



| 2022 LEAP CHALLENGE

LEAP Final Deliverable(s)

Project Host:

WVOB education for development

Fellows:

Annelise Eaton Buzaid, Research Fellow

Tsz Man (Bethany) Fong, Social Entrepreneur Fellow

Lea Mörsdorf, Research Fellow

Barbara Trudell, Research Fellow

TABLE OF CONTENTS

Executive Summary	4
Introduction	4
Organization's role & strength	4
Need summary	5
Solution summary & next steps	5
List of Acronyms	6
Deliverable 1 - Conceptual Framework	7
Introduction	7
Programme Overview	7
Conceptual Framework	8
Research Questions	9
Overarching Hypotheses	9
Subtopic 1: Instructional quality	10
Subtopic 2: School environment	11
Subtopic 3: Parent/community engagement	13
Deliverable 2: General Study Design	14
Introduction to study design	14
Rollout of the CPD programme	14
Planned study design	15
Planned data collection	16
Procedure	16
Measures	17
Deliverable 3: Data Analysis Plan	28
Introduction	28
General recommendations	28
Data analysis for the overarching hypotheses	29
Data analysis by subtopic	30
Bibliography / References	37

Executive Summary

Introduction

VVOB and the Ministry of Education in Rwanda are collaborating to institutionalize continuous professional development (CPD) for school leaders. The effective school leadership professional development programme aims to improve school leadership competences and practices among headteachers and ultimately the quality of teacher performance, and student learning outcomes. School leadership capacity-building programmes in the global North have been found to be effective in boosting student learning gains; but few programmes in sub-Saharan Africa have focussed on enhancing school leadership and its impact on student learning outcomes (Sampat et al., 2020). Through the LEAP project, the ministry and VVOB are aspiring to better understand and measure programmatic impact of school leadership CPD on teachers and students.

Organization's role & strength

VVOB has a mission of ensuring quality education for all learners. The organisation realises its mission by working with government officials focused on strengthening school leadership and advancing teacher professional development. VVOB invests in research to build the evidence base on what makes professional development effective. It also supports governments to build evidence-aligned programming and take effective innovations to scale.

VVOB's theory of change starts from a human rights-based approach, ultimately focused on every child's right to receive a quality education. VVOB works to strengthen the capacity of government institutions that are responsible for the professionalisation of teachers and school leaders. Ultimately, these actors have the greatest impact on the learning performance and the well-being of learners.

Many governments have policy plans in place to continuously improve their education systems in order to achieve SDG4. However, a major implementation gap lies between the existing legal guarantees and policies and the effective realisation of rights and SDG4 targets. VVOB puts these policy plans at the centre of its partnership with the Ministry of Education, and focuses on:

- Reinforcing the capacities of duty bearers, to enable them to guarantee the right to education of the citizens of the nation (rights holders) in a more qualitative and equitable way;
- Promoting equity (the right to non-discrimination) within governments' education policies and actions.

In Rwanda, the *Learning through Assessment and Data* (LEAD) multi-year programme for 2022-2026 aims to support SDG4 by ensuring that all Rwandan children can actively participate in quality education and successfully progress through primary education.

Need summary

In Rwanda, the Continuous Professional Development (CPD) programme for school leaders focuses on training and supporting newly assigned school leaders to effectively use data to improve teaching and learning in their schools. The programme is built on a strong evidence base suggesting that with adequate training and support, school leaders can improve the quality of teaching and conditions for learning.

The primary need that is being addressed by the LEAP Fellows is the design of a research methodology to demonstrate the impact of the CPD programme on distributed school leadership and student outcomes. In particular, the research will support VVOB's partnership with government partners in Rwanda, who seek to understand the return on investment of the CPD programme in terms of student learning outcomes. Investing in research is critical to increasing the credibility of the CPD programme in order to mobilise government financial, human, and policy resources to support program expansion.

Solution summary & next steps

In a first step, the LEAP Fellows conducted a literature search to develop a more fine-grained ToC. The major purpose of this step was to make the different components of the ToC testable (i.e., so that specific hypotheses could be derived). To this end, the LEAP Fellows focused on three main actions of school leaders: providing instructional support for teachers, creating a safe and positive school environment, and collaborating with parents and the community. These actions are hypothesised to improve student outcomes, such as lower dropout rates, better exam grades, and higher pass rates.

In a next step, the LEAP Fellows turned to the planned intervention rollout to suggest a potential impact study design. Because it cannot be guaranteed that the cohort-wise trained schools will be randomly assigned, the LEAP Fellows based their planning on a quasi-experimental design in which schools will be matched between cohorts based on certain criteria.

Finally, the LEAP Fellows provided recommendations on which variables should be assessed at which level (student, teacher, school leader) and how these variables can be used in statistical analyses to test the hypotheses.

List of Acronyms

ANOVA - Analysis of Variance

BLF - Building Learning Foundations

CPD - Continuous Professional Development

DDE - District Director of Education

DEO - District Education Officer

DHT - Deputy Headteacher

ESSP - Education Sector Strategic Plan

HT - Headteacher

LEAD - Learning through Assessment and Data

LEAP - Leveraging Evidence for Action to Promote change

NESA - National Examination and School Inspection Authority

PLC - Professional Learning Communities

REB - Rwanda Basic Education Board

SBM - School-Based Mentor

SEIs - Sector Education Inspectors

SL - School Leadership

SIP - School Improvement Plan

SDG - Sustainable Development Goal

ToC - Theory of Change

URCE - University of Rwanda – College of Education

Deliverable 1 - Conceptual Framework

Introduction

School leaders in the Global South face challenges in performing their roles, due to inadequate training (Sampat et al., 2020). School leaders in Rwanda have stated that more professional development is needed to prepare them for difficulties such as parental disengagement and unmotivated teachers (Karareba et al., 2019). Throughout this document, the term school leader refers to headteachers and deputy headteachers in Rwanda.

The Rwandan Ministry of Education acknowledges their challenges, and the need for capacity-building and support. One of the intended outcomes stated in the Education Sector Strategic Plan (ESSP) 2018-24 is to improve leadership in schools (Rwandan Ministry of Education, 2019). The ESSP highlights the need to transform school leaders from playing “an administrative role” to becoming “true leaders of their schools” who are required to support the professional development of teachers and overall school development. Hence, the Rwanda Basic Education Board (REB) developed five professional standards for effective school leaders (see below) and, in collaboration with VVOB and the University of Rwanda – College of Education (URCE), initiated a CPD programme on effective school leadership for school leaders.

Programme Overview

The CPD programme on effective school leadership for school leaders in Rwanda is led by VVOB, in partnership with the Rwanda Ministry of Education and operational partnership with REB, URCE and the National Examination and School Inspection Authority (NESA). The programme responds to the need to both support newly-assigned school leaders and strengthen their use of data. The programme aims to improve the quality of education by strengthening the competencies of newly-assigned school leaders to effectively use data to improve teaching and adequately address repetition, drop-out, and equity gaps in learning outcomes, with the support of sector and district officials. The CPD programme will be implemented in 650 primary schools in all 30 districts of Rwanda between 2022-2026.

