja teams carried out spot checks, monitored fieldwork progress and provided on-the-spot guidance throughout the period.

28. Data were entered in Microsoft Access, and a detailed checking and cleaning process was undertaken by a team of experienced survey analysts in liaison with the NPC who were responsible for data entry. The process of data checking and cleaning took much longer than expected for a number of reasons, including: (i) data entry was slower than expected, and a considerable number of questionnaires had to be re-entered; (ii) some of the identifier codes for the various units sampled, particularly teachers, were not completed correctly in the field, and had to be rectified manually; (iii) the design of the questionnaires did not include some standard features, including skip codes when questions were not applicable; the problems arising from this had to be fixed manually[8]. These issues, and others, will be documented as part of a review process to ensure that improvements are made for the 2014 CS round.

29. One other data issue to highlight is that the intended stratification of teachers in the phase 1 schools into two groups: teachers trained under the SIP and teachers not-trained under the SIP, was not possible in the field. The majority of teachers were unable to distinguish SIP training from other types of in-service training they have received during the same period. ESSPIN deliberately sought to ensure that the SIP training is part of the state in-service training delivery system, and thus did not attempt to tag it to the 'ESSPIN or SIP brand'. This appears to have been a successful strategy and again is characteristic of surveys conducted in applied programme intervention mode rather than pure research mode. In interpreting the findings on teacher competence in the phase 1 schools later in the report, it is important to bear in mind that only a maximum of six teachers in each school participated in SIP training directly. For small schools this would have been all teachers, whilst for large schools it was a very small percentage.

[8] The non-standard coding features of the questionnaires were deliberate. This decision was taking because the data collectors were not experienced survey enumerators, and it was decided that this type of coding would be too complex. Such polylemmas are characteristic of social survey research in contexts such as this.

Approach to survey analysis

Key estimates

30. The CS was explicitly designed to produce estimates of nine of ESSPIN's logframe indicators covering school-level outputs, outcomes and impact. These are set out in Annex C. The five **output indicators** relate to the five areas of the SIP set out in the previous section. A standard (or benchmark) has been set for each SIP area, namely:

    - competence of teachers to teach literacy and numeracy,
    - effectiveness of head teachers,
    - use of School Development Plans,
    - functioning School Based Management Committees, which
    - reflect women and children's concerns.

    The indicator is defined as the proportion of units (schools or teachers) which meet the standard. A new standard and indicator has also been developed which directly relates to the 'inclusive practices' SIP area. This complements the five logframe output indicators. Each standard has a number of criteria underpinning it, and these are clearly described in the relevant section in the Key Findings.

    The logframe **outcome indicator** measures overall school quality. The logframe **impact indicators** measure p2 and p4 learning achievement in literacy and mathematics and are expressed as the proportion of pupils in all public primary schools of the state who reach the designated standard. The exact test questions used to compute these indicators are also in Annex C.

Survey weights

31. In order to ensure that estimates are representative of the population of interest (for example in each state: public primary schools, public primary teachers, public primary p2 and p4 pupils), analytical weights were applied. Analytical weights are calculated as the inverse of the selection probability of each unit. Annex A sets out the formulae used to calculate the selection probabilities and the weights.

Sampling errors, confidence intervals and design effect

32. Estimates derived from samples are characterised by sampling errors. In other words, the fact that we do not obtain the information that we want from the entire population but from a random subset, means that statistical measures of interest, such as the mean, are not calculated with perfect precision but are likely to fall within a range of values called a confidence interval. In this report, mean estimates for logframe indicators are presented graphically with their 95% confidence interval[9]. Estimates presented for teachers and pupils in each state are based on well over 100 observations (and commonly many more), while estimates for school-level indicators are based on more than 30

[9] A dot represents the estimated mean, with lines called 'whiskers' extending a certain distance above and below the dot, to denote the confidence interval ie, the range of values between which there is a 95% likelihood that the true mean value for the population in question lies.

observations. Annex D contains mean estimates, standard errors (used to compute confidence intervals), and sample sizes, for all of the logframe indicators.

33. Extreme caution is needed in interpreting estimates which have very large confidence intervals. The estimates of the pupil learning indicators in Kano, and in some cases Kaduna, display large confidence intervals and are very imprecise. This is caused by an unexpectedly large 'design effect'. The design effect is the loss of effectiveness due to the use of cluster sampling rather than simple random sampling. In the case of the CS, primary pupils were not selected randomly from a list of all pupils in each state, because no such list exists. Instead, they were selected from within the sampled schools. In simple terms, selecting an additional unit from the same cluster (in this case, the same school) adds less information than a completely independent unrelated selection would. If intracluster correlation for the statistic in question is high then this drives up the design effect. This could happen if pupils within the sampled schools gave similar (or the same) answers. The design effect reduces the effective sample size by the 'DEFF factor'. In general for a well-designed study design effects would be less than 3. For some of the mean estimates of pupil learning for Kano, the DEFF is more than 50. For this reason, apart from presenting the logframe indicators for information, Kano has been excluded from the analysis of pupil learning in this report. It would require somewhat lengthy further analysis to determine what the cause of the high design effect in Kano was. ESSPIN will report the results of any such further investigations in due course, if the capacity to undertake them can be identified—not least to avoid a repetition of this problem in future rounds of the survey.

Statistical tests

34. In the analysis of the CS, it is of particular interest to know whether schools, teachers and pupils in phase 1 are performing better than those in the control group on the logframe indicators. In other words, we want to test whether the difference in the mean estimate between the two sub-populations is statistically significant. The results of these tests are denoted by an asterisk on the phase 1 estimates in the results tables which follow if there is a significant difference at 5% level of significance. (The data in Annex D also uses this notation.) For details of the standard formula used to compute this test, see Annex A. As highlighted in paragraph 18 above, statistically significant differences between phase 1 and control group estimates cannot necessarily be attributed to the impact of ESSPIN interventions at this stage. It would be reasonable to demand an equally high standard of evidence for any alternative theory of what had caused any such differences, though.

35. The second type of statistical test used in this CS analysis is a difference-in-difference test. The aim is to establish whether there is a statistically significant difference between changes over time in mean estimates of a particular indicator for two sub-populations. In the CS analysis we apply this test to the difference between the gains in the mean estimates of the numeracy logframe indicator between 2010 and 2012 for pupils in the phase 1 schools compared with the control schools. For details of the standard formula used to compute this test, see Annex A.

Findings from the composite survey

36. This section presents the key findings from the CS. It is structured according to the logframe hierarchy and the theory of change embedded in the ESSPSIN results chain, moving from an analysis of programme outputs, through outcome to impact indicators. For the output indicators, separate estimates are shown for each phase (phase 1, phase 2 and control) for five states combined (Enugu, Jigawa, Kaduna, Kano and Lagos). Estimates for Kwara are displayed separately because, as explained in the section above, the phasing of the programme has been different there (so there are no phase 2 or control schools). Separate estimates for each state are available in the State Briefs.

37. The CS fieldwork took place after two years of programme interventions and so it is reasonable to expect statistically significant differences between phase 1 and control group output indicator estimates at this stage, given the information we have on the way the phase 1 schools were selected, and the fact that the output indicators are based on criteria which have been explicitly targeted in the design of the SIP interventions.

38. The outcome indicator of school quality is actually a composite measure of some of the output indicators. For this reason, it is also realistic to expect a significant difference between phase 1 and control school estimates for this indicator. This section also includes an estimate of the outcome indicator for all six states combined, representing all public primary schools, as required in the ESSPIN logframe.

39. The analysis of pupil learning (impact indicators) is mostly based on five states combined (Enugu, Jigawa, Kaduna, Kwara and Lagos). Kano is excluded because its estimates are very imprecise due to the design effect explained in paragraph 35 above. The analysis of pupil learning indicators by phase therefore excludes both Kano and Kwara. This section also includes estimates of the impact indicators for all six states combined, representing all public primary schools (as required in the overall ESSPIN logframe). For transparency, all figures are reported in the Kano State Brief, irrespective of the reservations over the design effect.

40. Compared with the lower levels of the results chain, it is less likely that at this stage of programme implementation a significant difference between pupil learning in phase 1 and control schools would be observed[10]. ESSPIN's interventions are designed to ultimately contribute to improvement in pupil learning but this will take time, partly because the interventions work indirectly (with SSOs, SBMCs, head teachers and teachers, not pupils); partly because the pilot phase 1 covered only a small proportion of teachers in phase 1 schools; and partly because they take a number of years to have a cumulative effect measurable in terms of all pupils' learning. For example, p2 and p4 pupils assessed in phase 1 schools may not have been taught by a SIP trained teacher. Some SIP trained teachers should be present in the school to share knowledge and skills, but not all early grade children will receive the benefit of their training directly.

[10] As noted earlier, even though there is a statistically significant difference, this cannot necessarily be attributed to ESSPIN at this stage.

41. In the analysis which follows, the standards and underlying criteria for each indicator are clearly defined at the start of each subsection for easy reference. In the case of pupil data, the headline indicator of the proportion of children who have met the logframe standard is further analysed in terms of the distribution of test scores by domains of learning, to contribute to a more nuanced understanding of the educational implications of the data. Annex D contains the mean estimates, standard errors and sample sizes for all the data shown in the charts.

Teacher competence

42. The teacher competence logframe indicator is based on four criteria. These are summarised below.

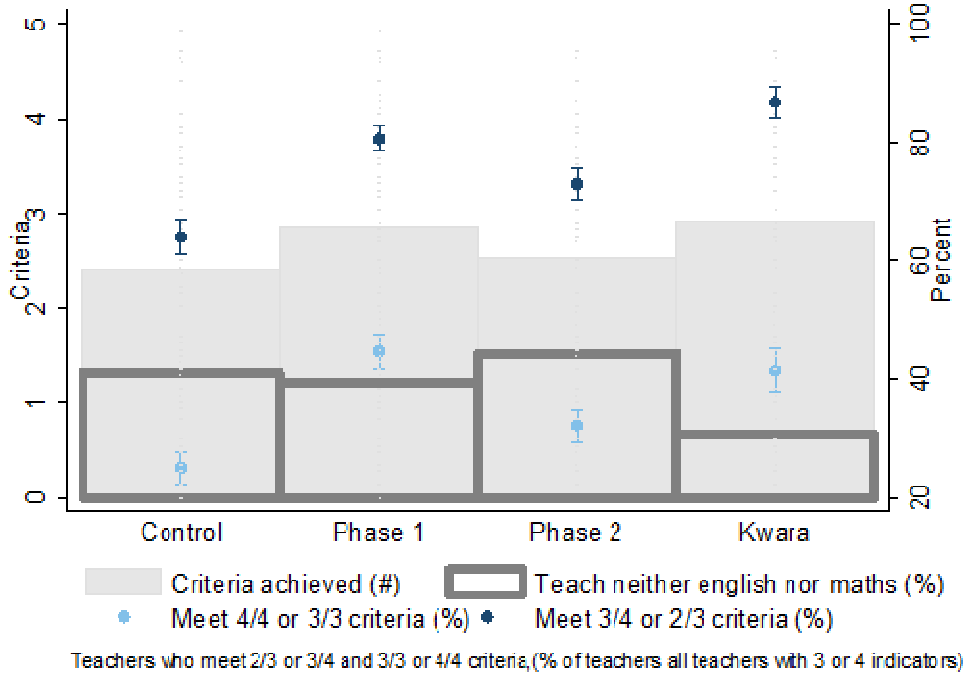
Logframe standard for teacher competence

A teacher must meet three out of four of the following criteria to meet the competence standard if he/she teaches English and/or maths. Teachers of other subjects must meet two out of three criteria (excluding 1 below).

1) *Knowledge of English or mathematics curriculum (based on interview)*
2) *Use of at least one teaching aid during lesson observation*
3) *Greater use of praise than reprimand during lesson observation*
4) *Class organisation: assigning individual or group tasks at least twice during lesson observation (or for two contiguous five-minute blocks)*

43. Across five ESSPIN states, an estimated 67% of teachers demonstrate competence based on the criteria set out above. Within this group of teachers, **the proportion who met the overall standard is significantly higher in phase 1 schools (80%) compared with teachers in control schools (63%).** The estimate for teachers in phase 2 schools is 72%. In Kwara 85% of teachers demonstrate competence. See Figure 1 below and Annex D for details of the estimates.

Figure 1: Proportion of teachers who meet the competence standard by phase for 5 states, and for Kwara (%)



Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of teachers who meet three out of four criteria (if they teach English and/or maths) or two out of three criteria (if they teach neither English nor maths) with a 95% confidence interval. This is the teacher competence indicator.
- The light blue dots represent the percentage of teachers who meet all criteria (4 out of 4 criteria if they teach English or maths, or 3 out of 3 criteria if they teach neither English nor maths), with a 95% confidence interval. This is the teacher proficiency indicator.
- The light grey bars indicate the average number of criteria met by teachers.
- The dark grey bars indicate the percentage of teachers who teach neither English nor maths.

44. Using the more rigorous **teacher proficiency** standard (as defined under the chart above), again across the five states, **teachers in phase 1 significantly outperform teachers in control schools**.

45. Looking at the results for the individual criteria which underpin the competence standard (Table 4); it is clear that **teachers in phase 1 are performing significantly better on all four criteria than teachers in control schools**. On two of the criteria: knowledge of the curriculum and use of praise/reprimand, teachers in phase 2 schools also perform significantly better than the control group.

Table 4: Proportion of teachers who meet each of the teacher competence criteria by phase for 5 states, and for Kwara (%)

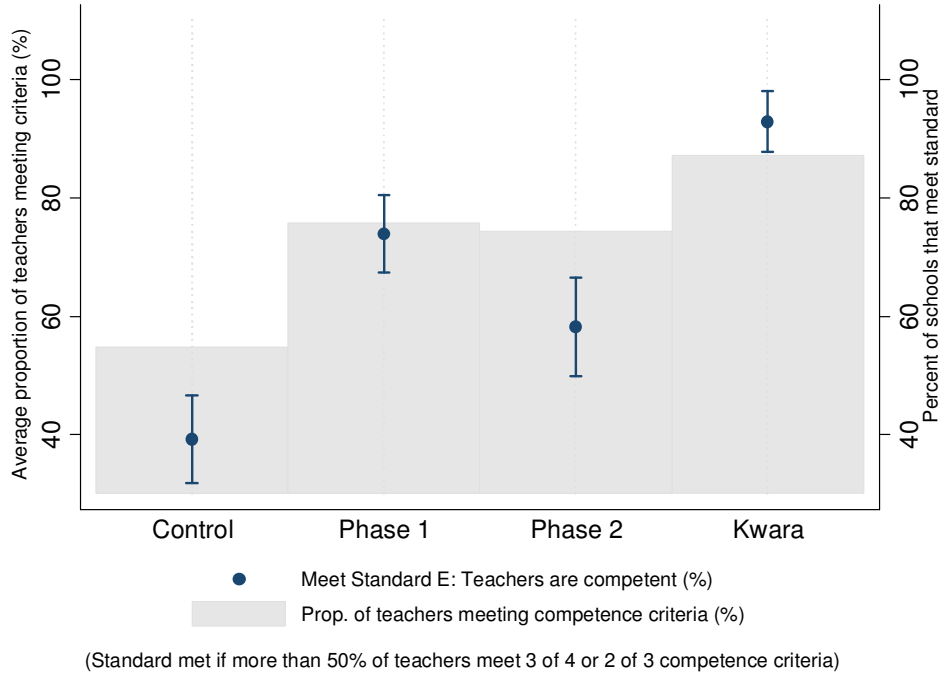
	5 States[1]			Kwara
Criteria for teacher competence	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
Knowledge of curriculum	50.5	61.0*	60.2*	79.0
Use of teaching aids	86.9	92.3*	85.7	89.4
Greater use of praise than reprimand	63.7	76.5*	76.2*	83.9
Class organisation	53.3	70.2*	53.4	61.1

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different from the mean estimate for the control group at the 0.05 level of statistical significance.

46. As well as tracking the proportion of teachers who meet the competence standard, it is also important to understand the distribution of these skills across schools. Teacher competence is an important element of the quality of a school, and, as such, ESSPIN's standard for school quality (an outcome indicator) includes this. For a *school* to meet the teacher competence standard: 'more than 50% of teachers in a school must meet the competence standard'. Figure 2 presents estimates for this school-level indicator (denoted by blue dots). The grey blocks in the chart repeat the teacher-level estimate of competence already presented.

47. For five states combined, some **74% of phase 1 schools meet the teacher competence standard compared with 39% in control schools, a difference which is statistically significant** at 5% level. Some 58% of phase 2 schools meet the teacher competence standard, while in Kwara the comparable figure is 93%.

Figure 2: Proportion of schools that meet the teacher competence standard by phase for 5 states, and for Kwara (%)



(Standard met if more than 50% of teachers meet 3 of 4 or 2 of 3 competence criteria)

Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools in which more than 50% of teachers meet three out of four criteria (if they teach English and/or maths) or two out of three criteria (if they teach neither English nor maths) with a 95% confidence interval. This is the school-level teacher competence standard.
- The light grey bars indicate the average proportion of teachers who meet the teacher competence criteria.

Head teacher effectiveness

48. The head teacher effectiveness logframe indicator is based on seven criteria. These are summarised below.

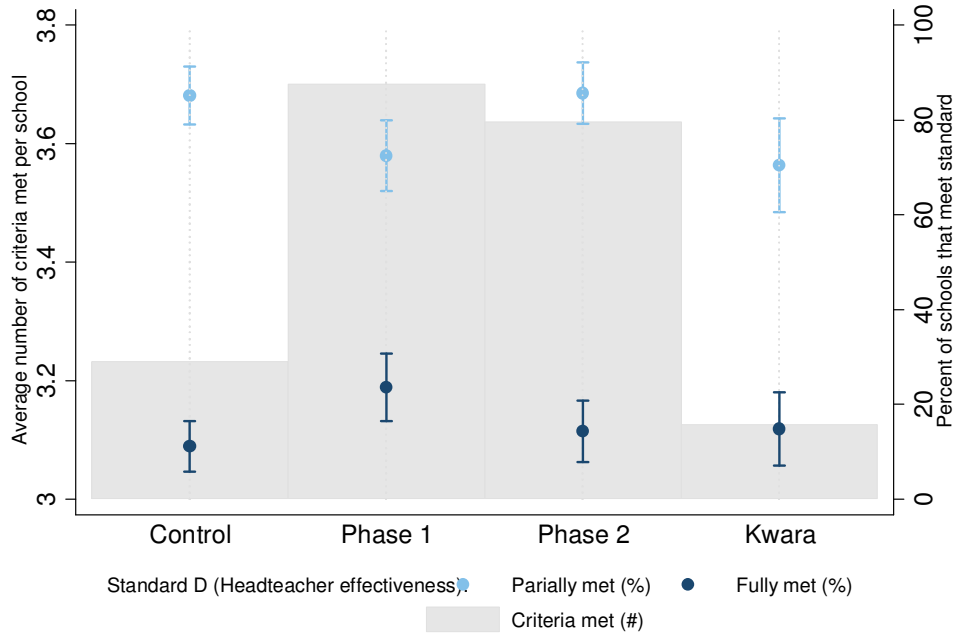
Logframe standard for head teacher effectiveness

A head teacher must ensure that five out of seven of the following criteria are met in order to meet the head teacher effectiveness standard

1) *Carry out two or more lesson observations in the past two weeks*
2) *Hold four or more professional development meetings since the start of the 2011-12 school year (NB: survey took place more than 9 months into the school year)*
3) *School has a teacher attendance book and head teacher recalls at least two actions taken to promote teacher attendance*
4) *Clear school opening time: more than 50% of pupils sampled agree on the school opening time and more than 50% of teachers sampled agree on the school opening time*
5) *More than 50% of classes are in their classroom with their teacher within 30 minutes of school opening time*
6) *Length of morning break is 35 minutes or less, except in Enugu when it must be 15 minutes or less*
7) *More than 50% of lessons observed finished within 5 minutes of a standard 35 minute lesson duration (i.e. between 30 and 40 minutes long)*

49. Across five ESPPIN states, an estimated 13% of all public primary schools meet the head teacher effectiveness standard. Within this group of schools, the **proportion who met the standard is significantly higher in phase 1 schools (24%) compared with control schools (11%)**. The estimate for phase 2 schools is 14%. In Kwara 15% of schools have an effective head teacher, based on the criteria above. See Figure 3.

Figure 3: Proportion of schools where head teacher meets the effectiveness standard by phase for 5 states, and for Kwara (%)



(Standard met partially if school meets 2/3/4 criteria, fully if school meets 5/more criteria)

Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools meeting 5 or more of the head teacher competence criteria, with a 95% confidence interval. These schools meet the head teacher competence standard.
- The light blue dots represent the percentage of schools meeting 2 to 4 of the head teacher competence criteria (not more), with a 95% confidence interval. These schools partially meet the head teacher competence standard.
- The light grey bars indicate the average number of criteria met by schools.

50. The head teacher competence standard is composed of seven different criteria listed in Table 5 below. The disaggregated results show that the difference between phase 1 and control schools is driven largely by a single criterion (#1), which assesses whether the head teacher carries out at least one lesson observation each week. Some **34% of phase 1 schools meet this criterion, compared to just 5% of control schools. The difference is statistically significant** at 5% level. Phase 1 schools also perform better than control schools on the remaining six indicators, although the difference is not statistically significant. Phase 2 schools also tend to perform better than control schools on all but two criteria (#2: head teacher holds at least 2 professional development meetings per term, and #4: More than 50% of pupils and more than 50% of teachers know the school opening time). However, these differences are only significant at the 5% level for two of the indicators (#3: head teacher has more than one strategy for promoting teacher attendance, and #6: the length of the long morning break was no more than 35 minutes or 15 minutes for Enugu state)

Table 5: Proportion of schools where head teacher meets each of the effectiveness criteria by phase for 5 states, and for Kwara (%)

	5 States[1]			Kwara
Criteria for effective head teacher	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
Lesson observation	5.1	33.6*	11	22.5
Professional development meetings	12.1	16.2	8.2	21.9
Teacher attendance	73.9	85.3	88.6*	57.2
School opening time	53.3	46.7	48.9	43.1
Timing of first lesson	68.8	77.8	81.1	85.2
Length of break	77.1	73.9	88.9*	86.6
Length of lesson	27.3	36	30	5

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different from the mean estimate for the control group at the 0.05 level of statistical significance.

School development planning

51. The school development planning logframe indicator is based on five criteria. These are summarised below.

Logframe standard for effective school development planning

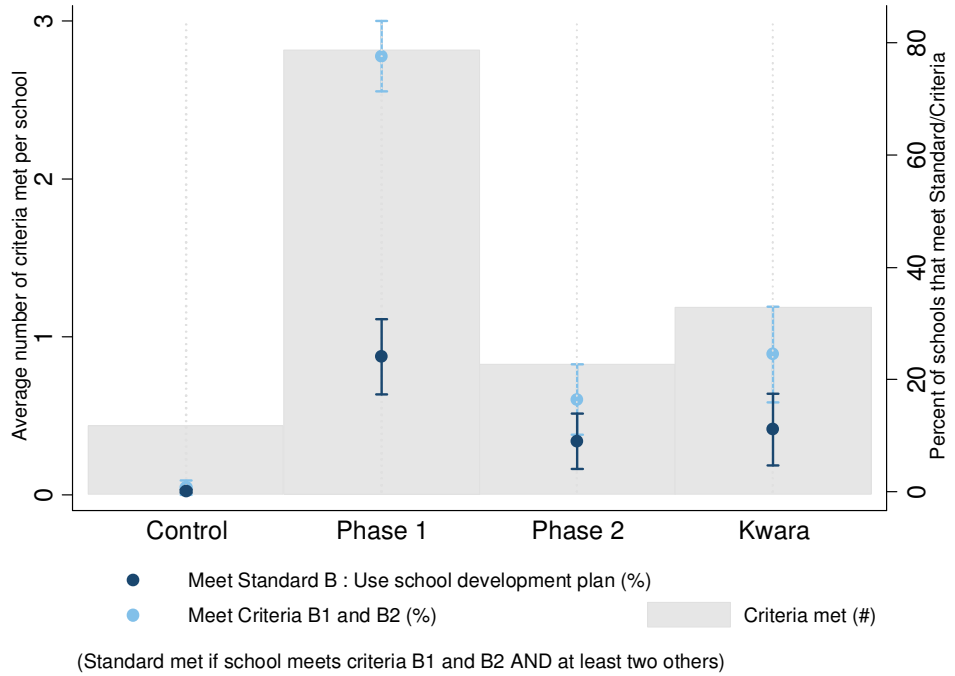
The school must meet criterion one and criterion two listed below and at least two out of three of the remaining criteria in order to meet the effective school development planning standard

1) *Written evidence of school self-evaluation process for 2011-12 school year*
2) *School Development Plan (SDP) for 2011-12 school year available*
3) *SDP contains three or more activities which aim to strengthen teaching and learning*
4) *Physical evidence of four or more activities from SDP having been carried out*
5) *Cashbook is up-to-date (balanced in the last 60 days)*

52. Across five ESSPIN states, only 3% of all public primary schools meet the school development planning standard. Within this group of schools, the **proportion who met the standard is significantly higher in phase 1 schools (24%) compared with control schools (0%)**. The estimate for phase 2 schools is 9%.

53. In Kwara an estimated 11% of schools provided evidence of effective school development planning, based on the criteria above. It has to be noted, however, that only a small proportion of schools (17%) in Kwara received school grants to *implement* their SDPs, which drives satisfaction of the criterion. For all other Kwara schools, the head teachers were trained but there were no state or programme resources to implement SDPs, and their SBMCs had not received training or support.

Figure 4: Proportion of schools which meet the effective school development planning criteria by phase for 5 states, and for Kwara (%)



Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools that meet criteria #1 (school self-evaluation involved the SBMC) and #2 (school development plan completed) and at least two other school development planning criteria, with a 95% confidence interval. These schools meet the standard for school development planning.
- The light blue dots represent the percentage of schools that meet criteria #1 and #2 (or more), with a 95% confidence interval.
- The light grey bars indicate the average number of development planning criteria met per school.

54. The standard for effective school development planning is composed of five different criteria presented in Table 6 below. The disaggregated results show that **phase 1 schools perform better than control schools on all five criteria. The difference between phase 1 and control schools is statistically significant at the 5% level for all but one criterion** (#5: the school has an up-to-date cash book). Phase 2 schools perform significantly better than control schools on two criteria (#1: School self-evaluation involved the SBMC, and #3: the school development plan contains at least two activities related to raising achievement).

Table 6: Proportion of schools which meet each of the school development planning criteria by phase for 5 states, and for Kwara (%)

	5 States[1]			Kwara
Criteria for effective school development planning	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
SBMC involvement in self-evaluation	10.0	77.6*	28.0*	25.4
SDP availability	16.5	93.9*	25.6	40.8
SDP has learning achievement activities	4.7	61.0*	17.5*	30.3
SDP activities completed	2.2	24.1*	4.8	11.1
School cashbook up-to-date	15.0	23.5	8.9	13.7

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different from the mean estimate for the control group at the 0.05 level of statistical significance. (3) In Kwara, where only 17% schools received SBMC support, any figure approaching (or exceeding) 17% is a positive indication.

