ESSPIN Composite Survey 3

Kaduna State Report

Jana Bischler

Acknowledgements

Oxford Policy Management Limited	Level 3, Clarendon House	Tel	+44 (0) 1865 207 300
	52 Cornmarket St	Fax	+44 (0) 1865 207 301
	Oxford OX1 3HJ	Email	admin@opml.co.uk
Registered in England: 3122495	United Kingdom	Website	www.opml.co.uk

Executive summary

The Education Sector Support Programme in Nigeria (ESSPIN) (2008–17) seeks to improve learning outcomes for children of basic education age in six Nigerian states: Enugu, Jigawa, Kaduna, Kano, Kwara, and Lagos. The aims of the ESSPIN Composite Surveys are to assess the effects of ESSPIN's integrated School Improvement Programme (SIP), and to report on the quality of education in the six ESSPIN-supported states. ESSPIN is funded by the UK Department for International Development (DFID) and managed by a consortium led by Cambridge Education. The Composite Surveys have been carried out for ESSPIN by Oxford Policy Management (OPM).

This report presents findings for Kaduna State from the first, second and third rounds of the Composite Survey (CS1, CS2 and CS3). These took place in 2012, 2014 and 2016, respectively. The surveys covered a wide range of indicators related to teachers, head teachers, School-Based Management Committees (SBMCs) and pupils. The aim is to understand how schools in ESSPIN states are changing over time, and whether schools which receive ESSPIN interventions are working better than those which do not. The main findings are as follows:

Head teacher effectiveness has not changed significantly over time in Kaduna. Only 3.4% of head teachers met ESSPIN's overall standard for head teacher effectiveness in 2016. This seems to be partly related to the difficult political and economic environment in which ESSPIN has been operating, particularly the recent delays in the payment of teachers' salaries. Head teachers in schools that have had more years of ESSPIN intervention are much more effective than those in schools which have received little ESSPIN intervention.

The effectiveness of **school development planning** has not changed significantly either. Less than 5% of Kaduna schools met the standard for effective school development planning in 2016. Schools which have had more years of ESSPIN intervention did not perform significantly better than those with less intervention.

Trends in **inclusiveness** – measured by indicators such as whether the head teacher has taken action on learners' attendance, and whether teachers engage boys and girls equally – depend on the exact measure used. Fewer schools in Kaduna met ESSPIN's overall inclusiveness standard in 2016 than in 2012, but a more nuanced continuous measure of behaviour related to inclusiveness has improved over this period. Schools with more years of ESSPIN intervention are not more likely to meet the overall standard than those with less intervention.

SBMCs in Kaduna are more functional than in 2012 or 2014, and they are slightly more inclusive of women and children. The SBMCs in schools which have received Output Stream 4 interventions are found to function better across some individual criteria than those in schools which received no interventions, but they are not more functional overall. They are also more inclusive of women, but there is no statistically significant effect of them being more inclusive of children.

Teachers in Kaduna have become more competent since 2014, although not compared with 2012. Teachers' test scores in English have not changed since 2014, while test scores in mathematics have improved significantly. Teachers trained through ESSPIN are not significantly more competent and do not have higher test scores than non-ESSPIN-trained teachers.

Only 5% of all schools in Kaduna reached the **school quality** standard, a composite measure based on head teacher effectiveness, school development planning, SBMC functionality and teacher competence. This is an increase from only 1% in 2012. This means that an estimated 175 more schools attained the quality standard in 2016 than in 2012 which translates into 58,000 more

pupils now learning in a better educational environment.[1] Furthermore, schools with more years of ESSPIN intervention are of a higher quality.

Children's **learning outcomes** in Kaduna are worse than in 2012, but are mostly better than in 2014. Learners from schools that received the most ESSPIN intervention had the highest scores in almost all tests (except for Grade 2 numeracy). However, the differences are not statistically significant. We also did not find any evidence that schools with more years of ESSPIN intervention improved faster over time. It must be noted that a large increase in enrolment numbers could have led to overcrowding of class rooms, which in turn would have put downward pressure on learning outcomes.

Table 1, Table 2, and Table 3 below summarise the key findings.

Table 1:   Kaduna: Change over time – Key indicators in 2012, 2014, 2016

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change 2012–16	Change 2014–16
Effective head teacher (%)	8.5	10	3.4	-5.1	-6.5
School development planning (%)	1.4	2.1	4.6	+3.3	+2.5
Inclusive (%)	23.1	8.9	3.6	**-19.6***	-5.3
Functioning SBMC (%)	28.0	25.8	46	+18.0	**+20.2***
Competent teachers (%)	76.2	49.6	62.9	-13.3	**+13.3***
Competent teachers (new measure, %)		18.6	21.2	n/a	+2.6
Good quality school (%)	1	2.2	5.1	+4.1	+3.0
Good quality school (new measure, %)		0.1	0.4	n/a	+0.3
Grade 2 literacy score	474.2	455.7	452.2	-22.0	-3.5
Grade 4 literacy score	460.5	419.5	447.1	-13.4	**+27.6***
Grade 2 numeracy score	524	442.2	454.9	**-69.1***	+12.6
Grade 4 numeracy score	493.9	431.3	459	**-34.9***	**+27.7***
Note. * indicates statistical significance (p < .05)					

---

[1] Calculations based on data from the Annual School Census.

Table 2:   Kaduna: Key indicators in 2016, by years of ESSPIN intervention

	Min. (one year)	Med. (two years)	Max. (four years)	Estimated effect of one year of full intervention
Effective head teacher (%)	0.3	13.1	31.5	**+3.7***
School development planning (%)	2.8	13.3	9.8	+2.6
Inclusive (%)	3.7	0.9	10.6	+0.8
Good quality school (%)	2.8	14.4	18.5	**+3.6***
Good quality school (new measure, %)	0	1	6.9	+0.6
Grade 2 literacy score	450.4	454	467.7	+5.2
Grade 4 literacy score	447.1	443.5	463.6	+3.0
Grade 2 numeracy score	458	445.8	456.1	-3.8
Grade 4 numeracy score	455.5	464.7	479.1	+8.2
Note. * indicates statistical significance (p < .05)				

Table 3:   Kaduna: Teacher competence, non-ESSPIN-trained versus ESSPIN-trained

	Non-ESSPIN-trained	ESSPIN-trained	Difference
Competent teachers (%)	60.7	65.2	+4.5
Competent teachers (new measure, %)	19.0	23.4	+4.4
Teachers' English scale	492	488	-3.3
Teachers' mathematics scale	488	501	+13.4
Note. * indicates statistical significance (p < .05)			

Table of contents

List of tables and figures

List of abbreviations

ACLED	Armed Conflict Location & Event Data Project
CAPI	Computer-assisted personal interviewing
CBO	Community-based organisation
CS1	Composite Survey 1
CS2	Composite Survey 2
CS3	Composite Survey 3
ESSPIN	Education Sector Support Programme in Nigeria
IRT	Item response theory
LGEA	Local Government Education Authority
LGSITs	Local Government School Improvement Teams
L2	Grade 2 literacy test
L4	Grade 4 literacy test
N2	Grade 2 numeracy test
N4	Grade 4 numeracy test
OPM	Oxford Policy Management
SBMC	School-Based Management Committee
SDP	School Development Plan
SIP	School Improvement Programme
SMO	Social Mobilisation Officer
SSIT	State School Improvement Team
SSO	School Support Officer
SUBEB	State Universal Basic Education Board

1 Introduction

ESSPIN (2008–17) seeks to improve learning outcomes for children of basic education age in six Nigerian states: Enugu, Jigawa, Kaduna, Kano, Kwara, and Lagos. The ESSPIN Composite Surveys seek to assess the effects of ESSPIN's integrated SIP, and to report on the quality of education in the six ESSPIN-supported states. ESSPIN is funded by DFID, and is managed by a consortium led by Cambridge Education. The Composite Surveys have been carried out for ESSPIN by OPM.

The first two rounds of the Composite Survey were carried out in 2012 and 2014. The surveys address five output indicators: teacher competence, head teacher effectiveness, school development planning, SBMC functionality and inclusive practices in schools. They also address one outcome indicator, school quality, and one impact indicator, pupil learning achievement. The third round of the Composite Survey (CS3) collects comparable data on these indicators in order to provide information on the extent to which key school-level indicators in the six states have improved during the course of the programme.

This report focuses on the Composite Surveys findings in Kaduna State. It presents the key findings from CS3, compares these to the findings of the previous rounds of the survey, and draws out the implications of these findings for ESSPIN's contribution to school-level outputs and outcomes in the state.

1.1 ESSPIN's SIP

ESSPIN aims to bring about better learning outcomes for children of basic education school age in six states, with a range of activities at the state, national, local and school levels. It has four output streams that focus on

- strengthening federal government systems;
- increasing the capability of state and local governments as regards the governance and management of schools;
- strengthening the capability of primary schools to provide improved learning outcomes; and
- improving inclusion policies and practices in basic education (ESSPIN, 2013b).

Under the third of these outputs, ESSPIN's SIP aims to provide and support the use of structured materials that ensure teachers can deliver quality instruction, to strengthen teachers' own understanding of literacy and numeracy concepts, and to improve academic leadership and school improvement planning by head teachers (Sanni, 2015). The SIP typically works through a two-year modular programme of workshops and school visits conducted by local government School Support Officers (SSOs), after which schools continue to receive school visits from government officers to help maintain and continue improving quality gains. At the same time, many of the same schools have been receiving interventions under the fourth output stream, facilitating community involvement and inclusion through SBMCs.

1.2 ESSPIN in Kaduna State

The work of ESSPIN in Kaduna State started in 2009. Initially, the programme focused on training teachers and head teachers, and on creating SBMCs which involve members from the local communities. In subsequent phases, the intervention continued to work with head teachers and SBMCs but increased its focus on strengthening teachers' abilities in teaching numeracy and literacy skills (ESSPIN 2013b).

ESSPIN's pilot phase (Phase 1) was implemented in two districts in the southern and central Senatorial District and three districts in the north. These districts were selected by the Education Secretaries of the relevant Local Government Education Authorities (LGEAs), ESSPIN 2013b. All schools in these districts were part of the pilot phase of ESSPIN's activities. The pilot lasted for two years and involved 165 schools.

Since then, ESSPIN has undergone rapid scale-up in Kaduna. After the pilot phase, the programme was gradually rolled out across the state (Table 4). At the time of the last Composite Survey (CS2) in 2014, 26% of Kaduna's primary schools had been benefitting from the full package of ESSPIN Output Stream 3 activities for at least one year (ESSPIN 2015). Up to that point the provision of activities had been continuous, so that once a school was in the programme it received training and school visits every year.

The following year, the programme was expanded to include the remaining 74% of schools, so that by 2015 all 4,225 primary schools in Kaduna had received at least one year of the full package of ESSPIN Output Stream 3 activities. The schools that had started to receive the programme in previous years did not continue to receive the full intervention package in 2014/15. School visits continued but there was no more leadership or teacher training, due to a lack of government funding.

In 2015/16, no schools in Kaduna received any ESSPIN SIP intervention because political issues prevented a budget from being produced and approved by the State Universal Basic Education Board (SUBEB). An overview of the roll-out of ESSPIN activities under Output Stream 3 between 2009/10 and 2015/16 in Kaduna is given in Annex A.

Table 4: **Proportion of schools receiving full package of ESSPIN Output Stream 3 interventions**

%	2009/10	2010/11	2011/12	2012/13	2013/14	2014/15	2015/16
Full package of Output Stream 3 interventions, including leadership training	4	4	14	19	11	74	0
Any Output Stream 3 intervention	4	4	14	19	26	100	100
Full package of Output Stream 4 interventions	0	0	0	0	0	15	15

Source: Author's calculations based on 2012/13 Annual School Census and intervention information provided by ESSPIN.

Note: Proportions are calculated relative to the total number of schools in the 2012/13 Annual School Census, and so the figures are not perfectly accurate for other years because the total number of schools changes slightly from year to year. Where census numbers are lower than ESSPIN's intervention tables, the information from ESSPIN is used, on the assumption that there are some missing data in the Annual School Census.

The expansion of the programme led to changes in the mode of implementation. Initially, the training was delivered with the help of the 24 State School Improvement Team (SSIT) members, which were hired by ESSPIN. In order to make delivery more sustainable, the management of programme delivery was shifted to the state government in Phase 2. The SSITs now train Local Government School Improvement Teams (LGSITs), who in turn train the SSOs, who, unlike the SSITs, are permanent local government staff. In turn, the SSOs train the head teachers and teachers, and, in partnership with CSO personnel, Social Mobilisation Officers (SMOs) train SBMCs. This allows for training to be located closer to schools and for the state to take control, which is necessary for scale-up and sustainability.

However, the change in model is likely to have affected the quality of implementation to some extent. Programme staff argue that locating training closer to the schools has longer-term benefits, but that in the shorter term the quality standards of the pilot programme might not be fully upheld as the new, much larger numbers of trainers, who typically have lower qualifications than those in the first wave, develop competencies.

Due to the staggered roll-out, different schools in Kaduna have experienced different levels of exposure to ESSPIN. For the CS3 analysis, schools will be categorised into different intervention groups depending on the number of years for which they have received the full package of ESSPIN activities (discussed further in Section 2.1.1). The more years a school has been part of the programme, the higher its expected improvement in output, outcome and impact indicators.

A summary of the characteristics of Kaduna primary schools according to the level of ESSPIN intervention can be found in Annex A. This shows that the schools from the different intervention groups differed in a number of aspects even prior to the ESSPIN intervention. The two groups that received ESSPIN intervention for the longest period (medium and maximum intervention groups) generally had lower pupil to teacher ratios (PTRs) and a greater number of teachers and class rooms. On average they are also closer located to local government authority headquarters and more likely to be situated in urban areas. These aspects could have a positive effect on pupils' learning outcomes and should be kept in mind when interpreting the CS3 findings in regard to comparing the results of schools in different intervention groups.

In addition to the SIP activities under Output Stream 3, schools in Kaduna received support under ESSPIN's Output Stream 4: improving inclusion policies and practices in basic education. ESSPIN has trained civil society members and government officers from the Department of Social Mobilisation, SMOs, to enable them to train and mentor SBMCs. SBMC members, in turn, have been trained on the roles and responsibilities of SBMCs, school planning and management, communication and leadership, change and relationships management, the participation of women and children in school improvement and education decision-making, resource mobilisation and financial processes, and child protection and participation. This has been complemented by follow-up mentoring visits by SMOs.

The Output Stream 4 interventions began in Kaduna in 2010/2011. However, to date there has only been one year in which 15% of schools received the full package of Output Stream 4 interventions. The Output Stream 4 interventions have therefore not yet been scaled up in Kaduna, with the majority of schools (2,583) not having received any Output Stream 4 intervention by the time of CS3 (see Annex C and Table 4).

1.3   Contextual factors and their implications for the SIP in Kaduna

This section describes some of the key aspects of the backdrop against which ESSPIN's implementation in Kaduna has taken place over the last couple of years. This is helpful in terms of interpreting the changes in school-level outputs and outcomes between CS1, CS2 and CS3. Changes over time in outputs and outcomes could have resulted from ESSPIN support, but equally they could have also been driven by other changes in the state over that period. This section considers the main developments in Kaduna that may have positively or adversely affected school-level outcomes in the state, or that may have interfered with SIP implementation.

Policy changes introduced in the aftermath of the 2015 elections in Kaduna appear to have had an unfavourable impact on ESSPIN's implementation in the state initially. Firstly, there were worries that if the new government does not work well enough through the civil service, structures built by ESSPIN as part of the organisational reform could be undermined.

There were reports that after the elections a large number of government staff trained by ESSPIN were replaced by untrained staff. This replacement could have negative effects on ESSPIN's operational efficiency. By the time of writing this report these worries had subsided.

Secondly, the sharp drop in oil prices and the consequent economic downturn had major fiscal repercussions for almost all Nigerian states, including Kaduna. Funding from federal allocations have decreased in the last two years, which led to some SIP activities not being conducted. In addition, there are also issues relating to the payment of teachers. Some teachers and LGEA staff have not been paid for eight months, while others have only been paid in part. This could negatively affect teachers' motivation, and could negatively impact ESSPIN's effectiveness by undermining teacher attendance, head teachers' perceptions of their own ability to influence attendance, and the extent to which teachers apply the new skills gained through ESSPIN training.

Furthermore, the new government has introduced a restoration plan, free school uniforms and a free school feeding programme. As a result of the free feeding and school uniform programme, enrolment figures have increased sharply. Stakeholders are concerned that this has led to an overcrowding of class rooms and a decline in teachers' control and students' attention. There are also reports of students leaving right after having received the food in the morning. In addition, some SSOs were assigned to monitor school feeding in some LGEAs, instead of focusing on their roles of supporting head teachers and making school visits. Since the new government did not fund the SIP, some of the SIP activities had to stop. All of these factors could undermine the success of some ESSPIN activities, and the outcomes they are trying to achieve.

Looking at the data, we can see indeed that enrolment in public primary schools in Kaduna has risen by 17%, from about 980,00 pupils in 2009/10 to close to 1,180,000 pupils in 2014/15, the time of the last available census (Table 5). The increase in enrolment was accompanied by an increase in PTR across all schools (see Annex A), from 36 students per teacher in 2009/10 to almost 50 students per teacher in 2014/15. Over this period, this could have had adverse effects on learning outcomes in Kaduna. Table 6 presents the numbers of male and female learners enrolled in 2014/15. For every 100 enrolled boys, there are only 87 enrolled girls.