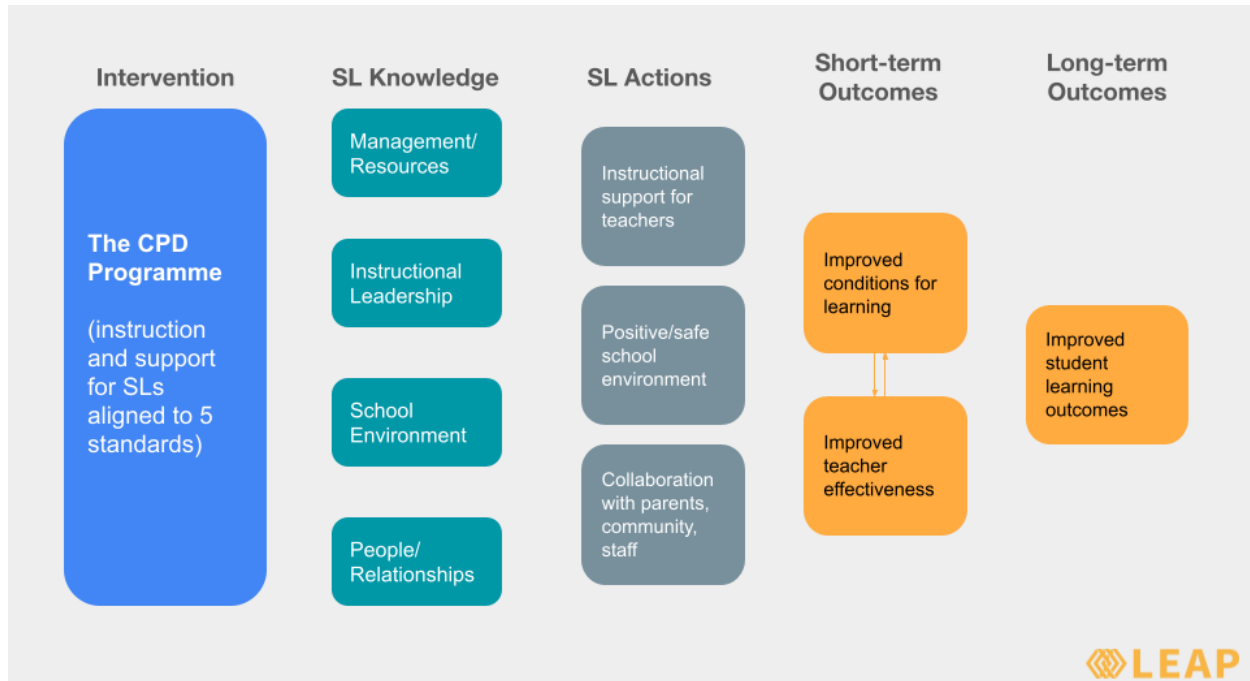
This document presents a conceptual framework, evaluation plan, and data collection plan to inform and facilitate research studies that connect effective professional development of school leaders with improved student learning outcomes. This in turn will facilitate scaling and institutionalisation of school leadership support in Rwanda.

Conceptual Framework

The graphic below illustrates the conceptual framework, demonstrating a hypothesised link between the CPD programme for effective school leadership and student learning outcomes. The framework is divided into three key categories: intervention (blue), outputs (teal and grey) and outcomes (orange).

Figure 1

Depiction of the conceptual framework



The diploma programme on effective school leadership consists of four modules based on the five professional standards for effective school leaders developed by REB. The five standards are:

- Standard 1: Creating strategic direction for the school
- Standard 2: Leading learning
- Standard 3: Leading teaching
- Standard 4: Managing the school as an organisation
- Standard 5: Working with parents and wider community

We hypothesise that participation in the CPD programme will increase headteachers (HT) and deputy headteachers (DHT)¹ knowledge related to resource management, instructional leadership, school environment, and people management. This knowledge will alter HT and DHT actions in the following ways:

¹ When talking about School leaders participating in the CPD programme we refer to Headteachers and Deputy Headteachers in charge of studies.

- School leaders will provide more frequent and better quality instructional support for teachers;
- School leaders will implement efforts to create a positive school environment;
- School leaders will strengthen partnerships with families and the community.

Research in the field indicates that the actions above improve the conditions for learning in a school and increase teacher effectiveness, thereby influencing student learning outcomes (Printy, S. 2010; Sampat, et. al., 2020; Grisson, et. al., 2021). In the sections below, we detail a proposed research design aligned to this conceptual framework.

Research Questions

The overarching research question of this study is ***What is the impact of a CPD programme for school leaders on student outcomes?*** Based on the conceptual framework, we focus on the following more specific research questions:

Table 1. List of research questions in this study

Sub-research questions	Related standards for effective school leaders				
	1	2	3	4	5
How does a CPD programme for school leaders influence the quality of instruction in schools?	*	*	***		
How does a CPD programme for school leaders influence the school environment?	*	*	*	***	
How does a CPD programme for school leaders influence parents' engagement in their children's learning?		*	*		***

*** Primary standard, * Secondary standards

Overarching Hypotheses

Based on the conceptual framework, we propose the following overarching hypotheses:

1. Students in schools participating in the CPD programme show lower student dropout rates compared to schools not participating in the programme.
2. Students in schools participating in the CPD programme show better national exam results than comparable students in schools that are not participating.
3. Students in schools participating in the CPD programme show higher pass rates than students in schools that are not participating.

4. Students in schools participating in the CPD programme show higher attendance rates than students in schools that are not participating.

Subtopic 1: Instructional quality

We hypothesise that the following mechanisms enable the CPD programme to positively impact instruction:

- School leaders set a clear vision for learning and create a data-driven instructional programme
- HTs/DHTs use data to identify school-wide and teacher-specific instructional needs
- HTs/DHTs increase instructionally-focused interactions with teachers (coaching, documented appraisals, professional learning, etc.)

Based on this, we derive the following hypotheses:

- A1. When compared to control schools, HTs/DHTs participating in the CPD programme will receive higher scores in setting a clear vision for learning and creating a data-driven instructional programme, as indicated by School Leadership Assessment Toolkit indicators 1.2.1 and 2.1.3.
- A2. When compared to control schools, HTs/DHTs participating in the CPD programme will be more likely to use data to identify school-wide and teacher-specific instructional needs, as indicated by a higher School Improvement Plan completion rate.
- A3. When compared to control schools, HTs/DHTs participating in the CPD programme will engage in more instructionally-focused interactions with teachers, as indicated by higher scores in School Leadership Assessment Toolkit indicators 2.1.4 and 3.1.1-3.1.8.

The CPD programme equips HTs/DHTs with the knowledge needed to assess teaching performance, identify areas in need of improvement, and provide teachers with actionable feedback. Prior research shows that when school leaders prioritise instructional quality and engage in targeted, instructionally focused interactions with teachers, the quality of teaching improves. In addition, teaching quality improves when HTs/DHTs are able to identify and arrange for high-quality professional development for teachers (Grissom et. al., 2021). Teacher quality is shown to be the single greatest contributor to student success among all school-level factors. Instructional quality influences student academic success far more than other school factors, such as facilities and curriculum (Chetty, et. al, 2014). Therefore, HT efforts to improve instructional quality have a direct link to student academic performance.

As a result, we hypothesise that changes in HT behaviour facilitated by the CPD programme will impact student learning outcomes in the following ways:

- B1. During their participation in the CPD programme, HTs demonstrate an increase in the frequency of instructionally-focused interactions with teachers, as measured by the number of classroom observations per teacher per term and the number of teachers that received Continuous Professional Development (School Improvement Plan).

- B2. During their participation in the CPD programme, HTs demonstrate an improvement in the quality of instructional feedback given to teachers, as measured by a teacher survey.
- B3. Due to an increase in support and feedback from HTs, teachers' instructional quality improves. If feasible, this can be measured via a student survey.
- B4. An improvement in teaching quality leads to academic benefits for students, as indicated by higher national examination results and student pass rates, when compared to control group schools.

Subtopic 2: School environment

School environment for teachers

We hypothesise that the following mechanisms enable the CPD programme to positively impact the working conditions for teachers:

- HTs/DHTs facilitate a collaborative learning environment among teachers
- HTs/DHTs provide instructional support for teachers (see subtopic 1)
- HTs/DHTs evaluate all teachers, appreciating the high performers

Based on this, we hypothesise that²:

- A1. Schools participating in the CPD programme receive higher scores in teacher evaluation (School Leadership Assessment Toolkit indicator 3.1.2) compared to control schools.
- A2. More HTs in the CPD programme will take steps to appreciate high-performing educators compared to control schools, as indicated by a higher percentage of participating schools providing motivational opportunities for school staff compared to control schools (indicator to be found in SIP - School Performance).
- A3. Schools participating in the CPD programme receive higher scores in facilitating collaborative learning environments for all teachers (School Leadership Assessment Toolkit indicator 3.2.1) compared to control schools.
- A4. Schools participating in the CPD programme facilitate more collaboration among teachers compared to control schools, as indicated by a higher number of functional department/ school-based in-service professional learning groups (SBI groups) per teacher (indicator to be found in SIP - School Performance).

