



## Learning Brief

# NECT Early Grade Learning Programme Evaluation

Brahm Fleisch  
January 2021

### Introduction

One of South Africa's flagship education initiatives in the last decade was the National Education Collaborative Trust's (NECT) Early Grade Learning Programme. As part of this important venture and as a vehicle to advance knowledge about how to improve early grade learning across the education system, the Zenex Foundation commissioned a mixed-method impact and process evaluation of this initiative. The evaluation served two interrelated purposes: first, it provided rigorous independent evidence to guide decision-makers about *if*, *how* and *what* to scale-up across the education system to improve early grade learning; second, the multifaceted impact evaluation would add to the body of knowledge on system-wide education improvement.

The intention of this Learning Brief, then, is to summarise selected elements of the evaluation reports, extract the lessons and learnings, and suggest the ways in which the initiative has contributed to advancing our understanding of how to improve policy and programme design, to improve early learning outcomes at-scale. Given the extensive data collection and analysis, it would not be possible to provide a comprehensive summary of all the aspects of the various evaluation reports. As such, this Learning Brief should be read in conjunction with the main report (February 2019).

The Learning Brief is divided into two parts. The first analyses the evaluation reports focusing on better understanding of the intervention models; the fidelity of implementation; and the interventions' impact on both curriculum coverage and learning outcomes. It begins with a summary description of the two interventions and a synopsis of the evaluation design. Close attention is then paid to interventions' initiation and implementation processes. This is followed by an examination of the findings of the interventions' impact, as measured by the amount of written work completed (as a proxy for curriculum coverage) and learning outcomes, on Early Grade Reading and Mathematics Assessments (EGRA and EGMA). The second part of the Brief

locates the NECT interventions and the evaluation reports within the wider context of South African and international literature on educational change.

The NECT Early Grade Learning Programme was a joint initiative of government and the NECT to advance knowledge about how to improve early grade learning outcomes at-scale across multiple and diverse contexts in South Africa. The Programme centred on testing two intervention models in six districts located in five different provinces. The various phases of the Programme began in 2012, with the final submission of the Technical Report in the middle of 2019. The two distinct intervention models that were implemented had in common the objective of improving curriculum coverage as the means to achieving substantial gains in early grade learning in reading (both in children's Home Language and English as a First Additional Language) and mathematics.

Both interventions used a structured pedagogy programme approach. UNICEF has defined this approach as orientated towards:

. . . a systemic change in education content and methods, delivered through comprehensive, coordinated programmes that focus on teaching and learning, with the objective of changing classroom practices to ensure that every child learns. (Chakera et al 2020)

Specifically, these structured pedagogy programmes included the use of detailed daily lesson plans and related curriculum trackers; the provision of quality teacher and learner materials linked to the lesson plans; and various kinds of training and ongoing teacher and school support. (For a rationale for the South African version of the approach see Fleisch, 2018.)

The two interventions, however, differed in both the scale at which they worked and their approach to supporting teachers and schools. The Whole District Model (WDM), implemented in Kwazulu-Natal, worked in all the primary schools in the district and relied on the district managers, subject advisors and heads of department in schools to carry the load of training teachers and ongoing monitoring and support. The Focused-District Model (FDM), formerly known as the Fresh Start Schools, was designed to work directly with a small group of schools and teachers within a district and provided training and support directly to teachers. The WDM was implemented in the King Cetshwayo District of the Kwazulu Department of Education. The FDM was implemented in six districts in four provinces: h Bohlabela (Mpumalanga), Mount Frere and Libode (Eastern Cape), Vhembe and Waterberg (Limpopo) and Bojanala (North West). Mount Frere and Bojanala were selected to be part of the FDM evaluation.

## Evaluation Design

The NECT Early Grade Learning Programme Evaluation was driven by three essential questions. First, to what extent and in what ways did the two implemented programmes change practices at the district, school department and classroom level? Second, what impact did they have on curriculum coverage? Third, what was their impact on learner outcomes? The evaluators adopted a complementary mixed-method approach, anchored in a quasi-experimental

comparison of learner academic performance in intervention and matching school groups. The evaluators selected three of the seven intervention districts and matched them (and schools within them) with schools in neighbouring districts (within the same provinces) to serve as counterfactuals. In addition to baseline (beginning of Grade 1) and endline testing (end of Grade 2), the evaluation team interviewed teachers, district officials and conducted 12 in-depth school case studies. The main instruments used in the learner testing were modified Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA), for Grades 1 and 2. To measure curriculum coverage, the evaluators used data generated from an analysis of the number of pages and exercises learners completed by the end of Term 2 in Grade 2, in the DBE workbooks and exercise books. Twelve case studies were also conducted.