School inclusiveness: meeting the needs of all pupils

55. The school inclusiveness indicator is based on four criteria. These are summarised below.

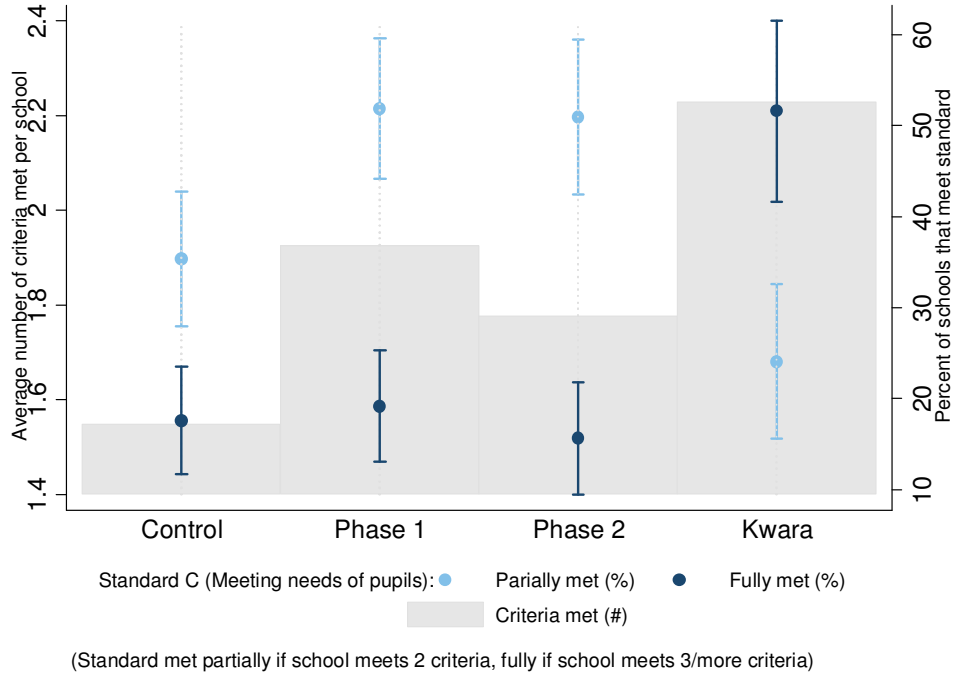
Standard for school inclusiveness (meeting needs of all pupils)

The school must meet at least three of the four criteria listed below in order to meet the school inclusiveness standard. The standard is partially met if two criteria are met.

1) *Head teacher states more than three actions that he/she has taken to improve pupil attendance*
2) *SDP contains two or more activities which aim to improve access*
3) *More than 50% of teachers observed provided evidence of using two or more assessment methods (marked class test, marked pupil workbook, or graded examination paper)*
4) *More than 50% of teachers observed met the spatial inclusion criterion (defined as engaging with at least one pupil from four different areas of the classroom during a lesson) and more than 50% of teachers observed met the gender inclusion criterion (defined as engaging with boys and girls proportionally to their presence in the classroom within a 10% margin. For example if the class contains 50% girls then teachers who engage with girls between 60% and 40% of total engagements will meet the criterion).*

56. Across five ESSPIN states, 17% of all public primary schools meet the school inclusiveness standard. Within this group of schools, the proportion who met the standard is similar in phase 1 schools (18%), control schools (19%) and in phase 2 schools (16%). The differences are not statistically significant. By contrast**, schools in Kwara are doing much better on this standard**, such that more than half of schools there are deemed to be inclusive, based on the criteria above.

Figure 5: Proportion of schools which meet the school inclusiveness standard by phase for 5 states, and for Kwara (%)



(Standard met partially if school meets 2 criteria, fully if school meets 3/more criteria)

Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools that meet at least 3 of the school inclusiveness criteria, with a 95% confidence interval. These schools fully meet the standard for meeting the needs of all pupils.
- The light blue dots represent the percentage of schools that meet 2 of the school inclusiveness criteria (not more), with a 95% confidence interval. These schools partially meet the standard for meeting the needs of all pupils.
- The light grey bars indicate the average number of criteria met by schools.

57. The school inclusiveness standard is composed of 4 different criteria. The disaggregated results presented in Table 7 below shows that Phase 1 schools perform better than control schools on 3 out of 4 inclusiveness criteria. However, the difference is only statistically significant at the 5% level for one of those indicators (#2: School development plans have more than one activity related to improving access). Phase 2 schools only perform significantly better than control schools with respect to one criterion (#3: The majority of teachers use more than one method to assess learning).

Table 7: Proportion of schools which meet each of the school inclusiveness criteria by phase for 5 states, and for Kwara (%)

	5 States[1]			Kwara
Criteria for school inclusiveness	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
Pupil attendance	53.9	62	54.7	73
SDP activities related to access	2.3	23.9*	6.7	21.4
Teacher assessment methods	61.8	69.4	89.9*	75.4
Teacher spatial and gender inclusiveness	36.3	36.1	26	53

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different from the mean estimate for the control group at the 0.05 level of statistical significance.

SBMC functionality and inclusiveness

SBMC functionality

58. The school based management committee functionality logframe indicator is based on ten criteria. These are summarised below.

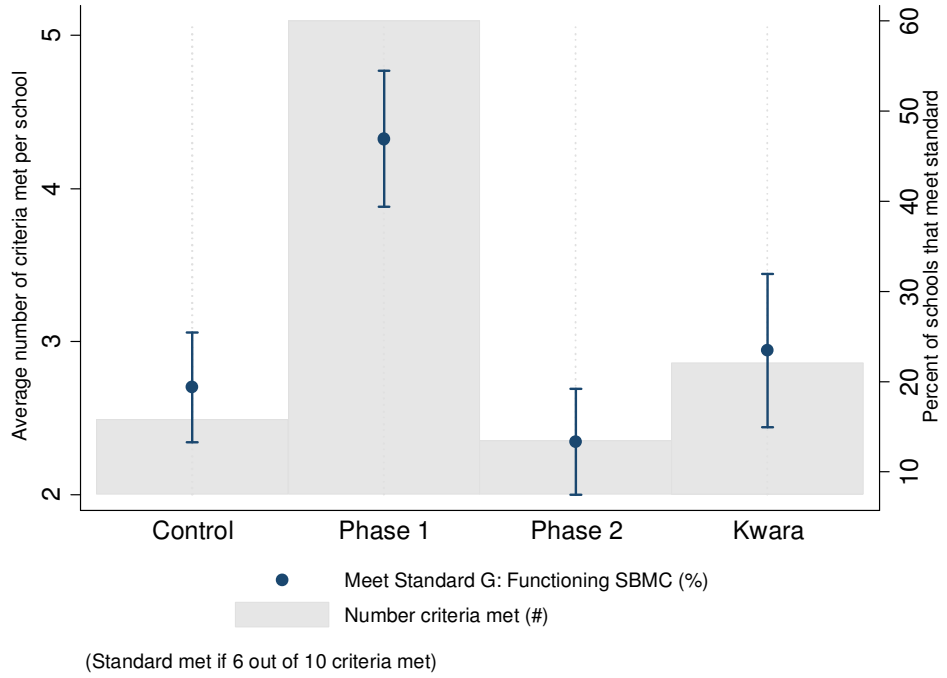
Logframe standard for SBMC functionality

The school must meet at least six of the ten criteria listed below in order to meet the SBMC functionality standard for the 2011-12 school year:

1) *Two or more SBMC meetings have taken place since the start of the 2011-12 school year (written evidence)*
2) *Two or more wider community members (not parents, teachers or pupils) involved in school development planning (written evidence)*
3) *SBMC conducted awareness raising activities (written or oral evidence)*
4) *SBMC took steps to address exclusion (written or oral evidence)*
5) *SBMC networked with CBOs, traditional or religious institutions, or other SBMCs (written or physical evidence)*
6) *SBMC interacted with local government education authorities on education service delivery issues (written or physical evidence)*
7) *SBMC women's committee exists (written or physical evidence)*
8) *SBMC children's committee exists (written or physical evidence)*
9) *SBMC contributed resources for the school (written or physical evidence)*
10) *SBMC chair visited the school at least three times since the start of the 2011-12 school year (written evidence)*

59. Across five ESSPIN states, about one-fifth of all public primary schools meet the SBMC functionality standard. Within this group of schools, the **proportion who meet the standard is significantly higher in phase 1 schools (47%) compared with control schools (19%)**. The estimate for phase 2 schools is 13%. In Kwara 23% of schools have a functioning SBMC, based on the criteria above. Phase 2 schools and schools in Kwara do not perform significantly better than control schools. However, it must be noted that only 17% of schools in Kwara were supported to establish SBMCs under the SIP.

Figure 6: Proportion of schools which have a functioning SBMC by phase for 5 states, and for Kwara (%)



Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools that meet at least 6 of the criteria for SBMC functionality, with a 95% confidence interval. These schools meet the SBMC functioning standard.
- The light grey bars indicate the average number of criteria met by schools.

60. The SBMC functionality standard is composed of ten separate criteria, which are listed in Table 8 below. **Phase 1 schools perform significantly better than control schools on all ten criteria**. There are no significant differences between estimates for phase 2 and control schools.

Table 8: Proportion of schools which meet each of the SBMC functionality criteria by phase for 5 states, and for Kwara (%)

	5 States[1]			Kwara
Criteria for SBMC functionality	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
SBMC meetings	29.8	71.5*	32.5	35.5
Community members involvement	13.3	53.6*	12.5	32.7
Awareness raising activities	34.2	59.4*	34.3	41.8
Addressing exclusion	34.9	55.0*	26.7	40.3
Networking	13.9	30.6*	11.5	16.9
Interaction with LGA	14.3	41.8*	27.3	22.8
Women's committee	13.1	47.8*	18.5	18.6
Children's committee	25.9	46.7*	14.2	12.3
Resource contribution	44	62.1*	34.4	39.3
SBMC chair school visits	24.4	43.9*	23.1	26.4

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the control group at the 0.05 level of statistical significance.

SBMC women's inclusiveness

61. The SBMC women's inclusiveness logframe indicator (SBMC reflects women's concerns) is based on four criteria. These are summarised below.

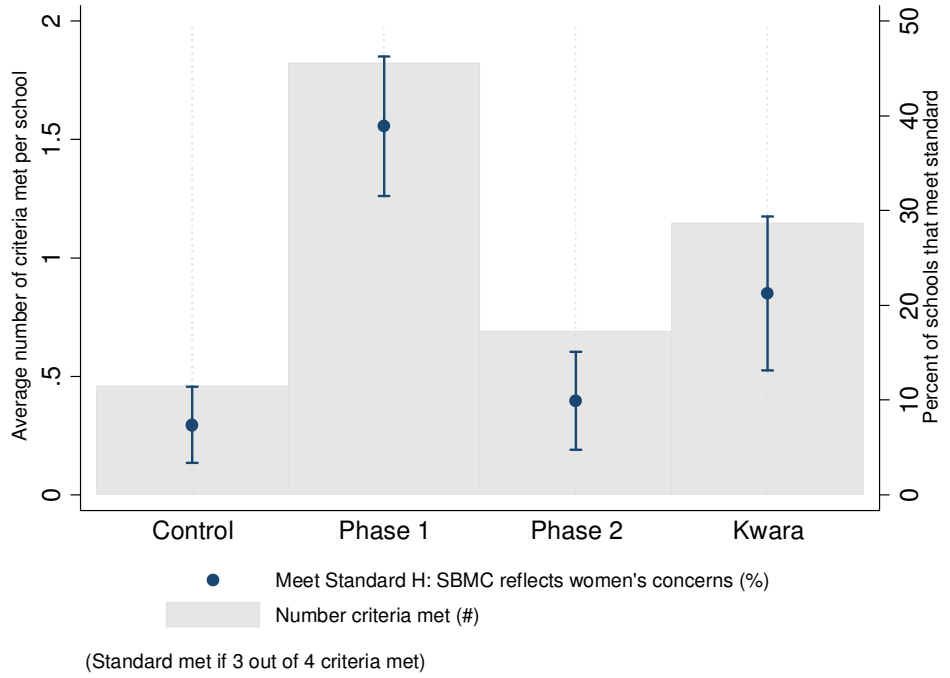
Logframe standard for SBMC women's inclusiveness

The school must meet at least three of the four criteria listed below in order to meet the SBMC women's inclusiveness standard for the 2011-12 school year:

1) *At least one woman attended two or more SBMC meetings (written evidence)*
2) *Female member of SBMC raised at least one issue at SBMC meetings (written evidence or oral evidence from female member of SBMC)*
3) *At least one issue raised by a female member at an SBMC meeting led to action (written, physical or oral evidence from female member of SBMC)*
4) *At least one SBMC women's committee meeting took place and committee has a female leader (written evidence)*

62. Across five ESSPIN states, one in ten of all public primary schools meets the SBMC women's inclusiveness standard. Within this group of schools, the **proportion who meet the standard is significantly higher in phase 1 schools (39%) compared with control schools (7%)**. The estimate for phase 2 schools is 10%. In Kwara 21% of schools have a SBMC which reflects women's concerns, based on the criteria above.

Figure 7: Proportion of schools where SBMC reflects women's concerns by phase for 5 states, and for Kwara (%)



Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools that meet at least 3 criteria for women inclusiveness, with a 95% confidence interval. These schools meet the standard for reflecting women's concerns .
- The light grey bars indicate the average number of criteria met by schools.

63. The SBMC women inclusiveness standard is composed of four different criteria, listed in Table 9 below. The results presented in this table show that **phase 1 schools perform significantly better than control schools on all four criteria**. The differences between phase 2 and control schools are not statistically significant at the 5% level. Schools in Kwara perform well on all but one criteria (#4: existence of a functioning SBMC women's committee supported by a female leader).

Table 9: Proportion of schools which meet each of the SBMC women inclusiveness criteria by phase for 5 states, and for Kwara (%)

	5 States[1]			Kwara
Criteria	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
Female attendance at SBMC meetings	5.9	50.5*	22.2	26.6
Issues raised by women	16.1	58.9*	24.2	41.6
Action taken on issues raised by women	19.3	48.9*	19.9	40.8
Women's committee meetings & leadership	4.7	24.6*	2.7	5.9

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the control group at the 0.05 level of statistical significance.

SBMC children's inclusiveness

64. The SBMC children's inclusiveness logframe indicator (SBMC reflects children's concerns) is based on four criteria. These are summarised below.

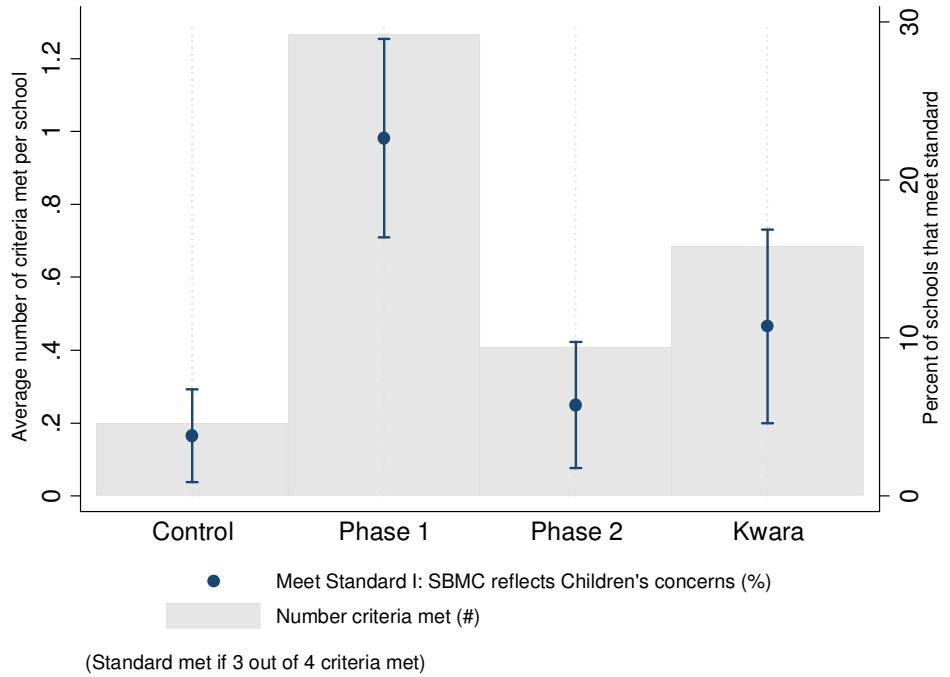
Logframe standard for SBMC children's inclusiveness

The school must meet at least three of the four criteria listed below in order to meet the SBMC 's children's inclusiveness standard for the 2011-12 school year

1) *At least one child attended two or more SBMC meetings (written evidence)*
2) *Child member of SBMC raised at least one issue at SBMC meetings (written evidence or oral evidence from child member of SBMC)*
3) *At least one issue raised by a child member at an SBMC meeting led to action (written, physical or oral evidence from child member of SBMC)*
4) *At least one SBMC children's committee meeting took place and committee has a trained facilitator (written evidence)*

65. Across five ESSPIN states, one in twenty of all public primary schools met the SBMC children's inclusiveness standard. Within this group of schools, the **proportion who met the standard is significantly higher in phase 1 schools (23%) compared with control schools (4%).** The estimate for phase 2 schools is 6%. In Kwara 11% of schools have a SBMC which reflects children's concerns, based on the criteria above.

Figure 8: Proportion of schools where SBMC reflects children's concerns by phase for 5 states, and for Kwara (%)



Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools that meet at least three criteria for child inclusiveness, with a 95% confidence interval. These schools meet the standard for child inclusiveness.
- The light grey bars indicate the average number of criteria met by schools.

66. The SBMC child inclusiveness standard is composed of four different criteria, listed in Table 10 below. The results presented in this table show that **phase 1 schools perform significantly better than control schools on all four criteria**. The differences between phase 2 and control schools are not statistically significant at the 5% level.

Table 10: Proportion of schools which meet each of the SBMC children inclusiveness criteria by phase for 5 states, and for Kwara (%)

	5 States[1]			Kwara
Criteria	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**
Child attendance at SBMC meetings	5.4	36.3*	14.5	18.9
Issues raised by children	7.1	40.7*	12.8	23.7
Action taken on issues raised by children	7.1	30.6*	11.2	21.7
Children's committee meetings & leadership	0.2	19.0*	2.1	4.1

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the control group at the 0.05 level of statistical significance.

School quality

67. The school quality logframe indicator is an outcome indicator, which is based on four of the output standards already presented. The definition is given below.

Logframe standard for school quality

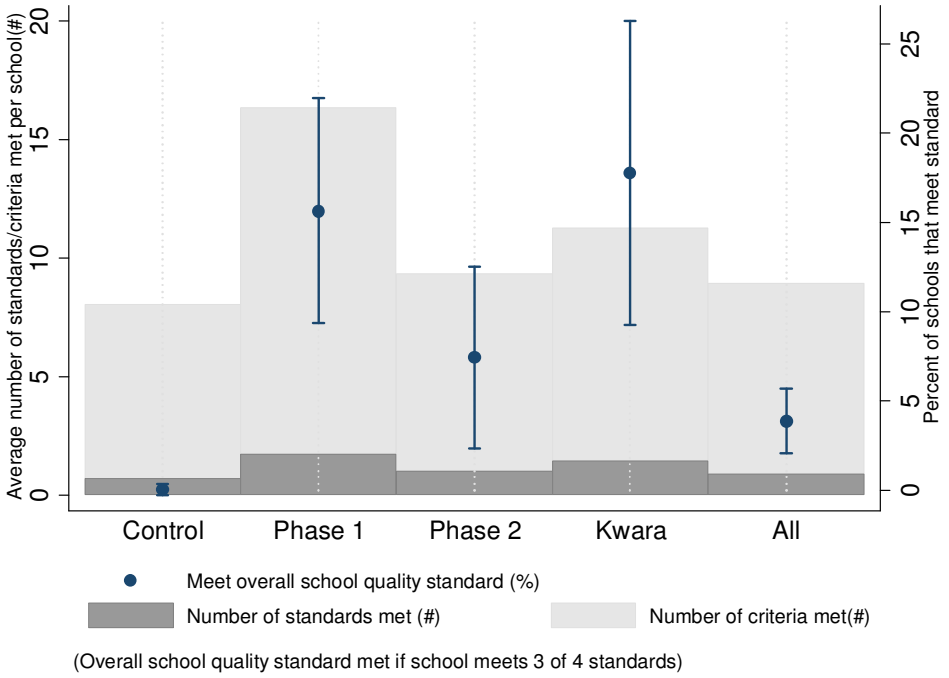
The school must meet at least three of the four output standards listed below in order to meet the school quality outcome standard, with teacher competence having to be one of those three.

1) *Teacher competence standard*
2) *Head teacher effectiveness standard*
3) *School development planning effectiveness standard*
4) *SBMC functionality standard*

As set out in boxes above, there are a total of 23 criteria underpinning the four output standards.

68. Across all six ESSPIN States, only 4% of all public primary schools met the school quality standard. Considering only the five ESSPIN States excluding Kwara, a **significantly higher proportion of phase 1 schools met the quality standard (15%) compared with the control schools (0%)**, whereas only 7% of phase 2 schools did. In Kwara 18% of schools met the quality standard.

Figure 9: Proportion of schools which meet quality standard by phase for 5 states, and for Kwara, and for all 6 states combined (%)



Source: Composite Survey 2012, see Annex D for details.

How to read the chart

- The dark blue dots represent the percentage of schools that meet at least three of the four standards, with a 95% confidence interval. These schools meet the overall school quality standard.
- The light grey bars indicate the average number of criteria met by schools.
- The dark grey bars indicate the average number of standards met by schools.

69. The breakdown of components of the school quality standard in Table 11 below shows the proportion of schools which met different combinations of the output standards. This table shows that about **one-third of the phase 1 schools that met the overall school quality standard had achieved all four output standards. By contrast, no control schools achieved 3 or 4 of the school output standards**. Among phase 1 schools, the most common combination of three standards met included **teacher competence, school development planning and effective SBMC**. Very few schools (only 1.0%) achieved the standards on head teacher effectiveness, school development planning and SBMC functionality simultaneously without *also* meeting the teacher competence standard.

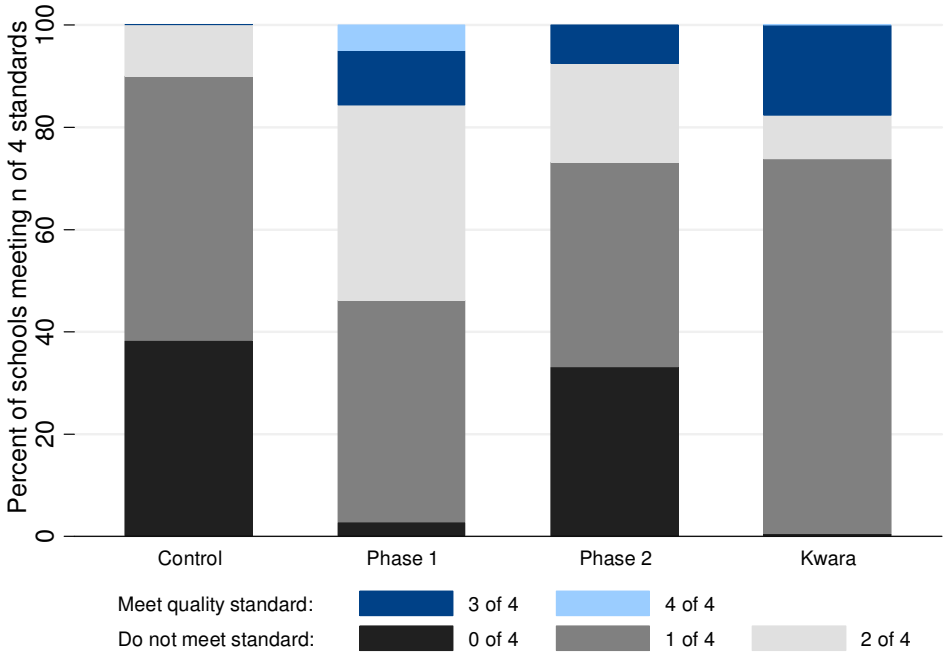
Table 11: Proportion of schools which meet 3 or more of the school quality standards by phase for 5 states, for Kwara and for 6 states combined (%)

	5 States[1]			Kwara	All 6
Criteria	**Control**	**Phase 1[2]**	**Phase 2[2]**	**All**	**All**
Meet standards TCH, HT, and SDP only	0	2.0*	0.7	2.1	0.4
Meet standards TCH, HT, and SBMC only	0	3.7*	2.6*	8.3	1.5
Meet standards TCH, SDP, and SBMC only	0	3.8*	4.2*	7.2	1.6
Meet standards HT, SDP, and SBMC only	0	1.0	0	0	0.1
Meet all four standards	0	5.1*	0	0.1	0.3

Note: (1) The 5 States are; Enugu, Jigawa, Kaduna, Kano, and Lagos. (2) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the control group at the 0.05 level of statistical significance.

70. Figure 10 shows the number of school performance standards achieved in five states by control, phase 1, phase 2 schools, respectively, as well as schools in Kwara. **98% of phase 1 schools (five states) and in Kwara met at least one of the output standards, while more than one third of control and phase 2 schools met none of the standards**. Close to 20% of phase 1 and Kwara schools met more than 2 output standards; and 6% of phase 1 schools met all four output standards. By contrast, no control schools managed to meet more than two output standards.

Figure 10: Proportion of schools which meet one, two, three or four output standards underpinning the overall quality standard, by phase for 5 states, and for Kwara (%)



Source: Composite Survey 2012.

How to read the chart

- The dark grey bars represent the proportion of schools that meet 0 of 4 standards. These schools do not meet the overall school quality standard.

- The medium grey bars represent the proportion of schools that meet 1 of 4 standards. These schools do not meet the overall school quality standard.
- The light grey bars represent the proportion of schools that meet 2 of 4 standards. These schools do not meet the overall school quality standard.
- The dark blue bars represent the proportion of schools that meet 3 of 4 standards. These schools meet the overall school quality standard.
- The light blue bars represent the proportion of schools that meet 4 of 4 standards. These schools meet the overall school quality standard.

Pupil learning achievement in English literacy and numeracy

71. The learning achievement indicators are impact indicators. Estimates are representative of children learning in all public primary schools across the six states, not just those learning in ESSPIN programme schools. In five of the six states, the coverage of SIP in phase 1 has been 10% or less of public primary schools. In Kwara, parts of the SIP have universal coverage, while the full set of interventions covered 17% of public primary schools in Phase 1. The baseline MLA survey in 2010 revealed that pupils' English literacy and numeracy skills were very weak in most states, and it is unlikely that this situation has changed markedly overall, even if the SIP interventions are making a positive contribution.

72. The test instruments were designed to capture more than the information needed for the logframe indicators. Each test has questions which span different grade levels, and also different learning domains. A mapping of the questions in each test to grade levels and learning domains is in Annex C.

73. The p2 and p4 literacy results are discussed first, followed by the numeracy results. All estimates refer to pupils in public primary schools (not private schools).

Literacy: primary 2 results

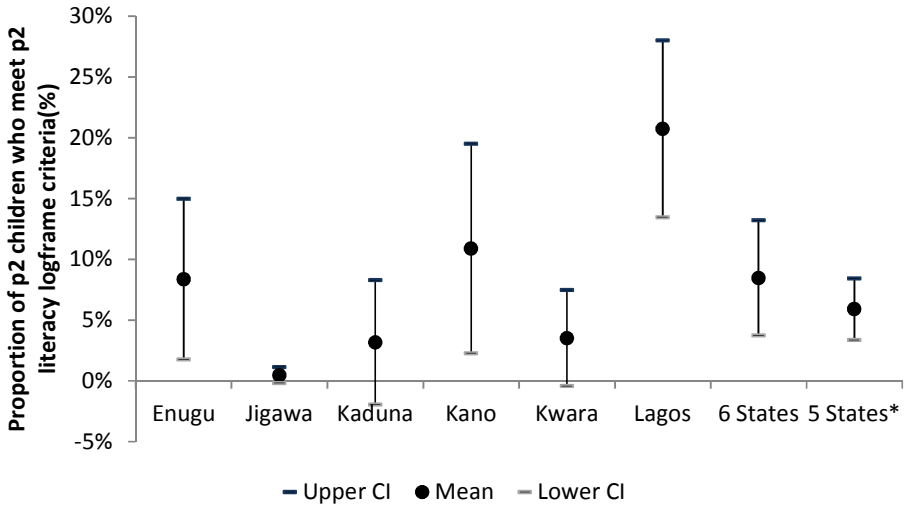
74. The ESSPIN logframe indicator for p2 literacy measures whether pupils have p2 level skills for reading comprehension. The definition is:

> *Proportion of p2 children who correctly answer a p2 curriculum level question on listening comprehension **and** correctly read a sufficient number of words from a p2 curriculum level passage. (See Annex C for full details)*

75. Across six states, an estimated 9% of all public primary school p2 pupils have skills for reading comprehension. This estimate ranges from 1% in Jigawa to 21% in Lagos (see Figure 11). As can be seen from the wide confidence intervals, some of these estimates are fairly imprecise. For five states excluding Kano, an estimated 6% of p2 pupils reach the grade appropriate literacy standard.