HTs can also increase teacher motivation and lower teacher turnover by cultivating a strong professional environment (Grissom et al., 2021). Teachers tend to leave their posts when they find teaching more challenging and less enjoyable, such as insufficient classroom space and having less access to technology or learning resources (Grissom, 2011). Several school leader practices are identified to be effective in cultivating a strong professional environment for teachers, namely maintaining order and discipline, supporting teachers via addressing their concerns and providing professional development, evaluating teachers regularly and building a positive and collaborative school culture (Grissom et al., 2021; Kraft & Papay, 2014). A study investigating public schools in Brazil reveals that an increase in operation management scores

² The hypotheses regarding the instructional support given by the HTs are suggested in subtopic 1.

is associated with a significant increase in the teacher motivation index (Leaver et al., 2019). It indicates that school leaders with better management are likely to increase teachers' motivation.

As a results, we hypothesise that these changes in HT behaviour impact student learning outcomes in the following ways:

- B1. When the school climate is more positive, as indicated by teacher perceptions of school climate, teacher job satisfaction will increase and turnover rates will decrease.
- B2. When teacher job satisfaction increases, fewer teachers will drop out, hence lower turnover rates.

Zeitlin (2021) also finds that the loss of a teacher by a school in Rwanda is significantly associated with a reduction in learning levels. It implies that teacher loss may result in a decline in learning outcomes. We believe the CPD programme can prevent learning loss by reducing teacher turnover.

School environment for students

We hypothesise that the following mechanisms enable the CPD programme to positively impact the school learning environment for students:

- HTs/ DHTs ensure that drinking water, toilets, and other necessary supplies are available for students
- HTs/ DHTs provide learning resources for all students
- HTs/ DHTs establish orders and measures to manage students' disruptive behaviours and monitor child safety
- HTs/ DHTs set high academic expectations that are transparent to both teachers and students

Based on this, we hypothesise that:

- A5. Schools participating in the CPD programme receive higher scores in building an environment where students feel safe and welcomed (School Leadership Assessment Toolkit indicators 2.2.1-2.2.8) compared to control schools.
- A6. Schools participating in the CPD programme receive higher scores in shaping a learning-conducive environment for all students (School Leadership Assessment Toolkit indicators 4.2.1-4.2.4 and indicators 6.1.1-6.1.8) compared to control schools.
- A7. Schools participating in the CPD programme receive higher scores in ensuring student attendance (School Leadership Assessment Toolkit indicator 2.1.7) compared to control schools.
- A8. Schools participating in the CPD programme receive higher scores in communicating their vision and goals to students and teachers (School Leadership Assessment Toolkit indicators 1.1.1 and 1.1.6) compared to control schools.

Participating school leaders learn how to develop an institutional environment where all students are welcomed to the school and feel safe in the CPD programme. Research has found that with this learning-conducive environment, students are more engaged in school activities and less likely to drop out (Grissom et al., 2021; Hopkins et al., 1997; Ma, 2003). Holding high

expectations for student achievement among teachers and students is also identified as having a significant influence on student academic performance (Griffith, 2000).

As a result, we hypothesise that these changes in HT behaviour impact student learning outcomes in the following ways:

- B3. When the school climate is more positive, as indicated by teacher perceptions of school climate, student attendance as well as engagement will increase, and dropout rates will decrease.
- B4. When student engagement increases, fewer students will drop out.

Subtopic 3: Parent/community engagement

We hypothesise that the following mechanisms enable the CPD programme to positively impact the degree to which parents actively support their children's learning:

- HTs/ DHTs set expectations for teachers to actively engage with parents on a regular basis, build teacher capacity for such engagement, and take leadership in modelling it themselves as well.
- HTs/ DHTs build teachers' ability and readiness to use their understanding of students' home contexts to teach the students more effectively.

Based on this, we derive the following hypotheses:

- A1. When compared to control schools, HTs participating in the CPD programme are more likely to build teacher capacity and readiness to engage with parents on a regular basis, as indicated by the School Leader Assessment Tool, indicators 5.2.1 through 5.2.6, and the number of teachers that participated in CPD activities on parental engagement.
- A2. When compared with control schools, HTs participating in the CPD programme are more likely to build teachers' ability to teach students more effectively, based on a better understanding of student home contexts, as indicated in the School Leader Assessment Tool indicator 5.2.2.

Further, we hypothesise that these changes in HT behaviour impact student learning outcomes in the following ways:

- B1. When teachers are able and expected to engage with parents in regular and positive ways, student attendance will increase and fewer students will drop out.
- B2. When teachers use their understanding of students' home contexts to teach them more effectively, student pass rates will go up.

Deliverable 2: General Study Design

Introduction to study design

The study plan below provides recommendations on how the impact study on the overall effectiveness of the CPD programme for school leaders could be conducted. We will first describe how the CPD programme will be rolled out, as this sets the boundaries for the study. Based on this, the idea for the study design will be presented, followed by concrete recommendations on data collection. Finally, the measures needed to assess all variables of interest will be introduced.

Rollout of the CPD programme

The CPD programme for school leaders will be rolled out stepwise in 4 cohorts, similar to the depiction in Table 2 below. A total of 650 ‘newly appointed’ primary HTs and 117 DHTs across the country will be trained between 2023-2026. ‘Newly appointed’ is defined as leaders who took on their role within the last three years. The training will last approximately 10 months and will include quarterly professional learning communities (PLC) sessions within the sector where the school is located. Also, school-based mentors (SBMs) and other school leaders like district education officers (DEOs), sector education inspectors (SEIs), and district directors of education (DDEs) will be trained, though in a less systematic manner. Schools will be assigned to cohorts, though a completely random assignment cannot be guaranteed.

Table 2. Illustration of the planned intervention rollout

Cohort	2022	2023	2024	2025	2026
4	No	No	No	No	Yes
3	No	No	No	Yes	Yes
2	No	No	Yes	Yes	Yes
1	No	Yes	Yes	Yes	Yes

Note: The first year of receiving the intervention is depicted in a darker green than the following years, because this is the year in which the training will take place, i.e., the intervention is strongest.

Planned study design

Because longer-term changes are of major interest to VVOB, we suggest to focus the analyses on cohorts 1 and 3 (see purple frames in Table 3). By focusing on these two cohorts, it will be possible to compare changes in the intervention group (cohort 1) to changes in the waiting control group (cohort 3) in the first two years after project start. Due to the non-random assignment to cohorts, an experimental design to compare cohorts 1 and 3 is not feasible. Therefore, we suggest a quasi-experimental design, in which the HTs of cohorts 1 and 3 are as comparable as possible.

It will be essential to match HTs of cohort 3 to HTs of cohort 1 with regard to their years of experience. As table 3 illustrates, HTs that can be assigned to cohorts when the project starts at t_0 , will have 0-1, 1-2, or 2-3 years of experience. Because HTs should have a maximum of 3 years of experience when they enter the CPD programme, only HTs with 0-1 years of experience can be assigned to cohort 3 at the start of the project at t_0 . These are the HTs that can be compared to the HTs in cohort 1 with 0-1 years of experience.

Apart from years of HTs experience, researchers should select matched schools in cohorts 1 and 3 that are as similar as possible based on the criteria below:

- Location of the school (urban/rural)
- Exposure to other programmes (e.g., BLF, FHI360)
- Sector (because PLC will be sector-based) and/or district
- Ratio of to-be-trained persons (amount of DHT vs. HT per cohort)
- School size
- School category (primary school with/without secondary school)
- Share of female students

This matching will be important to ensure that the schools in the intervention group (cohort 1) do not differ in potentially confounding variables from the schools in the waiting control group (cohort 3). For instance, if schools in cohort 1 were primarily from urban regions and schools in cohort 3 from rural regions, it would be likely that the cohorts would differ already at t_0 in relevant variables such as HT behaviour or student outcomes. Also, it could be that changes over time might occur because of the rural versus urban surrounding, and could not be clearly separated from intervention effects. Therefore, the two groups should be as similar as possible.