Table 5:   Number of schools and enrolment in the 2009, 2013 and 2014 Annual School Censuses

	Enrolment	Number of schools	Enrolment change (%)
2009/10	979,659	3,947	
2013/14	1,153,460	4,225	17.7
2014/15	1,180,039	4,225	2.3
Overall			17.0
Note: Enrolment is for Primary Grades 1–6.			
Source: Annual School Censuses.			

Table 6:   Male and female enrolment in 2014/15

	Male	Female	Girls enrolled for every 100 boys
Kaduna	631,287	548,752	87
Source: Annual School Censuses.			

In addition to the above factors, this report is also written in the context of growing insecurity in Nigeria, particularly in, but not limited to, the three states in the north-east, in which a state of emergency has been declared (Borno, Yobe and Adamawa). Across Nigeria the number of

recorded incidents of political violence and conflict have increased dramatically since 1997. In Kaduna, the incidents of violence have also increased during this time period, reaching their peak in 2012 and decreasing again slightly after that (Figure 1). In 2015, there were 38 recorded violent events in Kaduna, causing 432 fatalities. This is slightly less than the year before (Table 7).

There is reason to believe that in 2016 the downward trend will continue. Compared to the first half of 2015, the number of conflict incidents in the first half of 2016 was slightly smaller, both in Nigeria as a whole and in Kaduna in particular.

Figure 1:   Incidents of political violence in Nigeria and Kaduna
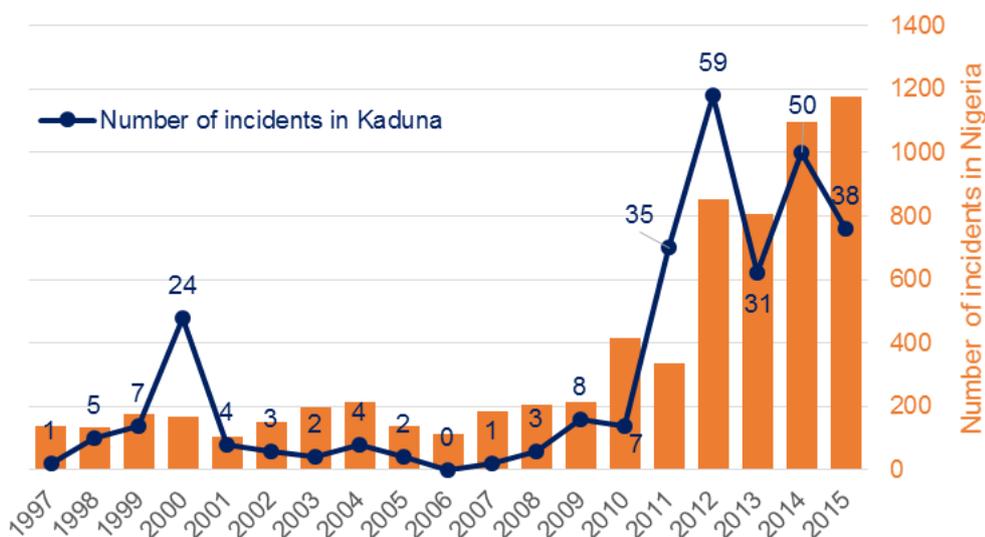
Source: Armed Conflict Location & Event Data Project (ACLED), Version 6 (1997–2015). Note, all events from ACLED are included except for those categorised as protests which did not involve a fatality.

Table 7:   Kaduna: Political violence – Incidents and fatalities, 2010–2015

Variable	2010	2011	2012	2013	2014	2015
Events	7	35	59	31	50	38
Fatalities	20	851	269	140	538	432

Despite a slight downward trend, violence and insecurity remain a problem in Kaduna and many stakeholders stated that this has adverse effects on the education sector. It was reported that armed robberies, attacks and kidnappings are common in certain areas in Kaduna. In some cases, whole communities are displaced as a result. Children are taken out of school by their parents as they move away, which disrupts their learning process. In addition, some schools have become more insecure and there have been reports of attacks not only on communities but also on schools directly. The security situation in Kaduna has led to the closure of a number of primary schools, some of which have remained closed during the whole year. In other instances it has made it impossible for head teachers and teachers to attend schools (e.g. Chikun, Giwa, Kaura, Sanga and Jema'a, to mention a few). Keeping in mind the security situation in other parts of Nigeria, it is also possible that a significant proportion of the 200,000 children that crowded into Kaduna's schools since 2009 (see Table 5) are internally displaced from conflict-affected communities in the north-east of Nigeria. This should be kept in mind when interpreting the findings of CS3, as it might have had some adverse effects on education outcomes in Kaduna.

2 Methodology and analysis

2.1 Evaluation strategy

2.1.1 ESSPIN intervention groups

ESSPIN was originally intended to be rolled out in a simple phased pattern across the six states, with schools falling into one of three groups: no intervention (control), Phase 1 (roll-out prior to the 2012/13 school year), and Phase 2 (roll-out in 2012/13 or 2013/14). In practice, Kaduna State decided to extend the programme over several phases, adding a group of schools each year after the pilot, based on the government's capacity and willingness to roll out the SIP. In 2014/15, the programme was expanded to include all remaining schools in Kaduna which had not previously received any ESSPIN intervention. In 2015/16, none of the primary schools in Kaduna received any ESSPIN school-level intervention.

For the purpose of this analysis, we categorise schools according to the number of years that they have received the full package of ESSPIN interventions under Output Stream 3. (See Annex A for details).[2] Schools that have received one year of Output Stream 3 intervention will be in the 'minimum' category; schools with two to three years will be in the 'medium' category; and schools with four to five years are in the 'maximum' category. This gives us three different intervention groups, which is necessary because there is no longer a 'pure' control group that has not received any ESSPIN intervention. Therefore, the analysis will not compare 'ESSPIN schools' to 'non-ESSPIN schools', as it did in CS2, but rather it will compare schools with different degrees of exposure to ESSPIN Output Stream 3 interventions.

Due to ESSPIN's recent expansion to all primary schools in Kaduna, around 81% of primary schools are in the 'minimum' intervention group. These schools only received one year of intervention in 2014/15. 15% of Kaduna primary schools are in the 'medium' intervention group and received two years of full intervention. The pilot schools which received a total of four years of intervention since the start of ESSPIN in Kaduna only amount to 4% of all primary schools in Kaduna (Table 8).

While it makes sense to compare the outcomes of schools with different levels of exposure to the intervention, two points must be kept in mind when interpreting the results: (i) there are spill-over effects between schools, which means that staff in minimum or medium intervention group schools might have already been exposed to ESSPIN ideas through informal communication, or deliberately, by LGEA personnel; (ii) sometimes there are quite extensive changes in school personnel within the state. Thus, just because a school has been exposed to the ESSPIN intervention in the past, this does not necessarily mean that its current teachers and head teachers have, and vice versa.

For certain indicators, we alter the classification scheme slightly according to the purpose of our analysis. For example, when examining teacher competence within the CS3 survey, we consider two different groups: teachers who are in schools that have received ESSPIN intervention but who have not themselves been trained by ESSPIN[3] and teachers who have been trained by ESSPIN.

---

[2] A companion report, *Composite Survey 3: Gender and Inclusion Report* (Daga and Cameron, 2016), focuses on ESSPIN's Output Stream 4 interventions, which run in parallel with Output Stream 3 and which aim to improve inclusion and community participation in schools.

[3] Three to six selected teachers within each school attended workshops delivered by SSOs. In some states the same group of teachers continued to receive training year after year, while in other cases attempts were made to spread the training to teachers who had not yet received any. However, teachers in ESSPIN schools are also expected to receive more support through other channels, and particularly through professional development meetings organised by the head teacher (RTI International, 2014; and personal communications from ESSPIN). We distinguish the teachers who

When examining SBMC functionality and inclusive practices of SBMCs we classify schools according to the amount of Output Stream 4 intervention received. Schools are classified as 'no intervention' (less than five days of Output Stream 4 intervention received), 'post-CS1' (started receiving intervention after CS1), and pre-CS1 (started receiving intervention prior to CS1). This means that the pre-CS1 schools have received the most Output Stream 4 intervention.

When we are looking at one point in time (cross-sectional analysis), the schools that have received more years of ESSPIN activities are expected to perform better. When we are looking at change over time, the schools that have received more years of ESSPIN activities are expected to have improved faster because of ESSPIN.

2.1.2   Types of analysis

The purpose of CS3 is to provide insights into the changes over time in Kaduna, and to evaluate whether the ESSPIN model is having an effect in the specific schools in which its school improvement and community inclusion interventions have operated. We are interested in a wide range of output indicators: teacher competence, head teacher effectiveness, school development planning, school inclusiveness, and the functionality and inclusiveness of SBMCs. Some of these indicators are also combined to give an overall indicator of school quality. Finally, ESSPIN's impact is measured in terms of improved pupil learning outcomes, which we ascertain through test scores in numeracy and English literacy in Grades 2 and 4. For each of these indicators, we present in the following chapter the following two main types of analysis:

1. Change over time between CS1 and CS3 and between CS2 and CS3, for Kaduna as a whole. The recent expansion of ESSPIN interventions means that the programme now has direct links with the majority of schools in Kaduna. By 2015 all schools have had at least one module of ESSPIN intervention, although a large proportion of schools have received the full package of ESSPIN activities for only one year, in 2014/15. Therefore we would expect that schools in CS3 have higher output, outcome and impact measures than schools in CS1 and CS2. However, differences in indicator performance between CS1, CS2 and CS3 cannot be attributed entirely to ESSPIN, since there are likely to be other reasons why schools may be improving (or deteriorating) over time. Some of these other factors have been discussed in Section 1.4.

In the sections below, we use statistical tests (t-tests) to give an indication of whether a difference in results over time is significant. Given that the Composite Surveys present findings for a sample of schools in the state, it is possible that differences in results are driven by the specific features of the sample of schools covered by the surveys, rather than broader trends across the entire population of schools. Significance tests provide an indication of whether a particular difference is likely to be driven purely by the specific features of the sample in question, rather than the population of interest. If a particular difference is statistically significant, this means that we can say with a high degree of certainty that a corresponding difference does exist in the population of interest.

2. Differences between the different levels of intervention categories (minimum, medium and maximum) within the CS3 results. We hypothesise that schools that have received more years of full ESSPIN intervention have higher output, outcome and impact measures than schools which have received fewer years of intervention.

To test this, we calculate the estimated effect of having received one additional year of intervention using a simple regression model. This approach allows us to come one step closer to estimating

---

received direct training ('ESSPIN-trained') from those who were not themselves directly trained, but are in ESSPIN schools and so are expected to have received support from their head teachers and colleagues ('not ESSPIN-trained').

the effect of ESSPIN's intervention. However, this will not be a conclusive indicator of ESSPIN's effect because there are also differences in school and pupil background characteristics within Kaduna. The previous section has shown that pilot schools (the maximum intervention group) are significantly closer located to local government headquarters and have consistently had lower PTRs than schools that received the intervention at a later stage (the minimum intervention group). Controlling for this fully is a more difficult statistical exercise, so we will only attempt this for our impact measure, pupil learning outcomes.

For an outcome indicator, overall school quality, and impact indicators, children's results in literacy and numeracy tests, we conduct additional analysis in order to understand what basis there might be for making causal attribution of ESSPIN's impact. This analysis is described in Sections 5 and 6.2.

2.2   Sampling, coverage and weights

For the first Composite Survey (CS1) 105 schools were sampled in Kaduna. For CS2 this number was increased to 140 schools in order to increase precision (ESSPIN 2015b). For CS3, the sample in Kaduna consisted again of 140 schools (Table 8). Most of these schools had already been visited as part of CS2, although some schools were replaced if they no longer existed or were ineligible. [4]

The number of schools sampled in each intervention group is shown in Table 8. As noted above, the intervention groups are based on the number of years in which the schools received the full package of ESSPIN activities. We would expect to see some improvement between 2012 (CS1) and 2016 (CS3) for all schools that have been part of ESSPIN's SIP at some point.

Between 2014 (CS2) and 2016 (CS3) only the group of schools which had previously not received any intervention received the full package of interventions for one year. In 2014/15 all other schools still received school visits, but not the full intervention package. In 2015/16, none of the primary schools in Kaduna received any SIP intervention. Nevertheless, we would still expect to see continued improvements in outcomes for all schools as school visits are meant to reinforce the effects of training.

Table 8:   Sample in CS3 and population of schools, by intervention group

Category for sampling purposes	Years of full intervention	CS1 sample	CS2 sample	CS3 sample	Population of schools
Minimum	1	28	61	61	3,361
Medium	2	42	42	42	601
Maximum	4	35	37	37	163
Total		105	140	140	4,125

In each of the sampled 140 schools in Kaduna the head teacher and the SBMC chairperson or deputy were interviewed. As in CS2, we also intended to sample six teachers per school, or all of the eligible teachers in schools with fewer than six teachers. The sample coverage among head teachers, SBMC chairs and teachers for interviews and lesson observations is 100% (Table 9). 97% of all sampled teachers completed the literacy test, and 96% completed the numeracy test.

---

[4] Full details of the sampled schools and replacements were submitted to ESSPIN as an annex to the CS3 Training Report and Fieldwork Plan.

Within the classes taught by the sampled teachers, eight pupils from P2 and P4 each were randomly sampled, four for the literacy test and four for the numeracy test. The sample coverage across all pupils is around 97.5%, which is only slightly short of the targeted sample size. Again, the deficit in sample coverage is mainly explained by the fact that some schools had fewer than eight pupils in P2 or P4.

Table 9: Kaduna: Survey instruments, respondents, sample size and coverage in CS3

	Respondents	Targeted sample size	Number of respondents covered	Sample coverage (% of targeted sample size)
Head teacher interview	Head teachers	140	140	100%
SBMC interview	SBMC chair person	140	140	100%
Teacher interview	Sampled teachers	653	653	100%
Teacher literacy test	Sampled teachers	653	632	97%
Teacher numeracy test	Sampled teachers	653	627	96%
Lesson observations	Sampled teachers	653	653	100%
L2	Sampled P2 pupils	560	552	99%
N2	Sampled P2 pupils	560	549	98%
L4	Sampled P4 pupils	560	542	97%
N4	Sampled P4 pupils	560	539	96%

Note. (1) In this table and throughout this report, L2 refers to the Grade 2 literacy test, L4 to the Grade 4 literacy test, N2 to the Grade 2 numeracy test, and N4 to the Grade 4 numeracy test.

(2) The 'targeted sample size' for teachers represents six teachers per school, or the number of eligible teachers in schools where this is less than six.

Comparing the number of schools sampled to the actual population it becomes clear that simple averages of the results from the Composite Survey data would not be representative of what is happening across the state (as Table 8 above shows). To solve this problem sampling weights are applied, giving a greater weight to the results in schools that are relatively under-represented in the survey. Sample weights were calculated for the CS1, CS2 and CS3 schools, teachers and pupils.

2.3   Fieldwork and instruments

Fieldwork for CS3 in Kaduna was conducted using computer-assisted personal interviewing (CAPI) between May and June 2016. We made a number of changes to instruments in response to some additional concerns and to make use of innovations introduced in other recent Nigerian school surveys (described in detail in the CS3 Overall Technical Report). At the same time we retained the questionnaire items required for comparability with previous rounds of the Composite Survey.

Data were collected on teacher competence, head teacher effectiveness, school development planning, inclusive practices in schools, SBMC functionality, teacher competence, teacher subject knowledge and learning outcomes of children in Grades 2 and 4 in English and mathematics. The following activities were carried out as part of the data collection:

- structured interviews with head teachers, SBMC chairpersons and teachers;
- teacher tests in English literacy and numeracy;
- lesson observations; and
- literacy and numeracy tests for pupils in Primary Grades 2 and 4.

The instruments were pre-tested over two days in Abuja during April 2016. State coordinators and monitoring officers collected the data on CAPI after they had been trained on the instruments. Minor revisions were made to the instruments in consultation with state coordinators.

As in CS2, pupil assessments in CS3 were administered using CAPI. Children were given a printed pupil book to read and write in. The interviewers made use of a tablet computer, which prompted them on the questions the children were to be asked orally, gave instructions on the administration of the different test items, including timing, and allowed them to input whether each part of each question was answered correctly or incorrectly (or not attempted at all) by the pupil. A number of changes were made to the CAPI systems and manuals for the administration of the pupil tests, to make them easier to train on and administer. This included a clear manual, with consistent instructions across questions of a particular type, automated timers for timed questions, and translations into Hausa, Igbo and Yoruba of text that did not need to be read in English.

3   School management and head teacher effectiveness

ESSPIN's interventions include leadership training for head teachers on managing the school and its teachers, planning for the school's development, advocating for more resources, and ensuring that the school is inclusive. ESSPIN also supports the development of SBMCs. This includes training and mentoring on how SBMCs can encourage the participation of women and children. This chapter examines how well schools in Kaduna are doing on each of these fronts.