Table 1: Summary of Key Elements of the Evaluation

Research Questions	<ol style="list-style-type: none"> <li>1. To what extent and in what ways did the two implemented programmes change practices at the district, school department and classroom level?</li> <li>2. What was the impact on curriculum coverage?</li> <li>3. What was the impact on learner outcomes?</li> </ol>
Approach	Complementary mixed-method approach anchored in a quasi-experimental comparison of learner academic performance in intervention and matching school groups.
Instruments	<ol style="list-style-type: none"> <li>1. Curriculum coverage -- number of pages and exercises learners completed by the end of Term 2 in Grade 2, in the DBE workbooks and exercise books.</li> <li>2. Learner outcomes -- modified Early Grade Reading Assessment (EGRA) and Early Grade Mathematics Assessment (EGMA) for Grades 1 and 2.</li> </ol>
Sample	Purposive selection of three of the seven intervention districts#. Matched these districts (and schools within them) with schools in neighbouring districts (within the same provinces) to serve as counterfactuals. 150 schools, 75 intervention and 75 control

The evaluators acknowledged a number of unavoidable design weaknesses. First, the evaluation was commissioned after the NECT Early Grade Learning Programme had been operational in the relevant districts for a number of years prior to the evaluation commencing. This made it difficult to understand the evolution of the designs and how critical design changes had been made during

the formative period of the interventions. Second, because districts and schools had already been selected, it was also not possible to randomly assign schools to treatment and control groups: the “gold-standard” design, if the aim is to establish an estimate of the causal effect size of the intervention model. To establish counterfactuals, the evaluation team used a mixed matching selection process to select comparative school groupings. Given that both interventions focused on the district level (one more than the other), the control schools had to be chosen from those in neighbouring districts to avoid contamination. As such, there remains uncertainty about the exact comparability of intervention and comparison schools.

## Initiation, implementation and impact on schools

Education improvement projects can be tracked through three broad phases: the initiation, implementation and impact/institutionalisation (Fullan, 2007). The ignition or initiation phase explores the conditions that give rise to an initiative and the ways it gets introduced to participants and stakeholders. The implementation phase explores how the project or initiative gets delivered; how closely the intervention plan translates into resources and capacity building on the ground (fidelity;) and how these get mobilised by teachers in actual classroom practices. The final phase focuses on the impact that the intervention has on key outcomes and the extent to which impactful practices get internalised and/or institutionalised.

Figure 1: Phases in the change process



Note: This has been adapted from Fullan (2007) framework. In the original formulation the final stage is institutionalisation.

### Initiation

As the evaluation began substantially after the initiation phase, the researcher relied on interviews with key project stakeholders. The interviews with teachers and district officials, in particular, revealed that both programmes achieved considerable stakeholder buy-in, or acceptance, in the initiation phase. Interviewees consistently expressed positive sentiments. In terms of programme activities, the success of the buy-in activities allowed for the successful implementation in line with the intended plan, with high levels of attendance at training sessions and widespread use of the core curriculum management tools.

Although both programmes managed the initiation phase successfully, innovations with the WDM warrants a full description. In relation to the WDM, the external agency began the buy-in

process with a formal district agreement, before triggering the start of the work. This was followed by an on-boarding workshop, with both schools and districts, to understand the current situation of teaching and learning and what the proposed intervention would do to strengthen current practices. This led to an agreement of a district charter, outlining principles on which the intervention was to be based. The stakeholders all agreed. Within the district office, the outside intervention team insisted on the election of intervention champions responsible for leading the intervention in the district as a whole and within the support team specifically. Finally, there were quarterly leadership sessions, in which the leadership of the district reviewed and planned for the intervention. The centrality of getting the district office to own the project was also reflected in the provision of district coaches, who helped facilitate the active role of the district office, not only at the outset, but also helping to build and sustain capacity throughout the duration of the intervention and beyond. In addition to these activities, the outside agency helped establish a district steering committee designed to have oversight of the entire process. This particular structure, however, never took hold after the first year of the intervention.

## Implementation

### Lesson Plans and Trackers

The cornerstone around which both the intervention models were constructed was the provision of a set of detailed daily lesson plans for teachers. The evaluators found that these lesson plans, provided in both the WDM and the FDM, were widely used and largely seen positively by teachers. The interviews revealed that teachers believed that they saved them time, provided structure to the curriculum and established appropriate daily routines.

The evaluators' observational data, however, suggested that the actual enactment of the lesson plans in classrooms was uneven. Most of the teachers did not complete all the lesson plans required for the term. Within the daily lessons themselves, few teachers managed to do all activities required in the plan for the day. The evaluators also found that, even with the activities that teachers did use in their lessons, the enactment was often superficial.

Notwithstanding these challenges and limitations, an analysis of the comparison schools suggests that, without lesson plans, teachers tended to teach only those activities and topics that they felt confident doing and/or thought were important, rather than covering the full range of content and mix of methodologies required by CAPS.

Like the lesson plans, the evaluators found that the curriculum trackers were widely used in intervention schools. This finding emanated from the interviews with teachers, case studies and the analysis of the tracker instruments themselves. While widely adopted, the evidence suggests that implementation began to drop off towards the end of the term. By the 10<sup>th</sup> week of Term 2, for example, only half of the HoDs were still signing the tracker tool. Underlying the falloff in the use of the trackers towards the end of the term were the difficulties teachers faced completing the packed national curriculum.

The evaluators also showed that some engagement with the tracker tools could be best described as “compliance-orientated”, rather than their intended purpose as a vehicle for genuine professional conversations about how to address challenges in curriculum coverage.

