

# Aprender +: Mozambique Coaching Pilot

## Section 1: Teachers' Continuous Professional Development (CPD) in Mozambique

There is a learning crisis in Mozambique, with less than half of 4<sup>th</sup> grade students being able to identify basic words, and less than 1 in 5 reading a paragraph in Portuguese (SDI 2018). Among the factors that are associated with low levels of learning, there is evidence of low instructional dosage (2hr and 5 mins per day), low teacher content knowledge and weak teaching practices in the classroom (Bold et al, 2017; SDI 2018).

In this context investing in strengthening teachers' knowledge and skills is crucial. While governments in the world invest large amounts to continuous professional development, there is a huge variability in teacher training programs effectiveness.<sup>1</sup>

Since independence (1975), the Ministry of Education and Human Development (MINEDH) in Mozambique has run over 15 different models for teacher's initial education and continuous professional development. In 2016, MINEDH approved a new strategy for professional development of primary education teachers and in 2018, a manual with the new strategy was published and started being distributed to schools (Figure 1).

Figure 1 – New Strategy for Teachers Continuous Professional Development



Source: MINEDH (2018)

In this new strategy, two teachers per school cluster - ZIP (*zona de influencia pedagogica*) - are trained in the teachers training institutes (IFP) for three days on the lesson study approach to improving teaching. This approach involves: (i) teachers within the ZIP identify a challenge they are experiencing, (ii) all the teachers teaching similar grades in the ZIP plan the lesson together; (iii) teacher uses the lesson plan to teach his/her students and other teachers within the ZIP observe the lesson; (iv) discuss what work/what did not work and how to improve; (v) a report that summarize the

<sup>1</sup> Popova, Evans and Arancibia – "Training Teachers on the Job", 2018

analysis done and conclusions. Teachers that received the training are then expected to train their peers within the ZIP and introduce this method. The content focus of the program is reading, speaking, writing and basic mathematics.

This new approach, which is currently under implementation, is based on some principles that are aligned with best international practices, namely the need for ongoing support to teachers rather than one-off lectures and use classroom observation and feedback in the process. Furthermore, this new strategy is proposing a substantial change in Mozambique education system in the way teachers develop their skills by recognizing that teaching is a “performance profession”, and as such requires continuous practice “on site” (Lemov et al, 2012).

Despite positive elements of the new teacher professional development model, some aspects can be strengthened by: (i) providing resources to simplify lesson planning; (ii) recognizing that lesson study methodology (reviewing lessons plans among peer teachers) may not be effective for teachers who do not master curriculum content or good pedagogical skills; (iii) providing more structure and intense teacher support ; (iv) recognizing that ZIP coordinators and teachers’ schedules and tasks have not been adjusted so they face time constraints to implement the new model.

*(i) Simplify lesson planning.* In a study commissioned by UNICEF in Zambézia and Tete about the new strategy, researchers noticed a weak lesson planning culture among teachers. For example, they observed that most teachers copy the lesson plan from the student textbooks. This is also highlighted in the latest SDI results, which show that more than 80% of teachers struggle to formulate a lesson plan (SDI 2018). In the current implementation of the model, neither coaches nor teachers received training on lesson planning. The proposal presented in this document includes providing lesson plans aligned with the curriculum (described in more detail in the next section).

*(ii) Lesson study may not be effective in this context if not complemented with further support.* The lesson study methodology, in which the new strategy is based, has been developed and successfully applied in Japan. To implement this methodology, it is critical that teachers master the content of the curriculum they teach, have high level of pedagogical skills and have a shared vocabulary to talk about teaching practices. In addition, teachers need to be well trained on what to observe in the classroom and on how to go from identifying a problem to finding a solution that would work for the observed teacher. Without having experienced coaches to help guide the observation and feedback, teachers implementing this methodology are flying blind, without a structure on how to approach their daily challenges. For example, a teacher may identify that his/her peer has a problem on classroom management during the observation. However, that in and of itself is not enough. In successful education systems coaches learn to identify what specific part of classroom management is the teacher struggling with (i.e. lack of clarity in instructions, lack of systems and routines, problems with redirecting misbehavior). Once the specific problem is identified, good coaches provide targeted strategies and practice with the teacher to make sure the teacher masters the skill and can implement it in his/her classroom thereafter. For instance, the coach may ask the teacher to practice simplifying his/her instructions, so students do not get confused or help the teacher establish routine to minimize transition time between tasks (Bambrick-Santoyo et al, 2016).