ESSPIN's logframe identifies and defines a number of indicators related to school management, inclusiveness and SBMCs. The logframe groups these indicators into a set of 'standards' or composite indicators. These are as follows:

- Head teacher effectiveness: A head teacher is deemed to be effective if they engage in a set of practices including observing teachers' lessons, holding professional development meetings with teachers, monitoring teacher attendance, keeping records and ensuring that the school adheres to a regular schedule.
- School development planning: As part of the SIP, schools are encouraged to carry out a self-review process involving the head teacher, teachers, SBMCs, parents and other community members. The aim of this process is to identify the school's strengths and weaknesses, and then list the steps that need to be taken to improve it in a school development plan (SDP). The SDP can also be used to request resources from local government or the community. The associated logframe standard assesses whether a self-evaluation has been carried out, whether the school has an SDP, and whether it has implemented the activities in its SDP.
- School inclusiveness: This refers to the extent to which the school makes an effort to include all learners, regardless of gender or socio-economic background. Inclusiveness is assessed on the basis of the steps listed in the SDP and actions taken to boost access, as well as the extent to which teachers encourage the participation of all children in the classroom.
- SBMCs' functionality and performance: The associated standards assess the extent to which SBMCs are functioning and active, and the degree to which they ensure that women and children are actively participating in their activities.

The rest of this section describes each of these standards, and then presents associated findings from the Composite Surveys.

3.1   Head teacher effectiveness

Box 1: Head teacher effectiveness: Key findings

- In 2016, around 3% of all head teachers in Kaduna primary schools met ESSPIN's head teacher effectiveness standard.
- This is not significantly more than in 2012 and 2014.
- There have been significant improvements in some criteria, but also very significant decreases in others.
- There is evidence that head teachers from schools that had received more than just one year of training and school visits by 2016 performed significantly better than those that had received just one year of training.

3.1.1   Criteria for assessing head teacher effectiveness

Head teacher effectiveness is based on seven criteria set out in the ESSPIN logframe (Box 2). These include actions taken by the head teacher, as well as behaviour by teachers and pupils. Head teachers must meet five of the seven criteria in order to be classified as effective.

Box 2: Logframe criteria 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) carried out two or more lesson observations in the past two weeks;

2) held four or more professional development meetings since the start of the 2014/15 or 2015/16 school year (NB: the survey took place more than nine months into the school year);

3) school has a teacher attendance book and the 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; and

7) more than 50% of lessons observed finished within five minutes of a standard 35-minute lesson duration (i.e. the lesson was between 30 and 40 minutes long).

In 2016, 3.4% of head teachers in Kaduna met ESSPIN's standard for head teacher effectiveness. This represents a small decline from the levels seen in 2012 and 2014 (Table 10).

Between 2012 and 2016, head teachers improved significantly in some criteria but worsened in others. The proportion of head teachers who had carried out more than two lesson observations increased from around 5% in 2012 to 25% in 2016. In addition, there has been progress in the proportion of head teachers who carried out at least one professional development meeting.

However, there were also some significant decreases. While in 2012 88% of all head teachers took actions on improving teacher attendance, in 2016 the figure was only 35%. The reasons for the worsening over time are not clear and the findings raise a number of questions for ESSPIN. In particular, it is worth exploring why head teachers have become less likely to take action on teacher attendance. One possibility is that they have limited efficacy to influence teacher attendance (partly because they lack the authority to take disciplinary action against teachers).

Another is that their motivation has been undermined by the fiscal crisis and associated delays in salary payments.

Schools have also worsened with respect to a clear opening time, which is defined by the extent to which teachers and learners agree on what the school's opening time is. This result seems to be mainly driven by a confusion among learners about the opening time. In fact, the proportion of schools where teachers agreed on the opening time increased. There are some doubts as to whether this is a good indicator of school management. Field observations suggested that children were confused about whether to consider the time that they arrived at the school, the time of assembly, or the time when lessons started as the school opening time.

There were similar developments in the proportion of schools in which teachers and pupils were in class on time in the mornings. This is arguably a better indicator of whether schools maintain a clear schedule, as it is based on direct observation.

Fewer schools also conformed to a 35-minute lesson length in 2016 than in 2012 and 2014 (measured as a lesson length of between 30 and 40 minutes). A length of 35 minutes was formerly considered the standard lesson length across the six states. However, schools have been encouraged to adopt 60 minute lessons, in line with ESSPIN lesson plans, which are intended to be taught over one hour. Longer lessons should therefore arguably be discounted as an indicator of school quality, as they may reflect a shift towards one-hour lessons in literacy and numeracy. We have therefore calculated a new indicator, defined as the proportion of schools in which at least half of the observed lessons are at least 30 minutes in length. On this indicator, schools also declined in 2016 compared to 2014, suggesting that lessons are in fact becoming shorter rather than longer. This may partly reflect the effect of observation on teachers. For example, they may be teaching components of lesson plans discussed during training, in a bid to impress the observer, but they may be unable to work these ideas into a full lesson. Even if this is the case, the short lesson times suggest that teachers have difficulty in planning lesson activities that fill a set duration. While disappointing, the lack of clear improvement in head teacher indicators is not surprising given that many teachers have not been paid for months. This might have significantly undermined the motivation of head teachers, as well as the extent to which their actions can have a clear impact on some of these indicators.

Table 10:   Kaduna: Head teacher effectiveness in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
(1) Lesson observations (%)	4.6	2	25.3	+20.8*	+23.4*
No. lesson observations in past two weeks	0.3	0.2	1.2	+1.0*	+1.0*
(2) Professional development meetings (%)	10.1	4.3	22.7	+12.6	+18.4*
No. professional development meetings last term	0.6	0.4	0.9	+0.3	+0.5
(3) Action on teacher attendance (%)	88.2	58.7	34.7	-53.4*	-24.0*
(4) Clear opening time (%)	49.8	36.4	4.5	-45.3*	-31.9*
Learners who agree on opening time (%)		59.2	27.5	n/a	-31.7*
Teachers who agree on opening time (%)		65.9	78.9	n/a	+13.0*
(5) In class on time in morning (%)	77.2	72.3	52	-25.1*	-20.2*
(6) Appropriate morning break (%)	85.4	83.7	70.5	-14.9*	-13.2*
(7) Appropriate lesson length (%)	26.6	46.1	21.1	-5.5	-25.0
Lesson not too short (%)		52.6	36.4	n/a	-16.2
Number of criteria fulfilled (out of seven)	3.5	3.1	2.3	-1.2*	-0.8*
Effective head teacher (five out of seven criteria met) (%)	8.5	10	3.4	-5.1	-6.5
* indicates change over time is statistically significant (p < .05)					

Looking at the findings from CS3 only, the results show that schools which by 2016 had received more years of ESSPIN intervention performed better in most head teacher effectiveness criteria (Table 11). For example, while 45% of all head teachers from maximum intervention group schools (five years of ESSPIN SIP) had conducted two or more lesson observations in the previous two weeks, only 19% from minimum intervention group schools (three years of ESSPIN SIP) did. The estimated effect of one year of full intervention is statistically significant, at 13 percentage points. Head teachers in schools with more years of intervention also held more professional development meetings, ensured that more learners agreed on an opening time and took actions to make sure that teachers were present in class on time.

The proportion of head teachers who met the head teacher effectiveness standard was the highest, at 32% of the maximum intervention group schools, compared to only 0.3% among minimum intervention group schools and 13% across medium intervention group schools. This suggests that schools which have only received one year of ESSPIN intervention so far might benefit from additional years of support.

Table 11:   Kaduna: Head teacher effectiveness in CS3, by intervention group

Intervention group	Min.	Med.	Max.	Estimated effect of one year of full intervention by 2016
(1) Lesson observations (%)	18.8	56.4	45.7	+13.2*
No. lesson observations in past two weeks	0.9	3	2.3	+0.9*
(2) Professional development meetings (%)	20.3	26.6	57.2	+8.8*
No. professional development meetings last term	0.8	1.2	1.6	+0.3
(3) Action on teacher attendance (%)	34.3	28.6	67.1	+6.4
School has a teacher attendance book (%)	98.2	100	100	0.0
(4) Clear opening time (%)	3	11.8	6.7	+2.1
Learners who agree on opening time (%)	25	39.7	32.6	+5.8*
Teachers who agree on opening time (%)	79.5	78.4	68.3	-3.1
(5) In class on time in morning (%)	48.7	59.9	90.8	+15.2
Classes where learners present on time (%)	99.4	99.4	98.6	-0.2
Classes where teachers present on time (%)	47.4	62.5	76.9	+11.1*
(6) Appropriate morning break (%)	68.2	83.8	64.9	+3.3
(7) Appropriate lesson length (%)	24.6	2	19.1	-10.1
Lesson not too short (%)	40.2	14.3	38.5	-7.8
Number of criteria fulfilled (/7)	2.1	2.7	3.5	+0.5*
Effective head teacher (five out of criteria met) (%)	0.3	13.1	31.5	+3.7*
Additional indicators				
In class on time after break (%)	46.3	39.3	85.4	+8.0
Classes where learners present on time (%)	99.6	99.8	97.1	-0.6
Classes where teachers present on time (%)	49.4	57.7	73.2	+8.0
Teacher absenteeism (%)	22.9	13.3	17.1	-3.8
* indicates estimated effect of one year of full intervention is statistically significant (p < .05)				

3.2   School development planning

Box 3: School development planning: Key findings
• In 2016, only around 5% of all primary schools in Kaduna met the standard for effective school development planning.
• Although there were significant improvements in many of the criteria, the proportion of schools that met the SDP standard was not significantly higher than in 2012 and 2014.
• There is some evidence that schools that had received more than just one year of training and school visits by 2016 were significantly more effective at school development planning.

ESSPIN's leadership training encourages and supports head teachers to review how the school is doing each year and to put together a plan for the development of the school, which can be used to advocate for more resources from local government or from the community. Ideally, the plan will not just include infrastructure improvement (e.g. a school fence, a toilet block), but also activities related directly to strengthening teaching and learning, and activities to improve access—particularly for children from disadvantaged backgrounds. Head teachers are also trained on using a cashbook to record the school's expenditures and income.

The effectiveness of school development planning is assessed with respect to five criteria set out by the ESSPIN logframe (   ).

Box 4: Logframe criteria for the effectiveness of school development planning
The school must meet criterion 1 and criterion 2 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 current school year;
2) SDP for current 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; and
5) cashbook is up-to-date (balanced in the last 60 days).

While the share of schools that meet ESSPIN's standard for school development planning effectiveness has not increased significantly since 2012, the average number of criteria that schools meet has increased significantly, from 0.6 in 2012 to 1.5 in 2016. As Table 12 highlights, improvements have been recorded for almost every indicator of SDP effectiveness between 2012 and 2016.

Most notably, the percentage of schools that had an SDP available rose from almost 20% in 2012 to 64% in 2016. The percentage of schools which included activities to strengthen teaching and learning in their SDP increased from 3% to around 22%, and there is evidence that more schools carried out the activities stated in their SDPs. This may reflect the roll-out of the ESSPIN intervention across Kaduna as a whole: school development planning appears to have become a much more widespread practice.

Table 12: Kaduna: SDP effectiveness in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
(1) Written evidence of school self-evaluation process (%)	26.3	10.8	38.3	+12.0	+27.5*
(2) SDP available (%)	19.5	11.3	64.2	+44.7*	+52.9*
(3) SDP contains three or more activities to strengthen teaching and learning (%)	3.4	2.4	21.5	+18.1*	+19.1*
No. activities in SDP to strengthen teaching and learning	0.3	0.2	1.3	+1.0*	+1.1*
(4) Evidence that four or more activities stated in SDP carried out (%)	1.4	2.9	12.8	+11.4*	+9.8*
No. activities in SDP carried out	0.3	0.3	1	+0.8*	+0.7*
(5) Cashbook up-to-date (%)	14.2	25.1	8.5	-5.6	-16.6
School has a cashbook (%)	39.9	62.8	76	+36.1*	+13.2
Number of SDP criteria fulfilled (/5)	0.6	0.5	1.5	+0.8*	+0.9*
School meets effective school development planning standard (%)	1.4	2.1	4.6	+3.3	+2.5
* indicates change over time is statistically significant (p < .05)					

Looking at the findings from 2016 only, and comparing schools from different intervention groups, the data show that more years of ESSPIN intervention is associated with more effective school development planning (Table 13).

As would be expected, the greater the number of years in which the schools received the ESSPIN intervention, the greater the proportion of schools that met the effective school development standard. While only 3% of schools from the minimum intervention group (one year of ESSPIN only) met the effective school development standard, 13% from the medium group (two years of ESSPIN) and 10% of the maximum group (four years of ESSPIN) met the standard. Therefore, there seems to be a benefit in relation to SDP effectiveness in receiving SIP activities for more than one year.

Table 13:   Kaduna: SDP effectiveness in CS3, by intervention group

Intervention group	Min.	Med.	Max.	Estimated effect of one year of full intervention by 2016
(1) Written evidence of school self-evaluation process (%)	34.9	53.1	52.9	+8.7
(2) SDP available (%)	56.8	98.6	90.6	+41.4*
(3) SDP contains three or more activities to strengthen teaching and learning (%)	20.3	25.4	30.4	+3.4
No. activities in SDP to strengthen teaching and learning	1.2	1.9	1.5	+0.3
(4) Evidence that four or more activities stated in SDP carried out (%)	10.5	23.6	18.5	+4.1
No. activities in SDP carried out	0.7	2.5	1.5	+0.6*
(5) Cashbook up-to-date (%)	5.1	22.8	26.3	+5.6*
School has a cashbook (%)	79.3	55.6	82.1	-4.6
Number of SDP criteria fulfilled (/5)	1.3	2.2	2.2	+0.5*
School meets effective school development planning standard (%)	2.8	13.3	9.8	+2.6

* indicates estimated effect of one year of full intervention is statistically significant (p < .05)

3.3   School inclusiveness

Box 5: School inclusiveness: Key findings

- In 2016, only around 4% of all primary schools in Kaduna met the full school inclusiveness standard. 10% were classified as partially inclusive.
- Many of the school inclusiveness criteria had significantly declined compared to 2012 and 2014.
- Schools that received more years of ESSPIN intervention were slightly but not significantly more inclusive in 2016 than schools that received fewer years of ESSPIN intervention.

The criteria on school inclusiveness measure the extent to which the school makes efforts to include all learners, including those from disadvantaged backgrounds. The overall standard for school inclusiveness in ESSPIN depends on four criteria (

Box 6: Standard for school inclusiveness

).

In order to meet the school inclusiveness standard, schools must fulfil three out of four criteria. A school that fulfils two out of the four criteria is classified as partially inclusive. Further detail on these is provided in the companion Gender and Inclusion Report.

Box 6: Standard for school inclusiveness

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 three or more 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); and

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 in between 60% and 40% of total engagements meet the criterion).

The share of schools that met ESSPIN's standard for inclusiveness fell significantly from 23% in 2012 to 4% in 2016 (Table 14). The largest decline was in the share of head teachers that took three or more actions on pupil attendance.

Similar trends were also recorded on two alternative measures of inclusiveness, suggesting that the findings above are not sensitive to the cut-off used in ESSPIN's definition of an inclusive school. In 2012, 55% of all schools were classified as partially inclusive, while in 2016 the figure was only 10%. Another alternative measure of inclusiveness yields similar results. We calculate a percentage score based on the number of actions to improve attendance, the number of activities in the SDP on access, the average number of assessment methods used, the average number of zones participating in each lesson observed (observers imagined the classroom as being divided into six zones), and a measure of the extent to which girls and boys participated equally in the class. This measure shows a decrease since 2012, and a smaller but still significant decrease since 2014.

Table 14:   Kaduna: School inclusiveness in CS 1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
(1) Three or more actions on learner attendance (%)	48.3	40.3	4.5	-43.9*	-35.9*
Number of actions on learner attendance	2.5	2.2	1.5	-1.0*	-0.7*
(2) Two or more activities in SDP on access (%)	3.6	5.3	4.6	+0.9	-0.8
Number of activities on access	0.2	0.2	0.4	+0.2	+0.2
(3) >50% of teachers use two or more assessment methods (%)	71.3	55.1	59.1	-12.2	+4.0
(4) >50% of teachers spatially inclusive and >50% are gender inclusive (%)	38.4	28.4	16.6	-21.8*	-11.8
Number of inclusiveness criteria fulfilled (out of four)	1.7	1.2	0.8	-0.8*	-0.4*
Weighted sum inclusiveness score	52.3	41.8	31.4	-20.9*	-10.4*
School fully met standard (three to four criteria) (%)	23.1	8.9	3.6	-19.6*	-5.3
School partially met standard (two to four criteria) (%)	54.9	44.8	10.1	-44.8*	-34.6*
* indicates change over time is statistically significant (p < .05)					

Furthermore, the maximum intervention group schools performed slightly better in terms of school inclusiveness than the medium and minimum intervention group schools, although these differences are not always statistically significant (Table 15).