Case study data also found that in most cases, the completion of the trackers was largely superficial, with the same basic reflections often being parroted by teachers week after week. (Technical Report, p. 36)

Notwithstanding drop off in use and the compliance orientation, the trackers and the conversation around them appeared to get teachers, school management and district staff to really talk about curriculum scope and sequence.<sup>1</sup>

### Training

Teacher training, either directly with teachers, or through a cascade model, was also one of the centrepieces of both intervention models. Attendance at planned training was very good in all three districts and teachers unanimously viewed the training they received very positively. In both models, the training programmes evolved into a substantial collaboration between the outside curriculum specialists and district subject advisors.

The core of the WDM training component was cascade just-in-time training for school Heads of Department. In practice, many of these sessions attracted ordinary teachers, rather than Heads. As such, the training programmes were adapted to address this reality. What stood out about this training was the extent to which it contributed to building trust between teachers, SMTs and the subject advisors. That said, the baseline evaluation report found that the actual amount of time teachers spent in the training sessions was substantially less than the time allocated and that the training methodology was less than effective in getting teachers to adopt new instructional approaches.

Given the emphasis on curriculum management at both school and district levels, a major part of the training programme was designed for school management, teams rather than teachers. The SMT training programme had nine modules designed to help building management capacity on a range of topics, such as team building. The focus, however, of SMT training was on curriculum monitoring and supervision.

In contrast to cascade training, FDM provided subject-specific training directly to teachers, who were trained in groups of between 30-50, which increased to groups of between 50-100 in 2018. Training was conducted by the external service provider’s curriculum specialists and subject advisors. Not unlike the training that occurred in the WDM, the actual amount of time in training in the FDM was substantially less than intended time for training. As such, in the FDM, teachers would have received approximately 2.5 hours of training per quarter for the three subjects: the Home Language, EFAL and mathematics. The evaluators also observed methodological

---

<sup>1</sup> “Scope refers to the breadth and depth of content and skills to be covered. Sequence refers to how these skills and content are ordered and presented to learners over time.” <http://www.ibe.unesco.org/en/glossary-curriculum-terminology/s/scope-and-sequence-curriculum> accessed 25 February 2021.

weaknesses within the training approach, characterising it as a trainer reading through the training script with occasional comments and explanations.

#### Onsite classroom/school support

The other major component of both models was the onsite classroom/school support. Like the differences in training, the approach to support was substantially different in the two models.

In the WDM, school coaches and SMT coaching was primarily aimed at supervisory level school staff. These coaches provided training workshops and individual coaching to SMTs. In terms of dosage, each school coach was assigned to approximately 20 schools, with visits twice per term. In practice, this was often done not with individual schools, but rather school coaches worked with clusters of schools together. In line with the strong emphasis on embedding the curriculum management model, the WDM also included coaching with district staff. This will be discussed below.

The FDM's coaching programme changed over time, with a strong emphasis on onsite teacher coaching in the initial Fresh Start School phase. In the early phase, coaches were assigned to 5-8 schools and teachers received between 2 to 4 visits per term. This was discontinued when the programme was rebranded as the Focus District Model.

Onsite support to teachers from district office was seen as an important component of both programmes in the second year of implementation. That said, the low level of staffing made this very difficult. The ratio of schools to subject advisor was very high, with 135 schools per subject advisor in Mount Frere; around 100 schools to subject advisor in the Bojanal, and about 150 schools per subject advisor in King Cetshwayo. Making matters worse, subject advisors often had many other responsibilities, including training, district meetings and checking promotion schedules.

Within the WDM, much of the work of the intervention, beside the provision of the lesson plans and trackers and train-the-trainers, was focused on upgrading the skill level of the district office. This involved training the subject advisors and curriculum management capacity-building across all functions. Core to this was the creation and use of the Curriculum Management Information Tool, the intention of which was to help track curriculum coverage and provide a signaling device to diagnose weak and strong school implementation. The information tool did take hold in the district to document coverage, but it was never really used to differentiate support to schools. That said, evaluators observed that the dashboard; the data collection needed to populate it; and the ranking of schools was useful in building a district focus on evidence-based planning.

Although there was little charismatic leadership within the district, after the initial flurry of excitement, the capacity building activities and the new way of thinking seemed to have endured after the external service provider's role was scaled down: a clear indication of the institutionalisation of new technical vocabulary and new practical tools within the district office.

## Impact

### Curriculum Coverage

Possibly one of the most positive findings in both models was that curriculum coverage was much higher in intervention schools, relative to the comparison schools. On average, learners in intervention schools significantly completed more pages in their DBE workbooks and exercise books than learners in comparison schools. Furthermore, the evaluators suggest that there is evidence that learners in the intervention schools were doing more cognitively complex tasks in their exercise books.

To assess curriculum coverage, the evaluators analysed the amount and quality of work learners did in both their exercise books and in the DBE workbooks, up to the end of Term 2 in 2018. They looked at both the number of pages completed, the number of exercises and the number of exercises per topic.