*(iii) Structured and Intense Teacher's support.* As exemplified above, successful approaches require the use of deliberate practice, meaning that coaches isolate actions that teachers do in the classroom to strengthen particular skills (i.e. economy of language) and then help them practice it to integrate those actions into their teaching toolkit. This change, while fundamental, requires big effort to implement effectively. Countries that have done this successfully have provided coaches and teachers more structured and intense training. Although there are several variants of teachers’ coaching programs, coaching models that deliver positive results are characterized by Kraft, Blazar and Hogan (2017) as “one where instructional experts work with teachers to discuss teaching practices in a way that is (a) individualized – coaching sessions are one-on-one; (b) intensive- coaches and teachers interact at least every couple of weeks; (c) sustained – teachers receive coaching over an extended period of time; (d) context-specific – teachers are coached on their practices within the context of their own classroom; and (e) focused – coaches work with teachers

to engage in deliberate practice of specific skills.” In addition, successful programs have had a detailed roadmap for coaches on how to implement this approach, together with several resources, including the coaching cycle, skills every teacher needs and how to coach teachers to acquire it, videos of good teaching isolating particular skills, tablets to register coaching session and agreed next steps, etc. This is especially relevant when the pool of coaches and teachers has low content knowledge, weak pedagogical skills and is not experienced in engaging in deliberate practice.

*(iv) Coaches and teachers' schedules.* Finally, ZIP coordinators and teachers within the ZIP don't have time to perform regular tasks at their schools and, also, organize meetings, observe teachers and provide feedback in a continuous manner, as established in the new strategy. The approach entails that several teachers observe one classroom and then meet to write the conclusions, which demands time from several teachers to do just one classroom observation. Thus, the model design needs to consider adjusting the mechanisms of implementation or the persons identified to perform different roles, to ensure that responsibilities are allocated jointly with time and resources needed to undertake them. This should be done, however, using existing structures (ZIPs, in the case of Mozambique).

## Section 2: What can we learn from other experiences in Africa?

This document presents a proposal to design a teacher's continuous professional development pilot, which adds some elements to the current strategy. The proposal is based on successful international experiences within Africa<sup>2</sup> and ongoing experiences in Mozambique.

In contexts like Mozambique's, where most teachers struggle to develop lesson plans, do not master the curriculum and don't have high pedagogical skills, research has found that successful programs tend to have supplementary reading materials and textbooks, focus on specific subjects and grades, and complement training with continuous support through follow-up visits. For example, in Kenya, Piper et al (2018a) analyze the PRIMR program (Kenya Primary Mathematics and Reading Initiative) and evaluate which elements are the most cost-effective to improve student learning. Their model includes the following interventions: (i) only training with ongoing support (coaching), (ii) students' books (1:1 ratio) and coaching, (iii) structured lesson plans, students' books and coaching. Findings showed that structured lesson plans, students' books and coaching not only produced the best results - improving 2.15 standard deviations (SD) and 1.82 SD in fluency in English and Kiswahili, respectively, in grade 2 - but also was the most cost-effective program. Coaching alone only improved outcomes modestly but adding the student books and the structured lesson plans increased the effect substantially. The program is not just a successful professional development program but is credited to be one of the most impactful and cost-effective interventions to improve students' outcomes (Conn, 2017; Popova, Evans, & Arancibia, 2016; Snilstveit et al, 2016). PRIMR was eventually scaled up (named Tusome, *Let's Read* in Kiswahili for the reading program and PRIEDE, the mathematics component) and the evaluation of the results are highly positive (Piper et al, 2018b). Anecdotal evidence suggests that using existing coaches (coaches who were already in the system) may have been critical in the success of the scale up in Kenya, compared to the equally successful pilot in Liberia (Piper and Korda, 2010) that failed to work at scale.

One question that is not answered by the previous study is whether it is better to train teachers in a centralized location (one-off training) or to provide ongoing support to improve student learning. Recently, another study in South Africa (Cilliers et al (2019)), studied this question and replicated the findings of other experiences in high income countries. The authors evaluated the Early Grade Reading Study program, which includes (i) scripted lesson plans, (ii) books and other instructional materials and (iii) centralized training versus in classroom coaching. This intervention was done in 230 schools in North West and the evaluation found that students exposed to two years of the program improved reading proficiency by 0.12 SD if their teachers received centralized training, compared to 0.24 SD if their teachers received in-class coaching. Furthermore, mediation analysis suggests that teachers in the in-class coaching treatment were more likely to use the group guided reading technique (splitting students into smaller reading groups) which enabled individualized attention and more opportunities to practice reading.

The remainder of this document builds on sections 1 and 2 to put present a continuous professional development (CPD) model tailored to Mozambique.