76. This is a very specific standard for pupils to reach. It does **not** mean that some 94% of pupils across the five states are failing to engage at all with the English curriculum. It shows that a small minority of pupils have skills appropriate to their grade level.

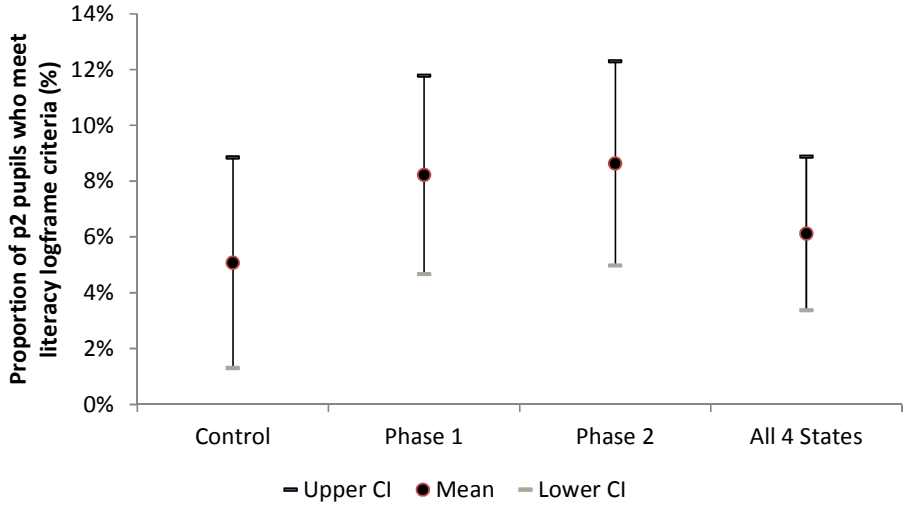
Figure 11: Proportion of p2 pupils with skills for reading comprehension by state, for 6 states and 5 states combined (%)



Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: * 5 States estimate excludes Kano because for this state estimate there is a very large design effect which makes its estimate less reliable.

77. There is no significant difference between the estimates of the p2 literacy logframe indicator for pupils in Phase 1 schools compared with those in control schools for four states combined (excluding Kwara and Kano), see Figure 12.

Figure 12: Proportion of p2 pupils with skills for reading comprehension for 4 states[1] (excluding Kwara and Kano) by phase (%)



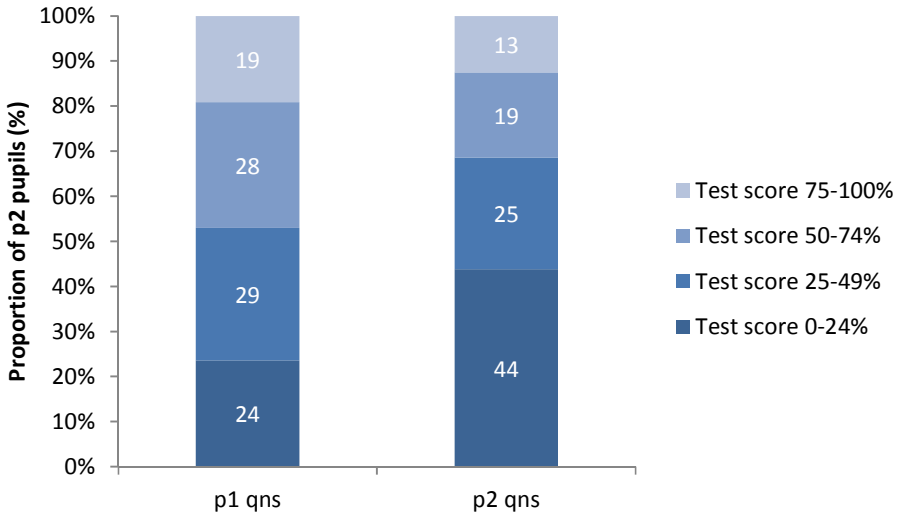
Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: (1) The 4 states are Enugu, Jigawa, Kaduna and Lagos. Kwara is excluded because it does not have comparable phase information. Kano is excluded because its estimates have a large design effect.

78. Analysis of overall test scores, rather than the narrow logframe measure, provides some useful insights to help guide future interventions. The distribution of overall p2 literacy test scores by grade level of questions shows how well p2 pupils are coping with literacy material at different grade levels. Figure 13 reveals that just over half of pupils in p2 across five states are still struggling with p1 questions, with

test scores of less than 50% on this material. This means that many children still have profound difficulties with English literacy. Without secure foundations in speaking, understanding, reading and writing, pupils will find it almost impossible to make progress at higher grade levels or to access the wider curriculum.

79. The analysis also suggests that the move from p1 to p2 level questions is a big jump for a sizable proportion of p2 pupils across the five states. Only one-third of pupils scored 50% or more on the grade appropriate questions.

Figure 13: Distribution of p2 English literacy test scores by grade level of question for 5 states[1] combined (%)

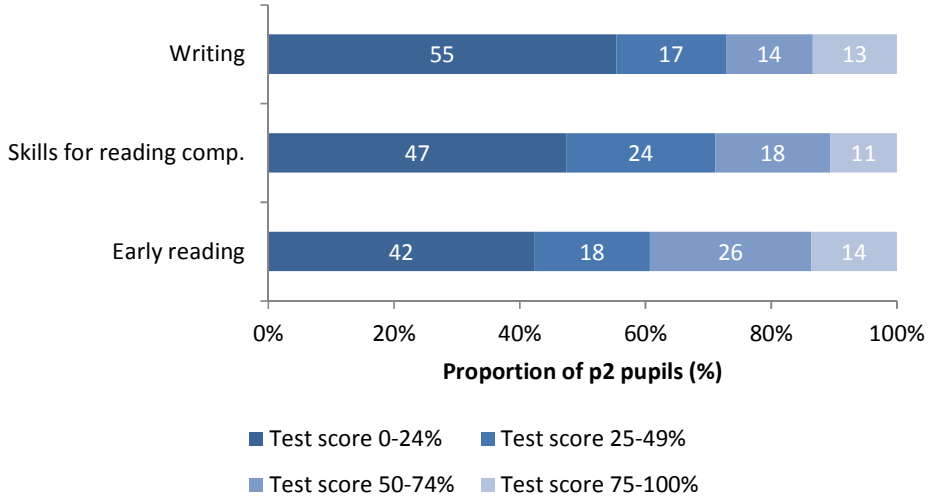


Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

80. Looking at the distribution of test scores by learning domain in Figure 14, the test score band pattern on the early reading bar is notably different to the others. A higher proportion of pupils (40%) have developed early reading skills than have skills for reading comprehension (29%) or writing skills in English (27%), based on scoring more than 50% on the relevant questions. This implies that there is a group of pupils who have foundational reading skills but are struggling to make the transition to develop skills for reading comprehension and writing.

81. 42% of p2 pupils are struggling to begin early reading (the recognition of letter sounds and simple words) with scores of less than 25% on this element of the test. Children need to master these early reading skills in order to have success with reading comprehension and writing.

Figure 14: Distribution of p2 English literacy test scores by learning domain for 5 states[1] combined (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

English literacy: primary 4 results

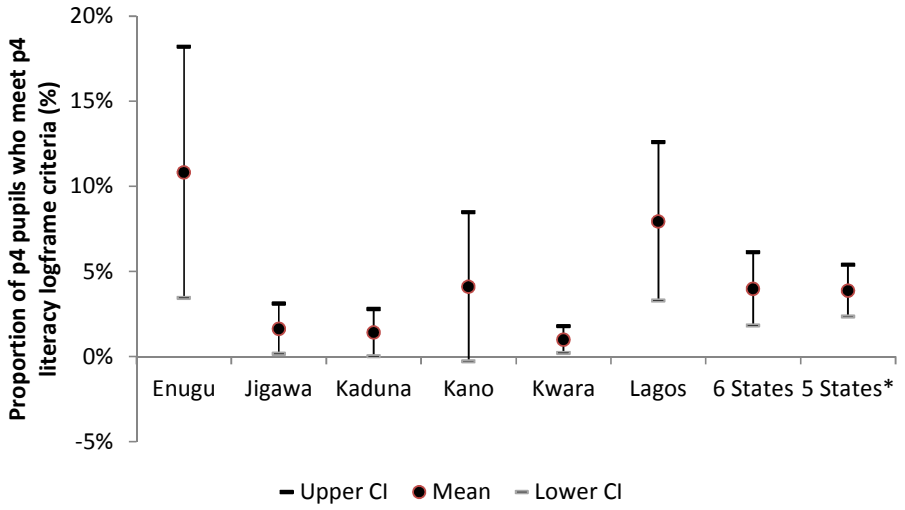
82. The ESSPIN logframe indicator for primary 4 (p4) literacy measures whether pupils are able to read a p4 text with comprehension. The definition is:

> *Proportion of p4 children who correctly read a sufficient number of familiar words at p4 curriculum level **and** correctly read a sufficient number of words from a p4 curriculum level passage **and** correctly answer at least four out of five reading comprehension questions. (See Annex C for full details.)*

83. Across six states, an estimated 4% of all public primary school p4 pupils are able to read with comprehension at grade appropriate level. This estimate ranges from 1-2% in Jigawa, Kaduna and Kwara to 11% in Enugu (see Figure 15). As can be seen from the wide confidence intervals, the estimates for Enugu, Kano and Lagos are fairly imprecise. For five states excluding Kano, an estimated 4% of p4 pupils reach the grade appropriate literacy standard.

84. This is a very specific standard for pupils to reach. It does **not** mean that some 96% of pupils across the five states are completely failing to engage with the English curriculum. It shows that a small minority of pupils are equipped with the skills needed to access the wider curriculum at the appropriate level, but the majority of children are not.

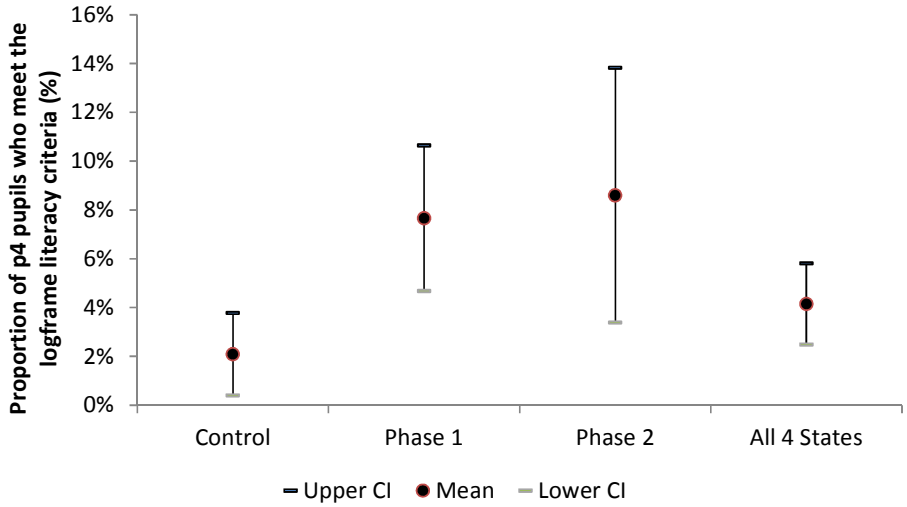
Figure 15: Proportion of p4 pupils able to read with comprehension by state, and for 6 states and 5 states combined (%)



Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: * 5 States estimate excludes Kano because for this state estimate there is a very large design effect which makes its estimate less reliable.

85. For four states (excluding Kwara and Kano), Figure 16 displays estimates for the p4 literacy logframe indicator by phase. The **proportion of p4 pupils able to read with comprehension is significantly higher in phase 1 schools (8%), and in phase 2 schools (9%), compared with control schools** where only 2% of pupils met the criteria.

Figure 16: Proportion of p4 pupils able to read with comprehension for 4 States[1] (excluding Kwara and Kano) by phase (%)

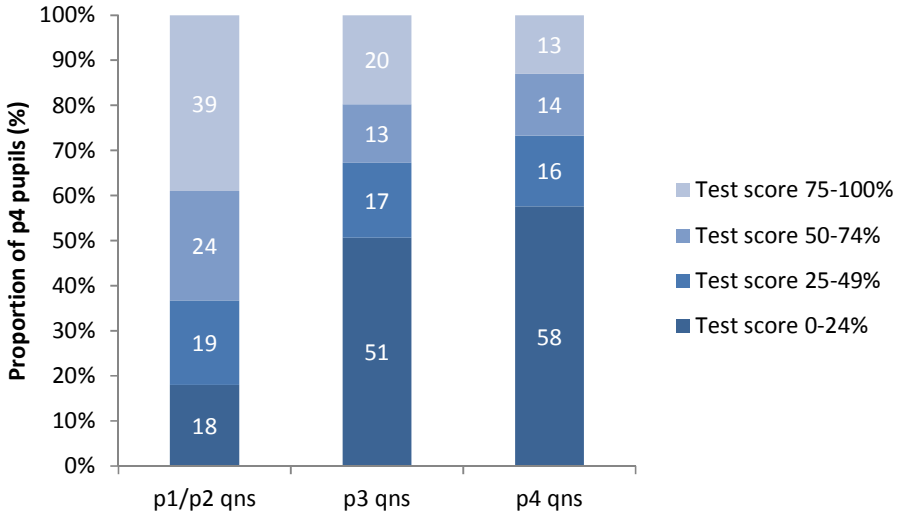


Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: (1) The 4 states are Enugu, Jigawa, Kaduna and Lagos. Kwara is excluded because it does not have comparable phase information. Kano is excluded because its estimates have a large design effect.

86. The distribution of overall p4 literacy test scores by grade level of questions shows how well p4 pupils are coping with literacy material at different grade levels from p1/2 to p4. The pattern of band-scores in Figure 17 is very different for the p1/p2 questions compared with both the p3 and the p4 bands which are quite similar. A lot of p4 pupils are finding the transition from p1/p2 work to p3 work difficult. Most p4 pupils (about two-thirds) are coping well with p1/p2 work, scoring above 50% on these questions, but only one-third or less of p4 pupils are coping with p3 and p4 work. There is little difference in the distribution of score bands on p3 questions compared with p4 questions which perhaps suggests that a top performing group of children grasp both p3 and p4 level work, but the majority of p4 children are still working at a level below their grade.

87. Less than 40% of p4 pupils across the five states are still struggling with questions two years or more below their grade level. These children are clearly not equipped to with the skills needed to access the English curriculum at grade 4.

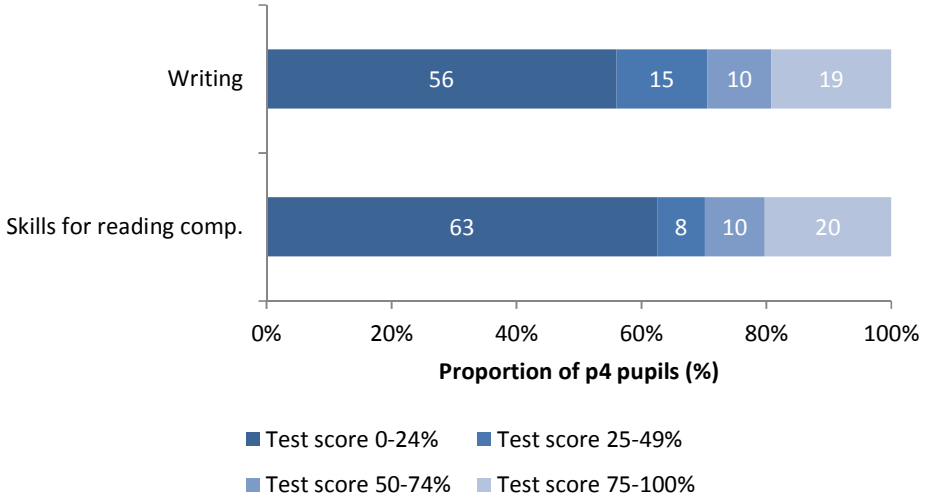
Figure 17: Distribution of p4 English literacy test scores by grade level of question for 5 states[1] combined (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

88. The pattern of p4 pupils achieving different test scores bands is broadly similar for reading comprehension and for writing as Figure 18 shows. The majority of P4 pupils in all public primary schools across the five states are profoundly struggling to develop both skills for English reading comprehension and writing skills. About six in every ten p4 pupils scored less than 25% on the questions which tested reading comprehension, and on those which tested writing.

Figure 18: Distribution of p4 English literacy test scores by learning domain for 5 states combined (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

Numeracy: primary 2 results

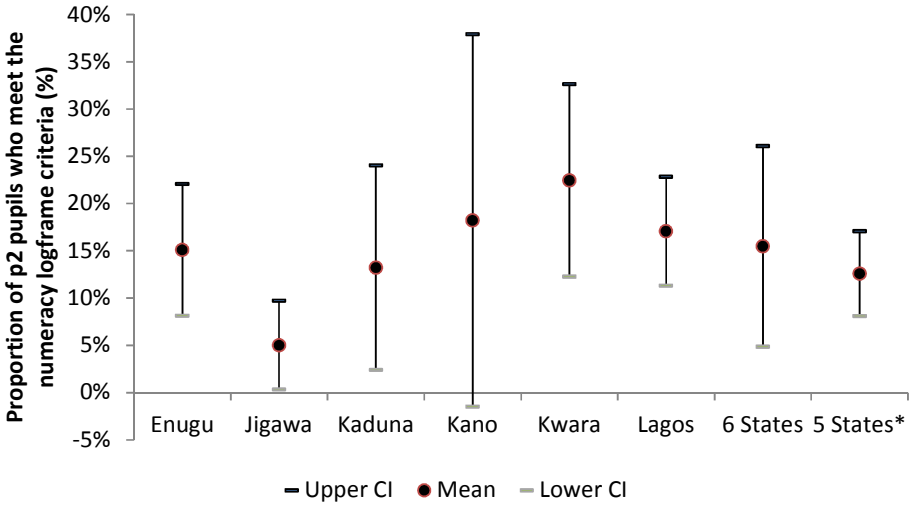
89. The ESSPIN logframe indicator for p2 numeracy measures whether pupils have p2 level arithmetic skills. The definition is:

> *Proportion of p2 children who correctly answer at least five out of six p2 curriculum level questions on addition and subtraction, **and** both multiplication questions. (See Annex C for full details.)*

90. Across six states, an estimated 16% of p2 pupils in all public primary schools can perform basic arithmetic calculations at the level expected in p2. This estimate ranges from 5% in Jigawa to 22% in Kwara (see Figure 19). As can be seen from the wide confidence intervals, many of these estimates are fairly imprecise. For five states excluding Kano, an estimated 13% of p2 pupils reach the grade appropriate numeracy standard.

91. This is a very specific standard for pupils to reach. It does not mean that some 87% of pupils across the five states are entirely unable to engage with the mathematics curriculum. It shows that a small minority of pupils are equipped with the skills needed to achieve at the appropriate level.

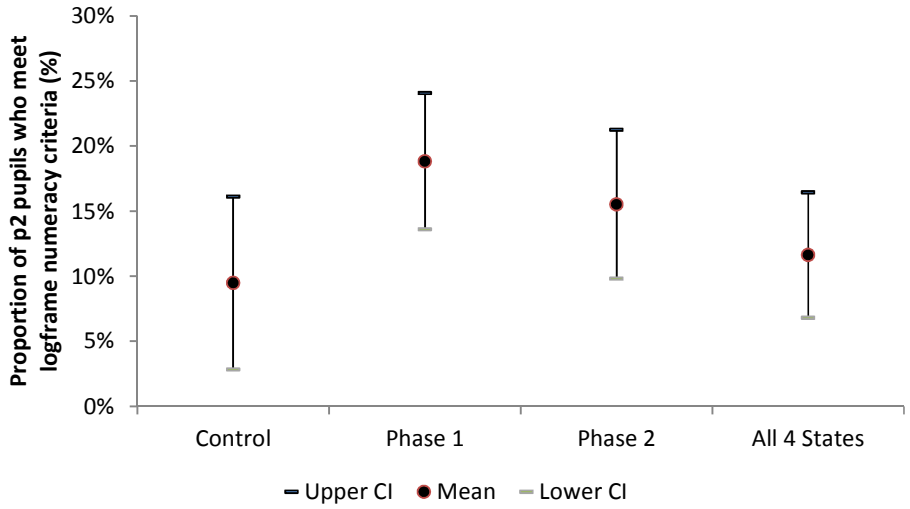
Figure 19: Proportion of p2 pupils able to perform basic arithmetic calculations at p2 level by state, and for all 6 states combined (%)



Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: * 5 States estimate excludes Kano because for this state estimate there is a very large design effect which makes its estimate less reliable.

92. For four states (excluding Kwara and Kano), Figure 20 displays estimates for the p2 numeracy logframe indicator by phase. The **proportion of p2 pupils able perform basic arithmetic calculations is significantly higher in phase 1 schools (19%) compared with control schools** where 9% of pupils met the criteria. In Phase 2 schools 16% of p2 pupils demonstrated ability to carry out p2 level arithmetic.

Figure 20: Proportion of p2 pupils able to perform basic arithmetic calculations at p2 level for 4 States[1] (excluding Kwara and Kano) by phase (%)

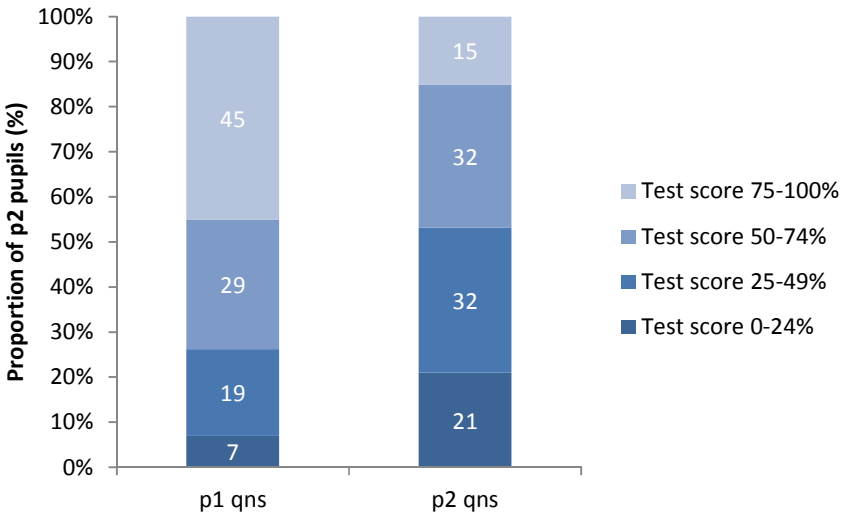


Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: (1) The 4 states are Enugu, Jigawa, Kaduna and Lagos. Kwara is excluded because it does not have comparable phase information. Kano is excluded because its estimates have a large design effect.

93. The distribution of overall p2 numeracy test scores by grade level of questions shows how well p2 pupils are coping with numeracy material at different grade levels. Figure 21 reveals that about three-quarters of p2 pupils are coping with p1 questions, scoring in the top two bands. The proportion of

pupils scoring 50% or more drops to about half for the p2 questions. But still, this suggests that many pupils have the basic foundations in numeracy to enable them to access the p2 mathematics curriculum.

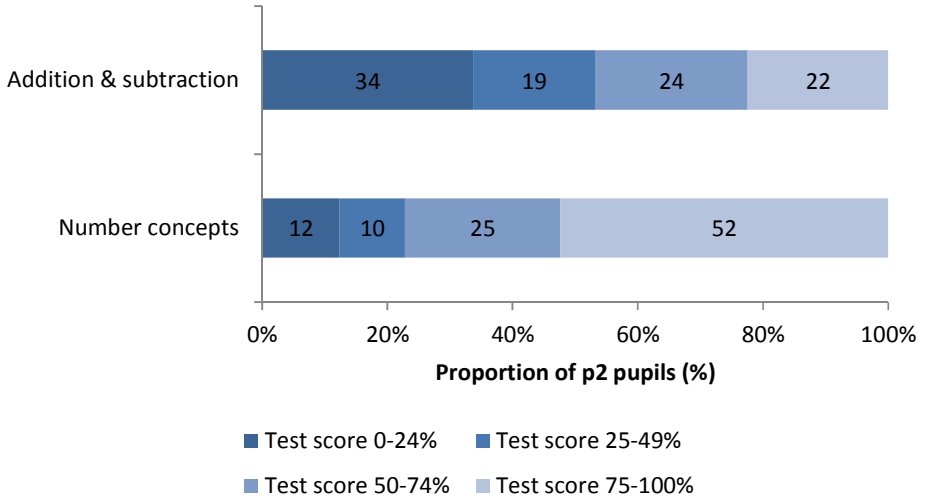
Figure 21: Distribution of p2 numeracy test scores by grade level of question for 5 states[1] combined (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state there is a very large design effect which makes its estimates less reliable.

94. P2 pupils find number concepts much easier than addition and subtraction, as the pattern of overall test scores by learning domain in Figure 22 shows. Nearly 80% of p2 pupils scored 50% or more on the number concept questions, with more than half of pupils demonstrating a secure understanding (scoring in the top band). But far fewer p2 pupils can use numbers to perform basic operations. In other words it appears that the majority of p2 pupils understand what numbers are, but teachers need to extend this understanding to enable pupils to manipulate numbers. About one in three p2 pupils scored in the lowest test band for the addition and subtraction questions.

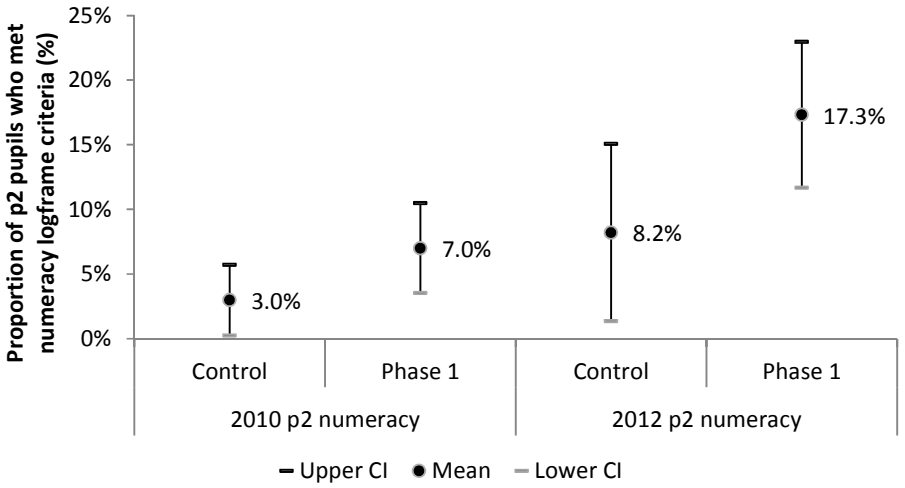
Figure 22: Distribution of p2 numeracy test scores by learning domain for 5 states[1] combined (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

95. A subset of questions in the 2012 CS p2 numeracy test was drawn from the 2010 MLA survey, and the marking was consistent. Based on these common questions, Figure 23 summarises trends in p2 numeracy achievement between 2010 and 2012 for three states combined: Lagos, Jigawa and Kaduna. The results reveal that in phase 1 schools, there is a significant gain in the proportion of p2 pupils able to perform arithmetic calculations at p2 level from 7% to 17%. There is also a gain in this measure of numeracy skills for p2 pupils in control schools, from 3% to 8% of pupils, but this gain is not statistically significant. There is no significant difference in the over-time gains in the numeracy indicator between phase 1 and the control schools. The analysis is limited to three states because there is insufficient 2010 MLA data for Enugu and Kwara, and Kano's results are too imprecise to include with confidence.

Figure 23: P2 numeracy logframe indicator by year and phase: 3 states combined (Lagos, Jigawa and Kaduna) (%)



Source: ESSPIN Composite Survey 2012. (1) The 2010 sample size is 1,183 pupils (543 control and 640 phase 1); the 2012 sample size if 904 pupils (511 control and 397 phase 1).