Table 3. Illustration of recruitment of HTs per cohort based on their experience in years

Cohort	2022	2023	2024	2025	2026
4			0-1	0-1 1-2	0-1 1-2 2-3
3		0-1	0-1 1-2	0-1 1-2 2-3	1-2 2-3 3-4
2		1-2	0-1 2-3	1-2 3-4	2-3 4-5
1		0-1 1-2 2-3	1-2 2-3 3-4	2-3 3-4 4-5	3-4 4-5 5-6
		t ₀	t ₁	t ₂	t ₃

Note: The black and yellow numbers in the cells of the table indicate HTs' years of experience at a given point in time. Yellow numbers indicate HTs that are appointed after the start of the project.

Planned data collection

Procedure

To be able to conduct pre-post comparisons, pre data needs to be collected before the training is introduced to the first cohort (i.e., the intervention group). At this measurement point t₀, data for the schools that will form cohort 3 (i.e., control group) should also be collected. This data will form the baseline, allowing control for potential pre-intervention differences between groups. To assess the effects of the intervention, multiple post-training measurement points should be included. For the comparison between cohorts 1 and 3, this would be t₁ and t₂, as indicated by the purple frames in the table above. Note that the t₂ assessments should also take place before cohort 3 receives training.

Apart from group comparisons (intervention vs. control group), it will also be possible to run regression analyses to investigate specific hypotheses of each subtopic with the collected data. For this, it would be ideal to assess all indicators in the available groups at all time points. If this is not feasible, compromises are possible (though not optimal).

Measures

In the following, we present the measures needed to test the overarching hypotheses and the specific hypotheses for each subtopic. We follow the same structure as in the presentation of the hypotheses.

In general, for all the group comparisons, the exposure to the training programme (yes/no) will serve as the independent/manipulated variable. The dependent variables will differ by hypothesis. Additionally, indicators for changes in leadership behaviours can be investigated as independent variables to predict certain student outcomes.

Overarching hypotheses

The overarching hypotheses consider the overall effect of the CPD programme. The crucial dependent variables are the indicators of long-term student outcomes, described in Table 4.

Table 4. Key dependent variables

Group of indicator	Name of indicator	Description	Comments
Long-term student outcome	Exam results	Results of the national exams	Consider average grades for different school subjects (e.g., English, Maths) separately
Long-term student outcome	Pass rates	SIP: Students' pass rate	Ensure data is assessed even if schools do not complete SIP (i.e., before training, in control schools). Consider average pass rates across school subjects
Long-term student outcome	Dropout rate	SIP: Students' drop-out rate per grade at all levels	Ensure data is assessed even if schools do not complete SIP (i.e., before training, in control schools)
Long-term student outcome	Attendance rate	SIP: Students' attendance rate at all levels	Ensure data is assessed even if schools do not complete SIP (i.e., before training, in control schools). If too unreliable: Consider regular school visits to collect data

Subtopic 1: Instructional quality

The main research question of Subtopic 1 is ***How does the CPD programme for school leaders influence the quality of instruction in schools?*** To assess the quality of instruction and its effects on student outcomes, the variables listed in Table 5 are needed.

Table 5. Variables for Subtopic 1

Group of indicator	Name of indicator	Description	Comments
Short-term HT outcome	Creating strategic direction for the school	SL Assessment Tool Rating 1.2.1	Must be collected for both intervention and control groups
Short-term HT outcome	Staff supervision and support	SL Assessment Tool Ratings 3.1.1-3.10; 3.2.4; 3.2.5	Must be collected for both intervention and control groups
Short-term HT outcome	Leadership for learning	SL Assessment Tool Ratings 2.1.3, 2.1.4	Must be collected for both intervention and control groups
Short-term HT outcome	Completion of SIP	Indicator whether school leader completed SIP (yes/no)	Variable needs to be generated by the research staff (based on whether SIP data is available or not) Must be collected for both intervention and control groups
Long-term student outcome	Exam results	Results of the national exams	Consider average grades for different school subjects (e.g., English, Math) separately. Must be collected for both intervention and control groups

Long-term student outcome	Pass rates	SIP: Students' pass rate	Consider average pass rates across school subjects. Must be collected for both intervention and control groups
Suggested extra assessment			
Short-term teacher outcome	Teacher skills and knowledge	Teacher survey	More details are provided below under teacher instructional quality. Must be collected for both intervention and control groups.
Short-term teacher outcome	Teacher skills and knowledge (student rating)	Student survey	More details are provided below under teacher instructional quality.
Short-term HT outcome	School leader leadership practices	Teacher survey	More details are provided below under school leader leadership practices.
Short-term HT outcome	School leader time use	SL Assessment Tool Extension: Observational Study	More details are provided below under school leader leadership practices.

School leader management and instructional practices

School leader management and instructional practices will be assessed via a teacher survey. Using a Likert scale, the survey will ask teachers to respond to a series of statements specifically aligned to VVOB expectations, such as:

- “My headteacher clearly articulates an instructional vision for the school” (School Leadership Assessment Toolkit indicator 1.1.1);
- “I understand my role to implement the School Improvement Plan” (School Leadership Assessment Toolkit indicator 1.2.3);
- “I have a positive relationship with my headteacher” (School Leadership Assessment Toolkit indicator 2.1.2);
- “My headteacher regularly reviews and provides feedback on my teaching plans” (School Leadership Assessment Toolkit indicator 2.1.4);

- “I have received professional development to implement the school’s instructional goals” (School Leadership Assessment Toolkit indicator 3.1.3).

In addition, if feasible, a school leader time-use observational study can be included as a component of the School Leader Assessment Toolkit. This would require onsite researchers to record school leader activities at regular intervals for the entire school day, over a series of days. In order to analyse this data, researchers would code school leader activities in particular categories, including instruction, administration, etc. School leader time spent on instruction would be compared between intervention and control group schools, to determine whether participation in the intervention leads to more school leader time spent on instruction.

Teacher instructional quality

There are a range of methodological challenges to measuring teacher instructional quality. There is no clear consensus in the field of education research on how best to measure instructional quality. Some of the most commonly used measures include classroom observations, student feedback surveys, and assessments of teacher impact on student learning. The LEAP fellow team does not recommend classroom observations at this time, given that the goal of this study is to determine the overall impact of CPD participation on student outcomes, rather than the specific impact of the program on instructional quality.

In the present study, the LEAP team recommends incorporating an assessment of teacher skills and knowledge in the teacher survey mentioned above. Survey questions should be designed to specifically assess instructional practices included in the LEAD programme. For example, mathematics-focused questions could include, “When introducing new content, I link new concepts to students’ prior learning” or “In all lessons, I ask students to explain their answers.” In constructing the teacher assessment of instructional practices, we recommend reviewing existing [TIMSS](#) (Trends in International Mathematics and Science Study) questionnaires.

If possible, this survey can be given in parallel to a student survey of teacher instructional practices. This survey should be given to students with sufficient literacy skills to understand survey questions, ideally those in primary grade levels 4-6. Aggregate classroom-level results from the student survey can support the validity of teacher survey findings. Questions should be aligned to the same practices assessed by the teacher survey, such as “My teacher connects new learning to things I already know.”

There is a risk of social desirability bias in both the teacher and student surveys recommended. To reduce this bias, it is critical that surveys are conducted anonymously, ideally in an online format in which research subjects have no interaction with researchers.

Subtopic 2. School environment

The main research question of subtopic 1 is ***How does CPD programme for school leaders influence the school environment?*** To assess the school environment and its effects on student outcomes, the variables listed in Table 6 are needed.