---

<sup>2</sup> Another approach that has been successful at improving student learning in India and now it is being used at scale in Africa is (i) grouping students by their ability level; (ii) providing materials; (iii) provided little initial training to teachers but a large amount of support throughout the year. This program focusses on foundational skills and does not follow a curriculum. Depending on where it was implemented, this is done for just one or a couple of hours in the school day. This program significantly increased student learning in both math and language, with the biggest gains among low-performing students (Banerjee et al., 2007). Other versions of the program include training people in the community to act as remedial support for students with equally positive results.

## Section 3: Proposed additions to the current model

This proposal seeks present the design of a pilot project that strengthens Mozambique's current CPD model in the following ways: (i) introduces structured lesson plans aligned to the curriculum and teachers' materials; (ii) has a narrower focus on specific area (i.e. early reading), (iii) trains teachers in the use of the new materials, including usage of appropriate instructional practices, and provides continuous support from trained coaches through observation and feedback. Coaches will also use structured guides, which provide support on how to conduct classroom observations and provide feedback, and what strategies to use for different challenges teachers experience; (iv) develops a system to monitor the implementation of the program that allows to find and correct problems in real time, as well as learn from bright spots. The proposal also includes impact evaluation of the pilot program to assess its cost-effectiveness and to learn how to best scale it up (See section 4).

The proposal is based on three principles:

### *1. Work within the system.*

This pilot is not intended to change the CPD program that the government has designed and started implementing, but rather adjust and intensify critical elements to achieve the proposed objective of the training, which is to improve early grade literacy in Portuguese. This pilot program will reinforce the current role of the ZIP, by hiring an *instructional leader* (*líder pedagógico*) who will have the specific task of providing in-class coaching to teachers within the ZIP in a continuous manner and will be trained for that assignment. This role will differ from the current model in that the coaches will not have other full-time tasks, which limit the time and capacity of providing sufficient support to the teachers receiving training. The instructional leader will be trained alongside selected *directores pedagógicos*, who will help the instructional leader to provide support to teachers in their respective schools. The *directores pedagógicos* that do not participate of the training will shadow the instructional leader when working on their school to learn, over time, how to conduct in-class coaching. To the extent possible, the training will be done at Teacher Training Colleges (*Instituto de Formação de Professores - IFPs*) in collaboration with the Provincial Department of Education and Human Development, and District Departments of Education, to allow knowledge transfer and build synergies among the institutions.

### *2. Build on what has already been done*

The pilot will use as much as possible existing resources. For example, *Vamos Ler!* (Let's read) program funded by USAID has develop structured daily lesson plans and reading materials for Grades 1 to 3 students that are aligned to the curriculum and have been validated by the Government, currently being implemented in near 1,500 schools in Nampula and Zambezia. We will build on those to develop the material for the program and collaborate with the teams implementing this program, so any improvement we made to the material could also be used in *Vamos Ler!*

### *3. Build an Integral Approach*

This program is designed to work on complementary aspects simultaneously.

- *Strengthening planning culture.* The program will introduce structured lesson plans that are easy to use and have built-in instructional practices that have shown to improve early grade reading. For example, each lesson will have a clear objective and several activities designed to support students in achieving that objective. It will also provide the teacher with ways to check for student understanding of the concepts and an exemplar answers the teacher can use to compare against students answers.