Numeracy: primary 4 results

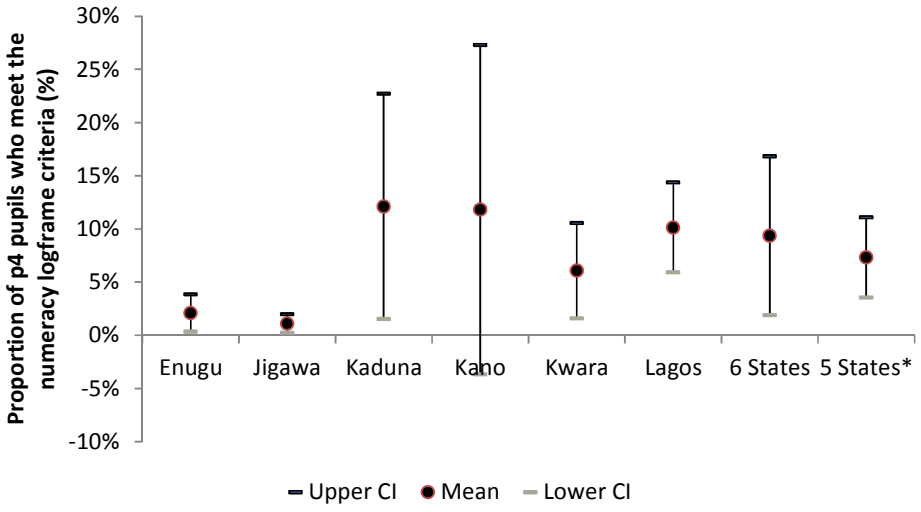
96. The ESSPIN logframe indicator for p4 numeracy measures whether pupils have p4 level arithmetic skills. The definition is:

> *Proportion of p4 children who correctly answer p4 curriculum level questions on addition and subtraction **and** multiplication **and** division. (See Annex C for full details.)*

97. Across six states, an estimated 9% of p4 pupils in all public primary schools can perform basic arithmetic calculations at the level expected in p4. This estimate ranges from 1-2% in Enugu and Jigawa to 10-12% in Kaduna, Kano and Lagos (see Figure 24). As can be seen from the wide confidence intervals, the estimates for Kano and Kaduna are very imprecise. For five states excluding Kano, an estimated 7% of p2 pupils reach the grade appropriate numeracy standard.

98. This is a very specific standard for pupils to reach. It does not demonstrate that some 93% of pupils across the five states are failing to engage with the mathematics curriculum entirely. It shows that only a small minority of p4 pupils are equipped with the skills needed to access the maths curriculum at the appropriate level.

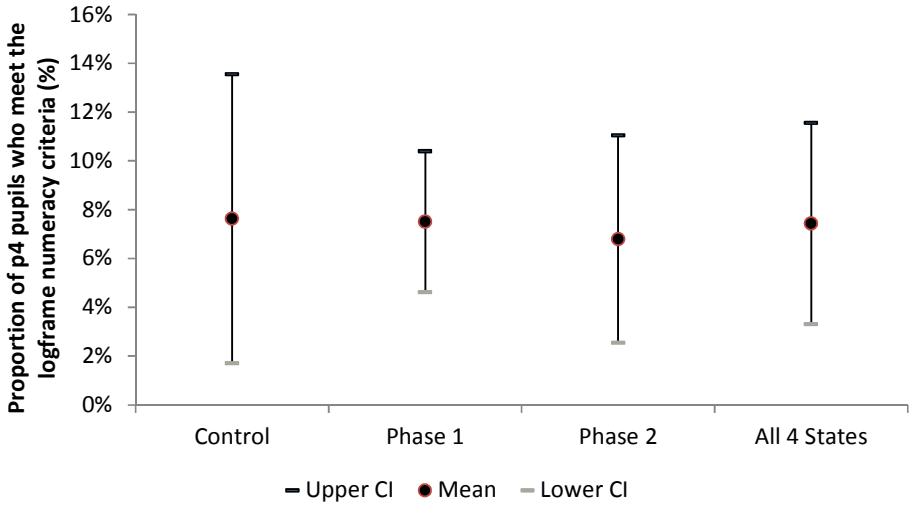
Figure 24: Proportion of p4 pupils able to perform basic arithmetic calculations at p4 level by state, and for 6 states and 5 states combined (%)



Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: * 5 States estimate excludes Kano because for this state estimate there is a very large design effect which makes its estimate less reliable.

99. For four states (excluding Kwara and Kano), Figure 25 displays estimates for the p4 numeracy logframe indicator by phase. The proportion of p4 pupils able perform basic arithmetic calculations is similar in each of the three groups of schools at 7-8%. There are no significant differences between estimates for the different phase groups.

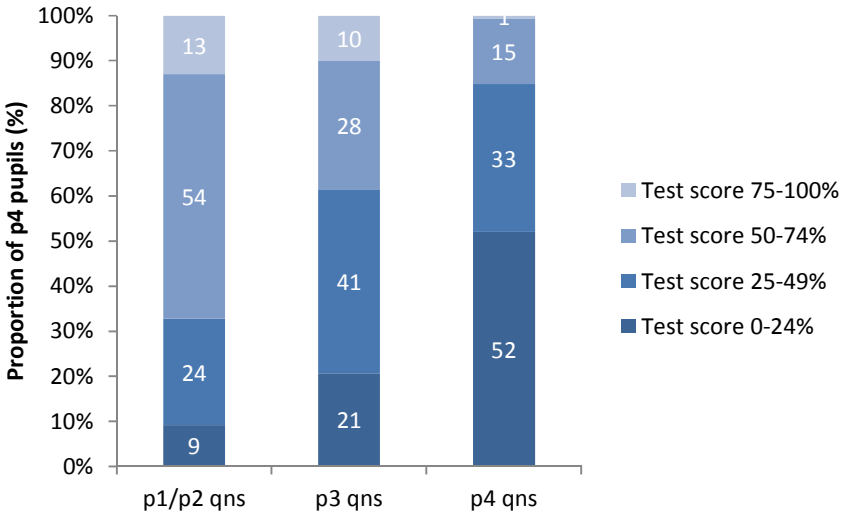
Figure 25: Proportion of p4 pupils able to perform basic arithmetic calculations at p4 for 4 States[1] (excluding Kwara and Kano) by phase (%)



Source: ESSPIN Composite Survey 2012, see Annex D for details. Note: (1) The 4 states are Enugu, Jigawa, Kaduna and Lagos. Kwara is excluded because it does not have comparable phase information. Kano is excluded because its estimates have a large design effect.

100. The distribution of overall p4 numeracy test scores by grade level of questions shows how p4 pupils are coping with numeracy material at different grade levels. It is clear from Figure 26 that the transition from p1/p2 questions to p3 questions is difficult for a sizable proportion of p4 pupils; this is also true for the transition from p3 to p4 questions. About two thirds of p4 pupils are able to cope with p1/p2 questions scoring in the top two bands, but this proportion drops by about a half when pupils are faced with p3 questions. Moving from p3 to p4 questions, again the proportion scoring 50% or more drops by about a half to 16%. This implies that more than 80% of p4 pupils do not have the skills to access the mathematics curriculum at p4 level.

Figure 26: Distribution of p4 numeracy test scores by grade level of question for 5 states1 combined (%)

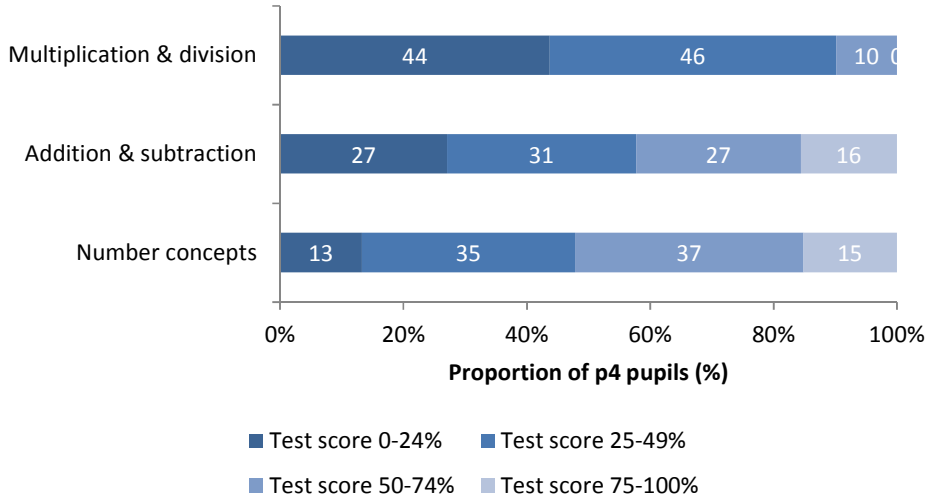


Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

101. P4 pupils find number concepts much easier than addition and subtraction. Multiplication and division poses even greater difficulty, as the pattern of overall test scores by learning domain in Figure 27 shows. It is striking that only 10% of p4 pupils are able to perform multiplication and division calculations; these pupils scored between 50% and 75% on these test questions, and no pupils scored 75% or more.

102. Almost half of p4 pupils are not coping well with basic number concepts. This is of serious concern given that these pupils have been in the school system for four years. Number concepts are foundational skills necessary for children to acquire the ability to do number operations.

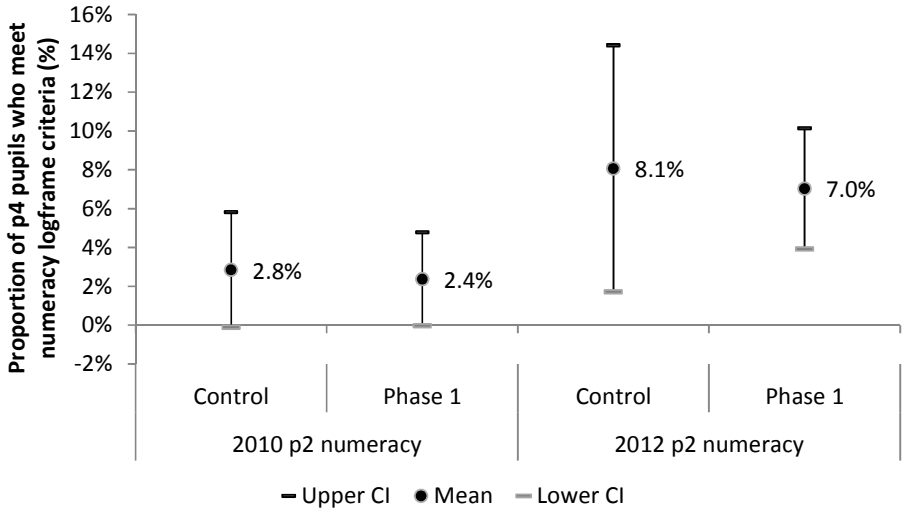
Figure 27: Distribution of p2 numeracy test scores by learning domain for 5 states[1] combined (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 5 states are Enugu, Jigawa, Kaduna, Kwara and Lagos. Kano is excluded because for this state estimate there is a very large design effect which makes its estimates less reliable.

103. A subset of questions in the 2012 CS p4 numeracy test was drawn from the 2010 MLA survey, and the marking was consistent. Based on these common questions, Figure 28 summarises trends in p4 numeracy achievement between 2010 and 2012 for three states combined: Lagos, Jigawa and Kaduna. The results reveal that **in phase 1 schools, there is a significant gain in the proportion of p4 pupils able to perform arithmetic calculations at p2 level** from 2% to 7%. There is also a gain in this measure of numeracy skills for p4 pupils in control schools, from 3% to 8% of pupils, but this gain is not statistically significant. There is no significant difference in the over-time gains in the p4 numeracy indicator between phase 1 and the control schools. The analysis is limited to three states because there is insufficient 2010 MLA data for Enugu and Kwara, and Kano's results are too imprecise to include with confidence.

Figure 28: P4 numeracy logframe indicator by year and phase: 3 states combined (Lagos, Jigawa and Kaduna) (%)



Source: ESSPIN Composite Survey 2012. Note: (1) The 2010 sample size is 1159 pupils (540 control and 619 phase 1); the 2012 sample size if 894 pupils (495 control and 399 phase 1).

104. Annex E contains a variety of more detailed analytical illustrations of the distribution of learning outcomes between Phase 1 and Control Schools in the states where such comparisons can legitimately be made on the strength of the data gathered. These include:

- Annex E.1: Lowest quartiles by Phase, State and Domain
- Annex E.2: Highest quartiles by Phase, State and Domain
- Annex E.3: Proportion of pupils in each quartile by Grade, Phase and State
- Annex E.4: Proportion of p4 pupils in each quartile by Phase and State on p1/p2 items only
- Annex E.5: Reduction of proportion of p2 pupils in bottom score band, Phase 1 cf. Control

105. The over-riding conclusion from each of these sets of charts is that the proportion of children in lower performance bands in Phase 1 schools is lower—often statistically significantly lower—than those in Control Schools (Annexes E.1, E.3, E.5). Conversely, there are larger proportions of children in the upper performance bands in Phase 1 schools than in Control Schools (Annexes E.2, E.3, E.4). These variances are strongest in Jigawa, the state with the lowest learning outcomes in absolute terms, and therefore the state where children are most in need of support (Annex E.5). In view of ESSPIN's intervention logic, it is most encouraging that the earliest effects of foundational pedagogical guidance and other school improvement activities seems to be reflected in the sharpest response coming from teachers and children who struggle most with the lowest levels of achievement. Item analysis shows that gains in p4 are largely attributable to better performance on p1/p2 level questions. This is as expected from the ESSPIN training and in-school support provided to date. Likewise, one would expect Enugu and especially Lagos teachers and children to benefit increasingly as the intervention moves up the curriculum levels to their own areas of relative weakness. This is summarised in Table 12 below:

Table 12: Test band quartile differences between Phase 1 and Control Schools

Differences in proportions of pupils in Phase 1 and Control Schools found in the lowest quartile (0-25%) of test scores

Subject	Grade	Domain	Enugu	Jigawa	Lagos
Numeracy	2	Number concepts	-0.8%	24.0%	1.0%
		Addition and subtraction	5.0%	21.0%	9.0%
	4	Number concepts	5.7%	44.7%	-2.5%
		Addition and subtraction	13.2%	44.0%	-3.7%
		Multiplication and division	35.0%	33.0%	-7.0%
Literacy	2	Early Reading	4.0%	27.0%	1.0%
		Reading comprehension	25.0%	11.0%	-5.0%
		Writing	7.0%	18.0%	-1.0%
	4	Reading comprehension	18.0%	30.0%	11.0%
		Writing	24.0%	31.0%	3.0%

Statistically significant positive impact	14
Statistically insignificant positive impact	10
Statistically insignificant negative impact	5
Statistically significant negative impact	1
	30

Out of 30 state/domain/grade categories, almost half (14) already reveal statistically significantly better pupil learning outcomes--and a further ten positive but non-significant results--in ESSPIN-supported schools compared with Control Schools, at the lowest end of the achievement spectrum. Also, the magnitude of those positive differences is much larger than that of the few negative differences observed. The positive differences are largest in Jigawa, the State with the lowest levels of achievement, which stands to gain most from the early emphasis on basic skills in English literacy and numeracy in the ESSPIN School Improvement Programme.

Differences in proportions of pupils in Phase 1 and Control Schools found in the highest quartile (75-100%) of test scores

Subject	Grade	Domain	Enugu	Jigawa	Lagos
Numeracy	2	Number concepts	5.0%	26.0%	8.0%
		Addition and subtraction	23.0%	19.3%	9.0%
	4	Number concepts	23.0%	12.7%	2.5%
		Addition and subtraction	22.0%	15.9%	3.7%
		Multiplication and division	*0.0%*	*0.0%*	*0.0%*
Literacy	2	Early Reading	24.0%	13.5%	-1.0%
		Reading comprehension	20.0%	12.9%	-4.0%
		Writing	10.0%	8.1%	-4.0%
	4	Reading comprehension	22.0%	18.7%	11.0%
		Writing	23.0%	14.7%	11.0%

Statistically significant positive impact	16
Statistically insignificant positive impact	8
Statistically insignificant negative impact	3
Statistically significant negative impact	0
	27

At the top end of the achievement scale, only three categories (all in Lagos State) fail to show a positive or significantly positive gain from pupils attending an ESSPIN-supported school. It is primarily Phase 1 School pupils' performance on p1/p2 questions which has driven the improvements at p2 and p4 levels, as would be expected from ESSPIN's intervention logic at this stage in the programme cycle. *NB: p4 multiplication and division contains missing values which are actively under investigation with a view to completing the analysis.*

106. In the absence of evidence of positive selection bias at the school level, it would seem reasonable to hypothesize that ESSPIN interventions are responsible for the pattern of findings in Table 12. However, definitive proof will require measurement of difference of differences in 2014.

Conclusion and implications of survey findings for ESSPIN programme

107. The ESSPIN programme is underpinned by a specific intervention logic (Annex B). This draws on research and experience in Nigeria and internationally to propose that children are most likely to attend regularly and complete their basic education if schools offer a safe, supportive and effective learning environment. ESSPIN's success will ultimately be measured in terms of its contribution to more children attending school and learning more whilst they are there. This *impact* is to be achieved through firstly piloting, and then supporting state roll-out of, an integrated School Improvement Programme. The SIP is designed to lead to the *outcome* of an increased number and proportion of good quality schools in each state. The SIP works through a number of *outputs* directly provided with UKaid funding in the phase 1 pilot stage in a limited number of schools, and in partnership with state authorities in the phase 2 roll-out to a significant fraction of all schools in the state. These SIP outputs include training, resources, professional support services and capacity development focused on teacher competence, head teacher academic leadership, school management, improved governance through SBMCs, fostering an inclusive ethos, and a constructive role in schools for local government officers and CSOs (amongst other interventions).

108. The first round of the Composite Survey provides plausible evidence that the ESSPIN intervention logic is valid, to the extent that it is possible to do so in the first round of a two-step evaluation methodology. In 2012 the pilot phase was drawing to a close and the transition to roll-out was underway, although as a state-based programme, major differences in approach and timing had evolved in each state in response to state authorities' priorities and capacity.

109. At the output level in the combined five states with phased interventions, ESSPIN phase 1 schools significantly out-performed control schools in six out of the eight indicators (see Annex D, Table D.5). Jigawa and Enugu also recorded significant differences between phase 1 and control schools in six out of eight output indicators at the individual state level. The exceptions in both states, and in the combined measure, were in the proportion of schools which 'meet the needs of all pupils' and head teacher effectiveness. A number of reasons for the former can be proposed; not least, that the concept of inclusiveness is difficult to capture in a survey of this kind, given the limitations in data enumeration in particular. Moreover, inclusiveness has not been split out as a specific focal area for intervention: it is woven throughout teacher, head teacher, SBMC, community, CSO and government capacity building efforts. And it is probable that only the best teachers and heads in *any* school system will actively teach and manage for inclusivity and individual differences. ESSPIN will look more closely at the data, review its response on this indicator and report back in due course. The detailed criteria on head teacher effectiveness also suggest that more time and coverage is needed, to see the raised levels of lesson observation by head teachers (which has been a focus of phase 1 ESSPIN training, Table 5) matched in other respects (such as school opening times, which has not). And a question mark still hangs over how the apparently impressive school development planning results can be sustained in an era when ESSPIN

no longer supplies Direct Funding of Schools to states. Focusing school development planning on academic support rather than infrastructure and physical investments is one approach which ESSPIN will encourage in future school self evaluation and development planning cycles.

110. Out of the 40 state-level phased output indicators measured (five states x eight indicators), 19 displayed significant positive differences between phase 1 schools and controls, and a further 16 were higher in the phase 1 schools (albeit not at the level of statistical significance chosen for the survey—see Annex D, Table D.6). These differences cannot be firmly attributed to ESSPIN until a difference in difference reading can be made in 2014. But in the absence of a stronger alternative explanation for these results, the CS evidence suggests it is reasonable to believe that ESSPIN is on track at the output level. For instance, it would be difficult to imagine what other than ESSPIN's interventions caused the finding that *all* 18 criteria for SBMC functionality, SBMC inclusion of women and SBMC inclusion of children are significantly more likely to be met in ESSPIN phase 1 schools than in control schools.

111. Encouragingly, there was also a significant difference at outcome level between the proportions of phase 1 schools meeting the school quality standard compared with control schools (see Annex D, Table D.5). This addresses the proof of concept requirement for the pilot phase. It also justifies the subsequent shift to programme mode in which increasing levels of ownership, leadership, funding and management for delivery of the SIP are taken by state ministries of education, SUBEBs, LGEAs and CSOs, with structures such as the SSITs, Advisory Services Units and CSO/Government Partnerships in place to sustain the intervention as a mainstreamed programme.

112. There is evidence from the output level that teachers in phase 1 schools are more competent than in control schools (Table 4). But as expected from the intervention logic, far more teachers will need to have been trained and supported to deliver change in their classrooms, and then work with children in a focused way for a longer time, whilst also benefiting from improved academic leadership, school management, governance and CSO/local government partnership support, before a significant impact on children's learning is discernable state-wide. This explains why it was found that no statistically significant difference in difference was observed between phase 1 and control schools on comparable items between MLA 2010 baseline maths and CS in 2012 (Figure 28), although the gain in phase 1 schools was statistically significant whilst that in control schools was not. P2 English skills for reading comprehension and arithmetic were both stronger in phase 1 schools than controls (with the latter significantly so); as was p4 reading comprehension, although no difference in p4 maths was found (Annex A, Table D.6). Again, none of these differences can be ascribed to ESSPIN at this stage, and at the impact level the case for suggesting they could be is necessarily more tenuous than at the output and outcome levels at which ESSPIN is currently working.

113. Evidence for monitors about whether the ESSPIN programme is on track to hit logframe targets is only part of the story. Much of the value of the Composite Survey lies in the detailed picture which can be built up regarding which aspects of the education system are showing good progress, and which remain cause for concern. These will inform adjustments and priorities for ESSPIN and state partners going forwards. The ESSPIN team has scheduled an internal review of the findings, to identify potential improvements to the SIP on a state-by-state and overall programme basis, for discussion with state, national and international partners during an up-coming dissemination and planning exercise for the

Composite Survey. ESSPIN will encourage further tailoring of the programme in each state to address the increasingly differentiated types of support requested by engaged political and school system leaders, and demonstrated as necessary according to the state-specific CS findings about children's learning.

114. Overall, the level of learning achieved by the majority of children in Nigeria's public primary schools remains unacceptably low, and this has been borne out in the Composite Survey 2012, as it was in the 2010 MLA. The CS has revealed in particular the significant proportions of children who are unable to achieve at grade appropriate levels in English literacy and maths, and the even-more worrying sections of the school population who struggle with even the most basic elements of p1 and p2 material by the time they have reached p2 and p4. These findings are not surprising, but it is important to provide clear evidence of the continuing scale of the problem for system accountability, diagnosis and advocating a coherent and relevant response. And this is before the question of out-of-school children is considered.

115. Further insights have been gained by breaking subject areas down into learning domains. There are similar patterns in English literacy and mathematics, in that children's acquisition of the building blocks of learning—such as skills for early reading and number concepts—is not translating readily enough into *applications* of those building blocks, such as reading with comprehension or solving maths problems. These in turn can potentially be traced back to teachers' limited understanding of the curriculum content, as well as their limited skill in scaffolding children to access the curriculum in turn. Furthermore, these findings play out differently state by state, and no doubt are subject to wide variations within states between LGAs and from school to school. ESSPIN will encourage a deeper understanding by teachers of the concept of children's progression through levels of knowledge and skills; how to relate these to their own classroom practice in areas such as differentiation between children according to their individual needs; relating lessons to the Learning Outcomes Benchmarks which are now available; and introducing a greater focus on formative classroom assessment to unpack some of these issues.

116. ESSPIN will focus on strengthening teachers further, by developing and extending the "best practice" initiative currently being piloted in 264 schools in Kano State (UBEC TPD funds). The best practice approach builds on and deepens generic teaching skills through a focus on the processes of learning, rather than teaching. In the Kano case, the emphasis is initially on basic English literacy and numeracy, but that restriction need not necessarily apply elsewhere. What is certain is that all teachers need help to be equipped to respond to the needs of children who have been left behind to date. The Composite Survey provides encouragement in that the School Improvement Programme is associated with fewer children in the bottom performance ranges, and more in the top.

117. ESSPIN will extend the use of p1-p3 lesson plans across all States, and introduce lesson plans into upper primary, initially by introducing p4 lesson plans throughout the upper primary grades, supported by the "best practice" model outlined above. However, distribution of teaching materials has never been thought of by ESSPIN as sufficient to improve learning: the programme will reinforce teachers' ability to understand and teach basic literacy and basic and intermediate numeracy through improved *use* of lesson plans. Applied or action research on how teachers use the plans will be undertaken to

design enhanced support. And a review of the teaching of higher order reading skills and writing (as opposed to copying) will be undertaken, with a programmatic response to the outcome.

118. Also in the investigative mode, ESSPIN will look into the different reasons (other than weak teaching or patchy attendance) why groups of children are not achieving foundation skills at p2 and p4 as shown in the CS, and respond accordingly. Such a review will need to be sensitive to gendered, socioeconomic and linguistic characteristics of pupils and teachers. Options to be explored include bringing teaching assistants into classrooms to work with the children most in need of extra help.

119. A new Challenge Fund will be proposed, with two specific learner participation objectives: (i) to support low cost, locally developed, innovative ways of putting reading or story material into classrooms; and (ii) community- and school-managed ways of ensuring all pupils in p1-p3 have something to write on and with. An evaluation of the use of Lifeplayer will be undertaken to inform the further roll out of this teaching resources facility.

120. ESSPIN plans to explore correlation relationships within the rich CS data set, which has been collected, cleaned and is about to be made publicly accessible, once confidentiality safeguards for respondents are in place. These would cast further light on the elements of the SIP model which go together, and are suggestive of an effective intervention in schools: similar to the finding in paragraph 69 above that very few schools achieved the three standards on head teacher effectiveness, school development planning and SBMC functionality without *also* meeting the teacher competence standard. ESSPIN uses its value for money data to constantly review the effectiveness of investments, by state and by programme output. The Composite Survey promises to add a new layer of insight to these management and M&E tools.

121. During 2012-13, there has been a sharp increase in the inter-state use of SSIT and state specialist team members to help develop and quality assure work of other states' teams and programmes. This acts as a two-way professional development process. It is an approach which ESSPIN will foster in future, for its professional effectiveness and excellent value for money, as well as being a key strategy for ensuring continued learning in and from initiatives in the northern states which remain inaccessible to international team members at present on security grounds.

122. ESSPIN will further develop professional leadership on pedagogy within and across the states, and in concert with incoming initiatives such as TDP and EDORE. This will embrace State Quality Specialists (SQS) themselves, SSITs, and explore the concept of a national SSIT. The capacity to both produce and use classroom research will be supported jointly with SQS, selected SSIT members, College of Education staff, officers from DPRSs, CSOs—a real classroom research partnership resource for each State. Such developments are required if the potential for stronger state-level leadership of school improvement is to be sustained beyond the ESSPIN lifetime.

References

DFID (November 2011) 'Annex 10: DFID learning indicator methodology'

ESSPIN (February 2011) 'ESSPIN: Logframe Indicator Handbook'.

ESSPIN (December 2010) 'Report on MLA testing in the 6 ESSPIN States June 2010'

ESSPIN (July 2012a) 'ESSPIN Composite survey: Revised summary concept paper'

Annex A: Sampling, survey weights and statistical tests

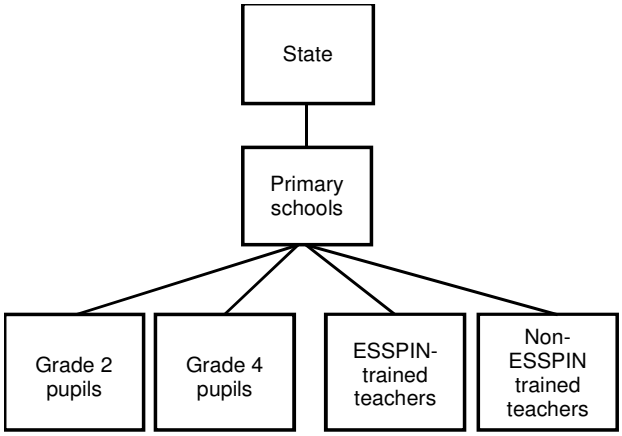
Sampling Strategy

This section outlines the sampling strategy and target sample sizes for each unit of observation for the 2012 ESSPIN composite survey conducted in the six focus states: Enugu, Jigawa, Kaduna, Kano, Kwara and Lagos.