Table 2: Average number of pages completed

	KC	Ilembe	Bojanala	KK	Mt Frere	Mbizana
EFAL pages in exercise books	<b>24</b>	18	<b>33</b>	17	<b>27</b>	14
EFAL pages DBE workbooks	<b>43</b>	38	<b>41</b>	45	<b>50</b>	35
Maths pages in exercise books	<b>59</b>	55	<b>51</b>	29	<b>40</b>	39
Maths pages in DBE workbooks	<b>79</b>	58	<b>77</b>	62	<b>61</b>	65
<b>Total pages</b>	<b>205</b>	169	<b>202</b>	153	<b>178</b>	153

Source: Technical Report p. 75

Learners in King Cetshwayo intervention schools completed significantly more pages in both their English and mathematics exercise books and in the DBE workbooks, compared to learners in comparison schools in Ilembe. Most striking is the substantial additional work in the English exercise books and Mathematics DBE workbooks. Learners in Bojanala completed almost double the number of pages in both their English and DBE workbooks. Learners in Mt Frere seemed to have been concentrating on English, as they did almost 60% more pages in English overall, but almost exactly the same as the comparison group in mathematics. They showed less reliance on the DBE workbook and did more work in the exercise books. This is likely to be a function of the lesson plans and is a good indicator of appropriate sequencing. In a related analysis, the evaluators found that, in all three intervention district schools, the English exercise books contained more complex phrases per sentences (all statistically significant) relative to the work that was being done in the control schools. This was also a real change and not simply a substitute effect, i.e. that intervention teachers concentrated on targeted subjects, to the detriment of non-priority subjects. That said, the evaluators do note that this average treatment effect was likely to have been driven by the top performing intervention schools.

### Learning Outcomes

The evaluators used versions of the EGRA (Home Language and English FAL) and EGMA instruments with a large sample of 2626 learners in 150 schools (75 intervention and 75

comparison schools), with the baseline at beginning of the Grade 1 school year in 2017 and the endline at the end of the Grade 2 year in 2018.

There was considerable interprovincial variability. Although there were positive gains in intervention schools, in many sub-tests in literacy and some in mathematics, in KwaZulu-Natal, only three sub-tests were statistically significant (e.g. HL letter recognition), regardless of which controls the evaluators used. In the Eastern Cape and North West province, the evaluators found no statistically significant difference between learner outcomes, when comparing intervention and control schools.

The evaluators offered a number of explanations for why learners in the intervention schools did not seem to do substantially better, in the various sub-tests, than learners in the control schools. They pointed to resource constraints, such as the shortage of district office staff and Head of Department at schools with heavy workloads. They identified curriculum overload, i.e. CAPS, which formed the basis for the lesson plans and trackers, as too academically demanding for the majority of learners. They even suggested that improved curriculum coverage could have an unanticipated negative effect on teaching and learning, as teachers speed up material to be covered and focus less on supporting learning at the level of the learners. The evaluators postulated that substantial learner backlogs, even in Grades 1 and 2, may also account for the lack of substantial impact. That is, by Grade 2, learners were already behind and therefore could not benefit from a curriculum programme already pitched at a level far too high for them.

Overall, on the key finding about real gains in curriculum coverage, but limited corresponding gains in learning outcomes, the evaluators offer an interpretation that curriculum coverage is a necessary, but not sufficient condition, for improving learning outcomes. The focus of both interventions was on improving teachers' and learners' curriculum coverage (which did take place), but this did not automatically lead to improved reading and mathematics outcomes.

On the overall district results, the evaluators indicated that, for two out of the three provincial sites, the results were inconclusive. For the Eastern Cape they note:

...it is thus inconclusive as to whether the programme itself causally improved learning outcomes for students in the treatment district (p. 96)

and for the North West it was observed:

A reliable estimate of difference between treatment and comparison is not possible. (p. 96).

The evaluators offered three reasons for the inconclusive findings. First, they found evidence of contamination. Experimental contamination occurs when decisions are taken that would mean schools, who were designated in the control group, receive features of the intervention. In 2018, there were efforts to implement the NECT programme in the comparison districts. As such, the absence of a difference between the intervention school outcomes and that of the comparison

schools may be the function not of weaknesses in the intervention models, but that both school groups benefitted to some degree. Second, and this was specific to the Eastern Cape, the evaluators found that there is a substantial language difference between Mt Frere and Mbizana. Most children in the Mt Frere district speak local dialects, while children in the Mbizana speak a version of the language that is very close to standard isiXhosa. If this is the case, slow progress in the intervention may, in part, be a function of a mismatch between children's everyday language and the formal language of the school learning.<sup>2</sup> Finally, the evaluators' geographic analysis suggests (but does not prove definitively) that the comparison district schools had different levels of rurality or proximity to the urban core. This geographic phenomenon would potentially have negatively impacted intervention schools in both the Eastern Cape and the North West.

## Lessons to be learnt from the evaluations

From the perspective of policy-makers, the single strongest lesson from the Early Grade Learning Programme Evaluation is that the results are simply not clear enough to make definitive decisions for or against the Whole District and the Focused District models. If improved learning outcomes is the outcome according to which effectiveness is gauged, the evaluation results do not permit an evidentiary warrant. An evidentiary policy warrant is justification based on credible and trustworthy evidence. In other words, it is not that the two programme models do not work, but rather researchers do not have sufficient credible evidence to make the main claim that they are either effective or ineffective.