- *Develop further teaching materials.* The program will provide additional reading materials for students. In addition, the program will provide materials teachers can use to instruct, such as pictures of the letters and common words and their pronunciation.
- *Train Coaches:* The program will introduce structured coaching methods, where the ZIP instructional leaders, who will act as coaches, receive material and extensive training on how to observe a classroom, what to look for, what the most frequent issues they will observe are, and how to conduct the feedback session. In particular, the coaches will receive material and support to use a 6-step protocol: (a) use a revised version of the *Teach* classroom observation tool to assess specific teacher practices lay out on the structure lesson plans (already implemented in Mozambique in the context of the SDI 2018); (b) based on the observation the coach will decide on one big takeaway, one thing that would be the easiest for the teacher to improve and would make the biggest change in the quality of their teaching practices. The coaching manual will provide a building blocks of skills teachers should have so the coaches know how to prioritize on where to start supporting the teachers as well as most frequent examples of things teachers struggle with and what are the best strategies to overcome them ; (c) introduce the technique/s to improve that practice by modelling it for the teacher, (d) ask teachers to practice the concept in a role-play and provide feedback as well as videos of local teachers performing this techniques for the teacher to practice independently, and (e) discuss and have teachers reflect on their application of the technique and create an action plan, (f) monitor teacher application of the new skill in the classroom and provide feedback using pre-designed rubrics for the specific skill; (g) start the cycle again. Observations and feedback will be conducted frequently, with a target of once per week in the first 3 months of the application of the program. The instructional leader will assess students' progress frequently to uncover problems faster and to learn from teachers' successes.
- *Reinforce the sense of community within the teachers in the ZIP.* During the program, the ZIP instructional leader will also conduct monthly clinics with all teachers within the ZIP, in which s/he would introduce/reinforce instructional techniques and support teachers to practice them, followed by group discussion on how to help each other to improve, fostering the creating of a community of practice within the ZIP.
- *Strengthen implementation.* Following the model used in *Vamos Ler!* program as well as the Kenya and South Africa programs discussed above, we suggest strengthening implementation by hiring an Implementation Unit (IU) within the MINEDH and having an external team of experts support the Implementation Unit. This team will work in close collaboration with the Teacher Training Department at MINEDH and with a *Program Coordinator* hired under MINEDH's Teacher Training Department for the program. The team of experts will help revise the existing material and develop additional material, as well as work with the IU within MINEDH to select the ZIP instructional leaders and trained them. It is envisioned that ZIP leaders will receive a 2-week training on how to use the material and how to provide ongoing support to teachers. Instruction leaders will also receive check-in visits from those that trained them. During these visits, master coaches will shadow the instructional leaders and observe their feedback meetings and then use the 6-step protocol discussed above to improve the way instructional leaders coach teachers.

## Section 4: Program scale and evaluation

To assess whether the program is effective or not in improving students' early reading and possibly inform a further scale up, this pilot will include an impact evaluation. The evaluation will also tell how to make the program more cost-effective. For example, can the *Directores Pedagógicos, who are already in the school*, with training and ongoing support become effective coaches? What is the optimal number of teachers to coaches?

**Scale.** As agreed with MINEDH, the program will be initially implemented in two provinces, tentatively Sofala and Gaza. This will allow the pilot to be part of the reconstruction effort after cyclone Idai devastated large areas of Sofala earlier this year, reinforcing the concept of “*building back better*” currently encouraged by cooperating partners in Mozambique. The timing of the pilot will be consistent with the rehabilitation and reconstruction efforts, as the actual training and implementation at the ZIP/school level is only expected for the next school year (starting in February 2020). Thus, most schools are anticipated to be functioning (probably under provisional infrastructure conditions) at that point. The program will also include a province that was not affected by the cyclone (such as Gaza or Maputo Province) to provide information on the impact and implementation of the program under a different (non-emergency recovery) setting. In each province the program will start with approximately 200 to 400 schools (as described below), distributed in approximately 50/100 ZIPs and 10 districts. The randomization will be done at the ZIP level, as the entire ZIP will belong to either the beneficiary or the control group. Districts can have beneficiary ZIPs, control ZIPs and ZIPs not participating in the program. In the first year of the program there will be one *instructional leader* per ZIP (as described below) and all first-grade teachers in the randomly selected ZIPs for treatment will receive the intervention package<sup>3</sup>. In the second year, all second-grade teachers in randomly selected ZIPs will receive the intervention package.

**Evaluation.** We propose to conduct two different “packages” of interventions in each province to assess the impact of the program rigorously using random assignment at the ZIP level. Specifically, the evaluation will answer the following research questions:

1. Compared to the current model, are the proposed adjustments including improved structured lesson plans and reading materials, and training and support of coaches, more effective in improving teaching practices of teachers and student learning outcomes? What is the cost effectiveness of each intervention? Does the effect differ by area (cyclone affected vs. not affected)?
2. Is it more cost effective to use new ZIP instructional leaders or existing *directores pedagogicos* to provide ongoing support to teachers?
3. Does the additional benefit of training and support of coaches differ by the number of teachers they coach? Since the number of teachers per coach is not exogenously assigned, we will rely on non-experimental variation in the number.<sup>4</sup>

*Sofala (cyclone affected areas)*

Group 1 will receive the intervention described in the “*Formação Continua*” Manual (control group)<sup>5</sup>;

---

<sup>3</sup> The team does not foresee spillovers into the control group as the program involves the use of materials non-participant schools will not receive and ongoing support that is tailor to the needs of the specific teacher and his/her classroom.

<sup>4</sup> Nonetheless, the number of teachers per coach will ultimately depend on the size of the ZIP, which is determined by the exogenous (for the purposes of this evaluation) rule approved in 2008 by a Ministerial Decree that created the ZIPs. The rule says that a ZIP should have at least three and a maximum of six schools, and that the radius between the “main/hub ZIP school” and the other ZIP schools cannot exceed 10 kms.