Table 15:   Kaduna: School inclusiveness in CS3, by intervention group

Intervention group	Min.	Med.	Max.	Estimated effect of one year of full intervention by 2016
(1) Three or more actions on learner attendance (%)	4.5	2.9	8.9	+0.6
Number of actions on learner attendance	1.5	1.8	1.7	+0.1
(2) Two or more activities in SDP on access (%)	4.1	4.1	15.5	+1.9*
Number of activities on access	0.3	0.7	0.7	+0.2*
(3) >50% of teachers use two or more assessment methods (%)	55.4	76.7	71.6	+10.7
(4) >50% of teachers spatially inclusive and >50% are gender inclusive (%)	18.9	3.2	17.5	-5.6
Number of inclusiveness criteria fulfilled (out of four)	0.8	0.9	1.1	+0.1
Weighted sum inclusiveness score	30.5	35.1	34.5	+2.1*
School fully met standard (three to four criteria) (%)	3.7	0.9	10.6	+0.8
School partially met standard (two to four criteria) (%)	9.8	7.6	26.7	+2.9
Additional indicators				
Enrolment increased since last year (%)	70.2	61.2	65.1	-3.3
Change in enrolment since last year	0.2	0.1	0	-0.1*
* indicates estimated effect of one year of full intervention is statistically significant (p < .05)				

3.4   SBMCs

Box 7: SBMCs: Key findings

- In 2016, all schools in Kaduna had an SBMC. Around 46% of all SBMCs were functional according to the standard used by ESSPIN, compared to 28% in 2012 and 26% in 2014.
- The more years of ESSPIN Output Stream 4 intervention a school has received, the more likely it is to meet some of the SBMC functionality criteria. However, this is not the case with regard to the overall standard.

In 2009 ESSPIN conducted qualitative research into SBMCs and community engagement in education in five ESSPIN states – including Kaduna (ESSPIN 2009). This research suggested that SBMCs were not functioning well: there was a lack of clarity and understanding regarding the SBMCs' role and responsibilities; they lacked the financial resources to support schools in the ways that LGEAs often expected them to; community members were sometimes excluded by local elites; and there was little participation by women and children, despite guidelines requiring their inclusion.

In this context, SBMCs were starting from a low base and with substantial sociocultural barriers to be overcome in order to achieve functionality and inclusive participation. ESSPIN has aimed to improve community involvement in schools through functioning SBMCs and increased women's and children's participation, with a number of interventions under its Output Stream 4 (see Appendix C).

To be counted as functioning well, SBMCs are expected to meet regularly, and work with the community, community-based organisations (CBOs), traditional or religious institutions, to raise awareness about the school and its needs, raise resources, and address exclusion. They are expected to have a women's committee and a children's committee, and to keep financial records, and the chairperson is expected to visit the school regularly. There are nine criteria in the standard for SBMC functionality (Box 8). In most cases, these require evidence to be presented, rather than just accepting the word of the respondent (usually the SBMC chairperson). Thus, they reflect the ability of the SBMC to keep good records of their activities, as well as to carry out the activities themselves. In addition to these criteria, we present statistics on a number of other measures from the Composite Survey SBMC interviews (Table 16).

Box 8: Logframe criteria for SBMC functionality

The school must meet at least five of the nine criteria listed below in order to meet the SBMC functionality standard for the current school year:[5]

1) two or more SBMC meetings have taken place since the start of the current school year (written evidence);

2) SBMC conducted awareness-raising activities (written or oral evidence);

3) SBMC took steps to address exclusion (written or oral evidence);

4) SBMC networked with CBOs, traditional or religious institutions, or other SBMCs (written or physical evidence);

5) SBMC interacted with LGEAs on education service delivery issues (written or physical evidence);

6) an SBMC women's committee exists (written or physical evidence);

7) an SBMC children's committee exists (written or physical evidence);

8) SBMC contributed resources for the school (written or physical evidence); and

9) SBMC chair visited the school at least three times since the start of the current school year (written evidence).

Over the past years, the proportion of schools with functional SBMCs in Kaduna has risen to around 46% (Table 16). Compared to 2014, on average, almost two more criteria of SBMC functionality were met by schools in Kaduna.

The developments across the individual indicators have been mostly positive. There have been significant improvements across certain criteria, such as the proportion of schools in which the SBMC held meetings over the past year, the proportion of schools in which the SBMCs conducted awareness training and the proportion of SBMCs that networked with CBOs, other institutions or other SBMCs.

---

[5] A slightly different standard, with 10 criteria, was used in CS1. The new standard, with nine criteria, was applied to both the CS1 and CS2 data.

Table 16: Kaduna: SBMC functionality in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
(1) Two or more meetings this school year (%)	18.9	6.5	36.5	+17.6	+30.0*
(2) Conducted awareness-raising (%)	33.7	20.1	61.1	+27.4*	+40.9*
(3) Addressed exclusion (%)	27.5	19.9	77.2	+49.7*	+57.3*
(4) Networked with CBOs/institutions/other SBMCs (%)	21.4	36.5	95.1	+73.3*	+58.7*
(5) Interacted with LGEA (%)	27.6	24.7	32.1	+4.5	+7.4
(6) Has women's committee (%)	17.3	31.5	24.4	+7.1	-7.1
(7) Has children's committee (%)	27.5	12.4	18.9	-8.6	+6.5
(8) Contributed resources for school (%)	48.1	49.2	42.1	-6.0	-7.1
(9) Chair visited school three or more times (%)	31.0	13.1	31.9	+0.9	+18.8
School meeting functioning SBMC standard (five out of nine criteria met) (%)	28.0	25.8	46	+18.0	+20.2*
Number of SBMC functionality criteria met (out of nine)	2.6	2.1	3.9	+1.3*	+1.8*
Additional indicators: inclusion and drop-out					
(A1) Action for commonly excluded groups (%)	2.1	11	4.8	+2.8	-6.2
(A2) Raised issue of children's exclusion (%)	3.9	17	2	-1.9	-14.9*
(A3) Raised cash to support vulnerable children (%)	-	-	20.1	n/a	n/a
(A4) Monitored drop-out or non-attendance (%)	-	-	57.1	n/a	n/a
(A5) Communicated with school or community about drop-out (%)	-	-	92.5	n/a	n/a
(A6) No. actions taken to address non-attendance	-	-	1.8	n/a	n/a
Additional indicators: organising and mobilising resources					
(A7) School has an SBMC (%)	91	100	100	+9.0	+0.0
(A8) Cashbook available (%)	35.3	24.7	59.1	+23.8*	+34.4*
(A9) Requested support from LGEA or SUBEB (%)	-	-	65.6	n/a	n/a
(A10) Raised cash to support school improvement (%)	29.8	38.6	10.9	-19.0	-27.7*
(A11) Mobilised non-cash resources (%)	28.9	40.5	38	+9.1	-2.5
(A12) Involved in making SDP (%)	-	7.6	66.9	n/a	+59.3*

* indicates change over time is statistically significant (p < .05)

Schools that received more years of ESSPIN Output Stream 4 intervention did not perform significantly better in terms of SBMC functionality than schools that received fewer years of

intervention (Table 17). Across most indicators, the estimated effect of a full year of intervention is not statistically significant, but for six out of nine indicators the effect is positive. Schools which received intervention pre-CS1 are significantly more likely to have a women's committee than the schools that had no intervention. However, overall, among schools that received the Output Stream 4 intervention, the proportion that meet the SBMC functionality standard is actually lower than among those that did not receive the intervention. The difference is not statistically significant.

Box 9: Asking SBMCs about inclusion and exclusion

A number of different criteria aim to measure the SBMC's inclusiveness and the actions it has taken on excluded children. These were based on the following questions addressed to the SBMC chairperson. As elsewhere, questions were asked in the local language, with instructions to use a language that the respondent could understand, but not to provide additional explanation or prompts.

Criterion	Question asked (with data collector instructions in blue)	Criterion met if…
(2) Conducted awareness-raising	Did the SBMC do anything to raise awareness about the value of education for all boys and girls in the community in the current school year?	Respondent answers yes and can present oral or written evidence
(3) Addressed exclusion	Did the SBMC do anything to address issues which prevent children from attending school or which cause drop-out in the current school year?	Respondent answers yes and can present oral or written evidence
(A1) Took action for commonly excluded groups	Did the SBMC do anything to **support commonly excluded groups** in the **current school year**? You can explain that commonly excluded groups could be orphans, nomadic children, girls, children with disability, ethnic or religious minorities, etc.	Respondent answers yes and can present oral or written evidence
(A2) Raised issues of children's exclusion	Did the SBMC **raise issues of children's exclusion** from school in the community, with the LGEA, or with the state government, in the **current school year**?	Respondent answers yes and can present oral or written evidence
(A3) Raised cash to support vulnerable children	Did the SBMC mobilise any cash to support vulnerable children in the current school year?	Respondent answers yes (no evidence required)
(A4) Monitored drop-out or non-attendance	**What actions were taken** to address issues which prevent children from attending school	Respondent answers yes to a previous question (asking whether any action was

| (A5) Communicated with school or community about drop-out

(A6) Number of actions taken to address non-attendance | or which cause drop-out in the **current school year**?
Do not prompt. This is a multiple response question **– SELECT ALL THAT APPLY**
• Monitoring drop-out
• Monitoring non-attendance
• Communicating with school about drop-out
• Communicating with community about drop-out
• Other (specify)
• Don't know / refused | taken to address these issues) and then provides this information in the follow-up question on what type of action and how many actions were taken. No specific evidence is required |

For two key indicators of SBMC action to make the school inclusive – whether the SBMC took action for commonly excluded groups, and whether it raised issues of children's exclusion (see , above) – around 5% of SBMCs could show evidence that they had met the criteria, and there was a positive difference between the schools with more years of Output Stream 4 intervention and those with fewer.

Table 17:   Kaduna: SBMC functionality in CS3, by intervention group

Intervention group	No intervention	Pre-CS1	Post-CS1	Estimated effect of one year of full intervention by 2016
(1) Two or more meetings this school year (%)	28.6	39.9	49.7	+9.0
(2) Conducted awareness-raising (%)	58.6	75.2	60	+4.7
(3) Addressed exclusion (%)	81.6	82.5	66.9	-2.7
(4) Networked with CBOs/institutions/other SBMCs (%)	97.2	97.1	90.3	-1.0
(5) Interacted with LGEA (%)	43.2	14.9	18.4	-13.1
(6) Has women's committee (%)	19.2	40.9	27.5	**+9.1***
(7) Has children's committee (%)	18.3	33.8	14.1	+5.1
(8) Contributed resources for school (%)	44.1	62.1	30.4	+0.8
(9) Chair visited school three or more times (%)	34.7	31.9	26.7	+0.3
Number of SBMC functionality criteria met (out of nine)	4	4.6	3.6	+0.2
School meeting functioning SBMC standard (five out of nine criteria met) (%)	53.6	38.2	35.6	-3.7
Additional indicators: inclusion and drop-out				
(A1) Took action for commonly excluded groups (%)	3.7	5.6	6.6	+0.8

(A2) Raised issues of children's exclusion (%)	0	11.5	2	**+1.7***
(A3) Raised cash to support vulnerable children (%)	22.9	16.3	16.4	-2.3
(A4) Monitored drop-out or non-attendance (%)	75.3	62.7	49.1	-7.8
(A5) Communicated with school or community about drop-out (%)	64.5	34.9	52.3	-9.6
(A6) No. actions taken to address non-attendance	96.5	96.5	83.7	-1.8
Additional indicators: organising and mobilising resources				
(A7) School has an SBMC (%)	100	100	100	n/a
(A8) Cashbook available (%)	59.3	49.1	62.7	+0.4
(A9) Requested support from LGEA or SUBEB (%)	75.3	62.7	49.1	-7.8
(A10) Raised cash to support school improvement (%)	7.7	17.4	14.2	+3.2
(A11) Mobilised non-cash resources (%)	39.1	52.6	30.3	+0.1
(A12) Involved in making SDP (%)	62.9	73.3	71.7	+3.8
*indicates estimated effect of one year of full intervention is statistically significant (p < .05)				

3.4.1 How inclusive are SBMCs of women and children?

Box 10: SBMCs' women and children inclusiveness: Key findings

- In 2016, 8% of all SBMCs met the women's inclusiveness standard and 7% met the children's inclusiveness standard. Both indicators have improved since 2012.
- There is evidence that schools that received some Output Stream 4 intervention pre-CS1 perform better in terms of SBMC women's and children's inclusiveness than schools that received no Output Stream 4 intervention.

As noted above, SBMCs are expected to have women's and children's committees. We also record a number of other measures of the extent to which SBMCs are inclusive of women's and children's concerns. In each case, there are four criteria and an overall standard (   ).

Box 11: Logframe standard for SBMCs' inclusiveness of women and children

The school must meet at least three of the four criteria listed below in order to meet the SBMC **women**'s inclusiveness standard:

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 a female member of the SBMC);

3) at least one issue raised by a female member at an SBMC meeting led to action (written, physical or oral evidence from a female member of the SBMC); and

4) at least one SBMC women's committee meeting took place.[6]

The school must meet at least three of the four criteria listed below in order to meet the SBMC **children's** inclusiveness standard:

1) at least one child attended two or more SBMC meetings (written evidence);

2) a 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); and

4) at least one SBMC children's committee meeting took place and committee has a trained facilitator.[7]

The proportion of SBMCs in Kaduna which met the women's inclusiveness standard rose from 6% in 2012 to 8% in 2016 (Table 18). This is a small and not significant change. Nevertheless, there have been strong increases across many of the individual criteria of SBMC women's inclusiveness over the past four years, such as the proportion of SBMCs where a female member raised an issue and the proportion of SBMCs whose women's committee met at least once.

The findings for SBMC children's inclusiveness criteria are similar to those for the SBMC women's inclusiveness criteria. The proportion of schools that met the children's inclusiveness standard rose from 0.5% to 7%, a statistically significant increase. In 2016 there were significant improvements in the proportion of schools in which issues raised by children led to action and in the proportion of schools in which more than two SBMC meetings were attended by at least one child.

---

[6] This criterion has been slightly altered since CS1, where it also required that the women's committee have a female leader.

[7] In CS1 this criterion required written evidence in the form of minutes of at least one children's committee meeting held in the past school year. This requirement was dropped for CS2 as it was considered unlikely that children's committees would keep good minutes, and that failure to keep minutes does not mean the committee is not functioning.

Table 18: Kaduna: SBMC inclusiveness of women and children in CS1, CS2 and CS3

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
Participation of women in SBMCs					
(1) At least one woman attended two or more meetings (%)	9.2	3.2	20.4	+11.2	+17.2*
(2) Female member raised an issue (%)	14.6	29.0	40.3	+25.7*	+11.2
(3) Issue raised by female member led to action (%)	31.1	8.5	22.7	-8.3	+14.2*
(4) Women's committee met (%)	1.6	24	18.4	+16.8*	-5.6
Number of criteria met (out of four)	0.5	0.6	0.8	+0.3	+0.2
Meets women's inclusiveness standard (three out of four criteria met) (%)	5.9	4.0	8.3	+2.4	+4.3
Participation of children in SBMCs					
(1) At least one child attended two or more meetings (%)	1.6	1.8	16.1	+14.5*	+14.3*
(2) Child raised an issue (%)	5.3	18.8	10.3	+5.1	-8.5
(3) Issue raised by child led to action (%)	1.3	2.8	16.7	+15.4*	+13.9*
(4) Children's committee met and has a trained facilitator (%)	0.4	15.1	6.7	+6.3*	-8.4
Number of criteria met (out of four)	0.1	0.4	0.5	+0.4*	0.1
Meets children's inclusiveness standard (three out of four criteria met) (%)	0.5	1.3	7.4	+6.9*	+6.1*
* indicates change over time is statistically significant (p < .05)					

In 2016 19% of schools in Kaduna which received some Output Stream 4 intervention pre-CS1 reached the SBMC women's inclusiveness standard and 8% of schools which received Output Stream 4 intervention post-CS1 did so. However, only 7% of schools which did not receive any intervention reached the same standard (

Table 19). The schools that received the intervention pre- and post-CS1 are also more likely to reach the children's inclusiveness standard, although the estimated effect of a full year of intervention is not statistically significant. However, schools that received some Output Stream 4 intervention perform significantly better across some of the individual children's inclusiveness indicators.

Table 19:   Kaduna: SBMC inclusiveness of women and children in CS3, by intervention group

Intervention group	No intervention	Pre-CS1	Post-CS1	Estimated effect of one year of full intervention by 2016
Participation of women in SBMCs				
(1) At least one woman attended two or more meetings (%)	15.3	24.6	28.3	+5.3
(2) Female member raised an issue (%)	48.4	53.4	25.7	-0.1
(3) Issue raised by female member led to action (%)	26.6	23.4	15.4	-0.4
(4) Women's committee met (%)	16	33.3	17	+6.7
Number of criteria met (out of four)	0.8	1.1	0.8	+0.2
Meets women's inclusiveness standard (3/4 criteria met) (%)	5.9	18.9	8.4	**+4.0***
Participation of children in SBMCs				
(1) At least one child attended two or more meetings (%)	9.7	21.4	25.6	**+5.2***
(2) Child raised an issue (%)	6	20.1	14.4	**+5.2***
(3) Issue raised by child led to action (%)	18	14.7	15.2	-0.7
(4) Children's committee met and has a trained facilitator (%)	0	23.5	12.5	**+5.2***
Number of criteria met (out of four)	0.3	0.7	0.6	**+0.2***
Meets children's inclusiveness standard (3/4 criteria met) (%)	4.4	11.6	11.3	+2.7

* indicates estimated effect of one year of full intervention is statistically significant (p < .05)

3.5   Summary and discussion

Kaduna recently expanded the ESSPIN intervention to all schools, so that by 2015 all primary schools in the state had benefited from at least one year of the full Output Stream 3 intervention package (see Annex A). Output Steam 4 intervention has also been scaled up recently and in 2014/15 and 2015/16 14% of school in Kaduna received the full package of the intervention. However, by 2016 the vast majority of schools in Kaduna had not yet received any intervention under Output Stream 4.