The second lesson relates to the use of curriculum tools. Most teachers used the lesson plans and the trackers – and evidence was offered to suggest that many teachers will likely continue to use them, as long as they continue to be provided. Although teachers generally thought of CAPS as an overly complex and overcrowded curriculum, if they were to use the lesson plans, it is likely that curriculum coverage would improve, even if only at a superficial level. From a management perspective, at the level of the SMT and the district office, the curriculum management tools made all staff much more aware of the importance of curriculum scope and sequence in teaching. At a school level, specifically, the case studies showed that some schools were thinking creatively about ways to increase time available to ensure more comprehensive curriculum coverage.

The third lesson relates to the importance of buy-in or acceptance. One of the critical weaknesses of many structured pedagogic models is that they tend to bypass the provinces and the district offices, in particular and they pay insufficient attention to school-based management. WDM intervention showed that the implementers could get substantial buy-in/acceptance using a variety of structures and processes.

---

<sup>2</sup> This is inconsistent with the Eastern Cape baseline data that suggested that the matching sites were comparable.

The fourth lesson relates to the power of simple technical terminology. WDM helped district officials adopt a new specialised vocabulary that shifts conversations to core aspects of instruction. These shifts helped change the dynamic relationship between district offices and schools. This new dynamic endured beyond the life of the specific intervention.

The fifth lesson points to the need for minimum standards for professional staffing of district offices and the need to revisit the job descriptions of Head of Department in schools. If there are too few district officials, with insufficient resources to visit schools and they do not have the capacity to prioritise based on real evidence of how well schools are actually performing, shifts in classroom practices are unlikely to be sustained. Similarly, if Heads of Departments are not promoted because of their expertise and experience, they lack a sophisticated understanding of the new practices and they do not have the time in the school day and year calendar, then their potential to add value would be limited.

## Learnings from comparison

What potentially valuable learnings surface when we locate NECT Early Grade Learning Programme and its evaluation within the South African and international research? Although the focus of much of this literature is on “what works”, as measured by standard deviations of improvement in learning outcomes, a whole range of other important insights are beginning to surface, all of which are germane to understand the challenges of system-wide educational change better. In particular, the Early Grade Learning Programme makes a valuable contribution to the knowledge base about problems of scaling up (moving from mid-size trials to whole province interventions), institutionalisation (how to ensure that the new practices are sustained after the interventions come to an end) and phasing in the change journey (understanding best practice for each phase in a multi-stage improvement process).

### Intervention Types and Structured Pedagogic Programmes

The first insight is that the broad contours of the design of the two NECT intervention models were consistent with what is emerging as best practice for education systems on the journey from poor to fair (Mourshed et al 2010): combined component structured learning, or structured pedagogy approaches. There is a growing consensus in the field that standalone intervention types (whether in the form of the provision of learning materials, capacity building/teacher training, financial incentives or accountability measures) are far less likely to be cost-effective than combined intervention models directed at the instructional core (McEwan, 2015). Snilstveit et al’s (2016) influential 3ie systematic review specifically pointed to the cross-country evidence of the effectiveness of structured pedagogy models when the goal is improved learning outcomes. Piper et al (2018) has provided a comprehensive cost-effectiveness analysis of combined ingredient structured models, including those with lesson plan guidance.

Why do combined component structured programmes work for systems on the journey from poor to fair? Kim and Anderson (2019) suggest that intervention models, that use a structured

learning approach, work, because they: (1) help teachers maximise instructional time; (2) enable teachers to make their instruction more explicit and systematic; (3) assist teachers to establish instructional routines; (4) provide better instructional scaffolding in the learning process; and (5) foster social and emotional engagement. Kim and Anderson's work provides insight not only that that structured learning programmes work, but, more importantly, it uncovers the instructional mechanisms behind the *how* and *why* they work, particularly in resource-constrained contexts like South Africa.

There is also an emerging understanding of both what the key components that need to go into these structured programme models are (Fleisch, 2018 and Piper, 2019) and the optimal dosages for each component. The core components or ingredients include: lesson plans and/or structured teacher's guides; quality learning materials at the right level; teacher training and classroom coaching (Nayak & Upreti, 2020; Kim & Davidson 2019; Chakera et al 2020).

RTI has recently published best practice guidelines for both structured teachers' guides (Piper, 2018) and for centralised training (Author, in press). These guidelines provide detailed information about what quantum have shown to works best for each ingredient. Two examples illustrate this. Zuilkowski et al's (2017) study provides useful guidelines into the cost-effectiveness of different teacher-coach ratios. Piper (2018) suggests that lesson plans/teacher guides should not contain more than a single page of text and illustrations per lesson.