<sup>5</sup> We will look at implementation fidelity as this could end up being a control group due to problems with implementation of the current program.

Group 2 will receive the *Aprender+* pilot described in section 3;

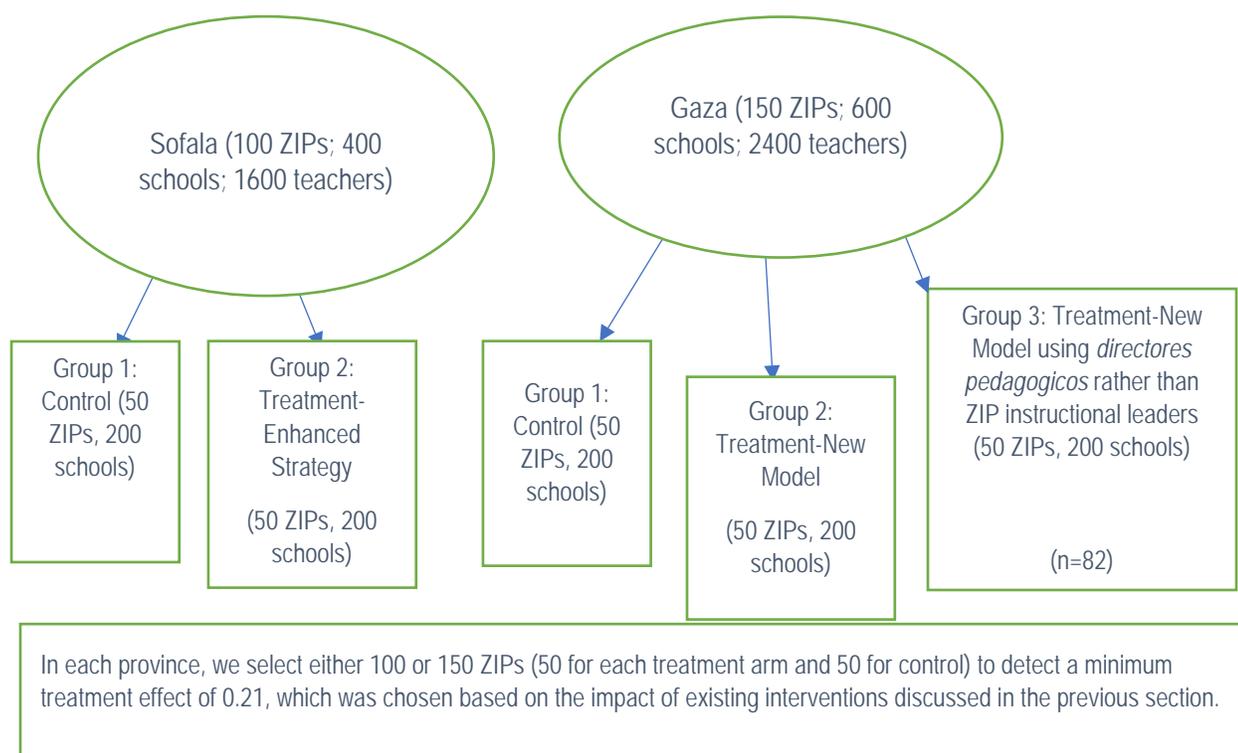
*Gaza (not affected areas)*

Group 1 will receive the intervention described in the "*Formação Continua*" Manual (control group);

Group 2 will receive the full *Aprender+* pilot described in section 3;

Group 3 will receive the *Aprender+* pilot, but without the ZIP instructional leader. The coaching will be conducted by *directores pedagogicos*, who will receive the same training as the ZIP instructional leaders under group 2<sup>6</sup>. Figure 2 presents the study design with sample size in each treatment and control group across the two provinces.

Figure 2: Proposed evaluation design in Sofala and Gaza



In addition to the quantitative evaluation, a qualitative evaluation will be conducted on sub-samples to contextualize the results from the quantitative evaluation.

**Timeline.** The project will be implemented over two years. The first six months will be dedicated to finalizing the design of the pilot and evaluation, hiring the data collection firm, reviewing and adapting existing materials, and preparing new materials as needed (especially materials for training and for supporting the instructional leaders, such as guides and videos). This period will also be important to communicate details of the program and its implementation to all actors involved – schools, teachers, directors, district authorities, school councils- so that objectives, expected outcomes, roles and responsibilities are clear. Over the following 24 months, planned interventions will be implemented - the training of instructional leaders and teachers will be conducted; instructional leaders will do regular follow up with their

<sup>6</sup> If the school does not have a "director pedgogico" the school director will receive the training.

assigned teachers, visiting schools, observing classes, providing feedback. The last six months of the pilot will be allocated to finalize the data collection for the impact evaluation and analysis.