Box 12 shows the average progress in the different standards of school management and head teacher effectiveness in Kaduna primary schools between 2012 and 2016. We would expect the proportion of schools that achieved the standards to be the lowest in 2012, and the highest in 2016, as ESSPIN was rolled out across more schools. The farther to the right a point on the graph is, the better the average performance of schools in that category.

Looking at the average proportion of schools in Kaduna which met the different effectiveness standards, little progress seems to have been made since 2012. There was a slight but not statistically significant decrease in the proportion of schools which achieved the head teacher effectiveness standard, a statistically significant decrease in the proportion of schools that achieved the school inclusiveness standard, and no statistically significant change in the proportion of schools with a functional SBMC.

While there were some improvements across school development planning and SBMC inclusiveness of women, these were not statistically significant. The only statistically significant increases across Kaduna primary schools between 2012 and 2016 were in the proportions of schools that had a functional SBMC and that had an SBMC that was inclusive of children.

These trends may be partly explained by the difficult context in which ESSPIN has been implemented in the state. Over the past two years, violence and armed attacks have been a continuing problem in Kaduna and many teachers have gone for months without being paid. Furthermore, the recent change in government in Kaduna led to a reshuffling of LGEA staff, including most head teachers, as well as a suspension of training from May 2015 to October 2016. These factors might have seriously undermined the motivation and capabilities of head teachers and community members in regard to improving school management.

Furthermore, it must be noted that over the last two years the implementation of the ESSPIN Output Stream 3 intervention in Kaduna was not continuous. While in 2014/15, the intervention was rolled out to all schools that had not previously received any intervention, schools that had previously received training and school visits did not receive any from 2014 onwards due to financial constraints. In addition, due to political problems, none of the schools received any training or visits in 2015/2016, the year before the survey. A large scale-up, such as that which took place in 2014/15, where over 3,000 new schools started receiving the intervention, could have been connected to organisational problems, and, thus, could have decreased the quality of the training and school visits delivered.

Box 12: Kaduna: Change in school management and head teacher standards over time
Comparing CS1 (2012), CS2 (2014) and CS3 (2016)
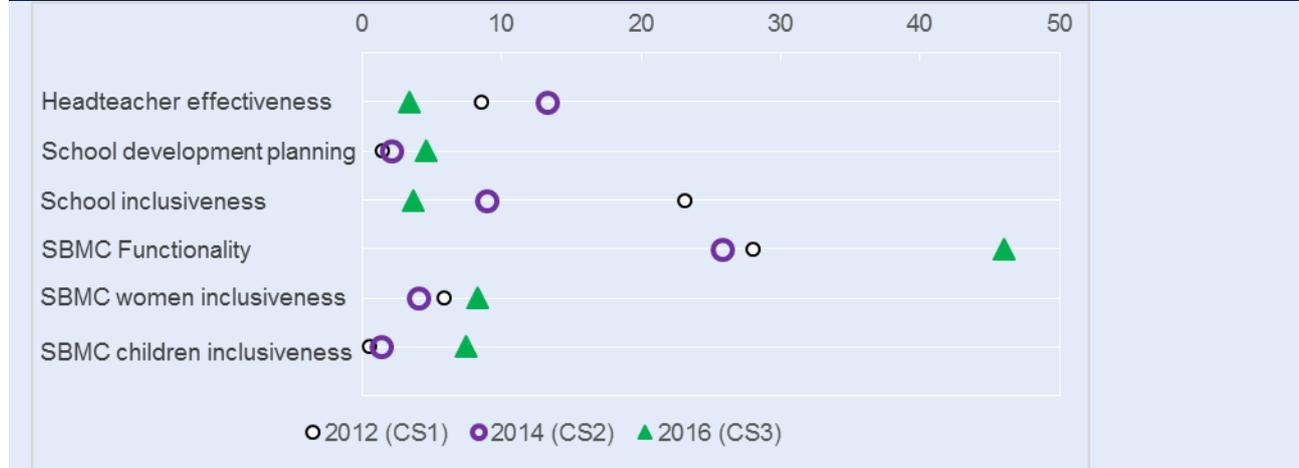
While the developments over time in Kaduna are disappointing, there is evidence that shows that the longer a school benefited from the ESSPIN intervention, the better it performed in most of these categories. In Box 13 we can see that among the schools that had received four years of intervention by 2016 (maximum), the proportion of schools that met the various standards is the highest in all but one category. Schools that received more years of ESSPIN intervention have significantly more effective head teachers. They also perform better in terms of school development planning and school inclusiveness, although we did not find a significant effect of a full year of intervention in these two categories. All in all, the results are positive and they suggest that the minimum intervention group schools, which have only received one year of intervention so far, could benefit from additional years of training and school visits. However, when interpreting these results it is important to keep in mind that the maximum intervention group schools were already systematically different from the other schools to begin with. Besides having received more

years of intervention, factors such as lower PTRs across all years could have contributed to their relatively better performance.

It must also be noted that only around 4% of all primary schools in Kaduna are part of the maximum intervention group, and around 15% are part of the medium intervention group. This means that over 80% of all schools in Kaduna had only received one year of ESSPIN intervention by 2016. Given that this is also the group of schools that performed the worst across all standards, the low averages across Kaduna in 2016 are unsurprising.

Box 13: Kaduna: Difference in school management and head teacher standards by ESSPIN intervention

Comparing minimum and medium intervention schools in 2016
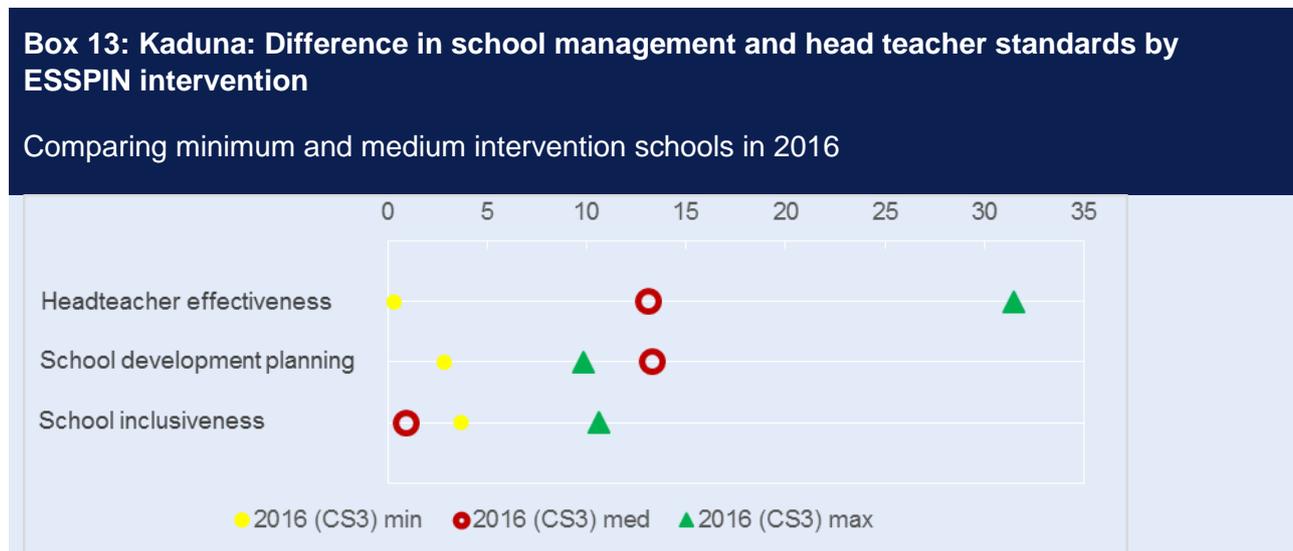
Schools which received more years of Output Stream 4 intervention (pre-CS1 and post-CS1) performed significantly better in terms of SBMC inclusiveness of children compared to schools that received no intervention. They are also more likely to meet the SBMC women's inclusiveness standard, although the effect of a full year of intervention is not statistically significant here. However, they are not more likely to meet the overall SBMC functionality standard than schools that received no intervention. Considering the limited roll-out of Output Stream 4 intervention post-CS1, this suggests that there may be some beneficial effects in regard to SBMC functionality and inclusiveness as a result of expanding Output Stream 4 intervention.

Box 14: Kaduna: Difference in SBMC standards by ESSPIN intervention

Comparing no Output Stream 4 intervention schools to pre-CS1 and post-CS1 Output Stream 4 intervention schools
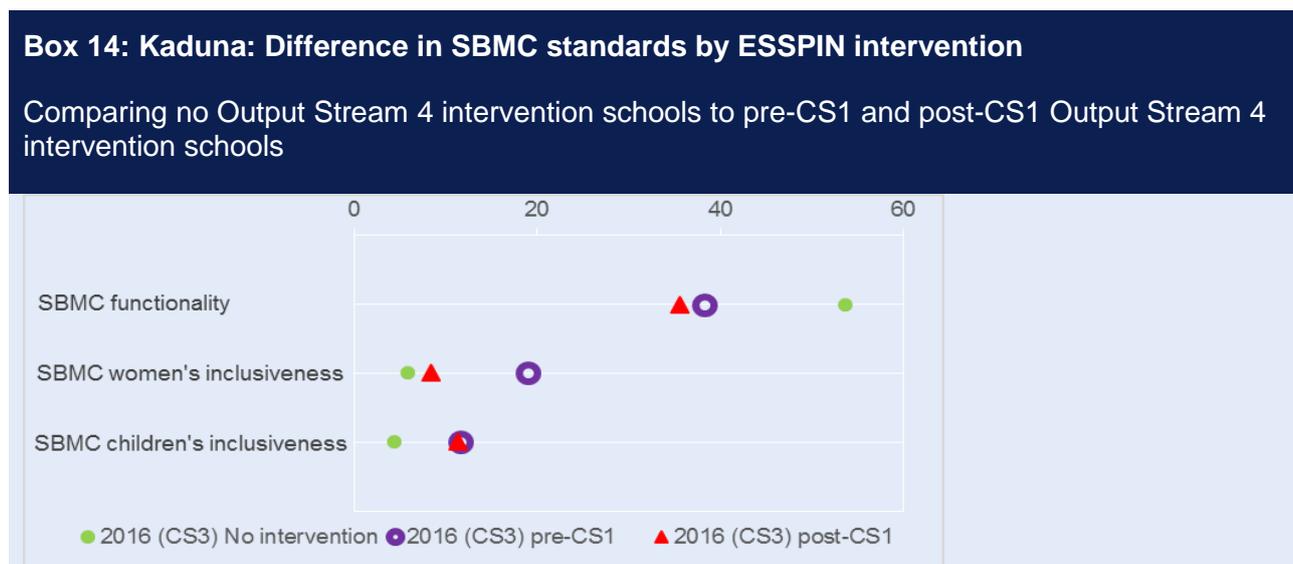
4 Teachers

ESSPIN's interventions include teacher training on teaching skills, including the use of teaching aids, participation and praise, and techniques for classroom organisation. The training also includes basic literacy and numeracy training. This chapter examines the changes in teacher competence over time and across different intervention groups. In this chapter, we first look at how teacher competence has changed according to the teacher competence logframe indicator that combines the various aspects on which teachers receive training. We then take a more in-depth look at teachers' performance on the literacy and numeracy content knowledge tests.

4.1 Teacher competence

Box 15: Teacher competence: Key findings

- Since 2012, teacher competence in Kaduna has decreased, but there have been some increases since 2014.
- In 2016 around 63% of all teacher met the competence standard and 21% met the stricter teacher competence standard.
- Teachers that had been trained by ESSPIN in 2014 performed better across most indicators in 2016 than those who had not been trained. However, these differences are not statistically significant.

Teacher competence is based on four criteria set out by the ESSPIN logframe (Box 16). Teachers who teach English and mathematics meet the competence standard if they fulfil three of the four criteria. Teachers who teach subjects other than English and mathematics are exempted from criteria 1 and therefore meet the competence standard if they fulfil two out of the three remaining criteria.

For CS2 and CS3, a stricter version of the competence indicator was developed. The criterion in regard to using at least one teaching aid during the lesson observation is changed to exclude reading from, writing on, or having pupils copy from, the blackboard, since this is considered poor use of a teaching aid that is less likely to enhance learning. In addition, a fifth criterion was added, based on teacher content knowledge test results. Teachers are defined as competent if they are competent according to the original criteria, and if they can also score at least 50% in primary school-level literacy and numeracy tests.

Box 16: Criteria 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 mathematics. 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 reprimands during lesson observation; and

4) in terms of class organisation: assigning individual or group tasks at least twice during lesson observation (or for two contiguous five-minute blocks).

For CS2 and CS3, stricter criteria for teacher competence were introduced. These modified (2) to exclude reading from or writing on, or having pupils copy from, the blackboard as a use of a teaching aid. A fifth criterion was added:

5) literacy and numeracy: scores at least 50% in both an English literacy and a numeracy test.

The percentage of teachers who meet the teacher competence standard decreased from 76% in 2012 to around 63% in 2016, although this change is not statistically significant (Table 20). However, this is a significantly improved result compared to 2014, when only 50% of all teachers reached the competence standard.

In terms of individual indicators, the changes were mostly positive between 2014 and 2016, but they were negative compared to 2012.

We also calculate a continuous 'competence score', based on the number of criteria met by each teacher. A teacher who meets all of the three or four criteria would score 100%, while a teacher who meets none of them would score 0%. The original teacher competence score as well as the stricter version have shown slightly positive but not significant improvements between 2012 and 2016.

When interpreting these results, it is important to note that it is possible that CS1 over-estimated the indicators of teacher competence. It was noted in the CS1 report that performance was surprisingly strong given the weak scores for teachers found in an earlier (2010) Teacher Development Needs Assessment (ESSPIN, 2013b). This could be because, for example, data collectors administered the question about English and mathematics curriculum benchmarks incorrectly.

Table 20:   Kaduna: Teacher competence in 2012–2016

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
(1) Knowledge of Eng./mathematics curriculum (%)	62.9	29.9	40.5	-22.4*	+10.6*
(2) Use of one or more teaching aid (%)	94.9	91.9	97.1	+2.2	+5.1
(2a) Use of one or more teaching aid, excl. blackboard (%)		74.4	77.0	n/a	+2.6
(3) Praise more than reprimand (%)	76.4	79.0	86.7	+10.3	+7.7
(4) Assigns two or more ind./group tasks (%)	60.2	32.4	29.2	-31.0*	-3.2
(5) Passes English and mathematics test (%)		33.6	36.9	n/a	+3.3
Teacher competence score (% of criteria fulfilled)	75.9	60.1	64.5	-11.4*	+4.4
Teacher competence standard fulfilled (three out of four criteria met) (%)	76.2	49.6	62.9	-13.3	+13.3*
Teacher competence score (% of criteria fulfilled; strict version)		55.5	57.9	n/a	+2.4
Teacher competence standard fulfilled (strict version: four out of five criteria met) (%)		18.6	21.2	n/a	+2.6

* indicates change over time is statistically significant (p < .05)

Note. The CS2 version of the competence score adds the teacher's performance in the literacy and numeracy tests to the number of other criteria met by the teacher. For example, a teacher who met all four original criteria and also scored 100% in the literacy and numeracy tests would receive a competency score of 100%.

Next, we examined how teachers who reported having received ESSPIN training performed compared to those that did not report having received ESSPIN training (Table 21).

Overall, 65% of all teachers who were trained by ESSPIN fulfilled the competence standard in 2016, compared to only 61% of teachers who were not ESSPIN-trained. ESSPIN-trained teachers performed better across most indicators compared to non-ESSPIN-trained teachers. While most of these differences do not reach the level of being statistically significant, the sizes of the effects are still positive and fairly large.

Table 21:   Kaduna: Teacher competence in CS3, ESSPIN-trained versus non-ESSPIN-trained

	Non-ESSPIN-trained	ESSPIN-trained	Difference in means
(1) Knowledge of Eng./mathematics curriculum (%)	35.9	45.2	+9.4
(2) Use of one or more teaching aid (%)	96.5	97.6	+1.1
(2a) Use of one or more teaching aid, excl. blackboard (%)	72.0	82.4	+10.3
(3) Praise more than reprimand (%)	90.2	83.0	-7.2
(4) Assigns two or more ind./group tasks (%)	23.8	34.9	+11.1
(5) Passes English and mathematics test (%)	37.0	36.8	-0.3
Teacher competence score (% of criteria fulfilled)	63.2	65.9	+2.7
Teacher competence standard fulfilled (three out of four criteria met) (%)	60.7	65.2	+4.5
Teacher competence score (% of criteria fulfilled; strict version)	55.8	60.2	+4.3
Teacher competence standard fulfilled (strict version: four out of five criteria met) (%)	19.0	23.4	+4.4
* indicates difference is statistically significant (p < .05)			

4.2   Findings from teacher content knowledge tests

The findings above suggest that teachers' content knowledge did not improve significantly between 2014 and 2016. There is also no statistically significant difference between teachers trained through ESSPIN and those who had not received ESSPIN training. Percentage scores in the teacher content knowledge tests provide a rough indication of teachers' test performance, but analysis using item response theory (IRT) provides more reliable learning scales that can also be interpreted more readily in terms of learning benchmarks (see Allen, 2016a). The teachers' results can be divided into four performance bands in literacy and five performance bands in numeracy. Review of the items that teachers in each band can mostly answer correctly then provides descriptors for each band (Table 22). For example, a teacher in Band 2 for literacy is one who shows knowledge of some basic phonics, can write a simple sentence, and can carry out basic comprehension of a passage, as well as satisfying the easier items – testing limited comprehension of simple passages, basic nouns and verbs – associated with a teacher in Band 1. The teacher in Band 2 cannot typically correctly answer the harder items associated with Bands 3 or 4, such as identifying simple antonyms.