Table 6. Variables for Subtopic 2

Group of indicator	Name of indicator	Description	Comments
Short-term HT outcome	Communication of school vision and goals	SL Assessment Tool Ratings 1.1.1; 1.1.6	Must be collected for both intervention and control groups
Short-term HT outcome	Building a safe environment	SL Assessment Tool Ratings 2.2.1-2.2.8	Must be collected for both intervention and control groups
Short-term HT outcome	Ensuring student attendance	SL Assessment Tool Ratings 2.1.7	Must be collected for both intervention and control groups
Short-term HT outcome	Teacher evaluation	SL Assessment Tool Ratings 3.1.2	Must be collected for both intervention and control groups
Short-term HT outcome	Facilitating collaborative learning environment for all teachers	SL Assessment Tool Ratings 3.2.1	Must be collected for both intervention and control groups
Short-term HT outcome	Building a learning-conducive environment	SL Assessment Tool Ratings 4.2.1-4.2.4; 6.1.1-6.1.8	Must be collected for both intervention and control groups
Short-term teacher outcome	Turnover of teachers	SIP Extension: Turnover of teachers	More details are provided below under teacher turnover rate. Must be collected for both intervention and control groups
Short-term HT outcome	Completion of SIP	Indicator whether school leader completed items on availability of motivational opportunities for school staff and the number of functional	Variable needs to be generated by the research staff (based on whether the SIP data is available or not)

		departments/ SBI groups in SIP	Must be collected for both intervention and control groups
Long-term student outcome	Dropout rate	SIP (Part 4, School performance): Student dropout rate	Must be collected for both intervention and control groups
Long-term student outcome	Dropout rate	SIP: Students' drop-out rate per grade at all levels	Must be collected for both intervention and control groups
Long-term student outcome	Attendance rate	SIP: Students' attendance rate at all levels	If too unreliable: Consider regular school visits to collect data. Must be collected for both intervention and control groups
Suggested extra assessment			
Short-term teacher outcome	Teacher job satisfaction	Teacher survey rating job satisfaction	More details are provided below under teacher job satisfaction. Must be collected for both intervention and control groups
Short-term HT outcome	Teacher perception of professional environment	Teacher survey	More details are provided below under school leader leadership practices. Must be collected for both intervention and control groups
Short-term HT outcome	Teacher perception of learning environment	Teacher survey	More details are provided below under school leader leadership practices. Must be collected for both intervention and control groups
Short-term HT outcome	Student perception of learning environment	Student survey	More details are provided below under school leader leadership practices.

			Must be collected for both intervention and control groups
Short-term student outcome	Student engagement	Student survey	More details are provided below under student engagement. Must be collected for both intervention and control groups

School leader leadership practices in developing...

... a strong professional environment

School leaders' abilities in developing positive working conditions will be assessed via a teacher survey. The survey can take reference to Kraft and Papay (2014)'s design to measure the overall professional environment at schools, with survey items such as: "Teachers are provided opportunities to learn from one another" and "Teacher performance evaluations are handled in an appropriate manner."

... a positive learning environment

The leadership practices will be assessed through a teacher survey. We can combine this survey with the one which measures the teachers' work environment, serving as an overarching measurement of the school climate. Survey items related to school safety and culture for academic excellence should be added, such as "School leaders or administrators consistently enforce rules for student conduct" and "I hold high expectations for individual student learning and behaviour" (Kraft & Papay, 2014; Leithwood & Jantzi, 1999).

If possible, students' perceptions of the school environment should also be sought, as their perceptions might differ from teachers (Liu et al., 2016). We can ask students to respond to statements regarding school safety and attitudes towards learning and schooling, such as "I feel safe at school" and "trying hard at school is important" (Ward, 2018).

Teacher job satisfaction

We also recommend including items to measure teacher job satisfaction in the teacher survey. We can ask teachers to respond to the statement(s) by selecting among four Likert-scale response options (*strongly disagree to strongly agree*). It can be a general statement such as "I am generally satisfied with my job at this school" (Grissom, 2011; Pettersson et al., 2015).

There is a risk of estimation bias as teachers might feel reluctant to report genuinely about their satisfaction, fearing the answers might be reported to HTs (Rawle et al., 2019). To reduce the bias, we can also formulate multiple statements to gauge their satisfaction more accurately and

set up another survey item to understand with which specific aspects of their jobs they are satisfied (e.g. [teacher questionnaire of the PISA for Development](#)).

Teacher turnover rate

To measure the teacher turnover rate, we recommend administering a survey to HTs or adding items to the School Improvement Plan to report the number of teachers leaving schools after an academic year, and for what reasons, such as: “transferred to another school within the district”, “transferred to another school across districts”, “left the teaching profession” (Grissom & Bartanen, 2019; Zeitlin, 2021).

We also recommend disaggregating the number of teachers who quit their job according to their years of teaching experience. Zeitlin (2021) found that in Rwanda, early-career teachers (e.g. with 0-1 year of teaching experience) are more likely to quit their job. Given Rwanda's high teacher turnover rate (Zeitlin, 2021), it would be helpful to further investigate how participating HTs influence teacher attrition according to the years of tenure.

Student engagement

Student engagement can be dissected into behavioural and affective components, which, for example, measure students' participation in school activities and their identification with school (Leithwood & Jantzi, 1999; Willms, 2003). We recommend administering a student survey to measure both components, with survey items such as "I put a lot of energy into my school work" and "I feel excited by my work at school" (Fredricks et al., 2005; Leithwood & Jantzi, 1999).

In order to reduce the social desirability bias, we can train enumerators to administer the survey to students in the absence of teachers and HTs.

Subtopic 3. Parent/community engagement

The main research question of Subtopic 3 is ***How does the CPD programme for school leaders influence parents' engagement in their children's learning?***. To assess parents' engagement and its effects on student outcomes, the variables listed in Table 7 are needed.

Table 7. Variables for Subtopic 3

Group of indicator	Name of indicator	Description	Comments
Short-term parent outcome	Scholastic materials	SIP: Percentage of parents who provide scholastic materials for their children	Must be collected for both intervention and control groups
Short-term parent outcome	Attendance of meetings and activities	SIP: Percentage of parents who attend school meetings and percentage of parents who attend other school activities	Average of both indicators. Must be collected for both intervention and control groups
Short-term parent outcome	School visits	SIP: Percentage of parents who visit their children at school	Must be collected for both intervention and control groups
Short-term HT outcome	Partnership with parents and community	SL Assessment Tool Ratings 5.2.1-5.2.6	Must be collected for both intervention and control groups
Long-term student outcome	Pass rates	SIP (part 3, student performance data): Students' pass rate	Consider average pass rates across school subjects. Must be collected for both intervention and control groups
Long-term student outcome	Dropout rate	SIP (part 4, school performance): Students' drop-out rate per grade at all levels	Must be collected for both intervention and control groups
Long-term student outcome	Attendance rate	SIP (part 4, school performance): Students' attendance rate at all levels	If too unreliable: Consider regular school visits to collect data. Must be collected for both intervention and control groups

Suggested extra assessment			
Short-term HT outcome	The number of teachers who received CPD activities on parental engagement	Can be added into SIP (part 4, school performance)	Must be collected for both intervention and control groups

Potential challenges and solutions

1. *What if we want to see longer-term effects of the programme, for instance after three years?*

In this case it might be an option to not train the control group in cohort 3 in 2025 but instead in 2026 (i.e., move them to cohort 4). The clear drawback is that by the time they will receive training, they will have 3-4 years of experience as a school leader.

A weaker alternative might be to analyse the data in a dose-response fashion: that is, let the control group take part in the training in 2025 and compare them to cohorts 1 and 2, which have received a “higher dose” of training (i.e., HTs in these cohorts have received the training earlier on and might be expected to continue to lead according to the training). However, since this is a strong assumption, we do not recommend this alternative.

2. *What if HTs change schools during or shortly after the intervention?*

If HTs change schools during or shortly after the training, try to include the new school into the intervention group. The biggest challenge is that you will probably not be able to collect pre-data of this new school. In case you know about the change before it takes place, the ideal strategy would be to collect data before the new HT arrives or immediately after they have arrived (before they can be expected to make major changes based on the training). The former school of this HT might be used in the control group instead (depending on whether the new HT will be trained or untrained). In case the situation is more complicated, the cleanest option would be to exclude both, the former and the new school of the HT. In addition, descriptive data of why HTs left their school (e.g., transfer to another school) could be collected to provide details on the reasons for instability in schools, especially when turnover is high. Also, if there is a critical amount of schools with HT changes, one might run an analysis to compare schools with versus without HT turnover regarding student outcomes.