In South African contexts, researchers are making a valuable contribution to the knowledge base about multi-component structured pedagogic programme models designed to improve early grade reading and mathematics. In the Eastern Cape, the Magic Classroom Collective (Ramadiro, 2018) has used a combined structured programme model that included a toolkit with workbooks, a phonics programme, a teacher storybook programme and an independent reading programme, combined with teacher guides, training and coaching. Ramadiro and Porteus' model has shown strong gains in learning outcomes, as measured in the Systemic Evaluations. The Funda Wandu, initiative, at the earliest stages of its impact evaluation (Ardington, 2019), is showing statistically significant learning gains, using a model that includes an accredited literacy course combined with structured teacher/learner materials and on-going onsite coaching. The Room-to-Read evaluation of the Sepedi reading programme, completed in 2016, showed substantial gains in early reading subtest items in Limpopo, using a multi-component structured pedagogy model. The Room-to-Read model included detailed lesson plans; classroom materials, including independent reading books, comprehensive teacher training; and ongoing monitoring and coaching of teachers (Room-to-Read, 2017). The Early Grade Reading Study I (2015-2017) in Setswana showed that the combined model, that included lesson plans, reading materials, centralised training and onsite coaching, is more cost effective than a similar model, that excluded coaching, and a model that focused on parent support (Cilliers et al, 2019). The Early Grade Reading Study II confirmed the findings of the basic model tested in the EGRS I study, but focused on English First Additional Language (EFAL) in the early grades. The Mpumalanga study demonstrated that statistically significant gains can be achieved in EFAL, using the standard combination of lesson plans, learner materials, training and onsite coaching (Cilliers, 2020).

Together with the Early Grade Learning Programme Evaluation, these various impact evaluation studies confirm that:

- (1) Improved early grade learning outcomes at scale is achievable and feasible.
- (2) Models to achieve these outcomes required combined component/ingredient interventions.
- (3) The interventions need to use structured pedagogic approaches.

**If implemented at meso-level scale (50 schools or so), if the intervention has all the key components and these are delivered using the right dosage, predictable learning gains can be expected, even with different service providers and in different contexts.**

While the evidence for the basic model is clear, consistent and convincing, a number of implementation challenges still remain. The Early Grade Learning Programme and the other early grade interventions provide useful sites to better understand how to overcome these. For example, the NECT Evaluation has illustrated the steps and processes that are needed to achieve “real” district buy-in, a precondition to helping interventions transition from the status of external projects to institutionalised practices. Both WDM and the FDM have shown the value of a spotlight targeted on curriculum coverage and the centrality of strengthening curriculum management at the provincial, district and SMT levels. The Room-to-Read intervention has illustrated how to incorporate independent reading into a structured programme. The Magic Classroom Collective has developed a new approach to teaching phonics in Nguni languages. The Funda Wande has shifted thinking about both the need for more rigorous accredited training and how to consolidate and print low cost reading materials. With robust evidence and the willingness to move away from protecting intellectual property rights, as a R&D community, we can learn to share powerful knowledge and continue to build the knowledge-base about system-wide education reform.

### **Scaling-up, Institutionalisation and Phasing**

Although the focus is on the basic model, little is known about three key additional challenges to improving learning outcomes system-wide. These challenges relate to scaling up, institutionalisation and phases of change. The Early Grade Learning Programme has made an important contribution to a better understanding of these challenges.

The first and most important challenge is taking an evidence-based model from the level of a trial and rolling it out at a large-scale, or system-wide. In most instances, the current wave of randomised trials have been undertaken at a meso scale, not very small, but seldom at the macro scale or system-wide. The dynamics, requirements and costing of an intervention at a macro scale are often very different from those at micro or meso levels. Even if the key cost drivers remain the same, questions of the availability of a scarce resource, such as quality coaches, can fundamentally undermine even the most robust basic change model when it goes to scale.

One of the biggest challenges in system change is the political demand for expansion of programmes without the requisite financial resources, to ensure that intervention models do not

get diluted. Under-dosing can change an effective model into an ineffective model. It can also have potentially negative consequences, as schools and teachers become demoralised and cynical, as they are expected to participate in watered down interventions that are unlikely to be impactful.

The second issue relates to how interventions get institutionalised, or bedded down, within the school, the district and the system as a whole. This is a question Piper and colleagues (2016) have grappled with, as they moved from a series of experimental trials to implementing the Tusome programme in over 24000 schools in Kenya. In the Kenya case, setting learning benchmarks and shifting national expectations of what children should be able to read at the end of Grade 2 was critical to institutionalising the new practices. The other key to institutionalisation in the Kenya case was building accountability systems that involved regular classroom visits and meaningful feedback to teachers. The focus on curriculum coverage and the tools and benchmarks to monitor it provides a new way of thinking about how to raise teaching expectations effectively system-wide and forms the basis for building an accountability system and providing feedback to teachers.

Another key element to scaling up relates to achieving system-level buy-in. The NECT Programme demonstrated considerable innovation in this regard, by working at the provincial, district and school management levels. It also recognised that buy-in is not a once-off event, but needs continuous renewal and that buy-in is best achieved by working through “champions” within each layer of the system.