Table 22:   Band descriptors based on IRT analysis

Band	Literacy	Numeracy
5		Understands conversion of fractions to decimals, and place values in decimals
4	Creates several sentences, shows knowledge of phonics, punctuation, formal letter layout, suffixes and alphabetical order	Understands ideas of area, nets, pictograms and rounding
3	Past/present of verbs, completes a sentence, extracts basic information from a passage, identifies simple antonyms, forms plurals	Understands basic sets, use of the number line to represent sums, conversion of units of time and mass, can complete word problems involving division
2	Shows knowledge of some basic phonics, writes a simple sentence, basic comprehension of a passage	Simple division, word problems involving addition, signs for arithmetic operations, integer comparisons and integer place values
1	Limited comprehension of simple passages, basic nouns and verbs	Simple addition with carrying over, simple subtraction, identify a fraction, counting, simple regular shapes

Within the literacy and numeracy tests, items can be grouped according to specific sub-domains of learning: reading, writing and grammar within literacy, and number concepts and calculation within numeracy.

In Kaduna, teachers' scale scores in English have not changed significantly since 2014, but there were some clear improvements in mathematics (Table 23). In English, the share of teachers in the higher proficiency bands increased slightly between 2014 and 2016. In mathematics, there were statistically significant increases in both sub-scales, number concepts and calculation. These increases were matched by a change in the distribution. Higher proportions of teachers are now found in the higher performing bands in mathematics and the proportions in the lower performing bands declined (Figure 2). However, the changes in the proportions per band were not statistically significant.

Table 23:   Kaduna: Teachers' test scores (IRT analysis) in 2014 and 2016

	2014 (CS2)	2016 (CS3)	Change: 2014 vs. 16
English IRT scale score (mean 500, s.d. 100)	489	490	+0.4
English Band 1 (%)	10	9	-0.7
English Band 2 (%)	45	37	-8.2
English Band 3 (%)	38	48	**+9.9***
English Band 4 (%)	7	6	-1.0
Reading (English sub-scale, mean 500, s.d. 100)	495	496	+0.5
Writing (English sub-scale, mean 500, s.d. 100)	483	477	-5.4
Grammar (maths sub-scale, mean 500, s.d. 100)	488	491	+3.5
Mathematics IRT scale score (mean 500, s.d. 100)	475	494	**+19.5***
Mathematics band 1 (%)	8	5	-2.6
Mathematics Band 2 (%)	39	30	-9.1
Mathematics Band 3 (%)	39	44	+4.6
Mathematics Band 4 (%)	12	15	+3.5
Mathematics Band 5 (%)	2	6	+3.6
Number concepts (maths sub-scale, mean 500, s.d. 100)	477	493	+16.4

Calculation (maths sub-scale, mean 500, s.d. 100)	474	496	+21.9*

* indicates change is statistically significant (p < .05)

Figure 2:   Kaduna: Proportion of teachers in each English and mathematics performance band, by year
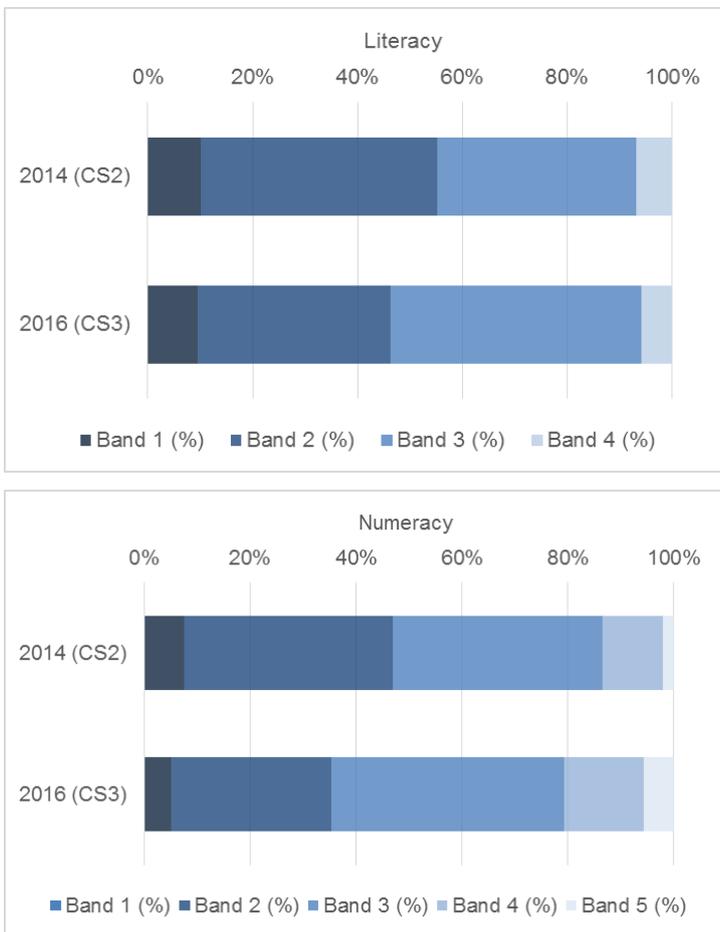
Teachers who had received ESSPIN training in 2014 did not have significantly higher test scores in 2016 than those who had not, in both English and mathematics (

Table 24). Overall scores in English are slightly lower for ESSPIN-trained teachers than for non-ESSPIN-trained teachers, although the difference is not statistically significant. ESSPIN-trained teachers have higher overall maths scores than non-ESSPIN-trained teachers, but again this is a result that does not reach the level of being statistically significant. This is also reflected in the distribution of teachers across proficiency bands (Figure 3). For English, the distribution across the bands between the two groups is fairly similar. In mathematics, significantly fewer ESSPIN-trained teachers can be found in the lower performing bands, but instead more can be found in the middle performing band. There was a statistically significant decrease in the proportion of teachers found in the highest band.

Table 24:   Kaduna: Teachers' test scores (IRT analysis) by ESSPIN training

	Non-ESSPIN-trained	ESSPIN-trained	Difference in means
English IRT scale score (mean 500, s.d. 100)	492	488	-3.3
English Band 1 (%)	10	9	-1.3
English Band 2 (%)	34	40	+6.7
English Band 3 (%)	50	46	-4.3
English Band 4 (%)	6	5	-1.1
Reading (English sub-scale, mean 500, s.d. 100)	496	495	-0.8
Writing (English sub-scale, mean 500, s.d. 100)	484	469	-15.0
Grammar (maths sub-scale, mean 500, s.d. 100)	490	492	+2.3
Mathematics IRT scale score (mean 500, s.d. 100)	488	501	+13.4
Mathematics Band 1 (%)	5	5	-0.4
Mathematics Band 2 (%)	36	24	**-12.8***
Mathematics Band 3 (%)	39	49	+9.6
Mathematics Band 4 (%)	16	15	-1.0
Mathematics Band 5 (%)	3	8	+4.6
Number concepts (maths sub-scale, mean 500, s.d. 100)	489	499	+10.3
Calculation (maths sub-scale, mean 500, s.d. 100)	488	504	+16.3
* indicates difference is statistically significant (p < .05)			

Figure 3:   Kaduna: Proportion of teachers in each English and mathematics performance band, ESSPIN-trained versus not ESSPIN-trained

Numeracy

Band 1 (%) Band 2 (%) Band 3 (%) Band 4 (%) Band 5 (%)
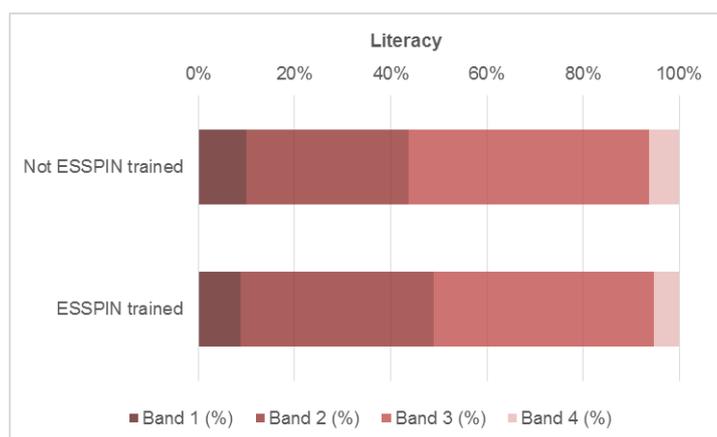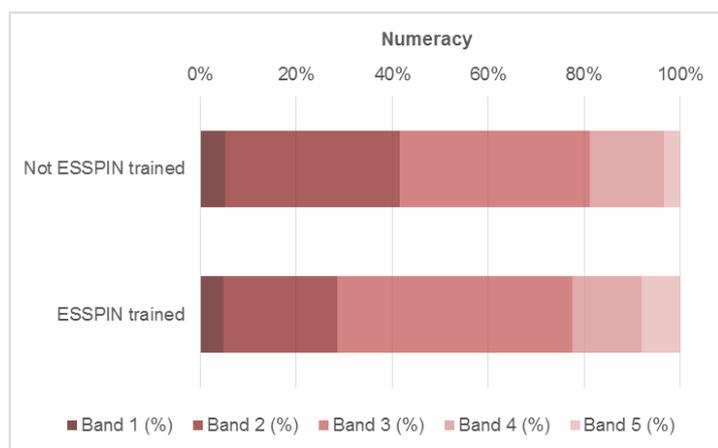
4.3   Teacher motivation

It has been proposed that teacher motivation is an important aspect which may sustain the effects of a training intervention. For example, as teachers acquire new skills through teacher training, their motivation may increase, as they feel more effective.

For this round of the survey (CS3), we included a measure of teacher motivation and teacher interaction using a scale that had been developed for the Nigerian context, and that had been used and tested in two previous school-based surveys. We define teacher motivation as the propensity of teachers to start and maintain behaviours that are directed towards fulfilling their professional goals, and in particular towards achieving better learning outcomes for the school's learners (Cameron, 2015b). Many existing instruments designed to measure teacher motivation focus exclusively on 'efficacy' – the extent to which teachers see themselves as able to influence their pupils' learning outcomes – which can also be seen as the 'can do' aspect of motivation (Bennell and Akyeampong, 2007). We wished to go beyond this to include measures relating more closely to teachers' willingness to work hard, their commitment, their effort and their enjoyment, which might together be labelled as 'will do' aspects of motivation.

The motivation scale we developed was incorporated into the teacher interview. Teachers were asked to what extent they agreed ('strongly disagree', 'disagree', 'Agree', 'Strongly agree') with a series of statements that measure different aspects of motivation. The scale consists of three sub-scales of teacher motivation (satisfaction, skills and engagement) and one scale of teacher–teacher interaction (collegiality). The three sub-scales of teacher motivation were combined into a composite motivation measure by calculating the mean of the three sub-scales[8]. The teacher motivation scale was also analysed using IRT.

Table 25 describes each of the different sub-scales and provides some examples of the items used to assess these.

---

[8] The three sub-scales were also combined into a composite measure using partially non-compensatory methods. These produced composite measures which were very highly correlated with the simple mean composite.

Table 25:   Teacher motivation and interaction scale and sub-scales

Scale	Description	Example of items
Collegiality	How I see the extent of commitment and collaboration among my colleagues ('teacher–teacher interaction')	• All of the teachers in my school trust each other • All teachers at this school are highly committed to their job
Satisfaction	The value I place on my role as a teacher ('interest and enjoyment')	• I always enjoy teaching very much • I like to spend a lot of energy to make my classes interesting
Skills	The perception I have of my competences and skills as a teacher ('self-efficacy')	• I believe I know how to teach well • I believe I have the skills needed to encourage my learners to always work hard
Engagement	How engaged and committed I feel I am in relation to my work as a teacher ('pressure/tension')	• It is difficult to manage learners in my classroom • Teaching is very tiring
Composite measure (mean of satisfaction, skills and engagement)		

Table 26 shows the levels of motivation among teachers as reported during CS3, comparing those who reported having received ESSPIN training and those that did not report having received ESSPIN training.

In Kaduna, teachers who were trained by ESSPIN do not appear to be more motivated in 2016 than teachers who were not. However, ESSPIN-trained teachers do report feeling more engaged at their school than non-ESSPIN-trained teachers, even if the results do not reach the level of being statistically significant.

While it may be disappointing to see that, overall, ESSPIN-trained teachers do not appear to be more motivated that those who have not been trained, these results have to be interpreted within the context of Kaduna. Interviews have shown that the state has had problems in regard to the payment of teachers in the past two years. Many teachers have not been paid for several months. In addition, enrolment in schools has increased due to the introduction of free feeding programmes and free school uniforms. Increased class sizes may have made it more difficult for teachers to apply good teaching methods. Coupled with a lack of payment, it might not be surprising that ESSPIN training did not have any effect on teacher motivation.

Table 26:   Kaduna: Teacher motivation and interaction by ESSPIN training

	Non-ESSPIN-trained	ESSPIN-trained	Difference in means
Collegiality	486.3	482.9	-3.4
Satisfaction	484.5	487.5	+2.9
Skills	483.2	467.7	-15.6
Engagement	476.6	497.3	+20.7
Composite motivation measure	481.8	484.1	+2.3

Note: All scores are normalised to have an average (mean) of 500 and a standard deviation of 100.

* indicates difference is statistically significant (p < .05)

4.4   Summary and conclusion

Teachers in Kaduna have become significantly more competent in 2016 compared to 2014, but they are less competent than in 2012. In 2016, 63% of all teachers fulfilled the original teacher competence standard and 21% fulfilled the stricter teacher competence standard, which also takes into account their results in the teacher knowledge tests. In summary, while a large proportion of teachers employs effective teaching methods, they still lag behind in terms of knowledge of the English and mathematics curriculum. Only 37% of teachers passed the content knowledge tests. However, it should be noted that the majority of teachers have not received any training since November 2014, and therefore noticeable positive effects were not expected.

More specifically, we did not find any evidence that teachers in Kaduna had improved their test scores in English over the last two years, but there were some significant improvements in teachers' mathematics test scores. In 2016, slightly more teachers could be found in the middle and high performing bands in English and mathematics than in 2014.

The results showed that teachers who were trained in Kaduna by ESSPIN were slightly more competent than teachers who were not trained, although these differences are not statistically significant. Furthermore, they did not perform better on the English tests, but they did have slightly higher test scores in mathematics than non-ESSPIN-trained teachers. There is no evidence that trained teachers are more motivated than their peers that are not trained.

5   Trends in school quality

Box 17: School quality: Key findings

- In 2016 only 5% of all schools in Kaduna met the school quality standard, and less than 1% met the stricter version of that standard.
- There is a very small but not statistically significant increase from 2012 and 2014.
- None of the schools from the minimum intervention group met the stricter school quality standard in 2016 but there seems to be a positive effect of an additional year of ESSPIN intervention on school quality. Among schools that received two years of intervention 14% met the original school quality standard, and among schools with four years 18% met the quality standard.

For a broader sense of school quality, how it differs between schools with different levels of ESSPIN intervention, and how it has changed over time, it is useful to define an overall measure of school quality. We do this using the standard developed as part of ESSPIN's logframe. This is a combination of the standards discussed above on teacher competence, head teacher effectiveness, school development planning, and SBMC functionality. A quality school is defined as one that meets the teacher competence standard and at least two of the other standards (   ). We also use a 'quality score' indicator, which is an average of the continuous indicators developed in the previous sections for teacher competence, head teacher effectiveness, school development planning and SBMC functionality. A school that meets all of the criteria under all of the standards will get 100%, while a school that meets none of the criteria will get 0%. The original version of these indicators, used in CS1, did not take into account teachers' content knowledge. For CS2 and CS3, however, we also present a 'strict' version of the standard, which takes into account the results in the teacher content knowledge tests (see Section 4.2 above).

Box 18: 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 (more than half the teachers sampled in each school must be competent);

2) head teacher effectiveness standard;

3) school development planning effectiveness standard; and

4) SBMC functionality standard.

The version of this standard used in CS1 did not rely on teacher content knowledge tests. For CS2, we introduce a second, more strict version of the standard, in which teachers must get above 50% in literacy and numeracy tests to be classed as competent (see Section 4.1 and Box 16 above).

We find that 5% of all schools in Kaduna met the standard for school quality in 2016. This is only slightly higher than in 2012 and 2014, and this difference is not statistically significant. The quality score decreased between 2012 and 2014, but increased again after 2014 so that in 2016 it is back around the same level as in 2012. Looking at the stricter version of the quality standard, we find that only 0.4% of all schools meet this standard in 2016. The stricter quality score shows some increase in school quality since 2014.