The analysis requires estimation of several indicators for each of the units of observation and where the 2010 MLA data and documentation allow it, attribution of any observed changes in the outputs and outcomes of interest over time to corresponding ESSPIN programme interventions; therefore, the sample of units was selected with rigorous scientific procedures in order that selection probabilities are known.

The school sample frame was constructed using information on school ESSPIN and 2010 MLA survey participation and school size from the Education Management Information System (EMIS). To enable the planned analyses a multi-stage sampling design was used as shown in Figure A.1

Figure A.1: ESSPIN composite survey 2012 sampling units and stages



The lines connecting the units of observation in Figure A.1 represent sampling stages. The six survey states were pre-determined as the ESSPIN programme operates in these states. In each focus state, public primary schools were selected (first stage), and then within each sample school, teachers and grade 2 and grade 4 pupils respectively (second stage) were selected.

In the first sampling stage, there is stratification in order to allow the observation of a minimum number of units in each stratum of various types of analytical importance such as ESSPIN phase 1, ESSPIN phase 2, and control (no ESSPIN interventions) schools. Table A.1 to Table A.6 present the strata defined for each sampling stage and the approximate expected sample sizes for each of the six focus states. The total intended sample across the six states was 595 public primary schools.[11]

[11] There was some downward adjustment to the total sample size if for example some selected schools could not be visited as part of the survey due to security or other access issues.

Table A.1: ESSPIN composite survey 2012 sampling strategy outline: Jigawa

Sampling units	Stage	Domain	Strata		Number of units selected	Selection method	Total sample size	
Primary schools	1	State	ESSPIN phase 1	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	105
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
			ESSPIN phase 2	2010 MLA	17 per phase (if less than 17 all)	Selected with certainty	7	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	28	
			Non-ESSPIN schools	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
Teachers	2	All 105 selected primary schools	Head teacher		1	Selected with certainty	105	~1155
			ESSPIN trained teachers		5 within each school	Simple random sampling	525	
			non-ESSPIN trained teachers		5 within each school		525	
Pupils	2	All 105 selected primary schools	Grade 2	Literacy	4 per grade and subject	Simple random sampling	420	~1680
			Grade 2	Mathematics	4 per grade and subject		420	
			Grade 4	Literacy	4 per grade and subject		420	
			Grade 4	Mathematics	4 per grade and subject		420	

Table A.2: ESSPIN composite survey 2012 sampling strategy outline: Kaduna

Sampling units	Stage	Domain	Strata		Number of units selected	Selection method	Total sample size	
Primary schools	1	State	ESSPIN phase 1	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	105
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
			ESSPIN phase 2	2010 MLA	17 per phase (if less than 17 all)	Selected with certainty	5	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	30	
			Non-ESSPIN schools	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
Teachers	2	All 105 selected primary schools	Head teacher		1	Selected with certainty	105	~1155
			ESSPIN trained teachers		5 within each school	Simple random sampling	525	
			non-ESSPIN trained teachers		5 within each school		525	
Pupils	2	All 105 selected primary schools	Grade 2	Literacy	4 per grade and subject	Simple random sampling	420	~1680
			Grade 2	Mathematics	4 per grade and subject		420	
			Grade 4	Literacy	4 per grade and subject		420	
			Grade 4	Mathematics	4 per grade and subject		420	

Table A.3: ESSPIN composite survey 2012 sampling strategy outline: Kano

Sampling units	Stage	Domain	Strata		Number of units selected	Selection method	Total sample size	
Primary schools	1	State	ESSPIN phase 1	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	105
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
			ESSPIN phase 2	2010 MLA	17 per phase (if less than 17 all)	Selected with certainty	3	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	32	
			Non-ESSPIN schools	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
Teachers	2	All 105 selected primary schools	Head teacher		1	Selected with certainty	105	~1155
			ESSPIN trained teachers		5 within each school	Simple random sampling	525	
			non-ESSPIN trained teachers		5 within each school		525	
Pupils	2	All 105 selected primary schools	Grade 2	Literacy	4 per grade and subject	Simple random sampling	420	~1680
			Grade 2	Mathematics	4 per grade and subject		420	
			Grade 4	Literacy	4 per grade and subject		420	
			Grade 4	Mathematics	4 per grade and subject		420	

Table A.4: ESSPIN composite survey 2012 sampling strategy outline: Kwara1

Sampling units	Stage	Domain	Strata		Number of units selected	Selection method	Total sample size	
Primary schools	1	State	ESSPIN phase 1	2010 MLA	17 per phase (if less than 17 all)	Selected with certainty	5	105
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	30	
			ESSPIN phase 2	2010 MLA	17 per phase (if less than 17 all)	Selected with certainty	2	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	33	
			Non-ESSPIN schools	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
Teachers	2	All 105 selected primary schools	Head teacher		1	Selected with certainty	105	~1155
			ESSPIN trained teachers		5 within each school	Simple random sampling	525	
			non-ESSPIN trained teachers		5 within each school		525	
Pupils	2	All 105 selected primary schools	Grade 2	Literacy	4 per grade and subject	Simple random sampling	420	~1680
			Grade 2	Mathematics	4 per grade and subject		420	
			Grade 4	Literacy	4 per grade and subject		420	
			Grade 4	Mathematics	4 per grade and subject		420	

Note: The phases have a different meaning in Kwara to the other states because ESSPIN head teacher and teacher interventions covered all public primary schools. These are termed 'Non-ESSPIN schools' or control group. The phase 1 schools are the select subgroup which also received other SIP components. The phase 2 schools are those schools which will receive other SIP components.

Table A.5: ESSPIN composite survey 2012 sampling strategy outline: Lagos

Sampling units	Stage	Domain	Strata		Number of units selected	Selection method	Total sample size	
Primary schools	1	State	ESSPIN phase 1	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	105
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
			ESSPIN phase 2	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
			Non-ESSPIN schools	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
Teachers	2	All 105 selected primary schools	Head teacher		1	Selected with certainty	105	~1155
			ESSPIN trained teachers		5 within each school	Simple random sampling	525	
			non-ESSPIN trained teachers		5 within each school		525	
Pupils	2	All 105 selected primary schools	Grade 2	Literacy	4 per grade and subject	Simple random sampling	420	~1680
			Grade 2	Mathematics	4 per grade and subject		420	
			Grade 4	Literacy	4 per grade and subject		420	
			Grade 4	Mathematics	4 per grade and subject		420	

Table A.6: ESSPIN composite survey 2012 sampling strategy outline: Enugu

Sampling units	Stage	Domain	Strata		Number of units selected	Selection method	Total sample size	
Primary schools	1	State	ESSPIN phase 1	2010 MLA	17 per phase (if less than 17 all)	Selected with certainty	3	70
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	32	
			Non-ESSPIN schools	2010 MLA	17 per phase (if less than 17 all)	Systematic equal probability sampling	17	
				2010 non-MLA	35 minus number of 2010 MLA schools	Probability proportional to size (measure of size is number of teachers)	18	
Teachers	2	All 105 selected primary schools	Head teacher		1	Selected with certainty	105	~805
			ESSPIN trained teachers		5 within each school	Simple random sampling	350	
			non-ESSPIN trained teachers		5 within each school		350	
Pupils	2	All 105 selected primary schools	Grade 2	Literacy	4 per grade and subject	Simple random sampling	280	~1120
			Grade 2	Mathematics	4 per grade and subject		280	
			Grade 4	Literacy	4 per grade and subject		280	
			Grade 4	Mathematics	4 per grade and subject		280	

Note: ESSPIN phase 2 schools had not been identified at the time of the ESPPIN composite survey and therefore no Phase 2 schools were included in the sample for Enugu.

Selection probabilities and survey weights

This section describes the implementation of the sampling stages shown in Table A.1 to Table A.6 and the selection probabilities of the selected units. To obtain representative estimates from each of the samples the observed values need to be analysed using **sampling weights equal to the inverse of the provided selection probabilities**.

Public primary schools

The major sampling strata (hereafter denoted with the subscript *h*) are the schools' participation in the ESSPIN programme: ESSPIN phase 1 schools, ESSPIN phase 2 schools, and control (no ESSPIN intervention) schools in each of the six states with the exception of Enugu, where there are no phase 2 schools. Each of the major strata is divided into two sub-strata, respectively composed of the schools selected and not selected for the 2010 MLA survey.

2010 MLA schools were selected in one of two ways depending on the total number of 2010 MLA schools in the 2010 MLA school sub-strata. If there were more than 17 MLA schools, 17 were selected using systematic equal probability sampling and if there were fewer than 17 MLA schools, all were selected with certainty.

The reason for using systematic equal probability sampling was that this method was used to select the school sample for the 2010 MLA survey combined with the need for a minimum number of 2010 MLA schools to be contained within the 2012 sample in order to enable analysis over time of any changes in pupil learning as measured by the MLA.[12]

The probability of selecting school *i* in stratum *h* is given by:

$$p_{hi} = \begin{cases} \frac{n_h}{N_h} & for\ 2010\ MLA\ schools \\ \frac{(N_h - m_{h)}}{N_h} \frac{(35 - n_h) t_{hi}}{T_h} & for\ non\text{-}2010\ MLA\ schools \end{cases} \quad \text{.....[1]}$$

where

m_h is the number of public primary schools that participated in the 2010 MLA survey in the stratum;

n_h is the number of public primary schools that participated in the 2010 MLA survey in the stratum selected by the ESSPIN project;

N_h is the total number of public primary schools in the state as reported in the school sample frame in the stratum;

t_{ih} is the number of teachers in each school as per the sample frame; and

[12] The possibility of a 2010 and 2012 comparison depends on whether the 2010 MLA sample, data, and documentation allow it.

T_h is the total number of teachers, as per the sample frame, in all public primary schools in the state that did not participate in the 2010 MLA survey in the stratum.

Table A.7 provides the values of N_h, m_h, n_h and T_h in the 17 strata.

Table A.7: Public primary schools by state, phase, 2010 MLA survey and ESSPIN participation

Stratum (State and phase)		Total number of public primary schools	Number of 2010 MLA public primary schools	of which, selected for ESSPIN	Number of teachers in non-2010 MLA public primary schools
h		N_h	m_h	n_h	T_h
Enugu	Controls	1,078	31	17	11,183
	Phase 1	115	3	3	1,080
Jigawa	Controls	1,272	30	17	8,133
	Phase 1	101	33	17	1,324
	Phase 2	267	7	7	2,826
Kaduna	Controls	3,238	35	17	27,721
	Phase 1	113	40	17	1,947
	Phase 2	279	5	5	4,792
Kano	Controls	3,902	31	17	31,359
	Phase 1	207	36	17	2,431
	Phase 2	575	3	3	6,656
Kwara[1]	Controls	971	28	17	9,508
	Phase 1	217	5	5	2,045
	Phase 2	73	2	2	1,119
Lagos	Controls	375	18	17	4,765
	Phase 1	61	31	17	987
	Phase 2	455	21	17	6,305
Total		13,299	359	212	124,181

Note: (1) The phases have a different meaning in Kwara from the other states because ESSPIN head teacher and teacher interventions covered all public primary schools. These are termed 'Non-ESSPIN schools' or control schools. The phase 1 schools are the select subgroup which also received other SIP components. The phase 2 schools are those schools which will receive other SIP components.

Head teachers and teachers

The head teacher was interviewed in all sample schools.

Five ESSPIN-trained and five non-ESSPIN-trained teachers were selected in each sample school using simple random sampling. The teacher sampling was conducted in schools by the enumerators who used a special form and random number tables (see *Sample selections in the field* below).

The probability of choosing a (ESSPIN-trained or non-ESSPIN trained respectively) teacher *j* in school *i* of stratum *h* is given by:

$$p_{hij} = p_{hi}\frac{5}{t_{hi}} \quad \text{.....[2]}$$

where

p_{hi} is the probability of selecting the school (given by Equation [1]); and

t_{hi} is the number of eligible teachers in the school at the time it was visited for the ESSPIN survey.

Pupils

Four grade 2 pupils and four grade 4 pupils were selected for each of the literacy and numeracy assessments respectively in each sample school using simple random sampling. The pupil sampling was conducted in schools by the enumerators who used a special form and random number tables (see *Sample selections in the field* below) similar to the teacher sampling.

The probability of choosing a (grade 2 or grade 4 respectively) pupil *k* in school *i* of stratum *h* is given by:

$$p_{hik} = p_{hi}\frac{4}{s_{hi}} \quad \text{.....[3]}$$

where

p_{hi} is the probability of selecting the school (given by Equation [1]); and

s_{hi} is the number eligible pupils in the school at the time it was visited for the ESSPIN survey.

Survey weights

The sampling weights for schools, teachers and pupils are the inverse of their respective selection probabilities.

Sample selections in the field

The teacher and pupil sampling was conducted in the field. The sampling selections delegated to the enumerators were conducted as a part of interviewing processes that had broader objectives. For this reason the selection processes were not supported by stand-alone forms but were instead integrated with the survey questionnaires and used as follows for pupils (the same procedure was used for teacher sampling):

- First, the enumerator used the school's pupil register to write pupil codes next to each pupil name starting with 1 for the first pupil listed up until the last pupil listed, which provided the largest pupil code.
- Second, the interviewer wrote down the largest pupil code in a box on the questionnaire.
- Third, the interviewer scanned the provided random number table according to the instructions provided to find the pupil codes of the eligible pupils to be selected.

Actual sample obtained

Table A.8: Sample units selected and interviewed by state

Unit	Intended sample	Total # of records	Highest # of records used in key estimates	lowest # of records used in key estimates	# of records/intended sample (%)	Highest # of records/total (%)	Lowest # of records /total (%)
	[1]	[2]	[3]	[4]	[2]/[1]	[3]/[2]	[4]/[2]
Public primary schools							
All 6 states	595	587	583	485	99%	99%	83%
Enugu	70	70	70	59	100%	100%	84%
Jigawa	105	103	103	87	98%	100%	84%
Kaduna	105	105	103	82	100%	98%	78%
Kano	105	104	104	87	99%	100%	84%
Kwara	105	102	101	85	97%	99%	83%
Lagos	105	103	102	85	98%	99%	83%
Teachers							
All 6 states	5950	4297	4121	3939	72%	96%	92%
Enugu	700	492	451	437	70%	92%	89%
Jigawa	1050	743	731	686	71%	98%	92%
Kaduna	1050	730	687	676	70%	94%	93%
Kano	1050	830	783	726	79%	94%	87%
Kwara	1050	737	731	727	70%	99%	99%
Lagos	1050	764	738	687	73%	97%	90%
Primary 2 and Primary 4 pupils							
All 6 states	9520	9106	8923	8923	96%	98%	98%
Enugu	1120	1080	1052	1052	96%	97%	97%
Jigawa	1680	1582	1518	1518	94%	96%	96%
Kaduna	1680	1640	1620	1620	98%	99%	99%
Kano	1680	1635	1603	1603	97%	98%	98%
Kwara	1680	1562	1545	1545	93%	99%	99%
Lagos	1680	1607	1585	1585	96%	99%	99%

Source: Composite Survey 2012.

Statistical tests

(i) Significance test for difference between means of independent samples

Calling μ_T and μ_{C0} the means for the Phase 1 treatment group and control group respectively; and σ_T and σ_c the corresponding standard errors, then the quantity

$$t = \frac{(\mu_T - \mu_C)}{\sqrt{\sigma_T^2 + \sigma_C^2}}$$

reveals a significant difference in differences if $t > t_\alpha$, where t_α is the normal deviate corresponding to the desired significance level α. In our case, α = 5% and t_α = 1.96.

(ii) Significance test for difference between differences of means of independent samples

Calling μ_{T0} and μ_{T1} the means for the Phase 1 treatment group in 2010 and 2012; μ_{C0} and μ_{C1} the means in the control group; and σ_{T0}, σ_{T1}, σ_{C0} and σ_{C1} the corresponding standard errors, then the quantity

$$t = \frac{(\mu_{T1} - \mu_{T0}) - (\mu_{C1} - \mu_{C0})}{\sqrt{\sigma_{T0}^2 + \sigma_{T1}^2 + \sigma_{C0}^2 + \sigma_{C1}^2}}$$

reveals a significant difference in differences if $t > t_\alpha$, where t_α is the normal deviate corresponding to the desired significance level α. In our case, α = 5% and t_α = 1.96.

Note: the second formula above was used in the CS analysis to test whether there is a significant difference between in the over-time gains in the numeracy logframe estimates between phase 1 and the control group at 5% level of significance. This is a prudent test of statistical significance in this case, because the 2010 and 2012 samples are not completely independent. There are some common schools. Strictly speaking, in this situation, the formula needs to be adjusted to take account of the covariance of the estimates for the common schools. This would reduce the denominator and increase the value of the test statistic, making a significant result more likely. In our analysis, the test statistic obtained from the unadjusted formula (for both the p2 and p4 numeracy logframe indicators) was so far from the critical value of 1.96 that the adjustment to account for common schools (note that the pupils are not common) would not alter the conclusion.

Annex B: ESSPIN Programme

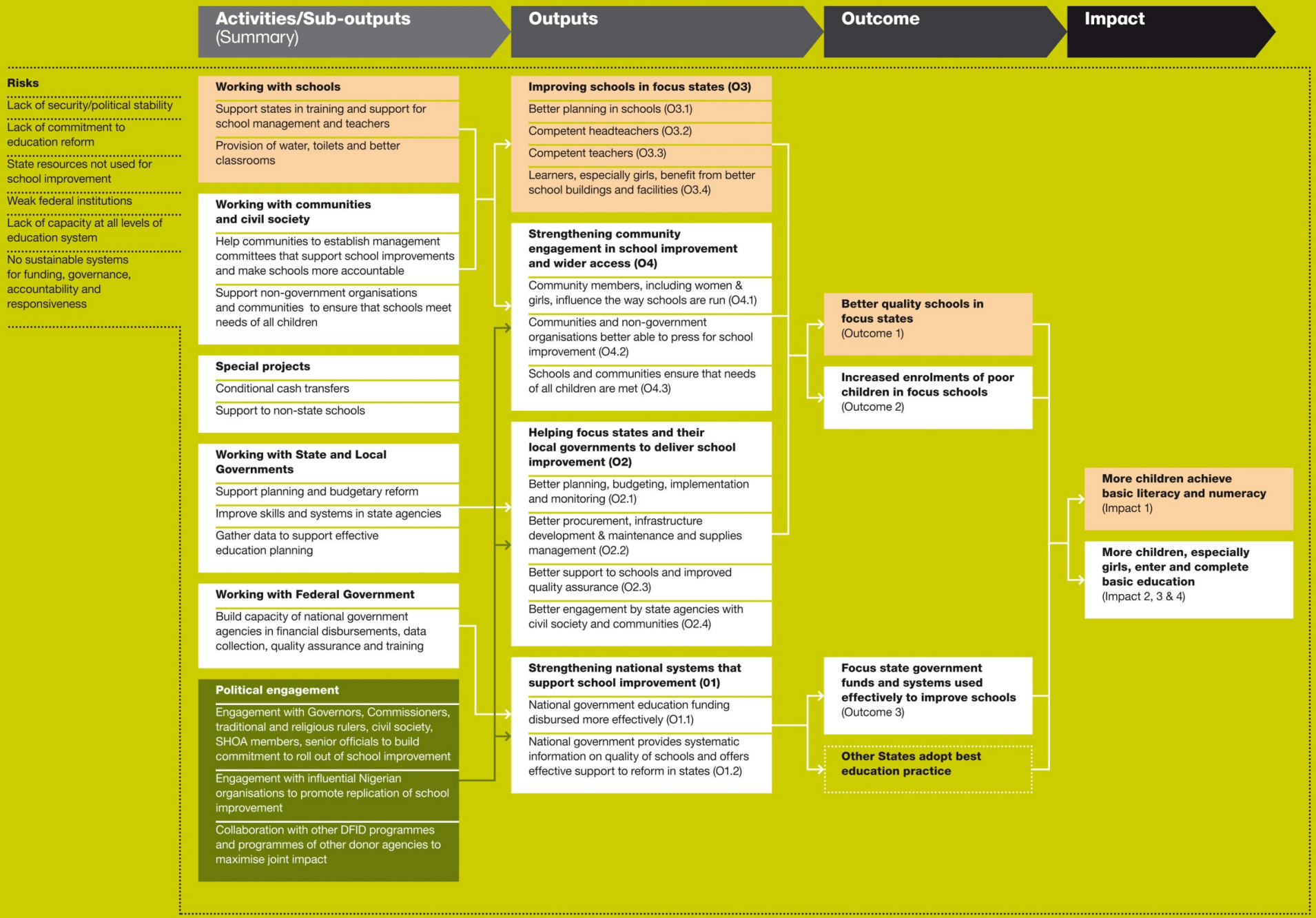
The diagram on the following page details the results chain underpinning ESSPIN's theory of change, which was revised in January 2012 following recommendations from the programme's mid-term review. The results chain sets out how inputs are expected to be translated into results: if a set of activities (interventions) is carried out and sets of external assumptions hold, then a set of identified outputs can be anticipated which in turn can be expected to contribute to a set of desired outcomes having, in turn, desired impacts.

The ESSPIN results chain gives the full programme context for the school- and community-level interventions and outputs, which are the focus of the composite survey. To meet the programme outcome and impact targets, the theory postulates that national and state-level system reforms are needed to reinforce the school- and community level interventions to contribute to the higher order results.

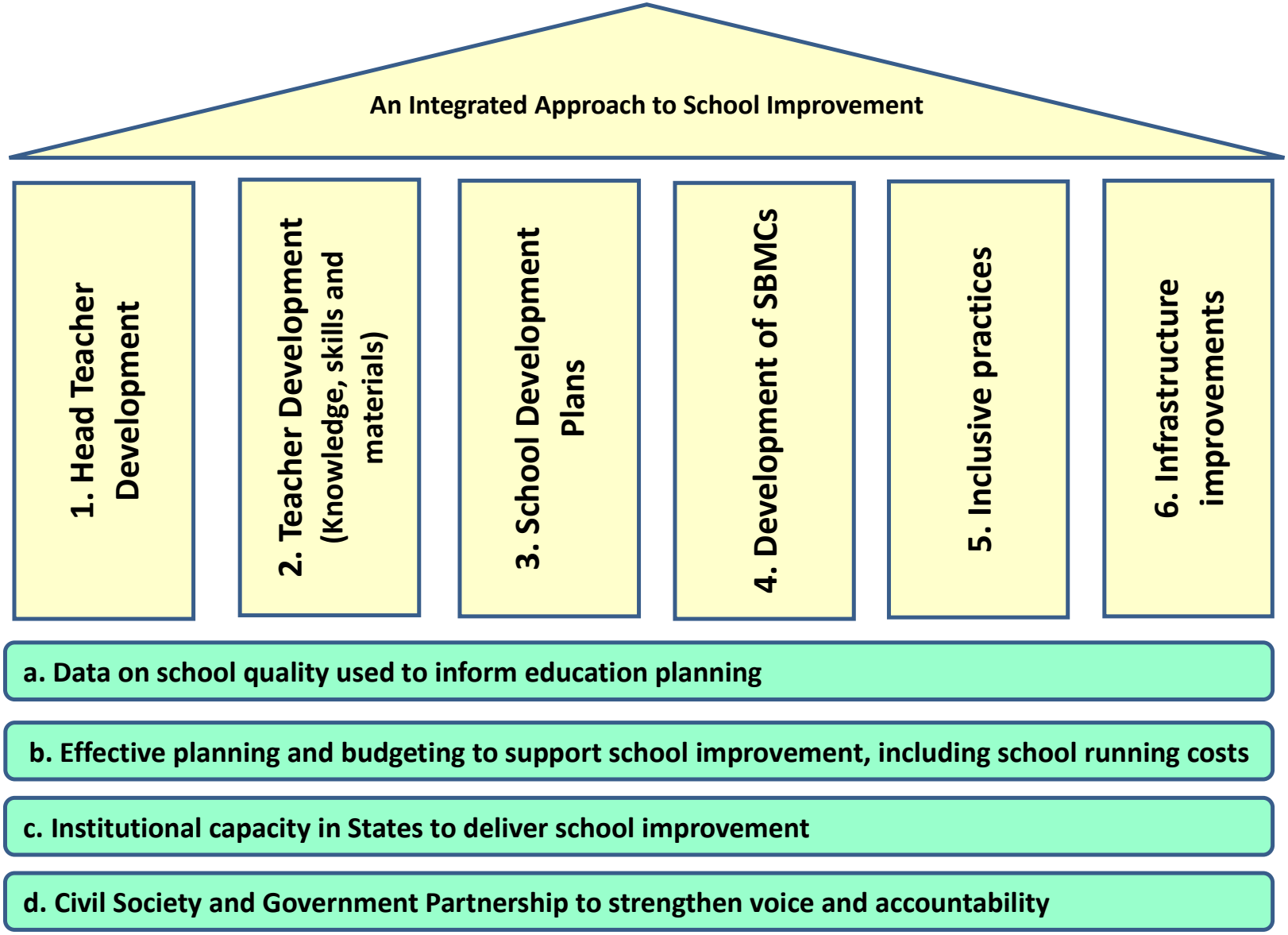
ESSPIN's integrated approach to school improvement is set out in a diagram on the page following the theory of change diagram below. This 'Greek Temple' model shows ESSPIN's school- and community-level interventions as vertical pillars, which are built on system-level interventions (state and national) shown as horizontal foundations. This integrated programme is known as the School Improvement Programme (SIP).

Theory of change and results chain

[Following pages…]


Activities/Sub-outputs (Summary)
Outputs
Outcome
Impact
Risks
Lack of security/political stability
Lack of commitment to education reform
State resources not used for school improvement
Weak federal institutions
Lack of capacity at all levels of education system
No sustainable systems for funding, governance, accountability and responsiveness
Working with schools
Support states in training and support for school management and teachers
Provision of water, toilets and better classrooms
Working with communities and civil society
Help communities to establish management committees that support school improvements and make schools more accountable
Support non-government organisations and communities to ensure that schools meet needs of all children
Special projects
Conditional cash transfers
Support to non-state schools
Working with State and Local Governments
Support planning and budgetary reform
Improve skills and systems in state agencies
Gather data to support effective education planning
Working with Federal Government
Build capacity of national government agencies in financial disbursements, data collection, quality assurance and training
Political engagement
Engagement with Governors, Commissioners, traditional and religious rulers, civil society, SHOA members, senior officials to build commitment to roll out of school improvement
Engagement with influential Nigerian organisations to promote replication of school improvement
Collaboration with other DFID programmes and programmes of other donor agencies to maximise joint impact
Improving schools in focus states (O3)
Better planning in schools (O3.1)
Competent headteachers (O3.2)
Competent teachers (O3.3)
Learners, especially girls, benefit from better school buildings and facilities (O3.4)
Strengthening community engagement in school improvement and wider access (O4)
Community members, including women & girls, influence the way schools are run (O4.1)
Communities and non-government organisations better able to press for school improvement (O4.2)
Schools and communities ensure that needs of all children are met (O4.3)
Helping focus states and their local governments to deliver school improvement (O2)
Better planning, budgeting, implementation and monitoring (O2.1)
Better procurement, infrastructure development & maintenance and supplies management (O2.2)
Better support to schools and improved quality assurance (O2.3)
Better engagement by state agencies with civil society and communities (O2.4)
Strengthening national systems that support school improvement (01)
National government education funding disbursed more effectively (O1.1)
National government provides systematic information on quality of schools and offers effective support to reform in states (O1.2)
Better quality schools in focus states (Outcome 1)
Increased enrolments of poor children in focus schools (Outcome 2)
Focus state government funds and systems used effectively to improve schools (Outcome 3)
Other States adopt best education practice
More children achieve basic literacy and numeracy (Impact 1)
More children, especially girls, enter and complete basic education (Impact 2, 3 & 4)

School Improvement Programme (SIP)



ESSPIN phase 1 and phase 2 school selection procedures by state

The school selection process for both phases was the responsibility of the State Ministries for Education and SUBEBs, assisted by the relevant ESSPIN state team. This is evident from the different selection models chosen by each state. ESSPIN is an education reform programme in which technical assistance and resources are used to support state administrations to improve education standards. It is not a research exercise and therefore random selection of schools for the pilot stage (Phase 1) or roll out (Phase 2) was not employed and would not be expected.