## Section 5: Cost and sources of funding

Below is the overall budget for this activity.

Activity	List of Activities	Cost	Source (TBC)
<i>1. Develop the materials</i>	This includes revised lesson plans and reading materials (using existing materials in Mozambique as a starting point) and adaptation of the material from the Coach program developed by GEDGE.	\$0.5M	Donors
<i>2 Support Implementation</i>	This includes support from the GEDGE Team as well as a local Team to help implement the program for the duration of the program.	\$1.5M	Donors
<i>2. Implement</i>	This includes Technical Assistance personnel within the Ministry to manage the program, new personnel in the education system (ZIP instructional leaders), and materials and logistics to support the implementation in both provinces during the duration of the program (two years of intervention).	\$10M	FASE
<i>3. Evaluation</i>	This includes finalizing the design, conducting and analyzing the results of the impact evaluation as well as conducting a qualitative evaluation.	\$1M	Donors
<b>Total</b>		<b>\$13M</b>	

## References

- Banerjee, Abhijit Vinayak, Shawn Cole, Esther Duflo, and Leigh Linden. 2007. "Remedying Education: Evidence from Two Randomized Experiments in India." *Quarterly Journal of Economics* 122 (3): 1235–64.
- Bambrick-Santoyo, Paul. 2016. *Get Better Faster: A 90-Day Plan for Coaching New Teachers*. John Wiley & Sons.
- Bold, Tessa, Deon Filmer, Gayle Martin, Ezequiel Molina, Christophe Rockmore, Brian Stacy, Jakob Svensson, and Waly Wane. 2017. "What Do Teachers Know and Do? Does It Matter?: Evidence from Primary Schools in Africa." Policy Research Working Paper 7956. World Bank.
- Conn, Katharine M. 2017. "Identifying Effective Education Interventions in Sub-Saharan Africa: A Meta-Analysis of Impact Evaluations." *Review of Educational Research* 87(5): 863-898. doi:10.3102/0034654317712025.
- Kraft, Matthew A., David Blazar, and Dylan Hogan. In press. The effect of teaching coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*.
- Lemov, D., Woolway, E., & Yezzi, K. (2012). *Practice perfect: 42 rules for getting better at getting better*. San Francisco: Jossey-Bass.
- Piper, B., & Korda, M. (2011). EGRA Plus: Liberia. Program Evaluation Report. RTI International.
- Piper, B., Zuilkowski, S. S., Dubeck, M., Jepkemei, E., & King, S. J. (2018a). 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.
- 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.
- Popova, Anna, David K. Evans, Mary E. Breeding, and Violeta Arancibia. 2018. "Teacher Professional Development around the World: The Gap between Evidence and Practice." Unpublished working paper, last modified August 22, 2018, World Bank.
- Snilstveit, Birte, Jennifer Stevenson, Radhika Menon, Daniel Phillips, Emma Gallagher, Maisie Geleen, Hannah Jobse, et al. 2016. "The Impact of Education Programmes on Learning and School Participation in Low and Middle-Income Countries: A Systematic Review Summary Report." 3ie Systematic Review Summary 7, International Initiative for Impact Evaluation, London. [http://www.3ie.org/media/filer\\_public/2016/09/20/srs7-education-report.pdf](http://www.3ie.org/media/filer_public/2016/09/20/srs7-education-report.pdf).

## Annex I: Power Analysis

We conduct power analysis to determine the sample size for evaluating the impact of the intervention. The sample size depends on the number of ZIPs in each treatment and control group, and the number of teachers and students within each ZIP. Note that teachers and students are nested within the ZIPs and therefore, we will account for the intra-class correlation (ICC) for each. The ICC will allow us to understand how much of the variation in teacher and student scores is accounted by the differences between ZIPs. If the ICC is 0, we know that the teachers/students nested within the ZIPs are no different from teachers/students in other schools and therefore, we can treat them as independent observations.

We use recently collected SDI 2018 data that contains information on Teach scores for 607 teachers and total test scores for 2500 students enrolled at primary level in 282 schools across Mozambique. Table A1 presents baseline estimates of mean and standard deviation of teacher and student scores, as well as estimates for the ICC. Teach scores are measured on a scale of 1-5, while student test scores on a scale of 1-100. Student test scores are a weighted average of scores in math, non-verbal and Portuguese. Because SDI is sampled based and we don't have enough information to compute the ICC at the ZIP level, we account for nesting within a ZIP by computing the ICC of students/teachers within a district. The ICC for teachers and students is moderately high, at 0.30 and 0.44 respectively. This suggests that differences between teachers and students are large between ZIPs.