Table 27:   Kaduna: School quality in 2012–2016

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
School meets quality standard (%)	1	2.2	5.1	+4.1	+3.0
Quality score (%)	38.3	33.7	41.9	+3.6	+8.2*
School meets quality standard (strict version) (%)		0.1	0.4	n/a	+0.3
Quality score (strict version) (%)		32.8	40.1	n/a	+7.2*
* indicates change over time is statistically significant (p < .05)					

Within CS3, the results suggest that schools which had more than one year of ESSPIN intervention were more likely to meet the school quality standard and had higher school quality scores. Among the minimum intervention group schools, which only received one year of intervention in 2014/15, a mere 2.8% met three out of four standards to reach the original school quality standard, and none of the schools in that group met the stricter school quality standard. On the other hand, 19% of all schools with four years of intervention met the original school quality standard and 7% met the stricter school quality standard. Looking at the quality score, we estimate a statistically significant effect of one year of full intervention of 6 percentage points.

Table 28:   Kaduna: School quality across different intervention groups

Intervention group	Min.	Med.	Max.	Estimated effect of one year of full intervention by 2016
School meets quality standard (%)	2.8	14.4	18.5	+3.6*
Quality score	40.2	47.7	55	+5.6*
School meets quality standard (strict version) (%)	0	1	6.9	+0.6*
Quality score (strict version) (%)	38.2	46.4	53.8	+5.9*
* indicates estimated effect of one year of full intervention is statistically significant (p < .05)				

Can the differences in quality between the intervention groups be attributed to the intervention, or are they associated with differences in the schools at baseline? One way of answering this question is to focus on the change over time in the different intervention groups.

Examining the difference between intervention groups in regard to the change over time between 2012 and 2016 shows that there is a difference: in schools with two to three years of intervention during the relevant time period, the average score increased by 10.5 percentage points, while in schools with no intervention, or one year of intervention, the score increased by only 2 percentage points (Table 29). While the difference is positive, suggesting that schools with more years of intervention improved faster, it does not reach the level of being statistically significant. Therefore, it is difficult to say with certainty whether the quality differences can really be attributed to the intervention.

Table 29: **Kaduna: Difference between intervention groups in change over time (2012 – 2016)**

	Intervention during 2011/12–2014/15		
	0–1 years	**2–3 years**	**Difference**
2012 (CS1)	38.2	38.7	0.5
2014 (CS2)	31.7	42.1	10.4
2016 (CS3)	40.2	49.2	9.0
Difference (2012-2016)	2	10.5	+8.5
* indicates the difference is statistically significant (p < .05)			

6 Learning outcomes

The ultimate aim of ESSPIN is to improve learning outcomes in government schools in the six states. In this chapter, we examine the trends in learning outcomes over time, differences in learning outcomes between schools that have received more or fewer years of ESSPIN intervention, and evaluate whether effects on learning achievement can be attributed to ESSPIN.

> **Box 19: Learning outcomes: Key findings**
>
> - Learners' performance on all tests has worsened significantly since 2012. Between 2014 and 2016, there were improvements in all tests, except for Grade 2 literacy. However, test scores still remain lower than they were in 2012.
> - Learners in schools that have received more years of ESSPIN intervention perform better on most tests (all except Grade 2 numeracy) than learners in schools with fewer years of ESSPIN intervention, although these differences are not statistically significant.

6.1 Pupil learning achievements in English literacy and numeracy

Learning outcomes were measures in literacy and numeracy at Grades 2 and 4, and analysed using IRT (see Allen, 2016b and Allen, 2016c). The analysis for each test produces a scale score which, by design, has an average (mean) of 500 and standard deviation of 100. This scale is also divided into bands, indicating the level of proficiency of the learner. For the Composite Surveys, bands have been designed to correspond to the levels of proficiency expected at each grade in the Nigerian curriculum. For example, a learner in Band 2 for literacy is one who is able to demonstrate knowledge and skills in at least some of the tasks that are considered to be within the range of Grade 2 proficiency. Table 30 and Table 31 list some examples of the tasks within each band.

Table 30: Examples of knowledge and skills that learners in each literacy band can demonstrate

Band 4: Grade 4 and above	Read and understand the grammatical structure of a sentence and complete a missing word using 'where', 'which', 'what' and 'who'
	Follow the conventions of letter-writing to complete a letter template. Completing grammatically accurate sentences, with correct spelling, and a greeting and sign off
	Read for meaning a short, simple text with a range of sentence structures independently
Band 3: Grade 3 literacy	Read phonically decodable two-syllable and three-syllable words that include common diagraphs and adjacent consonants
	Independently plan and write a grammatically correct simple sentence
	Read a simple sentence for meaning and complete a missing word using correct spelling
Band 2: Grade 2 literacy	Use phonic knowledge to say initial sounds of familiar animals
	Use knowledge of common inflections in spellings, plurals, to write the answer to a question
	Spell simple high frequency words accurately
Band 1: Emerging literacy	Verbally compose a short grammatically correct sentence in the continuous present tense in response to a question about a picture
	Listen to a short passage and remember specific details to respond verbally to a question
	Clearly shaped and correctly orientated copying of words with an understanding of space and full stops
Band 0: Pre-literacy	Understand and respond verbally with a grammatically correct sentence to a simple question about their age
	Understand and respond verbally with a grammatically correct sentence to a simple question about their name
	Use phonic knowledge to say initial sounds of familiar objects and animals

Table 31: **Examples of knowledge and skills that learners in each numeracy band can demonstrate**

Band 5: Grade 5 and above	Solve a word problem involving differences in time Determine which number rule was used to make one number into another Solve a simple algebra problem
Band 4: Grade 4 numeracy	Being able to gather information by interpreting simple graphs Calculate the area of a rectangle, multiplying a decimal number, to 1 decimal place, by a one-digit number, and record the answer in m2 Choose the most appropriate strategy to subtract a decimal number, to 2 decimal places and a two-digit number, involving measure
Band 3: Grade 3 numeracy	Multiply a two-digit number by a one-digit number Use short division; subtract a two-digit number from a two-digit number crossing the tens boundary Choose a strategy to add a three-digit number and a two-digit number crossing the tens boundary, involving money
Band 2: Grade 2 numeracy	Use non-standard units of measure to compare the capacity of three containers Subtract a two-digit number from a two-digit number Name common 2D shapes Extend counting past 800 and count in tens
Band 1: Emerging numeracy	Recognise and complete a sequence of three two-digit numbers that are multiples of five Subtract a one-digit number from a two-digit number 1–19 Read analogue clock to the hour
Band 0: Pre-numeracy	Compare the length of two straight lines Use non-standard units of measure to compare the capacity of three containers Count to 10

In Kaduna, the trend in learning outcomes between 2012 and 2016 has been mixed (

Table 32). Across both grade levels, literacy and numeracy scores decreased between 2012 and 2014. Since 2014, there has been no statistically significant change in Grade 2 literacy and numeracy but there have been some statistically significant improvements in Grade 4 literacy and numeracy. However, since the initial drop between 2012 and 2014 was quite large, scores across all categories are still lower in 2016 than they were in 2012.

As a result, compared to 2012, there are now larger proportions of students in the lower performing bands and smaller proportions of students in the higher performing bands for both literacy and numeracy. For example, while in 2012, 55% of all Grade 2 literacy students were performing at pre-school level, in 2016, 75% of Grade 2 students were performing at this level.

The results must be interpreted keeping in mind that between 2009/10 and 2014/15 Kaduna's education system had to cope with an increase in enrolment of about 200,000 pupils, many of whom are likely to come from disadvantaged backgrounds. It would be unrealistic to expect schools to absorb such a high number of additional pupils and still deliver progress in learning outcomes at the expected levels. That being said, to the extent that some of those additional children are studying in improving (ESSPIN) schools, more education has been delivered overall during the programme lifetime, which is an achievement in itself.

Table 32:   Kaduna: Learning outcomes in 2012–16

	2012 (CS1)	2014 (CS2)	2016 (CS3)	Change: 2012 vs. 2016	Change: 2014 vs. 2016
Grade 2 literacy score	474.2	455.7	452.2	-22.0	-3.5
Band 0: Pre-school (%)	54.6	69.9	74.6	**+20.1***	+4.7
Band 1: Grade 1 (%)	27.8	18.9	14.9	**-12.8***	-4.0
Band 2: Grade 2 (%)	17.7	11.2	10.4	-7.2	-0.7
Grade 4 literacy score	460.5	419.5	447.1	-13.4	**+27.6***
Band 1: Grade 1 (%)	66.4	86.4	79.6	**+13.1***	**-6.8***
Band 2: Grade 2 (%)	15.9	6.5	8.1	-7.8	+1.6
Band 3: Grade 3 (%)	5	4.6	4.1	-0.9	-0.5
Band 4: Grade 4 (%)	12.7	2.5	8.2	-4.5	+5.7*
Grade 2 numeracy score	524	442.2	454.9	**-69.1***	+12.6
Band 0: Pre-school (%)	2.4	19.2	5	+2.6	**-14.2***
Band 1: Grade 1 (%)	62.7	67.7	80.2	**+17.5***	+12.5
Band 2: Grade 2 (%)	34.9	13.1	14.8	**-20.1***	+1.7
Grade 4 numeracy score	493.9	431.3	459	**-34.9***	**+27.7***
Band 1: Grade 1 (%)	12.3	36.7	31.5	**+19.2***	-5.2
Band 2: Grade 2 (%)	38	43.6	37.3	-0.7	-6.3
Band 3: Grade 3 (%)	22.9	10.4	15.1	-7.8	+4.7
Band 4: Grade 4 (%)	16.4	6.2	9.9	-6.4	+3.8
Band 5: Grade 5 (%)	10.5	3.2	6.1	-4.3	+3.0

* indicates change over time is statistically significant (p < .05)

In Kaduna, learning outcomes appear to be slightly better for learners whose schools have received more years of ESSPIN intervention (

Table 33). For all tests, except for Grade 2 numeracy, the estimated effect of a year of full intervention is positive, but never statistically significant. Fewer students from maximum intervention group schools are located in the lower performing bands and more in the higher performing bands.

Table 33: Kaduna: Learning outcomes by ESSPIN intervention group in 2016

	Min.	Med.	Max.	Estimated effect of one year of full intervention
Grade 2 literacy score	450.4	454	467.7	+5.2
Band 0: Pre-school (%)	75.4	75.4	61	-3.2
Band 1: Grade 1 (%)	13.4	17.9	22.4	+3.0
Band 2: Grade 2 (%)	11.3	6.7	16.5	0.0
Grade 4 literacy score	447.1	443.5	463.6	+3.0
Band 1: Grade 1 (%)	80.5	79.4	66.2	-3.4
Band 2: Grade 2 (%)	7.6	8.7	12.7	+1.4
Band 3: Grade 3 (%)	2.4	8.8	6.5	**+1.9***
Band 4: Grade 4 (%)	9.5	3.1	14.6	-0.5
Grade 2 numeracy score	458	445.8	456.1	-3.8
Band 0: Pre-school (%)	5.9	1.2	9.8	-0.4
Band 1: Grade 1 (%)	78	87.7	73.7	+1.7
Band 2: Grade 2 (%)	16	11.2	16.5	-1.3
Grade 4 numeracy score	455.5	464.7	479.1	+8.2
Band 1: Grade 1 (%)	33.1	28.3	25.2	-3.4
Band 2: Grade 2 (%)	39.7	32.6	27	-5.3
Band 3: Grade 3 (%)	13.1	20.2	18.4	+2.9
Band 4: Grade 4 (%)	7.8	14.6	16.7	+3.2
Band 5: Grade 5 (%)	6.3	4.3	12.7	+0.9

* indicates estimated effect of one year of full intervention is statistically significant (p < .05)

Learners from minimum intervention group schools (blue bars) are disproportionately concentrated in the lower performing bands, especially for Grade 4 numeracy (Figure 4). The largest proportion of learners in the higher bands are from maximum intervention group schools (yellow bars) in each of the four categories.

Figure 4: Kaduna: Distribution of test scores by intervention group in 2016
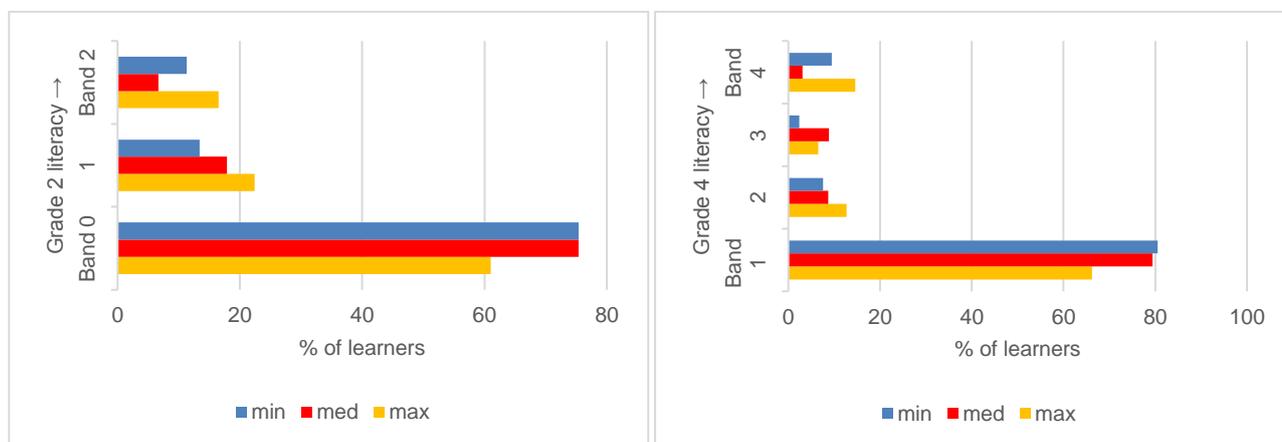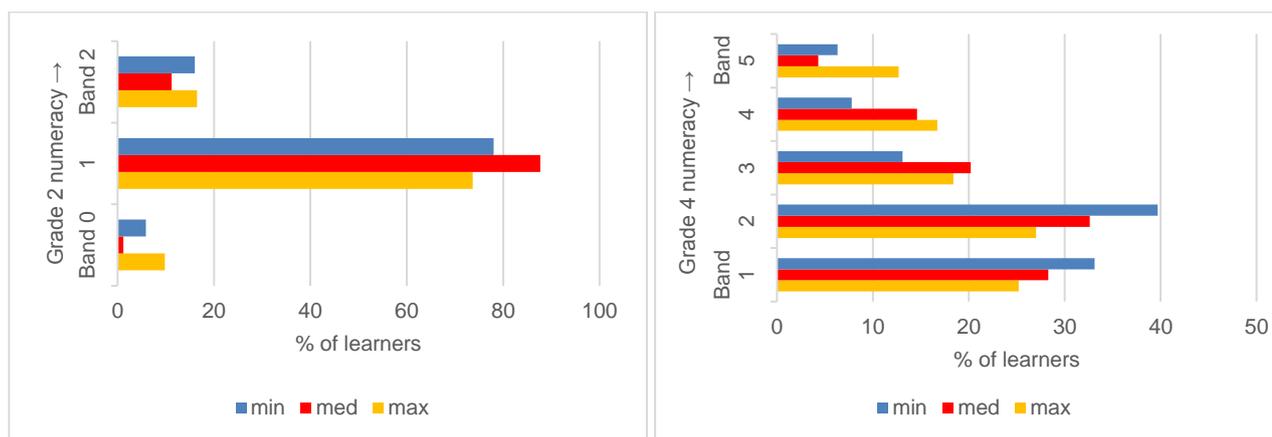
As noted above and as is shown in Figure 5, by 2016 learners from the maximum intervention groups had the highest scores among the intervention groups in all categories. While the scores of all groups have decreased since 2012, in some tests pupils from schools that had received more years of ESSPIN intervention did not worsen as much as pupils from schools that only had received one or two years of intervention. For Grade 2 literacy and Grade 4 numeracy, the scores of learners from maximum intervention group schools decreased by less than those of learners from minimum and medium intervention group schools. For Grade 4 literacy, learners from maximum intervention group schools appear to be the only ones who actually had higher scores in 2016 than in 2012. These patterns do not apply for Grade 2 numeracy, where the scores of learners from all intervention groups deteriorated at a similar rate.

In Section 6.2, we use regression analysis to examine more rigorously how change over time varies with ESSPIN intervention, while controlling for possible confounding variables such as school characteristics.

Figure 5:   Kaduna: Learning outcomes by test, year and ESSPIN intervention group
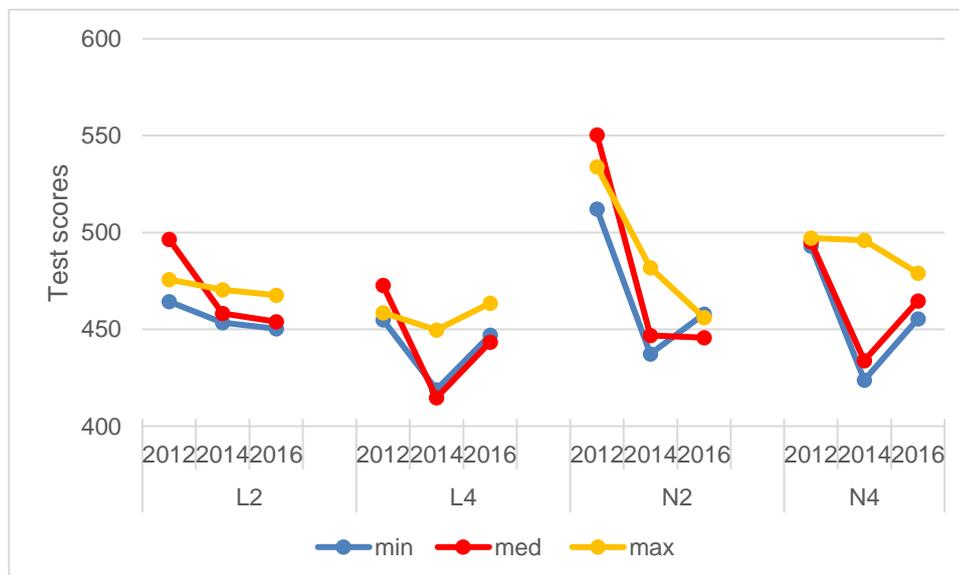
6.2   Controlling for school and pupil characteristics

6.2.1   Differences in background characteristics

Schools from different intervention groups in Kaduna have somewhat different background characteristics. In Kaduna, the schools from the maximum and medium intervention group tend to

be older and to be located closer to local government authority headquarters and urban centres. Furthermore, they are less likely to be nomadic or Islamic. Less remote, urban schools tend to have better learning outcomes than more remote, rural schools. In line with this, schools in the maximum intervention group may perform better than schools in the other two groups regardless of the level of ESSPIN support received. This highlights the need to control for background characteristics when estimating ESSPIN's impact on learning outcomes.