Finally, it is important to recognise that the basic intervention models are, at best, designed to ignite the education improvement journey, but are unlikely to continue to benefit teachers and classrooms over the long term. This is a key insight or learning, stressed by Mourshed et al (2010) and Hopkins (2016). We need to begin the research programme now that will develop models that will help schools and teachers reach higher levels of learning, once the improvement process gets underway. As the education systems and the professionals within them come to trust that improvement is achievable, and that teachers can institutionalise practices that work better, different components and different approaches will be required in the next phase of the change journey. It is likely that lesson plans and external onsite coaching will not be needed, but genuine communities of practice, better assessment tools and resources, like classroom libraries, may be what is needed most.

### Agility

What we have learnt is that the management and governance of interventions appear to work if led by agile organisations. The use of non-government agencies, working with and alongside the various government structures work, because of agile leadership, decision making and flexible structures. These organisations are better at communicating, mobilising commitment of a wide cross-section of stakeholders and bringing them together to collaborate. Specifically, the two agencies responsible for leading and managing the NECT Programme were involved in developing, reflecting, learning, inspiring, engaging, unifying, empowering and innovating.

## Evaluation

One of the critical weaknesses of the Early Grade Learning Programme was the failure to incorporate an external process and impact evaluation from the outset. These kinds of initiatives must be seen as part of the knowledge building process. Without building evaluation into the design, the credibility of the findings can be prejudiced. Along with upfront evaluations, built into the cost of the interventions, the education sector needs to use common and standardised measurements of learning outcomes to enhance comparability.

## Conclusion

What does the NECT Learning Programme Evaluation tell us? It was a large scale programme designed to improve early grade learning outcomes in Home Language, English as a second language and Mathematics. It was designed as a structured pedagogic programme, with a series of structured pedagogic tools, such as the lesson plans and trackers, at the core. In its simplest version, it had a clear theory of change: the provision of materials, structured tools and support would lead to greater curriculum coverage, which would be a catalyst for more and better teaching of the core content of the curriculum. More and better teaching would translate into learning gains and improvement of learning outcomes.

The big challenge that both interventions encountered was on the capacity building side. There were two core parts of the capacity building: the direct training on the structured programmes and the *in loco* support of the new practices associated with the lesson plans. The evidence for the baseline evaluation highlighted weaknesses in both.

With regards to direct training, the amount of actual training time was substantially less than the intended time. The training that was delivered was ineffective, either because it relied on a problematic OBE training model, that assumed that the knowledge already existed in the teachers and that discussions and sharing would bring it out, or, in the case of SSP/FDM, it was just about going over the training manual. In other words, the basic training on the core structured programme was neither long enough, nor provided teachers with an opportunity to really practise the new teaching practices embedded in the structured programme.

Second, the core onsite support and support to teachers on the new practices in the classrooms was problematic. The assumption in the WDM was that the circuit managers, subject advisors and Heads of Department could and would do the onsite support. The Evaluation revealed that there were simply too few subject advisors and circuit managers to play any meaningful role in classrooms. The school Heads of Department, although employed in all schools, had full teaching loads and had little experience or confidence in clinical classroom supervision.

What changed as a result of the interventions? Both interventions inserted a new technical vocabulary and technical tools into schools, either directly through the school coaches, or via the district office staff. Most teachers reported using the lesson plans and it is clear that the curriculum trackers were also in use. For some teachers, the lesson plans, educational materials

and the curriculum management tools had a substantive and lasting impact on their practices and, by extension, on learning outcomes. This is likely to have been only for a minority of teachers. Evidence suggests that, for a significant portion of teachers, the change in curriculum coverage was superficial.

Although the evidence from the test results is inconclusive, knowledge that the interventions and the evaluation has contributed is nonetheless useful. Two specific learnings need to be highlighted. First, both intervention models assumed that the interventions would need to be scalable and that providing intensive training and one-on-one external coaching would have been unaffordable. To compensate for this gap, they prioritised buy-in and capacity building of the provincial leadership, district office staff and school management (both the principal and the Heads of Department). Core to the capacity building was the introduction of a new technical vocabulary and a new set of management tools around the concept of curriculum coverage. Both the new technical vocabulary and the tools associated with them were embraced and shifted the ways in which both the district and school management staff worked. There is evidence that teachers engaged with the new concept of curriculum coverage. For some teachers, it had a powerful effect on their work in the classrooms.

The other contribution that needs to be highlighted relates to the institutional configuration of change. In the past, NGOs with weak relationships with provincial governments undertook change on the margins of schools. They often got little more than symbolic support for these interventions, from both district offices and the provincial head office. Too often, NGO interventions were frequently seen as a welcome substitute for overburdened district offices. When it worked well, the Early Grade Learning Programme was able to play the role of a network organisation, that directly and deliberately linked all the key stakeholder together and coordinated their contributions to enhance the capacity of the state.