3. *What if the sample size of the comparable parts of cohort 1 and cohort 3 is too low (i.e., if there are few HTs with 0-1 years of experience in 2023 when the project starts)?*

In case there will not be enough schools available, increasing the sample size with data from other cohorts might be an option. For the comparison of training and control group after the first year (i.e., at t1), data from HTs who had 1-2 years of experience and were included in cohorts 1 and 2 might be added to the sample (see blue dashed frames in

Table 3). For this, it is important that the HTs with 1-2 years of experience at the start of the project in 2023 are comparable between cohorts 1 and 2 (i.e., they should be matched, similarly to the HTs with 0-1 years of experience in cohorts 1 and 3). An option to increase the sample for comparisons across two years could be to add the groups marked with the green dotted frames in the table above to the sample. If this is desired, we recommend assigning all of the HTs with 0-1 years of experience in 2024 to cohorts 2 and 4, not 3, to maximise the sample size. Again, the schools of these HTs should be matched between cohort 2 and 4. In the analyses, this combination of samples can be considered by entering the different points of intervention as covariate.

4. *What if other programmes are started during our project that we cannot match cohorts for pre data collection?*

We recommend asking at every measurement point whether the respective school has taken part in any other programmes; if so, what the name and target content of this programme were. This information can then be analysed descriptively and if there are significant differences in the amount of schools that received other training between the intervention and control group, this information can be included in the analyses as covariate.

Deliverable 3: Data Analysis Plan

Introduction

This data analysis plan provides recommendations on how the collected data might be analysed to test the specified hypotheses for each subtopic. Therefore, the following section first provides some general recommendations before turning to each subtopic specifically. For each subtopic, we first suggest preliminary analyses regarding the implementation fidelity before providing ideas on how to test each hypothesis.

General recommendations

Overall, there are two major approaches to analysing the data VVOB will collect; both can be useful, depending on which questions one aims to answer.

The first approach focuses on group comparisons to answer questions, such as “Do schools with HTs that were trained in the CPD Programme have lower dropout rates than schools without trained HTs?”. This approach helps to answer overall impact questions. A classical type of analysis to answer such questions is the mixed Analysis of Variance (ANOVA). In (quasi-) experimental designs there is usually a between-subjects factor, that is, the group membership to either the intervention or the control group. The second factor is typically the measurement point (pre and post intervention). Then the mean scores of the variable of interest (e.g., dropout rate) can be compared based on these factors, to find out whether the intervention group had a significant decrease in dropout rate relative to before the intervention and compared to the control group. This type of analysis is very straightforward if only the ultimate outcomes are of interest and not so much the processes in between that led to these outcomes. Note that a mixed ANOVA requires certain assumptions to be fulfilled (e.g., homogeneity of variances). In cases where these are not met, the Friedman test might be a non-parametric alternative.

As an alternative to a mixed ANOVA, it is also possible to calculate difference scores of the pre- and post-assessments and to compare the difference scores of the two groups with t tests or non-parametric alternatives in case certain assumptions are not met.

Also, one can run the analysis in two steps: First, compare whether there are significant differences between the control and intervention group in the pre scores. If this is not the case, the post score can be compared between the two groups. If there are significant differences in the pre scores, these need to be taken into account by comparing the pre-post-difference scores between the two groups.

Note that these analyses are only feasible if pre and post data of both groups is available. If only post data of both groups is available, this can be compared between groups but with the limitations that it is unclear if group differences are due to the intervention or existed before. If pre and post data of only the intervention group is available, pre-post comparisons can indicate change but with the limitation that we do not know whether it is due to the intervention or other factors.

The second type of analysis focuses on relations between specific variables, for instance if changes in one variable can predict changes in another variable. This is useful to test hypotheses like “the more time HTs invest into instructionally-focused interactions with teachers, the better students rate teaching quality”. Different types of regression models can help to address such questions but in and of themselves they cannot prove causality. To examine if change in one variable leads to change in another variable (or vice versa), cross-lagged panel models and their extensions (Mulder & Hamaker, 2021) can be helpful. For this, measurements of the variables of interest across multiple time points are needed and the sample needs to be sufficiently large. If data on different levels (e.g., school, teacher, student level) is of interest, multilevel models should be used.

In order to gain a full picture of the effects in the programme, we recommend the use of both group comparisons and regression-type analyses. How exactly they might look will be detailed in the following section.

In case many HTs do not change their behaviour as expected (i.e., as described in the A hypotheses), it will be difficult to observe differences between the intervention and control group (because the intervention group did not implement the intervention as expected). If this occurs, it might be an option to split the intervention group into two groups: those schools with HTs who implemented what they learned in the programme and those who did not. Implementation fidelity can be estimated by School Improvement Plan completion, as well as receiving a score of 2 or 3 on all School Leader Assessment Toolkit indicators used to measure hypotheses A1 and A3. These two groups might then be compared regarding student outcomes, but also can be analysed to find out why one group of HTs was unsuccessful in putting the knowledge of the programme into practice.

Data analysis for the overarching hypotheses

The main aim of this analysis is to test whether students in schools participating in the CPD programme have lower dropout rates, higher attendance, higher pass rates, and better national exam results compared to students in non-participating schools. To this end, we recommend either running ANOVAs as described above or running t tests for independent samples to compare whether the difference scores of the national exam results of schools participating in the CPD programme are on average higher than in non-participating schools. We recommend considering the different school subjects separately to receive a more differentiated picture. However, we suggest adjusting the alpha level for multiple testing to avoid alpha error cumulation (Rubin, 2021). This can be done for instance by using the Bonferroni correction (e.g., if the planned alpha level is .05 and the planned number of tests is 5, the corrected alpha level is $.05/5 = .01$).

To compare the average pass rates, attendance rates, and drop out rates between schools with and without CPD programme participation, we also recommend either t-tests for independent samples on the pre-post-difference scores or mixed ANOVAs. Because these variables are

considered in individual hypotheses, no alpha level adjustment is necessary, so the standard alpha of .05 can be used.

Data analysis by subtopic

To provide insights into how the specific hypotheses of each subtopic could be analysed, we provide some examples and summarise them in the table below.

Table 8. Analysis methods for different hypotheses

Subtopic	Hypotheses	Variables	Analysis method
1	A1 + A3	Difference scores of the SL Assessment Tool Ratings (post minus pre measurement)	Compare difference scores of intervention and control group: t test
	A2	Difference scores of number of HTs that filled out the School Improvement Plan (post minus pre measurement)	Compare difference scores of intervention and control group: Pearson's chi-squared test
	B1	Difference scores of: (1) the number of classroom observations and (2) the number of teachers received CPD (post minus pre measurement)	Compare difference scores of intervention and control group: t test
	B2	Difference in the average scores in HT's quality of instructional feedback (post minus pre measurement)	Compare difference scores of intervention and control group: t test
	B3	Dependent: difference score in teachers' instructional quality (teacher and student survey; post minus pre measurement) Predictor: difference scores of: (1) the number of classroom observations and (2) the number of teachers received CPD (post minus pre measurement)	Linear regressions
	B4	Dependent: difference score of: (1) the national exam results OR (2) the pass rates (post minus pre measurement)	Linear regressions

		Predictor: difference score in teachers' instructional quality (teacher and student survey; post minus pre measurement)	
2	A1	Difference scores of the scores in teacher evaluation (post minus pre measurement)	Compare difference scores of intervention and control group: t test
	A2	Difference score of number of schools providing motivational opportunities for school staff (post minus pre measurement)	Compare difference scores of intervention and control group: Pearson chi-squared test
	A3	Difference scores of the scores in facilitating collaborative learning environments for all teachers (post minus pre measurement)	
	A4	Difference scores of the number of functional department/ SBI groups per teacher (post minus pre measurement)	Compare difference scores of intervention and control group: t test
	B1-B2	<p>Dependent: Difference score of teacher turnover rates (post minus pre measurement)</p> <p>Mediator: Difference score of teacher job satisfaction (post minus pre measurement)</p> <p>Predictor: Difference score of professional environment (post minus pre measurement)</p>	Mediation model (regression analyses)
	A5	Difference scores of the scores in building a safe environment (post minus pre measurement)	Compare difference scores of intervention and control group: t tests
	A6	Difference scores of the scores in building a learning-conducive environment (post minus pre measurement)	
	A7	Difference scores of the scores in ensuring student attendance (post minus pre measurement)	