In Kaduna, schools that had received more years of ESSPIN intervention also tended to have lower PTRs and, on average, more class rooms and teachers than schools that received less intervention. These characteristics could further impose a positive bias on our results. On the other hand, maximum intervention group schools in Kaduna experienced a disproportionate increase in enrolment and PTRs during 2009/10–2014/15. More specifically, enrolment in maximum intervention group schools increased by 39%, while in minimum intervention group schools enrolment only increased by 14% during the same period. PTRs in maximum intervention group schools increased by 69%, compared to 28% in minimum intervention group schools. As a result, schools may have difficulty coping with rapid enrolment increases, especially if new learners are from more disadvantaged backgrounds. Increases in PTRs would tend to reduce teachers' ability to ensure all students achieve good learning outcomes. This particular difference between intervention groups in Kaduna could bias our estimates of ESSPIN's intervention effects downwards.

In summary, there are a number of differences between the groups of schools that have had more years ESSPIN intervention and those that have had fewer, and taken together these could bias our estimates of ESSPIN's effect in either direction. We use a number of statistical methods to control for these differences in the following section.

6.2.2   Are learning outcomes better in schools with more intervention in 2016?

In this section we add statistical controls for these 'confounding variables' – characteristics of schools that might affect learning outcomes and make it harder to tell whether the intervention is having an effect or not. We also estimate a model which controls for pre-existing differences in test scores by adding test scores in CS1 as a confounding variable.

We use ordinary least squares regression analysis to estimate the models. Regression analysis estimates the correlation of learning outcomes with ESSPIN intervention, conditional on school characteristics.

The results are shown in Table 34. We can see that, before controlling for any of the factors that could bias the estimates, there appears to be a statistically significant effect of ESSPIN intervention on all test scores (Model 1). However, after controlling for a variety of different school characteristics, we do not find a statistically significant effect of more ESSPIN intervention on learning outcomes in Kaduna (Model 2). Introducing lagged school-level learning outcomes which also controls for pre-existing differences in test scores per school also does not find a statistically significant effect across any of the four tests (Models 3 and 4).

Table 34:   Estimates of the effect of ESSPIN intervention on learning outcomes in 2016

Model	L2		L4		N2		N4	
(1) Simple regression, clustered standard errors, no sample weights	26.22	*	25.82	*	32.96	*	29.02	*
(2) Full covariates	5.98		11.25		1.35		9.42	
(3) Lagged school-level learning outcomes	-1		1.24		0.22		0.32	
(4) Lagged outcomes and covariates	-5.01		1.12		-5.86		-0.3	
* indicates estimated effect is statistically significant (p < .05)								

7 Conclusions and implications of the Composite Surveys' findings for ESSPIN in Kaduna

This report has found that an overwhelming majority of the schools in Kaduna still do not meet ESSPIN's standards for a good school. In 2016, around 3% of all schools met the standard on head teacher effectiveness; 5% on school development planning; 4% on inclusion; 46% on functional SBMCs; and 5% or 0.4% on overall school quality (depending on which indicator we use). This might be partly due to the fact that around 80% of all schools in the state have only received one year of intervention thus far.

There has been little improvement over time across these standards and their indicators. Many of the standards were reached by a lower proportion of Kaduna primary schools in 2016 than in 2012. In particular, school inclusiveness has declined significantly since 2012. The lack of strong progress in Kaduna might at least partly be explained by rising enrolment numbers and PTRs, a decline in federal funding, payment issues surrounding teachers' salaries, as well as the ongoing security situation in the state. For example, between 2009/10 and 2014/15, enrolment rose by about 200,000 pupils, and PTRs increased from about 36 pupils per teacher to 49 pupils per teacher in that same period, However, there have been some increases since 2014, especially in terms of school development planning, SBMC functionality and, most importantly, school quality. While these improvements might be modest in terms of the share of schools meeting ESSPIN's quality standards, they are more significant when we consider the number of pupils affected. Compared to 2012, an estimated 50,000 more pupils now learn in a school of adequate quality, and an estimated 260,000 more pupils go to a school that has a functioning SBMC.

The good news is that we found that schools which have had more years of ESSPIN Output Stream 3 intervention have more effective head teachers in 2016. Schools with more years of Output Stream 4 intervention are more likely to have a SBMC that is inclusive of women and children than schools that have received no intervention. Given that the vast majority of schools has not received any Output Stream 4 intervention thus far, this suggests that it could be beneficial to expand Output Stream 4 activities to the remaining schools in the state. In terms of overall school quality, we also found that schools that had received more years of ESSPIN Output Stream 3 intervention performed better in 2016 and appeared to have improved faster than schools that received fewer years of intervention. However, these results do not reach the level of being statistically significant.

Teachers' competence in Kaduna appears to have worsened between 2012 and 2014, and then to have recovered in 2016, but the net result is no significant improvement between 2012 and 2016. Most teaching behaviours have not changed significantly, but teachers have improved their mathematics test scores since 2014. English test scores stayed the same. There are no statistically significant differences between ESSPIN-trained and non-ESSPIN-trained teachers, although a larger proportion of ESSPIN-trained teachers reached the competence standard. However, their test scores in both English and mathematics were lower than those of the non-ESSPIN-trained teachers. There was also no difference in teacher motivation between ESSPIN-trained and non-ESSPIN-trained teachers in Kaduna.

We found that pupil's learning outcomes in Kaduna are lower than in 2012, but higher than in 2014 (except for Grade 2 literacy). We found that there are some significant differences in background characteristics between the schools from the different intervention groups, which might be partly responsible for the differences in learning outcomes. Therefore, we used a regression model to control for all of these factors, as well as for pre-existing differences in the levels of test scores. However, we still did not find a significant effect of the ESSPIN intervention on test scores.

These findings must be interpreted in the light of the findings on teacher competence and knowledge. Although 63% of teachers were found to apply competent teaching methods, in 2016 only 9% of teachers fall within the top proficiency band in English and only 5% fall into the top band for mathematics. It is questionable how much impact one can expect to see from better teaching skills if teachers' subject knowledge levels are low. This highlights the difficult environment in which ESSPIN's interventions are being implemented. It also underlines the importance of identifying what types of interventions should be prioritised in the face of such severe weaknesses in content knowledge.

Bibliography

Allen, R. (2016a) CACSA report ESSPIN2.3: ESSPIN scaled scores and subscores – literacy. Unpublished note. Curriculum, Assessment and Certification Systems Architects

Allen, R. (2016b) CACSA report ESSPIN3.1: ESSPIN teacher survey – notes on subscales (updated). Unpublished note. Curriculum, Assessment and Certification Systems Architects

Allen, R. (2016c) CACSA report ESSPIN3.5: ESSPIN scaled scores and subscores – numeracy. Unpublished note. Curriculum, Assessment and Certification Systems Architects

Bennell, P., and Akyeampong, K. (2007). Teacher Motivation in Sub-Saharan Africa and South Asia (Department for International Development: Educational Papers). DFID.

Cameron, S. (2015b) Teacher motivation in the TDP baseline survey. Unpublished technical note. Oxford Policy Management.

ESSPIN (2009) School Based Management Committees in Policy and Practice, Volume 1: Research Synthesis Report. Report number ESSPIN 404.

ESSPIN (2013a) Overall findings and technical report of ESSPIN Composite Survey 1 (2012). Report number ESSPIN 060.

ESSPIN (2013b) Kaduna State report of ESSPIN Composite Survey 1 (2012).

ESSPIN (2013c) Extension of the Education Sector Support Programme in Nigeria, August 2014–January 2017. [Business case for DFID].

ESSPIN (2015) Kaduna State report of ESSPIN Composite Survey 2 (2014).

Megill, D.J. (2014) Final Sample Design and Weighting Procedures for the ESSPIN Second Composite Survey (CS2) in Nigeria. Unpublished note written for ESSPIN and Oxford Policy Management.

RTI International (2014) 'Nigeria reading and access research and activity (RARA): adaptation of Education Sector Support Programme in Nigeria's teacher capacity development strategy'. Document produced for the United States Agency on International Development (USAID).

Sanni, K, 2015, Taking School Improvement to Scale: The Education Sector Support Programme in Nigeria, Cambridge Education.

Annex A   School characteristics

The table below sets out summary statistics for Kaduna's schools, split by categories according to the level of Output Stream 3 intervention (minimum, medium, maximum). The data come from the Annual School Censuses from 2009/10, 2013/14 and 2014/15.

Kaduna's schools by level of ESSPIN intervention	Total	Min.	Med.	Max.	
Distance from Local Government Authority	28.0	29.9	19.9	18.5	*
Age of the school in 2014	23.4	22.4	27.9	28.1	*
Urban (%)	9.8	7.2	24.1	10.5	*
Nomadic (%)	5.6	5.8	3.9	2.5	*
Islamic (%)	1.0	1.3	0	0	
Double shift (%)	2.6	1.5	6.3	2.6	*
Had parent–teacher association in 2014/15 (%)	97.7	97.5	98.5	98.7	
Had SBMC in 2014/15 (%)	96.4	95.8	98.6	99.4	*
PTR in 2009/10	36.0	37.2	32.9	22.9	*
PTR in 2013/14	46.5	48.5	38.6	35.5	*
PTR in 2014/15	49.4	51.7	39.5	38.3	*
% change in PTR between 2009/10 and 2013/14	79.1	80.6	63.4	106.7	
% change in PTR between 2013/14 and 2014/15	14.5	15.6	9.8	10.6	*
Number of classrooms in 2014/15	5.0	4.6	6.7	7.3	*
Number of teachers in 2014/15	8.2	6.7	14.4	14.5	*
Primary enrolment in 2009/10	248.5	224.9	365.4	285.2	*
Primary enrolment in 2013/14	279.2	250.2	410.3	395.8	*
Primary enrolment in 2014/15	285.7	256.8	415.3	395.8	*
% change in enrolment 2009/10–2014/15 (%)	46.9	46.2	41.0	83.8	*
% change in enrolment 2013/14–2014/15	9.5	9.8	8.3	7.6	
% of teachers with academic diploma/degree	60.4	60.2	60.4	64.9	*
% of teachers with PGDE, BEd or MEd	4.7	4.7	4.6	4.8	*
% of teachers with NCE, Grade II or equivalent	80.9	81.4	79.1	80.0	
School has a power source (grid/other)	4.8	4.2	7.1	8.0	*
% of classrooms with enough seating	25.9	24.4	30.6	29.2	*
% of classrooms with a good blackboard	48.5	46.7	53.6	55.7	*
% of classrooms in good condition/minor repairs	59.8	57.5	66.4	68.2	*
School has at least one toilet (%)	75.5	78.6	68.6	38.7	*
Number of schools	4,125	3,361	601	163	

Notes: (1) * indicates a significant coefficient when running a linear or logistic regression of the variable of interest (dependent variable) on the number of years of ESSPIN intervention (independent variable); (2) the 'total' column includes schools that do not have an intervention code; (3) the PTRs shown in the table are calculated as the average PTRs for schools in the state ($\frac{\sum P_i/T_i}{N}$) and not the PTR for the state as a whole ($\sum P_i \big/ \sum T_i$).

Annex B   ESSPIN Output Stream 3 interventions

The table below shows the ESSPIN Output Stream 3 interventions delivered to date in Kaduna State. In order to make the variation in interventions across and within states manageable for analysis, each combination of interventions was categorised as minimum, medium, or maximum, according to the number of years of continuous intervention.

Expected impact	Number of schools	2009/10			2010/11			2011/12			CS1	2012/13			2013/14			CS2	2014/15			2015/16			CS3
		L	T	SV	L	T	SV	L	T	SV		L	T	SV	L	T	SV		L	T	SV	L	T	SV	
Minimum (1)	3,068																		8	2	3				
Minimum (2)	293														6	3	9				9				
Medium (1)	170											6	3	9	6	3	9				9				
Medium (2)	431							6	3	9		6	3	9	3		9				9				
Maximum	163	5*	5*	9*	10*	5*	9*	6	3	9		6	3	9	3		9				9				

Note: L = days of leadership training; T = days of teaching training; SV = school visits; * = pilot.

Annex C   ESSPIN Output Stream 4 Interventions

The table below shows the days of Output Stream 4 intervention in Kaduna under different headings: SBMC training; women and children participation training; and mentoring visits.

Level of Output Stream 4 intervention	Number of schools	2010/11			2011/12			CS1	2012/13			2013/14			CS2	2014/15			2015/16			CS3
		S	P	M	S	P	M		S	P	M	S	P	M		S	P	M	S	P	M	
No intervention	2,583																					
Post-CS1	1,150										7										5	
Pre-CS1	480				7		4		r		4		6	4*		5	2	8	1	2	10	
Pre-CS1	165	7		4	r		4			6	4*			4*		5	2	15	1	2	17	

Note: S = SBMC training; P = women and children participation training; M = mentoring visits; r = one-day refresher; mentoring visits were by civil society–government partnership teams, except those marked with an asterisk, which were by SMOs.

Annex D   Regression results – pupil learning outcomes

Test	Treatment variable	Model	Coefficient	SE	P value	N	R-squared
L2	pu_exposure	Simple model with survey weights	1.87	2.56	0.466	550	0.002959
L2	pu_exposure	No survey weights but clustered SEs	11.56	1.43	0	2836	0.064225
L2	intervention_binary	Binary exposure variable	26.22	9.6	0.007	550	0.031711
L2	pu_exposure	Full covariates, survey weights	-0.27	2.49	0.915	473	0.203299
L2	pu_exposure	Full covariates, no weights	-0.53	1.76	0.764	473	0.200407
L2	intervention_binary	Full covariates	5.98	9.97	0.55	473	0.201308
L2	pu_dexp13	Lagged school-level learning outcomes	-1	3.18	0.754	408	0.194271
L2	pu_dexp13	Lagged outcomes and covariates	-5.01	2.93	0.091	364	0.333148
L4	pu_exposure	Simple model with survey weights	0.62	1.48	0.679	542	0.001044
L4	pu_exposure	No survey weights but clustered SEs	6.78	0.64	0	3202	0.098513
L4	intervention_binary	Binary exposure variable	25.82	8.99	0.005	542	0.030739
L4	pu_exposure	Full covariates, survey weights	0.18	0.9	0.845	465	0.229019
L4	pu_exposure	Full covariates, no weights	0.88	0.77	0.252	465	0.224924
L4	intervention_binary	Full covariates	11.25	7.81	0.152	465	0.225662
L4	pu_dexp13	Lagged school-level learning outcomes	1.24	0.94	0.193	411	0.223858
L4	pu_dexp13	Lagged outcomes and covariates	1.12	0.83	0.177	365	0.322722
N2	pu_exposure	Simple model with survey weights	-1.97	5.82	0.736	546	0.001374
N2	pu_exposure	No survey weights but clustered SEs	11.62	1.47	0	2801	0.057978
N2	intervention_binary	Binary exposure variable	32.96	14.18	0.022	546	0.025853
N2	pu_exposure	Full covariates, survey weights	-6.47	3.65	0.079	469	0.290236
N2	pu_exposure	Full covariates, no weights	-0.07	2.74	0.981	469	0.21976
N2	intervention_binary	Full covariates	1.35	14.66	0.927	469	0.219791
N2	pu_dexp13	Lagged school-level learning outcomes	0.22	5.28	0.966	405	0.181243
N2	pu_dexp13	Lagged outcomes and covariates	-5.86	5.12	0.256	361	0.340064

N4	pu_exposure	Simple model with survey weights	1.46	2.23	0.512	539	0.002911
N4	pu_exposure	No survey weights but clustered SEs	6.9	0.7	0	3177	0.085963
N4	intervention_binary	Binary exposure variable	29.02	11.12	0.01	539	0.0247
N4	pu_exposure	Full covariates, survey weights	0.31	1.47	0.835	456	0.173296
N4	pu_exposure	Full covariates, no weights	0.95	0.98	0.334	456	0.173955
N4	intervention_binary	Full covariates	9.42	11.45	0.412	456	0.173155
N4	pu_dexp13	Lagged school-level learning outcomes	0.32	1.3	0.804	411	0.128582
N4	pu_dexp13	Lagged outcomes and covariates	-0.3	1.33	0.821	357	0.273591