This does though have implications for how Composite Survey 1 findings should be interpreted. Where better performance is observed in Phase 1 schools than controls (or Phase 2), it is not possible to attribute that performance unambiguously to the effect of ESSPIN, since inherent characteristics of those schools could determine that difference. In the absence of evidence of purposive selection to include the 'best' schools in Phase 1, there is no reason to believe that ESSPIN's work was not responsible for those results, but no such claim can reliably be made at this time. Composite Survey 2, in 2014, is intended to address this limitation, by measuring the difference in differences between change observed in control schools and change observed in Phase 1 schools and change observed Phase 2 schools—thereby controlling for any possible selection bias.

The paragraphs below set out the selection processes used in each state, to inform the reader's judgement about the likelihood of systematic bias being responsible for any observed differences in the findings reported for Phase 1, Phase 2 and control schools.

Enugu Phase 1: All the public primary schools came from one LGEA: Udi LGA, which was selected in consultation with the state government, from three shortlisted LGAs identified by the 2009 ASC as having the lowest education performance indicators. 30 low-fee paying Mission schools were also selected on the basis of criteria which sought to address social exclusion issues. ESSPIN, supported by Enugu-based civil society organizations, worked with the three missions – Catholic, Anglican and Methodist – to agree on the number and identity of the schools. Therefore, although the schools were not selected at random, there is no reason to believe that Enugu Phase 1 schools have inherent characteristics which would positively bias their performance in comparison with control schools. Phase 2 not applicable.

Jigawa Phase 1: SUBEB selected nine LGEAs (one from each of the training zones) and then selected approximately 25% of the schools from each of these LGEAs. Whilst the selection procedure was not random, neither was the purposive selection process used informed by an attempt to select the 'best' schools. Phase 2 schools were not part of the survey.

Kaduna Phase 1: All schools in two or three districts selected by the ESs (only Kaduna have districts) of seven LGEAs (two from the southern and central Senatorial District and three from the north). Phase 2: Similar, from the remaining LGEAs that had not been part of SESP (although one of these did sneak in).

Kano Phase 1: Out of the 44 LGAs in Kano State, one urban, one peri-urban and one rural LGA were selected; and within those, 69% of public primary schools and 55% of junior secondary schools were included in the pilot stage of the programme. Phase 2 schools were not part of the survey.

Kwara Phase 1 (those which had SBMC and HT training) all schools in four non SESP LGAs spread across the 3 senatorial districts to give approximately 200 schools.

Lagos Phase 1: Each state senatorial district to be represented. All LGAs eligible: 20 constitutionally recognised LGEAs. Same number of schools included per LGEA for sake of equity. Schools eligible only with an adequate level of infrastructure supplied by government. LGEA s and SUBEB prepared list of schools from which five were selected within each LGEA. Phase 2: 25 further schools from each LGEA.

Annex C: Definition of ESSPIN logframe indicators; test items

Logframe indicators

Impact (in six focus states)

1. Proportion of primary 4 (p4) and primary 2 (p2) pupils in public primary schools in focus states who:

   (a) Demonstrate skills for reading comprehension (applies to p2); Demonstrate ability to read with comprehension (applies to p4) (%)

   (b) Demonstrate ability to do basic arithmetic calculations at grade appropriate level (%)

Outcome (in six focus states)

1. Proportion of primary and junior secondary schools that meet benchmarks for a good quality school (%)

Outputs (across group of schools where interventions have taken place in six focus states)[13]

3.1 Proportion of primary and junior secondary schools using a school development plan (%)

3.2 Proportion of head teachers in primary and junior secondary schools operating effectively (%)

3.3 Proportion of teachers in primary and junior secondary schools who can deliver competent lessons in literacy and numeracy (%)

4.1a Proportion of primary and junior secondary schools with functioning SBMCs (%)

4.1b Proportion of communities served by primary and junior secondary schools where SBMCs reflect concerns of women and children (%)

[13] For the purpose of the logframe, the estimates of output indicators will cover the schools which have received interventions (phase 1 schools at the time of the composite survey). The survey will also generate estimates of these indicators for a group of non-intervention (control) schools, for the purpose of evaluation.

Test items used to calculate the ESSPIN logframe pupil learning indicators

The relevant questions are given below. The *[Italics]* underneath each question sets out the criterion used to classify a pupil as demonstrating the particular skill being tested.

P2 Literacy

Proportion of p2 children who demonstrate skills for reading comprehension:
Proportion of p2 children who correctly answer a p2 curriculum level question on listening comprehension (Q11) and correctly read a sufficient number of words from a p2 curriculum level passage (Q13).

P2 Q11 Listening comprehension

This is an oral question and answer. **Do not** show this page to the pupil.

Ndi has two brothers. Their names are Paul and Raymond. They are older than Ndi. Ndi likes to go to school, because she has many friends there.
Ask the pupil:
11a) How many brothers does Ndi have?
11b) Why does Ndi like to go to school?

[Pupil must answer both parts correctly].

P2 Q13 Reading a passage aloud

Read the following passage aloud:

Good morning. My name is Fatima. I am	*8*
seven years old. My brother's name is Sam.	*16*
He is five years old. I also have a sister. Her	*27*
name is Nandi. We like to read stories. We go	*37*
to the market every Saturday. My mother sells	*45*
fruit at the market in town.	*51*

[Pupil must get 26 or more words correct].

P4 Literacy

Proportion of p4 children who demonstrate ability to read with comprehension: Proportion of p4 children who correctly read a sufficient number of familiar words at p4 curriculum level (Q21) and correctly read a sufficient number of words from a p4 curriculum level passage (Q23) and correctly answer at least 4 out of 5 reading comprehension questions (Q23).

P4 Q21 Pronunciation (Reading)

Ask the pupil to read the words as quickly and carefully as they can, going along the row.

1	*2*	*3*	*4*	*5*	
back	glass	quick	small	start	5
fall	vary	shot	bird	miss	*10*
animal	calendar	beginning	introduce	medicine	15
chicken	their	carry	handle	rhyme	*20*
apple	banana	grass	yellow	orange	25
mistake	sugar	tangle	hospital	through	*30*

[Pupil must get 15 or more words correct].

P4 Q23 Reading with comprehension

My name is Umar. I live on a farm with my mother, father and sister Fatima. 16

Every year the land gets very dry before the rains come. We watch the sky and wait. 33

One afternoon as I sat outside, I saw dark clouds. Then something hit my head, lightly at first and then harder. 54

I jumped up and ran towards the house. The rains had come at last. 68

[Pupil must get 34 or more words correct and answer at least 4 out of 5 comprehension questions correctly].

P2 Numeracy

Proportion of p2 children who demonstrate ability to do basic arithmetic calculations: Proportion of p2 children who correctly answer at least 5 out of 6 p2 curriculum level questions on addition and subtraction (Q14) and both multiplication questions (Q15).

P2 Q14 Addition and subtraction of two and three digit numbers

The pupil will write the answers to the sums on this page in the spaces.

14a) 32 + 16 = _____

14b) 25 + 7 = _____

14c) 234 + 342 = _____

	H	T	U
	2	3	4
+	3	4	2

14d) 19 – 6 = _____

14e) 16 - 8 = _____

14f) 49 – 22 = _____

[Pupil must get at least five out of six sums correct].

P2 Q15 Multiplication of single digit numbers

The pupil will write the answers to the two sums on this page in the spaces provided.

Multiply these two numbers together.

15a 3 x 2 = _____

15b 4 x 4 = _____

[Pupil must get both sums correct].

P4 Numeracy

Proportion of p4 children who demonstrate ability to do basic arithmetic calculations: Proportion of p4 children who correctly answer p4 curriculum level questions on addition and subtraction (Q25) and multiplication (Q26) and division (Q27).

P4 Q25 Addition and subtraction

a)

$$\begin{array}{rrrr} & 3 & 2 & 3 \\ & 2 & 1 & 4 \\ + & 1 & 6 & 1 \\ \hline & & & \\ \hline \end{array}$$

b)

$$\begin{array}{rrrr} & 3 & 4 & 0 \\ & 6 & 4 & 3 \\ + & 6 & 3 & 4 \\ \hline & & & \\ \hline \end{array}$$

c)

$$\begin{array}{rrrr} & 4 & 3 & 2 \\ - & & 6 & 1 \\ \hline & & & \\ \hline \end{array}$$

[Pupil must get all three sums correct].

P4 Q26 Multiplication of numbers

a)

$$\begin{array}{rrr} & 1 & 4 \\ x & & 7 \\ \hline & & \\ \hline \end{array}$$

b)

$$\begin{array}{rrr} & 3. & 42 \\ x & & 2 \\ \hline & & \\ \hline \end{array}$$

[Pupil must do both multiplications correctly].

P4 Q27 Dividing numbers

a) 4 / 68 = _____

[Pupil must do division correctly].

Mapping of pupil learning test items to grade level and to learning domains

The tables below show a mapping of questions from the composite survey tests (available separately) to the four grade levels (p1, p2, p3 and p4), and to the different literacy and numeracy learning domains.

P2 Literacy (questionnaire 7)

Learning Domain	**Question no.**	
	Grade level	
	P1	***P2***
Early reading	3, 4, 7	
Skills for reading with comprehension		11, 13
Writing		9, 14, 16
Other	1, 2, 5, 6, 17	8, 10, 12, 15

P4 Literacy (questionnaire 8)

Learning Domain	**Question no.**		
	Grade level		
	P1 & P2	***P3***	***P4***
Early reading	3		
Skills for reading with comprehension	8	11, 12, 13, 15, 16	21, 23a, 23b
Writing	9	18, 19, 20	25
Other	1, 2, 4, 5, 6, 7, 10	14, 17	22, 24,

P2 Numeracy (questionnaire 9)

Learning Domain	**Question no.**	
	Grade level	
	P1	***P2***
Number concepts	2, 3, 4	11, 12, 13
Addition and subtraction	5, 6, 7,	14, 16
Other	8, 9, 10	15, 17, 18, 19, 20

P4 Numeracy (questionnaire 10)

Learning Domain	**Question no.**		
	Grade level		
	P1 & P2	***P3***	***P4***
Number concepts	2, 3, 4	13, 14	23, 24
Multiplication and division	7	16, 17, 18	26, 27, 28, 29
Addition and subtraction	5, 6, 8	15	25
Other	9, 10, 11, 12	19, 20, 21, 22	30, 31, 32, 33

Annex D: Data

Table D.1: ESSPIN Logframe output indicators: estimates for phase 1 schools and 'all' schools, 2012

Logframe indicators		Enugu		Jigawa		Kaduna		Kano		Kwara	Lagos		6 states
Output-level	Stat	Phase 1[1]	All	Phase 1[1]	All	Phase 1[1]	All	Phase 1[1]	All	All	Phase 1[1]	All	All
		v2	v3	v5	v7	v9	v11	v13	v15	v17	v19	v21	v26
Proportion of competent teachers (%)	**mean**	**76.3***	**51.1**	**87.1***	**62.2**	**81.5**	**76.2**	**76.7**	**66.6**	**84.7**	**79.2***	**71.1**	**69.4**
	SE	3.0	2.4	2.1	1.9	2.6	1.6	2.6	1.8	1.3	2.7	1.7	0.7
	N	201	437	260	686	232	676	260	726	727	231	687	3939
Proportion of schools with competent teachers (%)	**mean**	**75.8***	**42.9**	**87.5***	**25.0**	**91.8***	**65.1**	**50.1**	**30.1**	**92.9**	**87.1***	**64.7**	**47.9**
	SE	7.3	6.0	5.7	4.3	4.7	4.7	8.6	4.5	2.6	5.8	4.8	2.1
	N	35	70	35	103	35	103	35	104	101	34	102	583
Proportion of schools where headteacher is effective (%)	**mean**	**30.0**	**6.9**	**39.5**	**15.6**	**9.4**	**9.4**	**16.5**	**15.6**	**14.9**	**37.3**	**11.0**	**12.7**
	SE	9.2	3.4	9.4	3.9	5.3	3.3	7.1	4.0	3.9	9.3	3.4	1.5
	N	26	58	28	87	31	81	28	82	85	28	85	478
Proportion of schools where school development planning is effective (%)	**mean**	**87.5***	**8.5**	**21.9***	**1.6**	**8.2**	**1.4**	**7.9**	**0.5**	**11.1**	**21.5***	**13.0**	**3.4**
	SE	5.8	3.4	7.6	1.3	4.9	1.2	4.8	0.7	3.2	7.5	3.4	0.8
	N	33	68	31	96	32	96	32	98	97	31	98	553
Proportion of schools which meet the needs of all pupils (%)	**mean**	**22.9**	**7.8**	**24.2**	**24.1**	**15.9**	**25.0**	**7.1**	**10.5**	**51.6**	**45.0***	**22.9**	**20.4**
	SE	7.9	3.4	7.3	4.2	6.5	4.4	4.5	3.1	5.0	8.8	4.3	1.7
	N	29	63	35	103	33	98	34	101	101	33	98	564
Proportion of schools with functioning SBMC (%)	**mean**	**49.0***	**5.3**	**52.5***	**20.4**	**55.4***	**20.0**	**50.1**	**26.3**	**23.4**	**17.3**	**13.4**	**20.7**
	SE	8.6	2.7	8.6	4.0	8.8	4.1	8.6	4.4	4.3	6.6	3.4	1.7
	N	35	69	35	101	33	98	35	102	99	34	99	568
Proportion of schools where SBMC reflects women's concerns (%)	**mean**	**57.4***	**6.1**	**46.4***	**9.9**	**32.8***	**2.6**	**35.2**	**14.3**	**21.2**	**27.8**	**19.4**	**10.7**
	SE	8.5	2.9	8.8	3.0	8.1	1.6	8.2	3.5	4.1	7.8	4.0	1.3
	N	35	69	33	98	35	98	35	103	101	34	99	568
Proportion of schools where SBMC reflects children's concerns (%)	**mean**	**32.0***	**3.5**	**41.1***	**3.8**	**9.6**	**0.6**	**18.9**	**8.8**	**10.7**	**17.4**	**11.0**	**5.7**
	SE	8.0	2.2	8.4	1.9	5.1	0.8	6.7	2.8	3.1	6.6	3.2	1.0
	N	35	69	35	102	35	101	35	103	101	34	99	575

Source: Composite Survey 2012. Note: (1) The asterisk * signifies that the mean estimate is significantly different from the mean estimate for the **control school group** (estimate not shown in the table, see next table) at the 0.05 level of statistical significance.

Table D.2: ESSPIN Logframe output indicators: estimates for control schools and phase 2 schools, 2012

Logframe indicators		**Enugu**		**Jigawa**		**Kaduna**		**Kano**		**Lagos**
Output-level	Stat	Control	Control	Phase 2[1]	Control	Phase 2[1]	Control	Phase 2[1]	Control	Phase 2[1]
Proportion of competent teachers (%)	**mean**	**47.4**	**53.7**	**64.5**	**73.4**	**84.4***	**65.3**	**65.5**	**57.1**	**75.9***
	SE	3.3	3.9	3.0	3.3	2.3	3.4	2.9	3.6	2.6
	N	236	164	262	185	259	196	270	190	266
Proportion of schools with competent teachers (%)	**mean**	**39.1**	**16.5**	**35.6**	**62.1**	**86.0**	**27.7**	**37.4**	**40.9**	**80.4***
	SE	8.4	6.6	8.2	8.3	6.1	7.7	8.4	8.8	6.7
	N	35	33	35	35	33	35	34	32	36
Proportion of schools where headteacher is effective (%)	**mean**	**4.9**	**11.6**	**25.5**	**9.3**	**10.6**	**15.7**	**15.1**	**8.3**	**7.6**
	SE	3.9	6.3	7.8	5.8	6.4	7.8	6.5	5.4	4.9
	N	32	27	32	26	24	23	31	27	30
Proportion of schools where school development planning is effective (%)	**mean**	**0.2**	**0.0**	**0.0**	**0.0**	**13.2***	**0.1**	**0.0**	**0.0**	**23.4***
	SE	0.7	0.0	0.0	0.0	6.3	0.4	0.0	0.0	7.3
	N	35	32	33	34	30	33	33	32	35
Proportion of schools which meet the needs of all pupils (%)	**mean**	**6.3**	**24.5**	**22.1**	**26.6**	**11.9**	**11.8**	**2.8**	**8.4**	**30.7**
	SE	4.2	7.6	7.1	7.9	5.7	5.7	2.9	5.1	8.0
	N	34	33	35	32	33	33	34	31	34
Proportion of schools with functioning SBMC (%)	**mean**	**0.2**	**18.1**	**15.5**	**19.4**	**7.1**	**27.1**	**8.9**	**3.6**	**21.5**
	SE	0.8	6.9	6.3	6.9	4.7	7.7	5.0	3.4	7.1
	N	34	32	34	34	31	34	33	31	34
Proportion of schools where SBMC reflects women's concerns (%)	**mean**	**0.2**	**6.3**	**10.2**	**0.6**	**9.1**	**14.6**	**1.0**	**14.3**	**21.9**
	SE	0.8	4.4	5.3	1.4	5.3	6.1	1.8	6.4	7.2
	N	34	31	34	33	30	35	33	31	34
Proportion of schools where SBMC reflects children's concerns (%)	**mean**	**0.2**	**0.0**	**4.1**	**0.0**	**2.6**	**9.3**	**0.0**	**3.6**	**16.3**
	SE	0.8	0.0	3.4	0.0	2.8	5.0	0.0	3.4	6.4
	N	34	32	35	34	32	35	33	31	34

Source: Composite Survey 2012. Note: (1) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the **control school group** (estimate not shown in the table, see next table) at the 0.05 level of statistical significance.

Table D.3: ESSPIN Logframe outcome and impact indicators: estimates, 2012

Logframe indicators		Enugu	Jigawa	Kaduna	Kano	Kwara	Lagos	All
Outcome-level								
Proportion of schools meeting quality criteria (%)	**mean**	**3.3**	**2.8**	**1.3**	**0.7**	**17.8**	**7.1**	**3.9**
	SE	2.4	1.9	1.3	1.0	4.3	2.9	0.9
	N	56	81	71	78	81	80	447
Impact-level								
Proportion of p2 pupils with skills for reading comprehension (%)	**mean**	**8.3**	**0.5**	**3.1**	**10.9**	**3.5**	**20.7**	**8.5**
	SE	3.4	0.3	2.6	4.4	2.0	3.7	2.4
	N	269	388	408	407	388	400	2260
Proportion of p4 pupils able to read with comprehension (%)	**mean**	**10.8**	**1.6**	**1.4**	**4.1**	**1.0**	**7.9**	**4.0**
	SE	3.8	0.7	0.7	2.2	0.4	2.4	1.1
	N	257	380	405	402	385	396	2225
Proportion of p2 pupils able to perform basic arithmetic calculations at p2 level (%)	**mean**	**15.1**	**5.0**	**13.2**	**18.2**	**22.4**	**17.1**	**15.5**
	SE	3.6	2.4	5.5	10	5.2	2.9	5.4
	N	269	380	406	401	388	394	2238
Proportion of p4 pupils able to perform basic arithmetic calculations at p4 level (%)	**mean**	**2.1**	**1.1**	**12.1**	**11.8**	**6.1**	**10.1**	**9.4**
	SE	0.9	0.4	5.4	7.9	2.3	2.1	3.8
	N	257	370	401	393	384	395	2200

Source: Composite Survey 2012.

Table D.4: ESSPIN Logframe outcome indicator: estimates for control schools, phase 1 and phase 2, 2012

Logframe indicators		Enugu		Jigawa			Kaduna			Kano			Lagos		
Outcome-level	Stat	Control	Phase 1[1]	Control	Phase 1[1]	Phase 2[1]	Control	Phase 1[1]	Phase 2[1]	Control	Phase 1[1]	Phase 2[1]	Control	Phase 1[1]	Phase 2[1]
Proportion of schools meeting quality criteria (%)	**mean**	**0.2**	**40.3***	**0.0**	**20.8***	**8.6**	**0.0**	**8.7**	**12.1**	**0.0**	**5.9**	**1.8**	**0.0**	**9.7**	**12.6**
	SE	0.8	10.0	0.0	8.1	5.2	0.0	5.5	7.5	0.0	4.7	2.5	0.0	5.8	6.5
	N	31	25	25	26	30	24	27	20	23	26	29	26	27	27

Source: Composite Survey 2012.

Table D.5: ESSPIN Logframe output and outcome indicators: estimates for 5 states combined by phase, 2012

Logframe indicators		5 States combined : Enugu, Jigawa, Kaduna, Kano, Lagos			
		Control	Phase 1[1]	Phase 2[1]	All
Output-level					
Proportion of competent teachers (%)	**mean**	**62.8**	**80.1***	**72.4**	**67.0**
	SE	1.6	1.2	1.4	0.8
	N	971	1184	1057	3212
Proportion of schools with competent teachers (%)	**mean**	**39.3**	**73.9***	**58.2**	**43.8**
	SE	3.8	3.3	4.2	2.3
	N	169	174	138	481
Proportion of schools where headteacher is effective (%)	**mean**	**11.1**	**23.6**	**14.3**	**12.5**
	SE	2.7	3.6	3.3	1.7
	N	135	141	117	393
Proportion of schools where school development planning is effective (%)	**mean**	**0.0**	**24.0***	**8.9**	**2.7**
	SE	0.2	3.4	2.5	0.8
	N	165	159	131	455
Proportion of schools which meet the needs of all pupils (%)	**mean**	**17.6**	**19.2**	**15.6**	**17.4**
	SE	3.0	3.1	3.1	1.8
	N	163	164	136	463
Proportion of schools with functioning SBMC (%)	**mean**	**19.4**	**47.0***	**13.3**	**20.5**
	SE	3.1	3.8	3.0	1.9
	N	164	172	132	468
Proportion of schools where SBMC reflects women's concerns (%)	**mean**	**7.4**	**38.9***	**9.9**	**9.7**
	SE	2.1	3.7	2.6	1.4
	N	163	172	131	466
Proportion of schools where SBMC reflects children's concerns (%)	**mean**	**3.8**	**22.6***	**5.7**	**5.2**
	SE	1.5	3.2	2.0	1.0
	N	165	174	134	473
Outcome-level					
Proportion of schools meeting quality criteria (%)	**mean**	**0.0**	**14.6***	**7.4***	**2.1**
	SE	0.2	3.1	2.6	0.8
	N	129	131	106	366

Source: Composite Survey 2012. Note: (1) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the control school group at the 0.05 level of statistical significance.

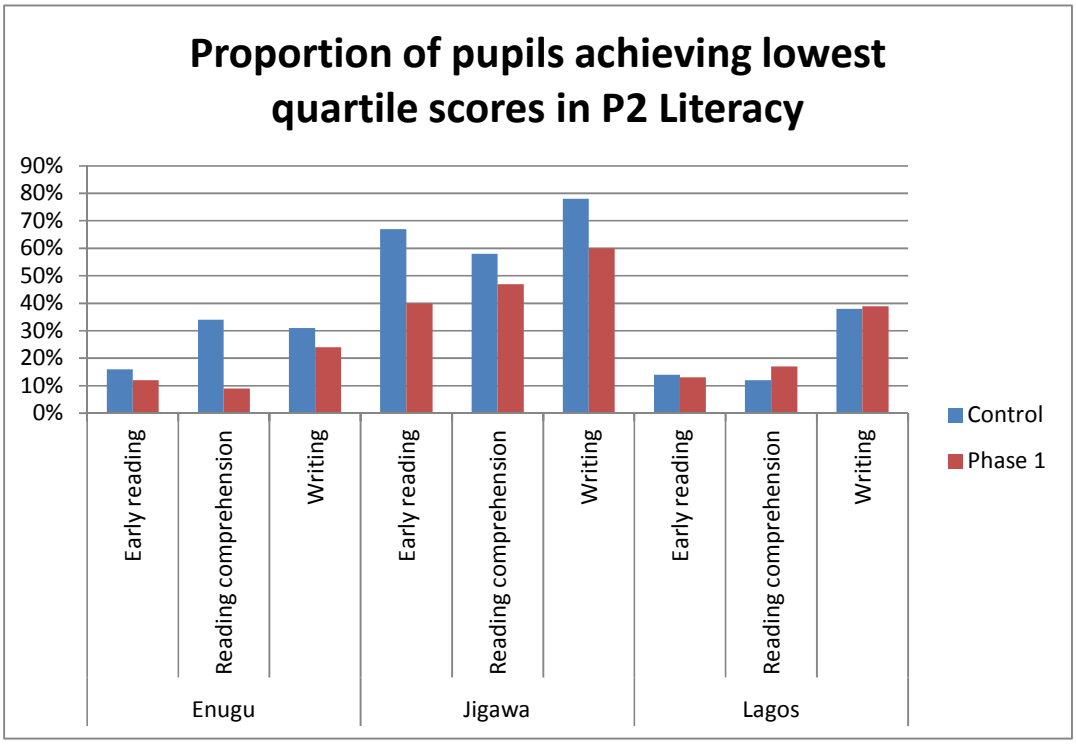
Table D.6: ESSPIN Logframe impact indicators: estimates for 4 states combined by phase, 2012

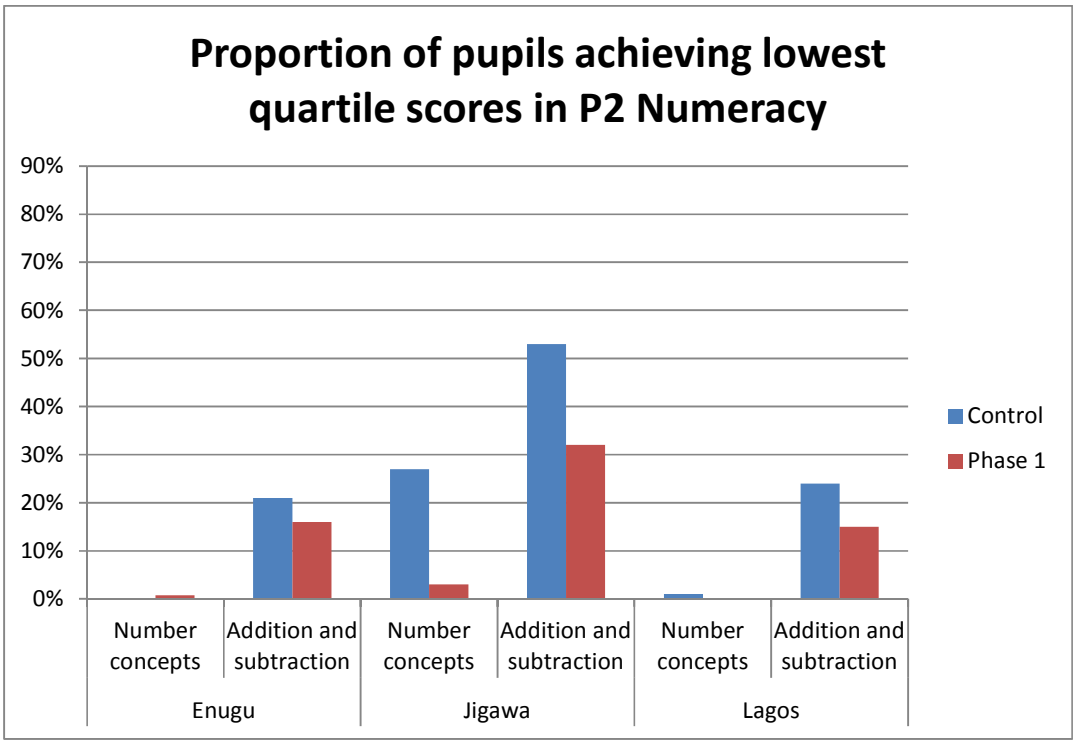
Logframe indicators		4 States combined : Enugu, Jigawa, Kaduna, Lagos			
		Control	Phase 1[1]	Phase 2[1]	All
Impact					
Proportion of p2 pupils with skills for reading comprehension (%)	**mean**	**5.1**	**8.2**	**8.6**	**6.1**
	SE	1.9	1.8	1.9	1.4
	N	514	541	410	1465
Proportion of p4 pupils able to read with comprehension (%)	**mean**	**2.1**	**7.7***	**8.6***	**4.1**
	SE	0.9	1.5	2.7	0.8
	N	495	535	408	1438
Proportion of p2 pupils able to perform basic arithmetic calculations at p2 level (%)	**mean**	**9.5**	**18.8***	**15.5**	**11.6**
	SE	3.4	2.7	2.9	2.4
	N	514	529	406	1449
Proportion of p4 pupils able to perform basic arithmetic calculations at p4 level (%)	**mean**	**7.6**	**7.5**	**6.8**	**7.4**
	SE	3.0	1.5	2.2	2.1
	N	505	521	397	1423

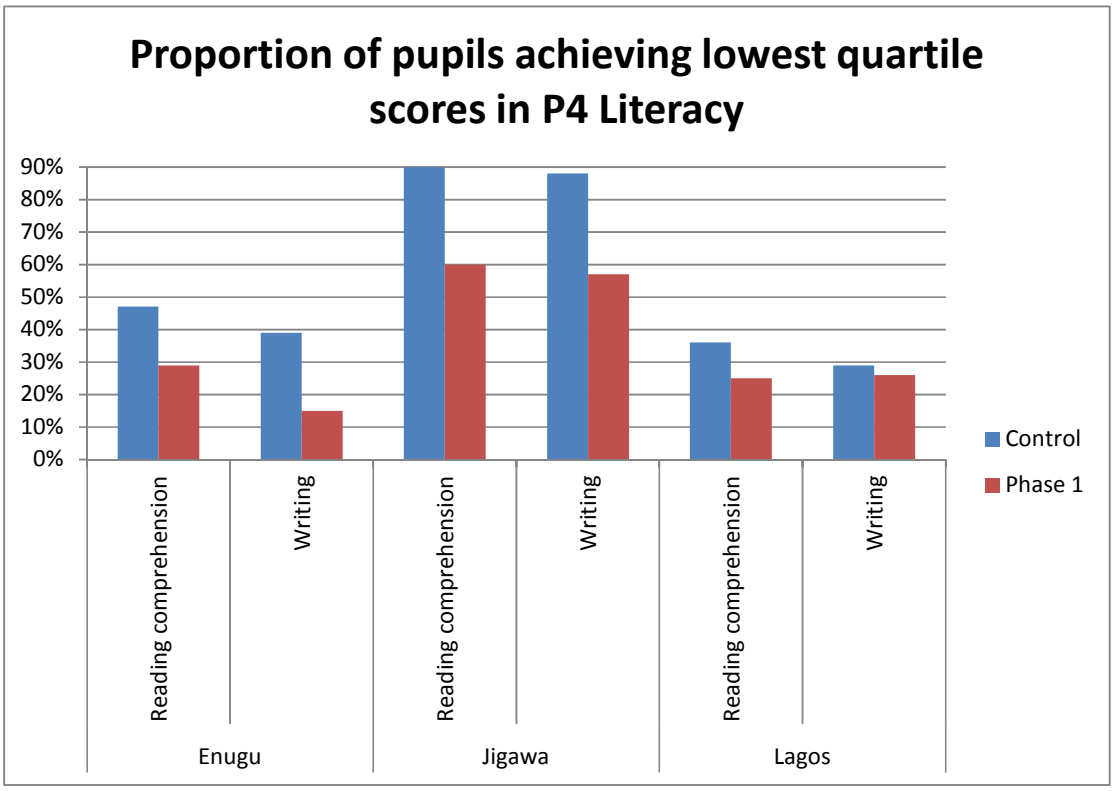
Source: Composite Survey 2012. Note: (1) The asterisk * signifies that the mean estimate is significantly different to the mean estimate for the control school group at the 0.05 level of statistical significance.