**Table A1: Baseline estimates of mean, standard deviation and ICC for teacher and student test scores**

Variable	Obs	Mean	SD	ICC
Teacher Teach Score	607	2.93	0.68	0.30
Student Test Score	2500	32.25	23.44	0.44

Further, based on school administrative records, there are on average 119 first-grade students per school in Sofala and on average 83 first-grade students per school in Gaza while there are 3-5 first-grade teachers per school in both provinces.

Using this information, we conduct two sets of power analysis<sup>7</sup>. We first look at teachers (level 1) nested within ZIPs (level 2). We make standard assumptions of power at 0.8 and statistical significance at 0.05, and vary ICC between 0.1 and 0.3. For a minimum detectable effect size (MDES) of 0.21 (which corresponds to a 0.14 point change in Teach score), ICC equal to 0.15, and a cluster size i.e. number of teachers per ZIP equal to 16<sup>8</sup>, we need to select at least 121 ZIPs; however, as ICC increases, we need to select a larger number of ZIPs. For a MDES of 0.31, and ICC at 0.3, we need to select at least 101 ZIPs. Table A2 presents a summary of the results.

---

<sup>7</sup> Power analysis for teachers nested within ZIPs has been performed in Optimal Design and Stata using `clustersampsi` function. Power analysis for students nested within ZIPs has been performed in Optimal Design only to account for the repeated measures.

<sup>8</sup> This is dependent on the number of schools and teachers within schools in our sample; we expect to sample 4 schools per ZIP and 4 teachers per school.

**Table A2: Power analysis for teachers nested within ZIPs**

<u>Without Repeated Measures (Midline)</u>			
<u>MDES</u>	<u>ICC</u>	<u>Cluster Size (No. of Teachers)</u>	<u>Number of Clusters (ZIPs)</u>
<u>0.21</u>	<u>0.10</u>	<u>16</u>	<u>96</u>
<u>0.21</u>	<u>0.15</u>	<u>16</u>	<u>121</u>
<u>0.21</u>	<u>0.30</u>	<u>16</u>	<u>172</u>
<u>0.31</u>	<u>0.10</u>	<u>16</u>	<u>50</u>
<u>0.31</u>	<u>0.15</u>	<u>16</u>	<u>62</u>
<u>0.31</u>	<u>0.30</u>	<u>16</u>	<u>101</u>

Next, we look at students (level 1) nested within ZIPs (level 2). Note, we will follow students over two subsequent years and therefore, will take repeated measured into account for the power analyses. Given standard assumptions of power and statistical significance, and ICC set at different values, we see a trade-off between MDES and required number of clusters. For example, for a MDES of 0.21, ICC of 0.1, and cluster size of 100<sup>9</sup>, we need to select 71 ZIPs; however with ICC at 0.4, we need to select 225 ZIPs. Similarly, for a MDES of 0.31 and ICC at 0.4, we need to select 100 ZIPs.

Based on prior literature on teacher training programs in Africa, we want to detect an effect size of at least 0.21 SD in student learning outcomes. However, we believe that the 0.3-0.4 ICC in SDI data is not representative for ZIPs in Sofala and Gaza. Based on information collected through experts' interviews and from our own field visits we understand that there is high heterogeneity within schools in a given ZIP. Therefore, we propose to set ICC at 0.15 and randomly select approximately 100 ZIPs (50 in treatment and 50 in control). Further, within each ZIP, we will select 4 schools, and 4 teachers and 25 students per school, on average. Note we will consider blocking by district and income level in the selection of schools to reduce the ICC and obtain more precise measures. Table A3 presents a summary of the results.