	A8	Difference scores of the scores in communication of school vision and goals (post minus pre measurement)	
	B3-B4	<p>Dependent: Difference score of student dropout rates/ student attendance rates (post minus pre measurement)</p> <p>Mediator: Difference score of student engagement (post minus pre measurement)</p> <p>Predictor: Difference score of learning environment (post minus pre measurement)</p>	Mediation model (regression analysis)
3	A1	Difference scores in SL Assessment Tool 5.2.1-5.2.6; number of teachers receiving CPD on parental engagement	Compare difference scores of intervention and control group: t tests
	A2	Difference scores in SL Assessment Tool 5.2.2	
	B1	<p>Dependent: Difference scores in student dropout rates/ student attendance rates (post minus pre measurement)</p> <p>Predictor: Difference scores in SL Assessment Tool 5.2.1-5.2.6 (post minus pre measurement)</p>	Linear regression
	B2	Difference scores in pass rates (post minus pre measurement)	Compare difference scores of intervention and control group: t tests

Examples for Subtopic 1

The potential of CPD participation to influence instructional quality is contingent on HTs putting their newly gained knowledge into practice in the school building. To test whether they do so, hypotheses A1-A3 measure whether the changes in number of HTs who completed the School Improvement Plan (hypothesis A2) and the changes in the scores received in the SL Assessment Tool Ratings (indicators 1.2.1, 2.1.3, 2.1.4, 3.1.1-3.1.10, 3.2.4, 3.2.5; hypotheses A1 and A3) are higher in schools that received training than in schools that did not receive training.

To this end, we suggest running a Pearson’s chi-squared test to compare the changes in numbers of HTs with CPD programme participation that filled out (vs. did not fill out) the School Improvement Plan (post minus pre intervention) to the changes in number of the HTs without programme participation that filled out (vs. did not fill out) the SIP (post minus pre measurement);

hypothesis A2). For the SL Assessment Tool Ratings (hypotheses A1 and A3), we recommend calculating difference scores from the post and pre measurements and then to run t-tests for independent samples to compare these difference scores between the group of schools participating in the CPD programme to the group of non-participating schools.

To test whether HTs improve their instructionally-focused interactions with teachers in the course of the CPD programme to a larger degree than HTs that do not participate in the programme (hypotheses B1 and B2), we recommend calculating difference scores and comparing them between the two groups of HTs. Similar to above, the difference score should be calculated as, for example: Number of classroom observations after taking part in the CPD programme minus number of classroom observations before taking part in the CPD programme³. For the schools without CPD programme participation the same two measurement points will be subtracted. These difference scores can then be compared in a t test for independent samples. The same procedure can be applied to test changes in the number of teachers that received CPD, and changes in the quality of instructional feedback. For the latter, we recommend calculating an average score per school leader and measurement point based on the teacher survey and to calculate the difference score based on these average scores.

To test whether improvements in teachers' instructional quality are associated with increased support from HTs (hypothesis B3), we recommend running a linear regression on difference scores: The difference between HTs pre and post training support and feedback to teachers can be used to predict the difference in teachers' instructional quality pre and post training. To this end, both aggregated student survey data and teacher survey data can be included as dependent variables in two different regression models.

To test hypothesis B4, whether improved teaching quality is associated with increases in national examination results and student pass rates, we suggest running a linear regression using the difference score of the aggregate student survey data by school to predict the difference score of the national exam results and the difference score of the pass rates of the respective school.

This can be done for the whole sample and separately for the groups of schools with versus without CPD to disentangle whether improvements in teaching quality always lead to better student outcomes or whether this effect is specific to schools participating in the CPD programme.

More granularly, a similar test can be run at the teacher level. Teachers in schools participating in the CPD programme would be split into two groups: those with a positive pre-survey/post-survey difference score versus those with a neutral or negative difference score. In order to run this test, student survey results, pass rates, and national examination

³ Note that the difference score can be based on the difference between pre intervention and post intervention year 1 or year 2. We recommend exploring both because some changes might not become visible after one year.

results will need to be matched to the teacher.⁴ This test will enable VVOB to understand how changes in instructional quality impact changes in student outcomes by differentiating between teachers who improve their practice based on school leader feedback versus those that do not.

If a student survey is not feasible, VVOB can use results from the teacher survey to approximate changes in instructional quality. If VVOB elects this option, we recommend focusing only on questions in the teacher survey that are related to instructional improvement. Using these survey findings, we suggest running a linear regression using the difference score of the aggregate teacher survey data by school to predict the difference score of the national exam results and the difference score of the pass rates of the respective school.

Examples for Subtopic 2

School environment for teachers

The potential of CPD participation to influence teacher satisfaction and turnover is contingent on HTs applying their newly gained knowledge to develop the working conditions. To find out if this is the case, we recommend testing hypotheses A1-A4 again using t tests for independent samples to compare the pre-post-difference scores of the group of schools that received CPD compared to the schools that did not regarding their A1) scores⁵ in teacher evaluation, A2) percentage of schools providing motivational opportunities for school staff, A3) scores in facilitating collaborative learning environments for all teachers, and A4) number of functional department/ school-based in-service professional learning groups (SBI groups) per teacher.

To test hypotheses B1 and B2, we recommend running a mediation model according to Baron and Kenny (1986). As a first step, a linear regression with the pre-post-difference score of teacher-perceived professional environment predicting the pre-post-difference score of turnover rates needs to be run. Then, a linear regression with the pre-post-difference score of professional environment predicting the pre-post-difference score of teacher job satisfaction needs to be computed. Finally, both the the pre-post-difference score of professional environment and the pre-post-difference score of teacher job satisfaction need to be entered in a regression model to predict the pre-post-difference score of teacher turnover rates. If the prediction power of the the pre-post-difference score of professional environment significantly decreases or entirely disappears once the pre-post-difference score of teacher job satisfaction is added to the model, the latter (partially) mediates the relationship between the the pre-post-difference score of professional environment and the pre-post-difference score of teacher turnover rate.

⁴ If disaggregating student data by teacher is not feasible, VVOB can ignore the teacher-level analysis. Aggregate data at the school level will serve the same purpose of measuring this hypothesis.

⁵ The scores mentioned in A1 -A8 in subtopic 2 refer to the difference scores computed according to the method mentioned in Subtopic 1.

School environment for students

The potential of CPD participation to influence student engagement and dropout is contingent on HTs applying their newly gained knowledge to build a conducive learning environment. To find out if this is the case, we recommend testing hypotheses A5-A8 again using t-tests for independent samples to compare the pre-post-difference scores between the group of schools that received CPD compared to the schools that did not regarding their A5) scores in building a safe environment, A6) scores in building a learning-conducive environment, A7) scores in ensuring student attendance, A8) scores in communication of school vision and goals.

To test hypotheses B3 and B4, we recommend running a mediation model according to Baron and Kenny (1986). As a first step, linear regressions with the the pre-post-difference score of learning environment predicting the pre-post-difference score of dropout rates and the pre-post-difference score of student attendance rates needs to be run separately. Then, a linear regression with the the pre-post-difference score of learning environment predicting the pre-post-difference score of student engagement needs to be computed. Finally, both the pre-post-difference score of learning environment and the pre-post-difference score of student engagement need to be entered in regression models to predict the pre-post-difference score of student dropout rates and the pre-post-difference score of student attendance rates respectively. If the prediction power of the the pre-post-difference score of learning environment significantly decreases or entirely disappears once the pre-post-difference score of student engagement is added to the models, the latter (partially) mediates the relationship between the pre-post-difference score of the learning environment and the pre-post-difference score of student dropout or the pre-post-difference score of attendance rate.

Examples for Subtopic 3

The potential of CPD participation to influence student pass rates and dropout is contingent on HTs applying their newly gained knowledge to engage with the parents and community. To find out if this is the case, we recommend testing hypotheses A1 and A2 again using t-tests for independent samples to compare the group of schools that received CPD compared to the schools that did not regarding the difference scores in the mentioned indicators (A1: SL Assessment Tool 5.2.1-5.2.6; number of teachers receiving CPD on parental engagement; A2: SL Assessment Tool 5.2.2).