## References

- Arlington, C and Meiring, T (2020) Midline 1: Impact Evaluation of the Funda Wande Coaching Intervention Midline Findings.
- Author (in press) Effective teacher training for foundational literacy and numeracy: Results of a 17-country teacher training study in low and middle-income countries
- Chakera, S., Haffner, D., Harrop, E., (2020) UNICEF Eastern and Southern Africa Region Working Paper – Structured Pedagogy: For Real-Time Equitable Improvements in Learning Outcomes. UNICEF: Nairobi.
- Cilliers, J., Fleisch, B., Kotze, J., Mohohlwane, N., Taylor, S., & Thulare, T. (2020). Can Virtual Replace In-person Coaching? Experimental Evidence on Teacher Professional Development and Student Learning in South Africa. *Unpublished Working Paper*.
- Cilliers, J., Fleisch, B., Prinsloo, C., & Taylor, S. (2020). How to improve teaching practice? An experimental comparison of centralized training and in-classroom coaching. *Journal of Human Resources*, 55(3), 926-962.
- Cuadros-Rodriguez, L., R. Romero, and J. M. Bosque-Sendra. "The role of the robustness/ruggedness and inertia studies in research and development of analytical processes." *Critical reviews in analytical chemistry* 35.1 (2005): 57-69.
- DNA Economics and Social Surveys (2017) Baseline Report: Evaluation of the NECT Learning Programme. A report to the Zenex Foundation Version 1.2. 13 November 2017.
- Dowd, A. J., Friedlander, E., Jonason, C., Leer, J., Sorensen, L. Z., Guajardo, J., ... & Pisani, L. (2017). Lifewide learning for early reading development. *New Directions for Child and Adolescent Development*, 2017(155), 31-49.
- Fleisch, B. (2018). *The education triple cocktail: System-wide instructional reform in South Africa*. UCT Press/Juta and Company (Pty) Ltd.
- Fleisch, B., Schöer, V., Roberts, G., & Thornton, A. (2016). System-wide improvement of early-grade mathematics: New evidence from the Gauteng Primary Language and Mathematics Strategy. *International Journal of Educational Development*, 49, 157-174.
- Fullan, M. (2007). *The new meaning of educational change* (4th ed.). New York, NY: Teachers College Press
- Hopkins, D (2006) *Every School a Great School: Meeting the Challenge of Large Scale, Long Term Educational Reform*.
- Kim, Y. S. G., Lee, H., & Zuilkowski, S. S. (2020). Impact of Literacy Interventions on Reading Skills in Low-and Middle-Income Countries: A Meta-Analysis. *Child development*, 91(2), 638-660.
- Kim, Y.-S. G., & Davidson, M. (2019). Promoting successful literacy acquisition through structured pedagogy: Global Reading Network Critical Topics Series. Prepared by University Research Co., LLC. (URC) under the Reading within Reach (REACH) initiative for USAID's Building Evidence and Supporting Innovation to Improve Primary Grade Assistance for the Office of Education (E3/ED).
- Kotze, J., Fleisch, B., & Taylor, S. (2019). Alternative forms of early grade instructional coaching: Emerging evidence from field experiments in South Africa. *International Journal of Educational Development*, 66, 203-213.

- McEwan, P. J. (2015). Improving learning in primary schools of developing countries: A meta-analysis of randomized experiments. *Review of Educational Research, 85*(3), 353-394.
- Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better*. McKinsey.
- Nayak, A. & Upreti, p (2020) Demystifying the Science of Teaching: A 'Structured Pedagogy' Approach to Improving Foundational Learning.  
<https://centralsquarefoundation.org/articles/demystifying-the-science-of-teaching-a-structured-pedagogy-approach-to-improving-foundational-learning.html>
- NORC (2019) Story Powered Schools Impact Evaluation. (Powerpoint presentation only)
- Piper, B., Destefano, J., Kinyanjui, E. M., & Ong'ele, S. (2018). Scaling up successfully: Lessons from Kenya's Tusome national literacy program. *Journal of Educational Change, 19*(3), 293-321.
- Piper, B., Sitabkhan, Y., Mejía, J., and Betts, K. (2018). Effectiveness of Teachers' Guides in the Global South: Scripting, Learning Outcomes, and Classroom Utilization. RTI Press Publication No. OP-0053-1805. Research Triangle Park, NC: RTI Press. <https://doi.org/10.3768/rtipress.2018.op.0053.1805>
- Piper, B., Zuilkowski, S. S., Dubeck, M., Jepkemei, E., & King, S. J. (2018). Identifying the essential ingredients to literacy and numeracy improvement: Teacher professional development and coaching, student textbooks, and structured teachers' guides. *World Development, 106*, 324-336.
- Ranadiro, B and Porteus, K (2018) Policy Brief. Early Grade Literacy and Mathematics. Placing the African Language Speaking Child at the Centre. Magic Classroom Collective.
- Room to Read (2017) Sepedi Evaluation Results, South African Literacy Programme, 2016.
- Snilstveit, Birte, et al. "The impact of education programmes on learning and school participation in low-and middle-income countries. 3ie Systematic Review Summary 7." *London: International Review for Impact Evaluation (3ie)* (2016).
- Zenex Foundation (2019) Evaluation of the NECT learning Programme (Foundation Phase).
- Zuilkowski, S. S., & Piper, B. (2017). Instructional coaching in Kenya: Supporting teachers to improve literacy outcomes. In *International handbook of teacher quality and policy* (pp. 505-516). Routledge.