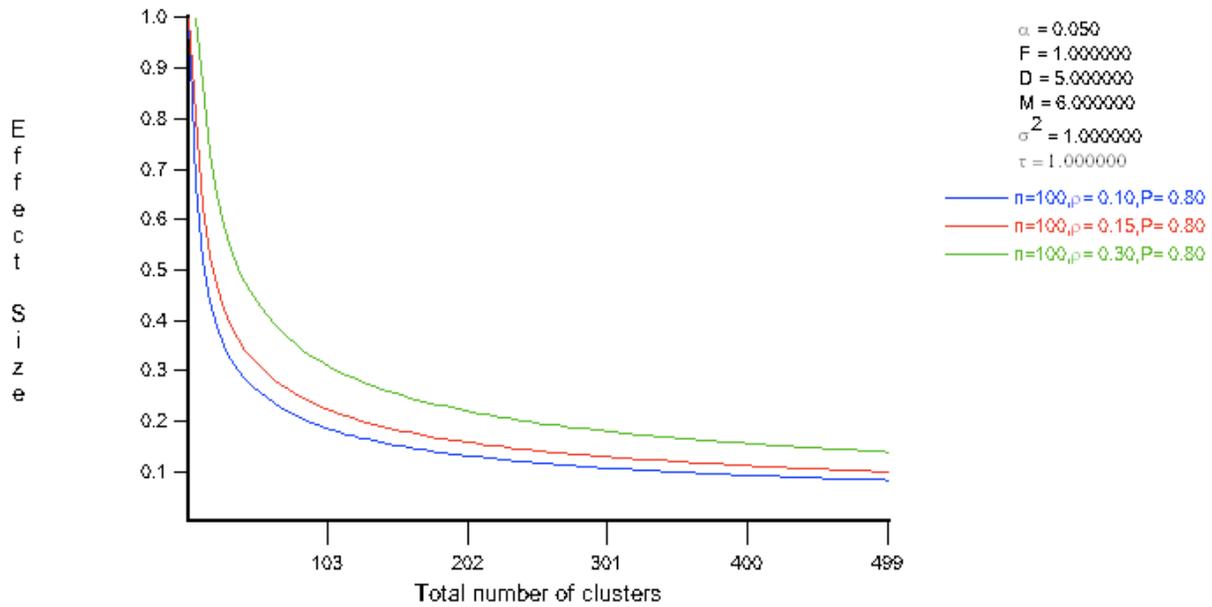
**Table A3: Power analysis for students nested within ZIPs**

<u>Repeated Measures (Midline + Endline)</u>			
<u>MDES</u>	<u>ICC</u>	<u>Cluster Size (No. of Children)</u>	<u>Number of Clusters (ZIPs)</u>
<u>0.21</u>	<u>0.10</u>	<u>100</u>	<u>71</u>
<u>0.21</u>	<u>0.15</u>	<u>100</u>	<u>101</u>
<u>0.21</u>	<u>0.30</u>	<u>100</u>	<u>181</u>
<u>0.21</u>	<u>0.40</u>	<u>100</u>	<u>225</u>
<u>0.31</u>	<u>0.10</u>	<u>100</u>	<u>35</u>
<u>0.31</u>	<u>0.15</u>	<u>100</u>	<u>50</u>
<u>0.31</u>	<u>0.30</u>	<u>100</u>	<u>90</u>
<u>0.31</u>	<u>0.40</u>	<u>100</u>	<u>116</u>

<sup>9</sup> This is dependent on the number of schools and children within schools in our sample; we expect to sample 4 schools per ZIP and 25 children per school

Figure A1 presents an illustration of the tradeoff between MDES and required number of clusters at  $\alpha = 0.05$ , power = 0.80, ICC=0.10, 0.15, 0.30 and cluster size (n) = 125.

Figure A1: MDES vs. number of clusters (students nested within ZIPs)



We do not expect attrition at level 2 over time. However, we do expect attrition at level 1 to be no more than 3-5%, which is considered within the conservative range and does not invalidate the results (IES, 2017).

**Analysis:**

We will use a two-level multi-level model for students nested within ZIPs. The specifications for the two-level model are as follows:

Level 1 model:

$$Y_{ij} = \beta_{0j} + \beta_{1j}X_{ij} + r_{ij}; r_{ij} \sim N(0, \sigma^2)$$

Where  $X_{ij}$  is the vector of student, teacher and school covariates (gender, age, baseline scores, private/public, teacher supports, etc.)

Level 2 model:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}Z_j + \mu_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}Z_j + \mu_{1j}; \begin{pmatrix} \mu_{0j} \\ \mu_{1j} \end{pmatrix} \sim N \left[ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{pmatrix} \right]$$

Where  $\beta_{0j}$  is the ZIP-level random intercept,  $\beta_{1j}$  is the ZIP-level random slope and  $Z_j$  is a vector of ZIP-level covariates (ZIP-dummies, demographic factors, etc.)

Overall model:

$$Y_{ij} = \underbrace{\gamma_{00} + \gamma_{01}Z_j + \gamma_{10}X_{ij} + \gamma_{11}Z_jX_{ij}}_{\text{Fixed Effect}} + \underbrace{\mu_{0j} + \mu_{1j}X_{ij} + r_{ij}}_{\text{Random Effect}}$$