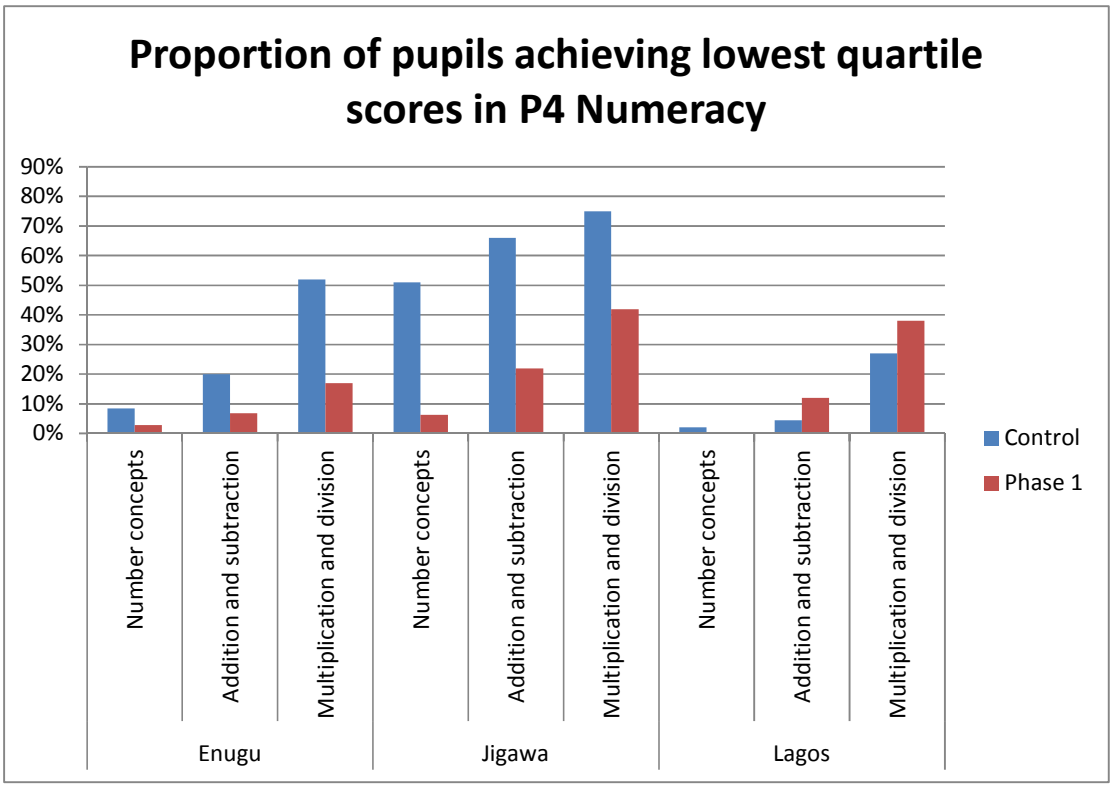
Annex E: Learning outcomes distributions by state, Phase, quartile and domain

Annex E.1: Lowest quartiles by Phase, State and Domain

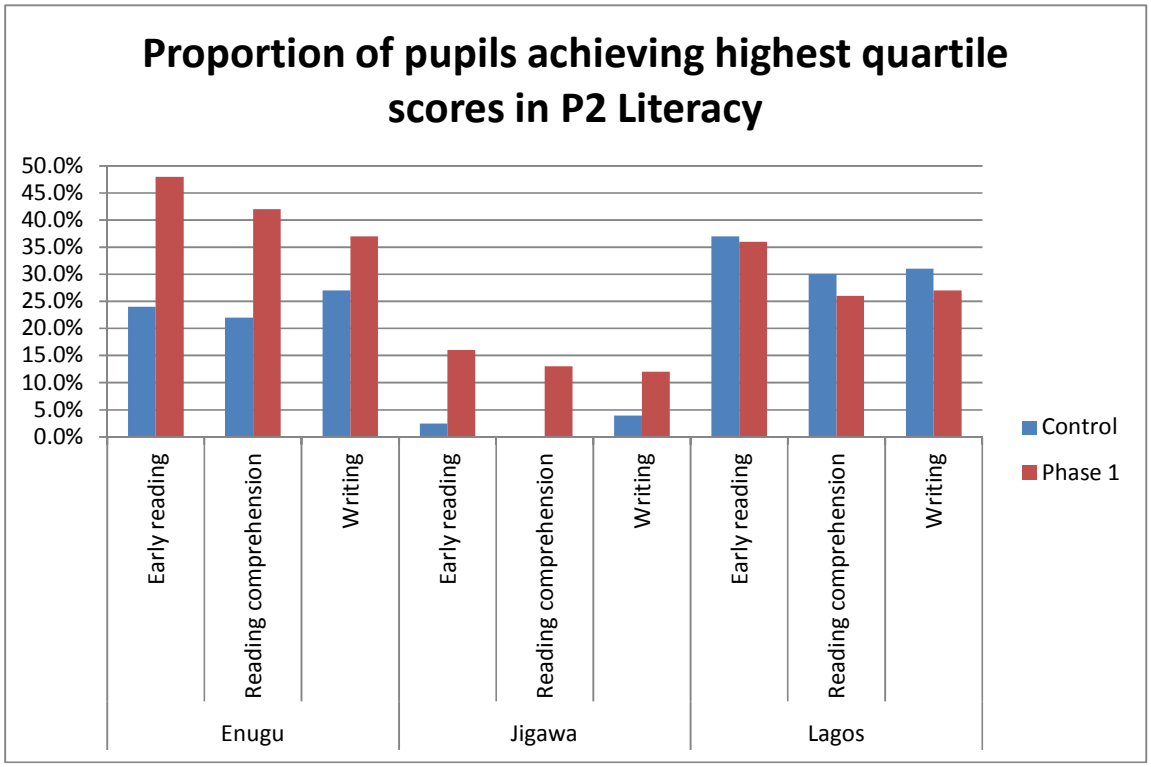


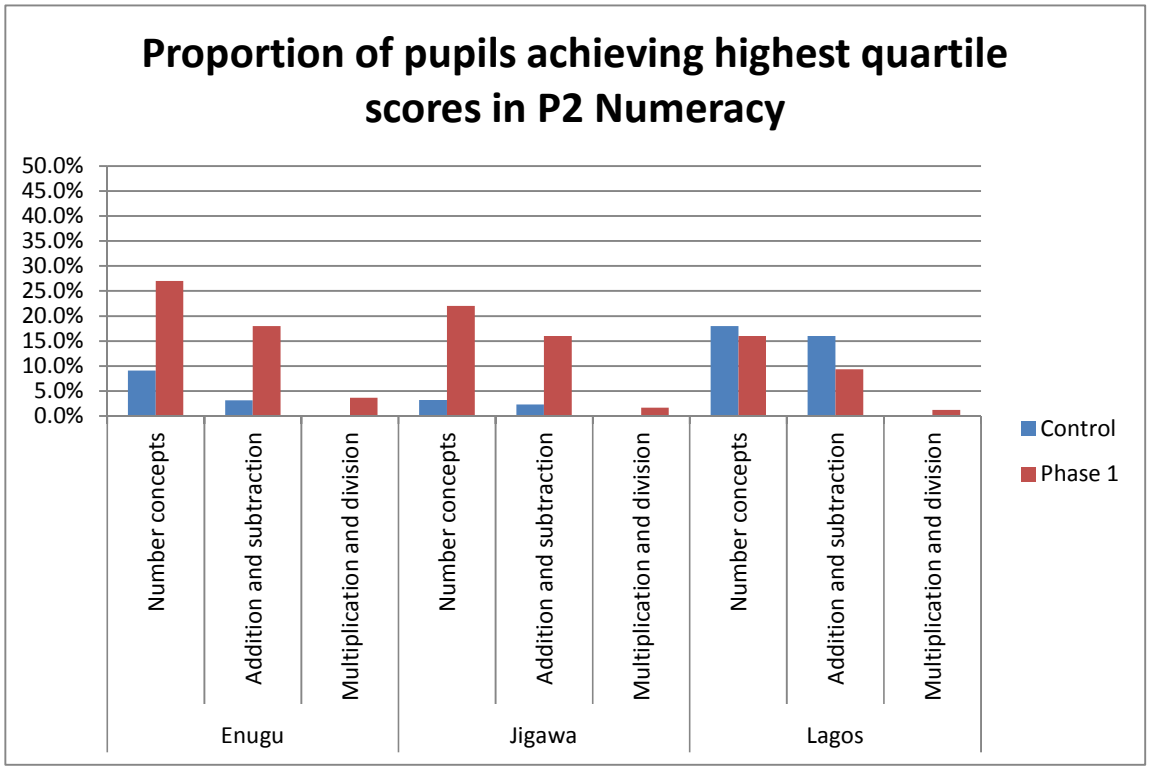


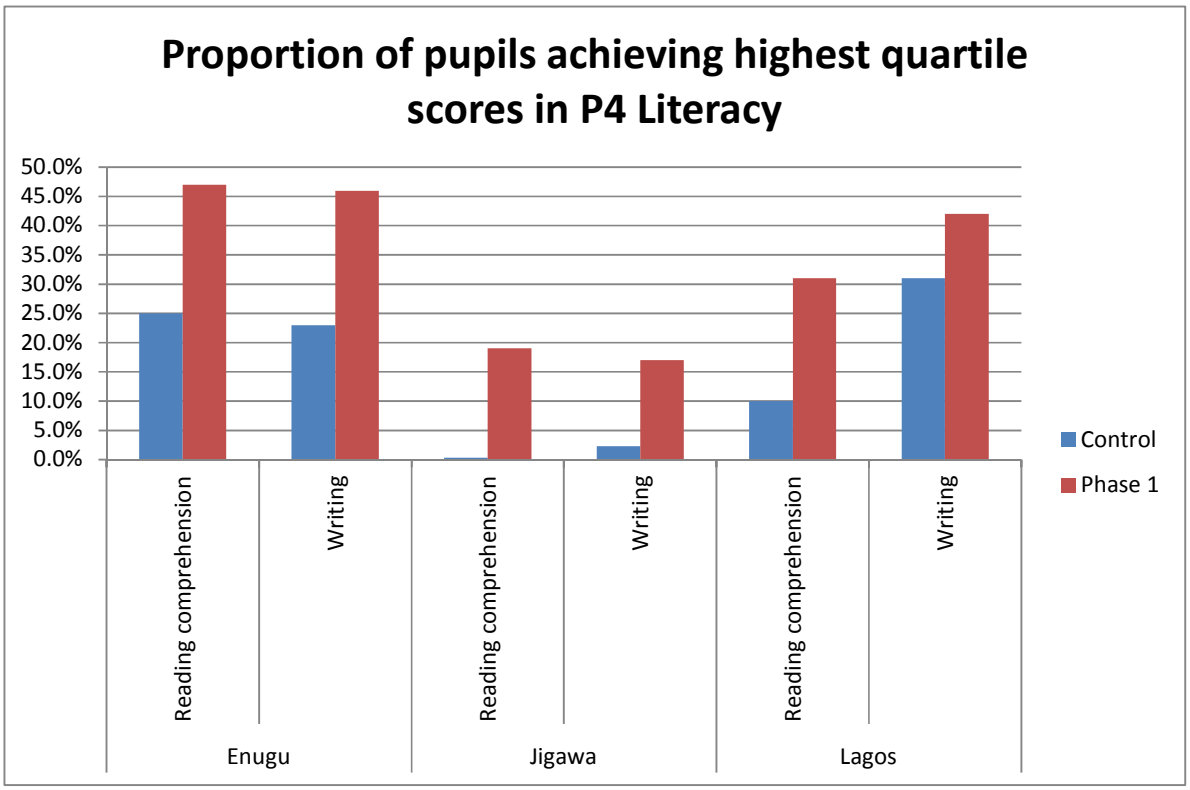

Proportion of pupils achieving lowest quartile scores in P4 Literacy
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Reading comprehension
Writing
Reading comprehension
Writing
Reading comprehension
Writing
Enugu
Jigawa
Lagos
Control
Phase 1

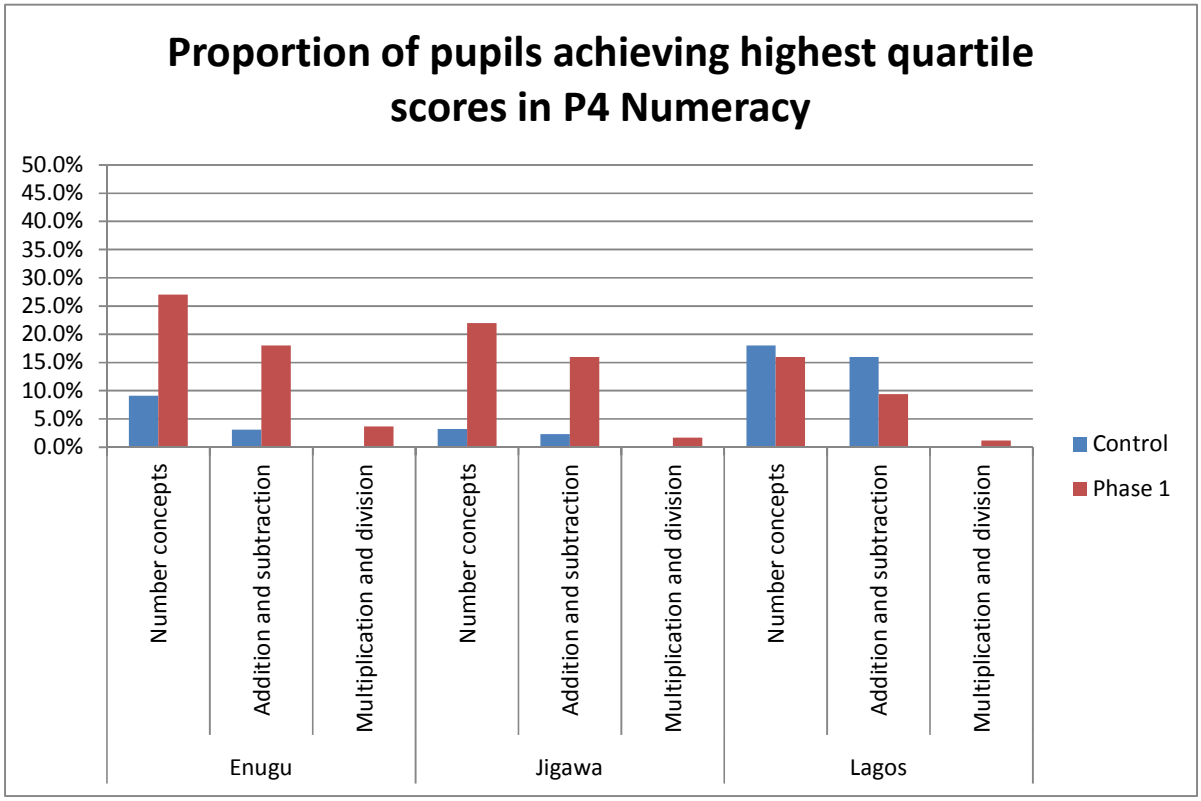

Proportion of pupils achieving lowest quartile scores in P4 Numeracy
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Number concepts
Addition and subtraction
Multiplication and division
Number concepts
Addition and subtraction
Multiplication and division
Number concepts
Addition and subtraction
Multiplication and division
Enugu
Jigawa
Lagos
Control
Phase 1

Annex E.2: Highest quartiles by Phase, State and Domain

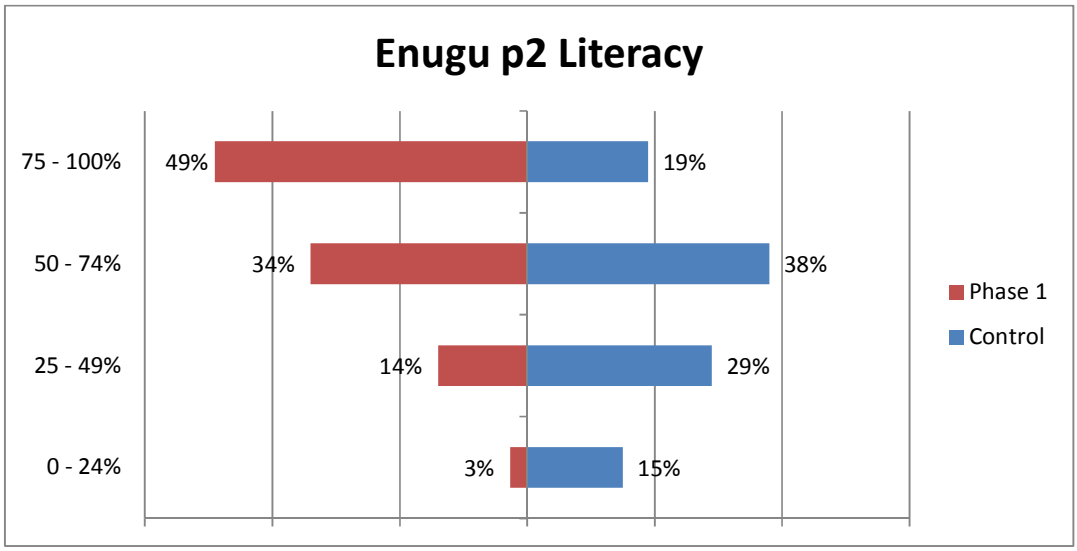


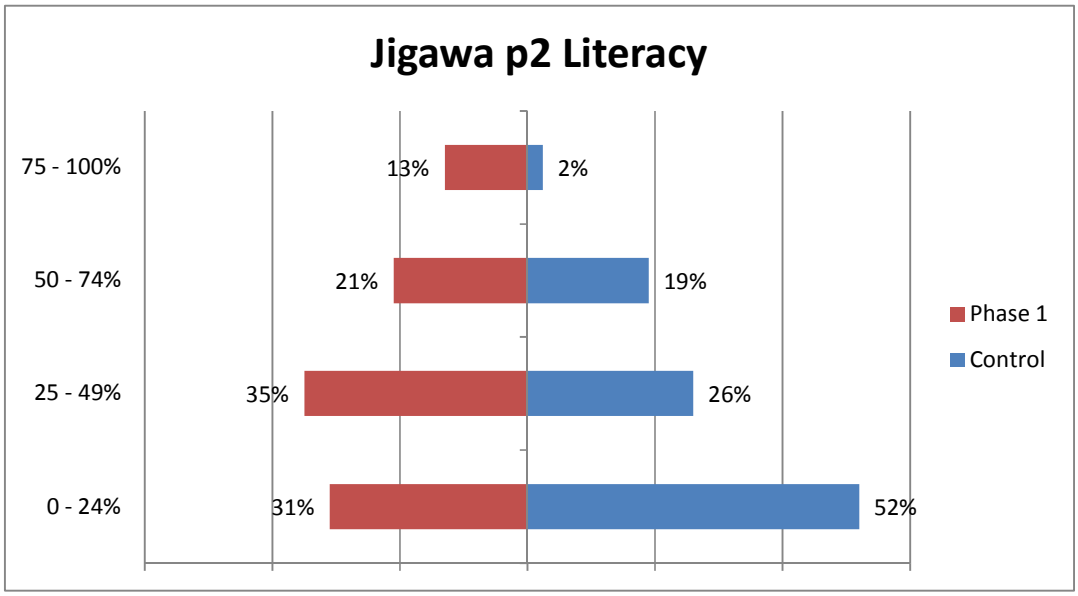


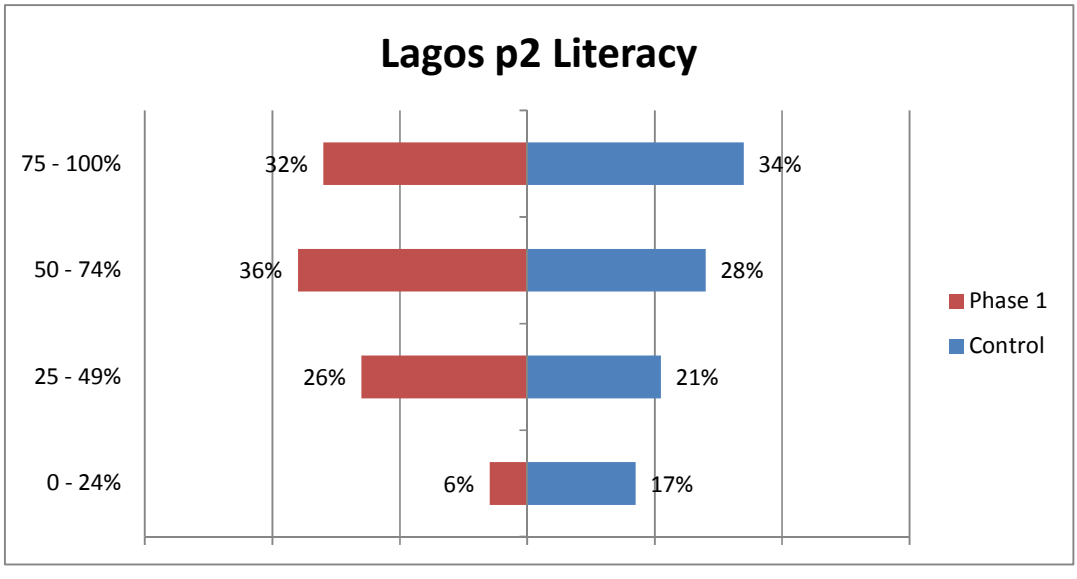

Proportion of pupils achieving highest quartile scores in P4 Literacy
50.0%
45.0%
40.0%
35.0%
30.0%
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
Reading comprehension
Writing
Reading comprehension
Writing
Reading comprehension
Writing
Enugu
Jigawa
Lagos
Control
Phase 1

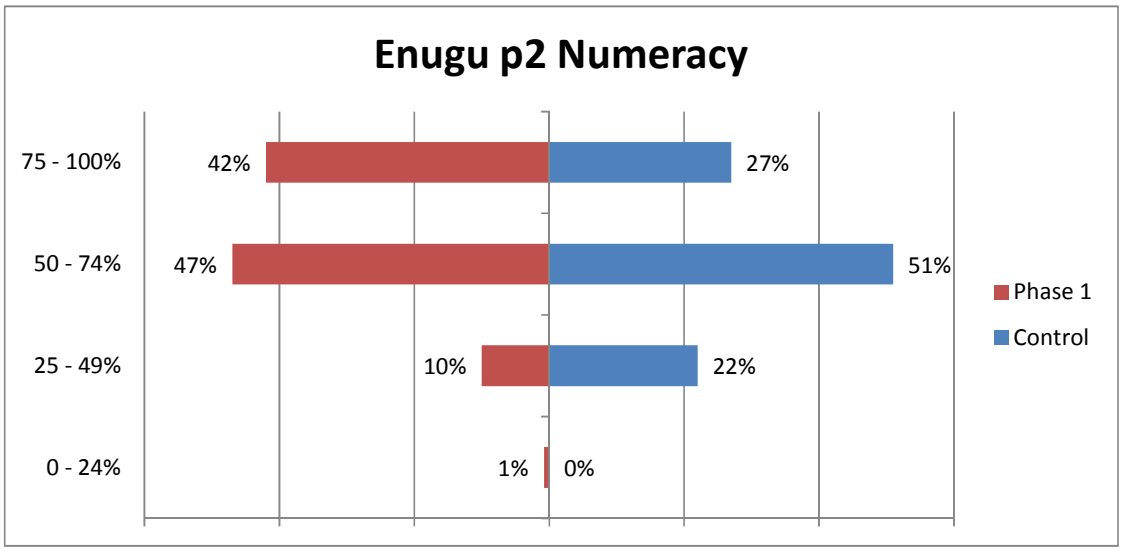

Proportion of pupils achieving highest quartile scores in P4 Numeracy
50.0%
45.0%
40.0%
35.0%
30.0%
25.0%
20.0%
15.0%
10.0%
5.0%
0.0%
Number concepts
Addition and subtraction
Multiplication and division
Number concepts
Addition and subtraction
Multiplication and division
Number concepts
Addition and subtraction
Multiplication and division
Enugu
Jigawa
Lagos
Control
Phase 1