Similar to the analyses above, we recommend running linear regressions to predict changes in attendance and dropout rates from changes in the SL Assessment Tool indicators by using pre-post-difference scores (5.2.1-5.2.6; i.e., hypothesis B1). Regarding hypothesis B2, we recommend for now to only investigate overall changes in pass rates (i.e., as suggested in the analyses of the overarching hypotheses above). As this is not the main focus of the current project, we recommend to reserve more detailed investigations for future research.

Conclusion

This proposed study design is crafted to enable VVOB to understand the overarching impact of the CPD programme on student outcomes, as well as the mechanisms of that impact. The study design uses a range of methods to test each hypothesis included in the conceptual framework. These analyses will support VVOB in gaining a clearer understanding of the areas in which the programme is having a particularly strong impact, as well as areas that can be adjusted in future iterations of the CPD programme. A review of the research literature demonstrates that the CPD programme is exceptionally well-aligned with the evidence base on effective school leadership programmes. As a result, the LEAP Fellow team hypothesises that results from the proposed study will demonstrate a positive impact on student outcomes.

The LEAP Fellow team would like to express our sincere appreciation to the VVOB team for their partnership throughout this engagement. We are inspired by your tremendous programming and your deep commitment to ensuring a quality education for every child.

Bibliography / References

- Abdalla, M., Mwingi, M., Wachira, N., Okoko, J. M., & Webber, C. F. (2019). School Leadership Preparation in Tanzania. *Preparation and Development of School Leaders in Africa*, 111. http://scholar.google.com/citations?view_op=view_citation&hl=en&user=x5zOw98AAAAJ&sortby=pubdate&citation_for_view=x5zOw98AAAAJ:RGFaLdJalmkC
- Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Chetty, Raj, Friedman, N., & Rockoff, J.E., “Measuring the Impact of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood,” *American Economic Review* Vol. 104, No. 9, May 2014
- Fredricks, J. A., Blumenfeld, P., Friedel, J., & Paris, A. (2005). School Engagement. In K. A. Moore & L. H. Lippman (Eds.), *What Do Children Need to Flourish? Conceptualizing and Measuring Indicators of Positive Development* (pp. 305–321). Springer US. https://doi.org/10.1007/0-387-23823-9_19
- Fryer, R. G. (2017). *Management and Student Achievement: Evidence from a Randomized Field Experiment*.
- Griffith, J. (2000). School Climate as Group Evaluation and Group Consensus: Student and Parent Perceptions of the Elementary School Environment. *The Elementary School Journal*, 101(1), 35–61. <https://www.jstor.org/stable/1002334>
- Grissom, J. (2011). Can Good Principals Keep Teachers in Disadvantaged Schools? Linking Principal Effectiveness to Teacher Satisfaction and Turnover in Hard-to-Staff Environments. *Teachers College Record*, 113(11), 2552–2585. <https://www.tcrecord.org/Content.asp?ContentId=16144>
- Grissom, J. A., & Bartanen, B. (2019). Strategic Retention: Principal Effectiveness and Teacher Turnover in Multiple-Measure Teacher Evaluation Systems. *American Educational Research Journal*, 56(2), 514–555. <https://doi.org/10.3102/0002831218797931>
- Grissom, J., Egalite, A., & Lindsay, C. (2021). *How Principals Affect Students and Schools. A Systematic Synthesis of Two Decades of Research*. 136.
- Hopkins, D., Harris, A., & Jackson, D. (1997). Understanding the School’s Capacity for Development: Growth states and strategies. *School Leadership & Management*, 17(3), 401–412. <https://doi.org/10.1080/13632439769944>
- Jacobson, S. L., Brooks, S., Giles, C., Johnson, L., & Ylimaki, R. (2007). Successful Leadership in Three High-Poverty Urban Elementary Schools. *Leadership and Policy in Schools*, 6(4), 291–317. <https://doi.org/10.1080/15700760701431553>
- Karareba, G., Clarke, S., & O’Donoghue, T. (2019). Leading Rwandan Primary Schools: Some Deliberations on the Past, Present, and Future. *International Journal of Educational Reform*, 28(1), 79–98. <https://doi.org/10.1177/1056787918824195>
- Kraft, & Papay, J. P. (2014). Can Professional Environments in Schools Promote Teacher Development? Explaining Heterogeneity in Returns to Teaching Experience. *Educational Effectiveness and Policy Analysis*, 36(4), 476–500.

- Leaver, C., Lemos, R., & Scur, D. (2019). *Measuring and explaining management in schools*: 83.
- Leithwood, K., & Jantzi. (1999). The Relative Effects of Principal and Teacher Sources of Leadership on Student Engagement with School. *Educational Administration Quarterly*, 35(5), 675–830. <https://doi.org/10.1177/0013161X99355002>
- Liu, K., Springer, J., Stuit, D., Lindsay, J., & Wan, Y. (2016). The Utility of Teacher and Student Surveys in Principal Evaluations. Stated Briefly. REL 2016-132. In *Regional Educational Laboratory Midwest*. Regional Educational Laboratory Midwest. <https://eric.ed.gov/?id=ED611587>
- Ma, X. (2003). Sense of Belonging to School: Can Schools Make a Difference? *The Journal of Educational Research*, 96(6), 340–349. <https://doi.org/10.1080/00220670309596617>
- Mulder, J. D., & Hamaker, E. L. (2021) Three Extensions of the Random Intercept Cross-Lagged Panel Model, *Structural Equation Modeling: A Multidisciplinary Journal*, 28:4, 638-648, DOI: [10.1080/10705511.2020.1784738](https://doi.org/10.1080/10705511.2020.1784738)
- Pettersson, G., Rawle, G., Outhred, R., Brockerhoff, S., Nugroho, D., Jasper, P., Kveder, A., & Beavis, A. (2015). *EQUIP-Tanzania Impact Evaluation*. 174.
- Printy, S. (2010). Principals' influence on instructional quality: insights from US schools, *School Leadership & Management*, 30:2, 111-126, DOI: 10.1080/13632431003688005
- Rawle, G., Binci, M., Pettersson, G., Harb, J., Jasper, P., Khan, S., Medardi, D., Romarri, A., Rorich, M., & Ruddle, N. (2019). *EQUIP-Tanzania Impact Evaluation Endline Quantitative Technical Report, Volume I*. Oxford Policy Management. <https://www.opml.co.uk/files/Publications/8383-assessing-equip-t/opm-ie-el-quant-report-vol-i-final-sent-rev.pdf?noredirect=1>
- Rubin, M. When to adjust alpha during multiple testing: a consideration of disjunction, conjunction, and individual testing. *Synthese* 199, 10969–11000 (2021). <https://doi.org/10.1007/s11229-021-03276-4>
- Rwandan Ministry of Education. (2019). *Education sector strategic plan 2018-2024—Rwanda*. <https://www.globalpartnership.org/content/education-sector-strategic-plan-2018-2024-rwanda>
- Sampat, S., Nagler, N., & Prakash. (2020). *Evidence Review Report: A Review of Empirical Research on School Leadership in the Global South*. Global School Leaders.
- Stander, A. S., & Stander, M. W. (2016). RETENTION OF EDUCATORS: THE ROLE OF LEADERSHIP, EMPOWERMENT AND WORK ENGAGEMENT. *International Journal of Social Sciences and Humanity Studies*, 8(1), Article 1. <https://dergipark.org.tr/en/pub/ijsshs/issue/26211/275963>
- Ward, M. (2018). *PISA for Development: Results in Focus* (PISA in Focus No. 91; PISA in Focus, Vol. 91). <https://doi.org/10.1787/c094b186-en>
- Willms, J. D. (2003). *Student engagement at school: A sense of belonging and participation: results from PISA 2000*. OECD.
- Zeitlin, A. (2021). Teacher Turnover in Rwanda. *Journal of African Economies*, 30(1), 81–102. <https://doi.org/10.1093/jae/ejaa013>