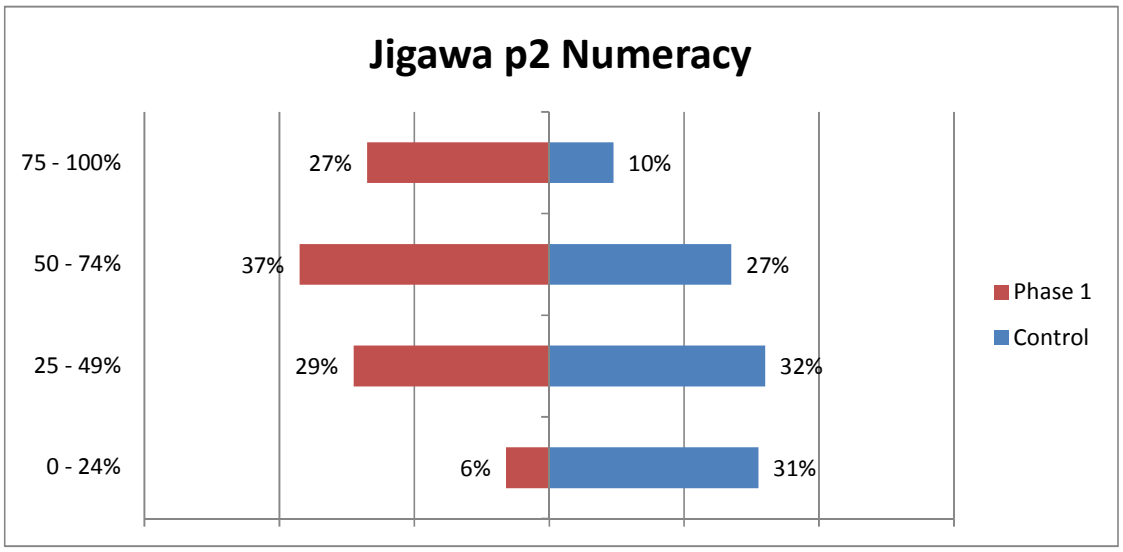
Annex E.3: Proportion of pupils in each quartile by Grade, Phase and State

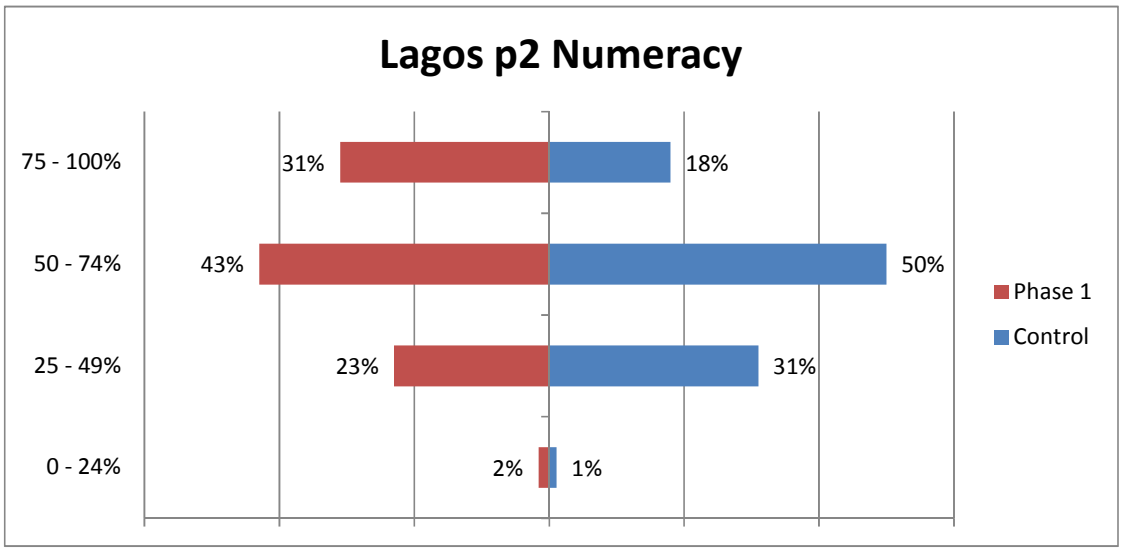


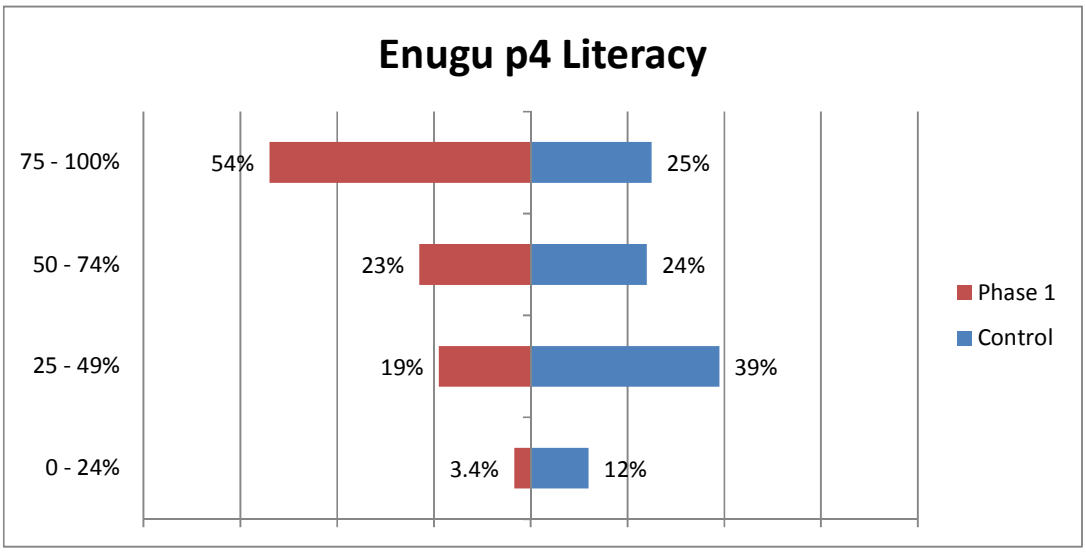


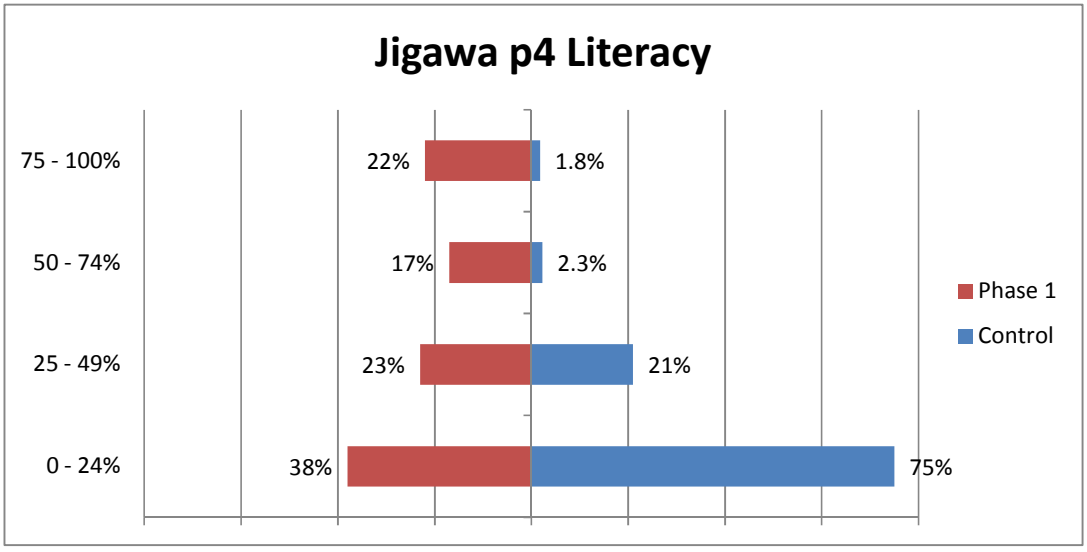


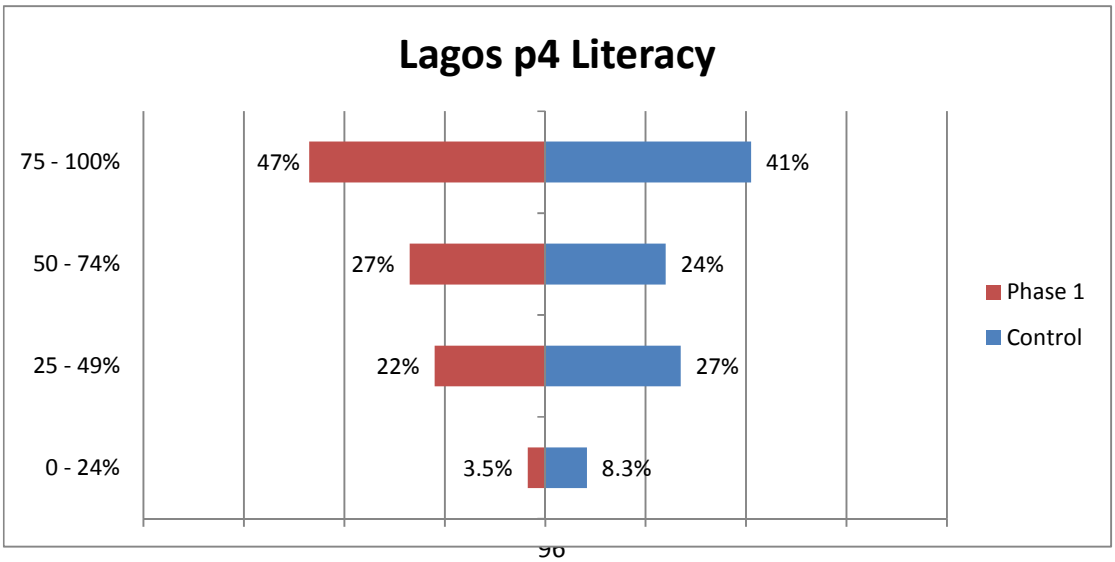

Enugu p2 Numeracy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
42%
27%
47%
51%
10%
22%
1%
0%
Phase 1
Control

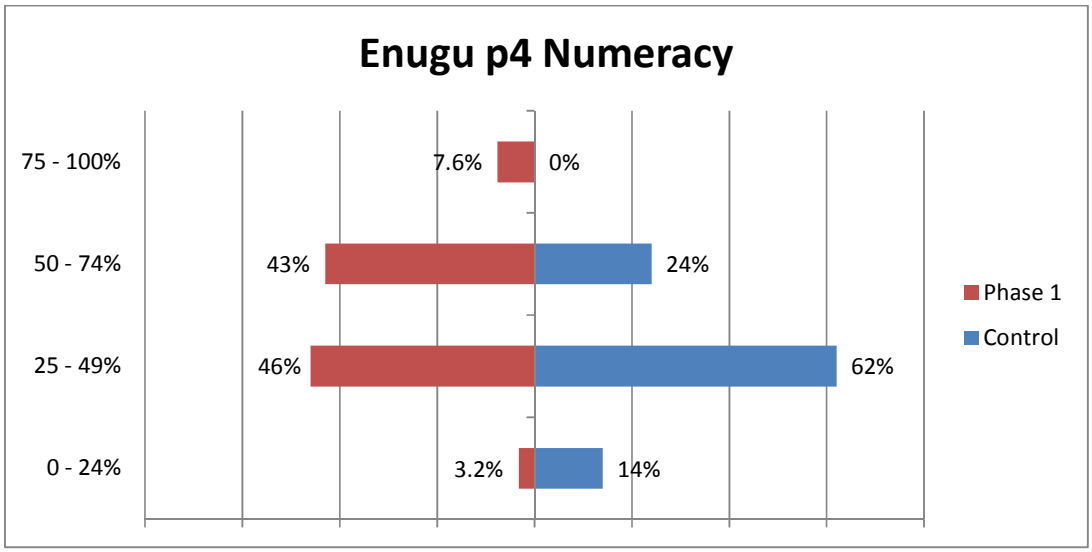

Jigawa p2 Numeracy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
27%
10%
37%
27%
29%
32%
6%
31%
Phase 1
Control

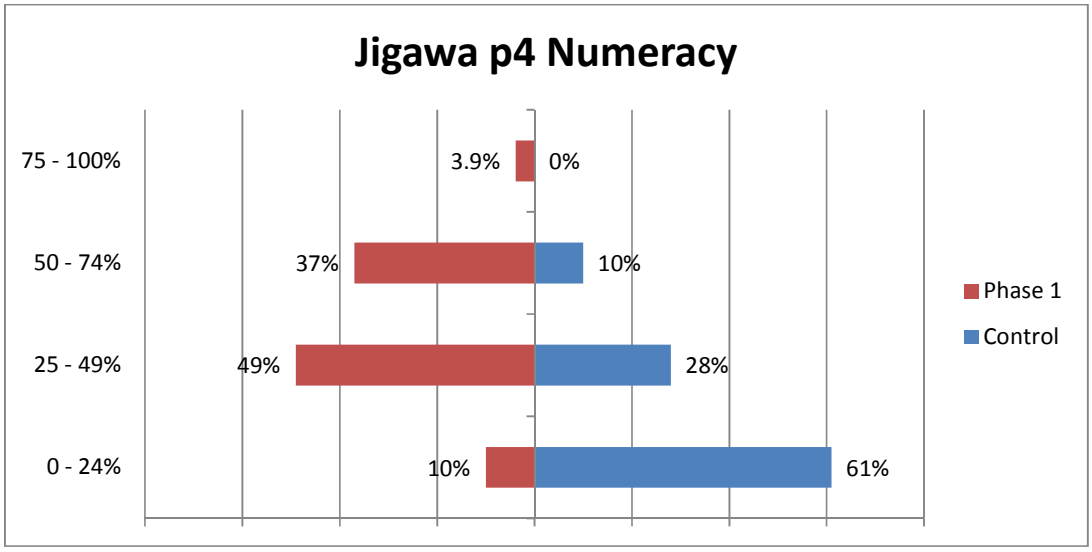

Lagos p2 Numeracy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
31%
18%
43%
50%
23%
31%
2%
1%
Phase 1
Control

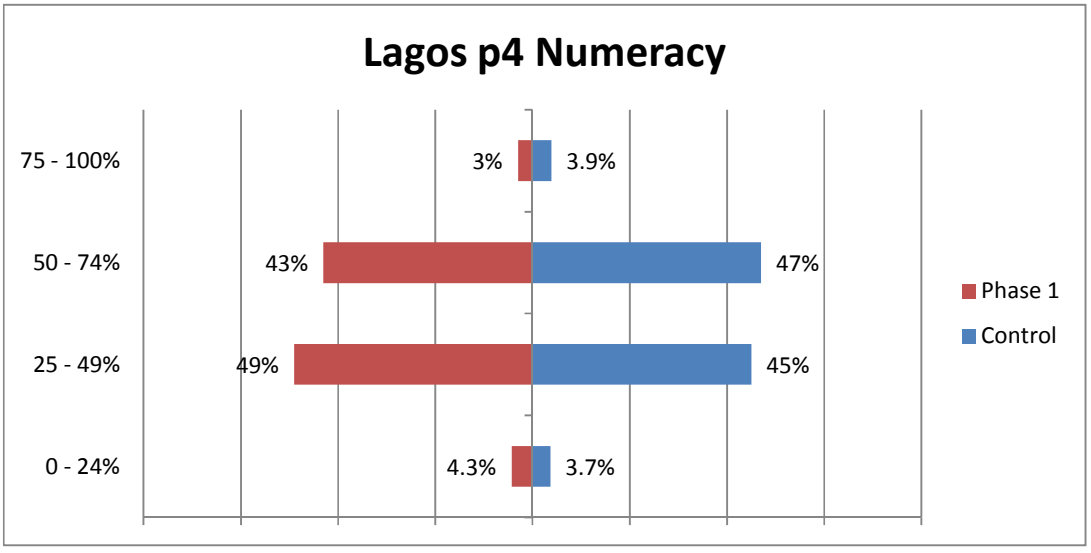

Enugu p4 Literacy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
54%
25%
23%
24%
19%
39%
3.4%
12%
Phase 1
Control


Jigawa p4 Literacy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
22%
1.8%
17%
2.3%
23%
21%
38%
75%
Phase 1
Control

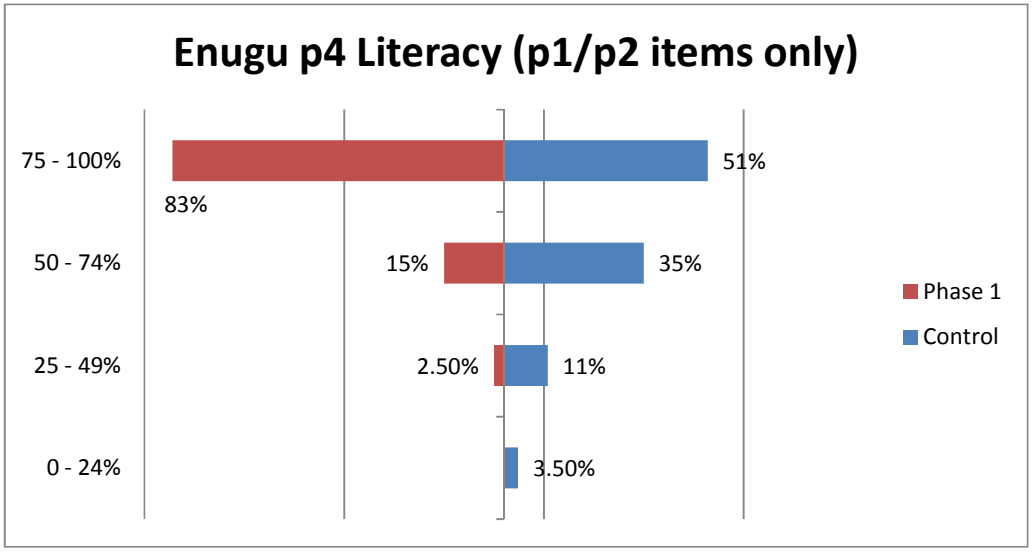

Lagos p4 Literacy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
47%
41%
27%
24%
22%
27%
3.5%
8.3%
Phase 1
Control

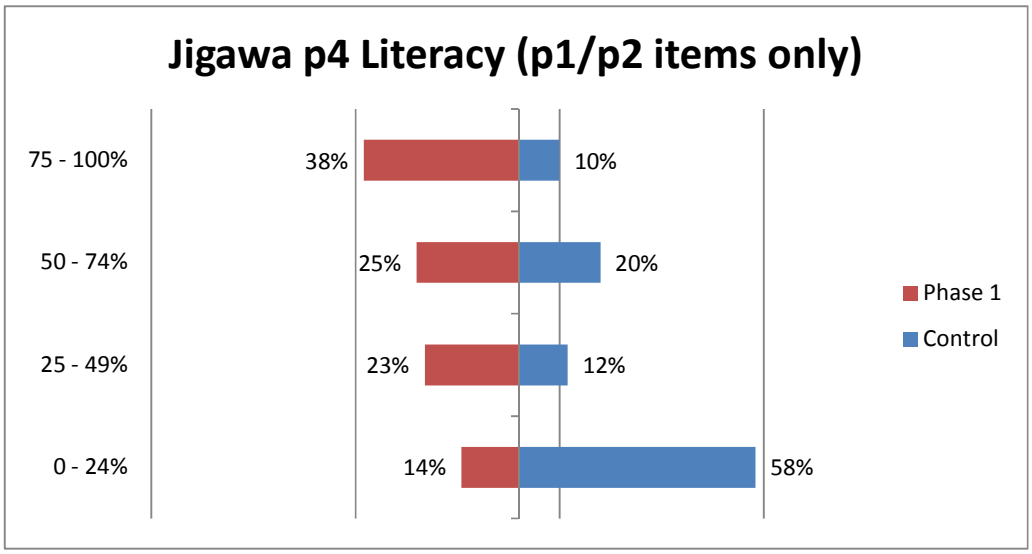

Enugu p4 Numeracy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
7.6%
0%
43%
24%
46%
62%
3.2%
14%
Phase 1
Control

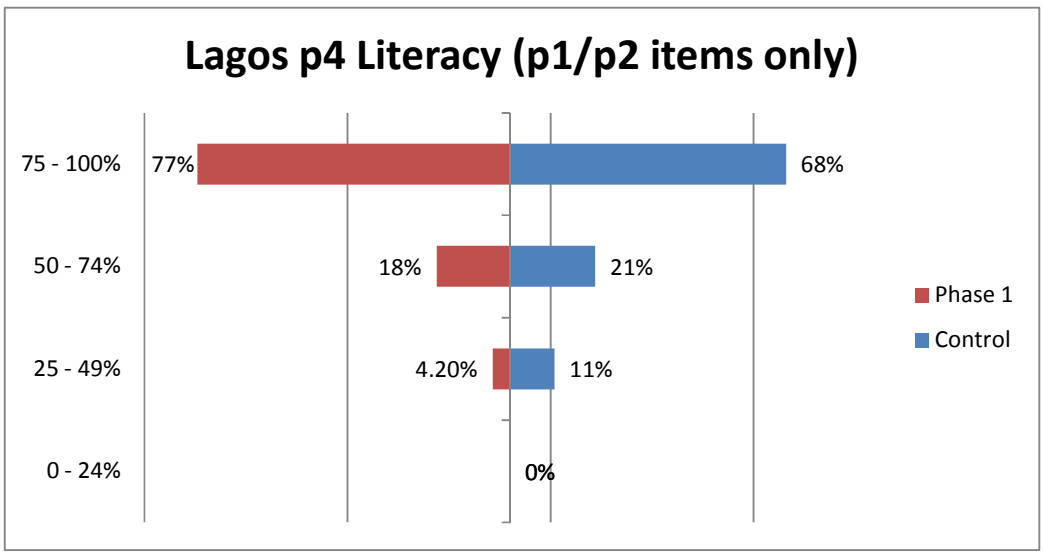

Jigawa p4 Numeracy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
3.9%
0%
37%
10%
49%
28%
10%
61%
Phase 1
Control

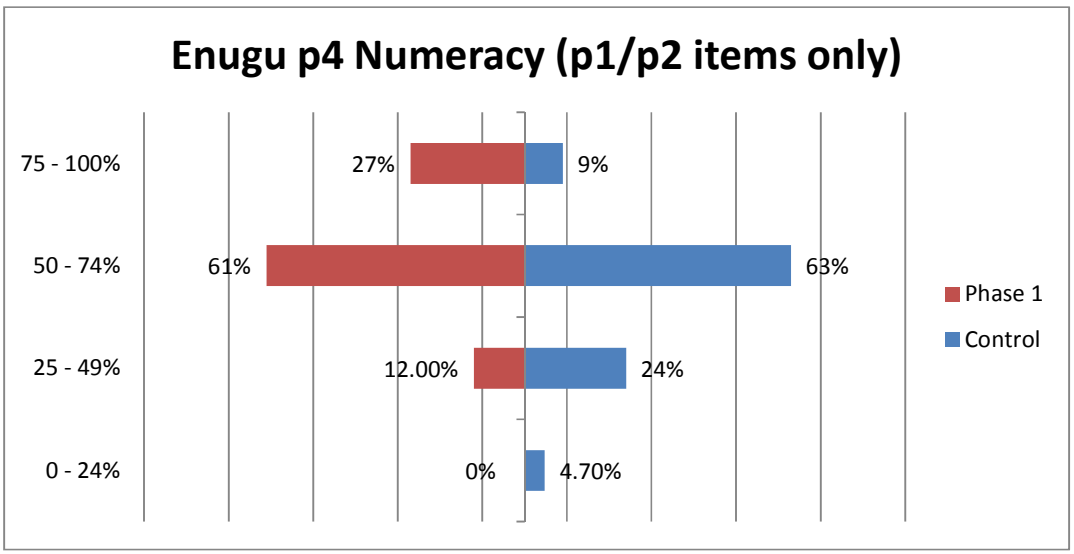

Lagos p4 Numeracy
75 - 100%
50 - 74%
25 - 49%
0 - 24%
3%
3.9%
43%
47%
49%
45%
4.3%
3.7%
Phase 1
Control

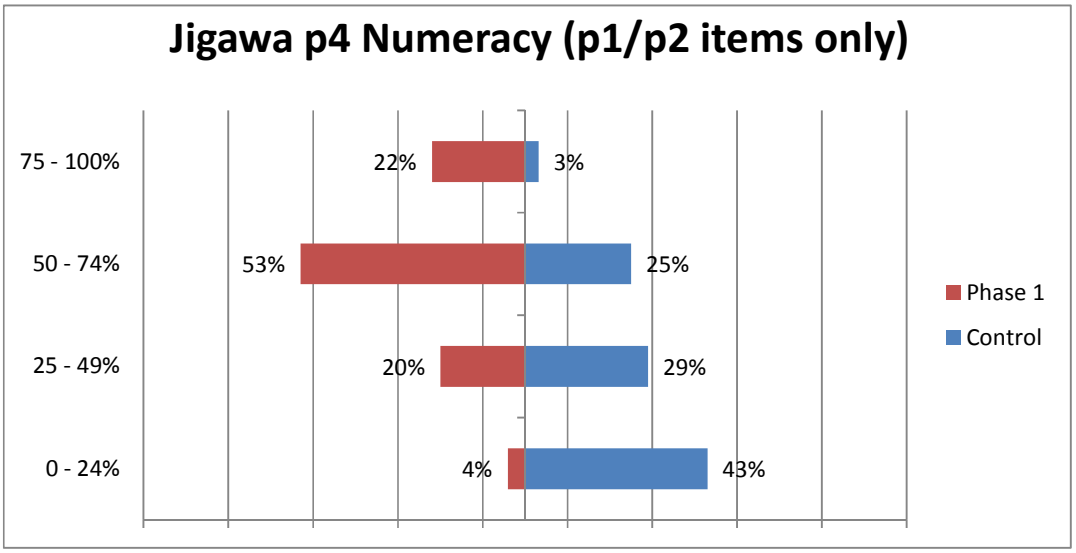
Annex E.4: Proportion of p4 pupils in each quartile by Phase and State on p1/p2 items only

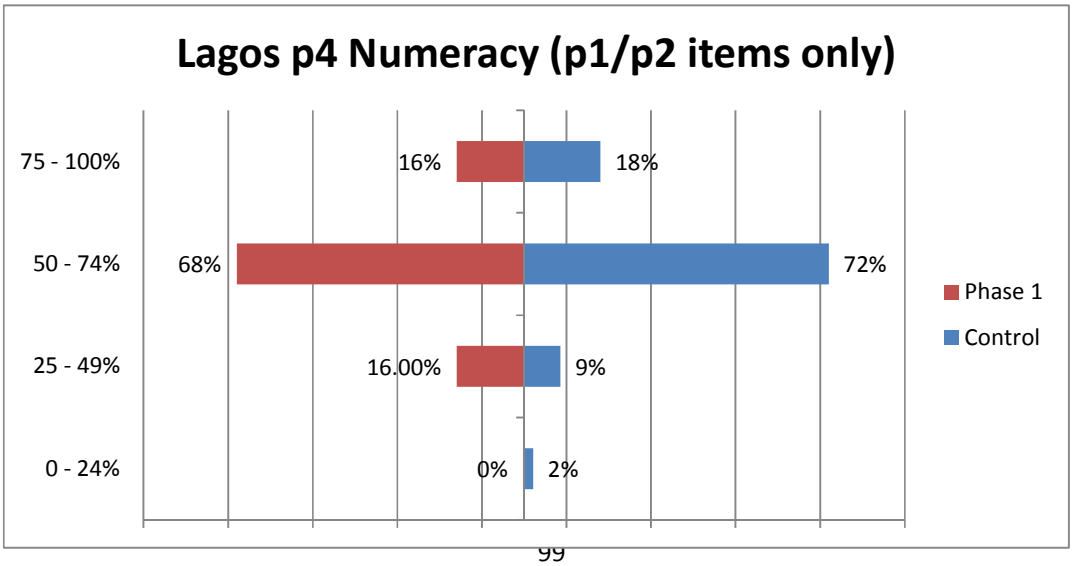







Enugu p4 Numeracy (p1/p2 items only)
75 - 100%
50 - 74%
25 - 49%
0 - 24%
27%
9%
61%
63%
12.00%
24%
0%
4.70%
Phase 1
Control


Jigawa p4 Numeracy (p1/p2 items only)
75 - 100%
50 - 74%
25 - 49%
0 - 24%
22%
3%
53%
25%
20%
29%
4%
43%
Phase 1
Control


Lagos p4 Numeracy (p1/p2 items only)
75 - 100%
50 - 74%
25 - 49%
0 - 24%
16%
18%
68%
72%
16.00%
9%
0%
2%
Phase 1
Control

Annex E.5: Reduction of proportion of p2 pupils in bottom score band, Phase 1 cf. Control