

VVOB BETTER Midterm Evaluation

Executive Summary

By Rachel Outhred, Shefali Rai and Lydia Marshall

14 January 2025



www.oxfordmeasured.co.uk
Rachel.outhred@oxfordmeasured.co.uk

Acronyms

Acronym	Full form
ABC Catch Up	Achieving Basic Competencies through Catch Up
BETTER	Better Education through Teacher Training, Evidence and Research
CMO	Context-Mechanism-Outcome
CPD	Continuous Professional Development
DGD	Belgian Directorate-General Development Cooperation and Humanitarian Aid
DOET/BOET	Department/Bureau of Education and Training
EGRI	Early Grade Reading Instruction
ESC	Education Scalability Checklist
FGD	Focus Group Discussion
INSET	In-Service Education and Training
KII	Key Informant Interview
LEAP	Learning Entrepreneurship and Agriculture Practically
MEL	Monitoring, Evaluation, and Learning
MoES	Ministry of Education and Sport
MOET	Ministry of Education and Training
MoE	Ministry of Education
MoU	Memorandum of Understanding
PIRLS	Progress in International Reading Literacy Study
QA	Quality Assurance
REB	Rwanda Basic Education Board
SACMEQ	Southern and Eastern Africa Consortium for Monitoring Educational Quality
SDG	Sustainable Development Goal
SMILE	Strengthening Early Grade Maths through Inclusive and Level Appropriate Education
TALK	Preschool Teachers Apply Language-rich Teaching Skills and Knowledge
ToC	Theory of Change
¡VAMOS!	Linking Agro-learning for a Sustainable World

Executive Summary

Background

The BETTER programme, funded by the Belgian Directorate-General for Development Cooperation and Humanitarian Aid (DGD), aims to institutionalise and scale education initiatives (impactful practices for teaching and school leadership and associated professional development (PD) to support delivery of solutions). VVOB's efforts are designed to align with the global Theory of Change (ToC) and contribute to SDG 4 (quality education for all) by strengthening government capacity, promoting effective teaching and leadership, and scaling inclusive, equitable education systems.

Key objectives include institutionalising scalable education practices and fostering systemic change across eight diverse country programmes.

VVOB's approach is based on a generic ToC aimed at improving learning outcomes and well-being, contributing to SDG 4. The ToC includes two main tracks:

1. **Professional Development (Results Area 1):** Focuses on strengthening the capabilities of teachers and school leaders.
2. **Institutionalisation and Scaling (Results Area 2):** Aims to create the environment necessary for scaling and embedding initiatives.

The evaluation specifically focuses on **Results Area 2**, which emphasises institutionalising and scaling effective education practices. The overall ToC revolves around three pillars:

1. **Reinforcing Government Capacity:** VVOB collaborates with governments to help them implement education policies effectively, promoting equity and ensuring governments meet their education-related obligations.
2. **Investing in Effective Teaching and Leadership:** VVOB works to improve teacher and school leadership quality through evidence-based approaches, as high-quality teaching and leadership significantly impact student outcomes.
3. **Scaling Inclusive and Equitable Education:** VVOB emphasises the need to scale effective teaching and leadership practices. They partner with governments to develop, test, and scale initiatives, ensuring they are integrated into the education system.

In **Results Area 2**, VVOB engages in activities such as capacity development, research, monitoring and evaluation (M&E), and advocacy to:

1. Increase the credibility of initiatives through strong evidence.
2. Build strong coalitions for change.
3. Adjust the policy environment to facilitate the adoption of these initiatives.
4. Develop clear scaling strategies with sustainable funding.

VVOB also uses an **Education Scalability Checklist (ESC)** to identify and address obstacles in scaling education initiatives. Since 2021, this tool has been used by country teams, with opportunities to exchange experiences in an internal Community of Practice (COP) and share lessons learned to refine scaling strategies across VVOB's eight country programmes. Occasionally, the COP also invites external experts to deepen the reflection and share best practices and, hence, support country teams in their scaling process.

We have sought to use the following terms in line with their usage within VVOB.

- **Solution:** Sets of impactful practices for teaching or school leadership.
- **Professional development:** Training for teachers, school leaders or other cadres of staff to support the effective delivery of these solutions.
- **Initiative or PD package:** These consist of solution and associated professional development (PD).

VVOB seeks to support the effective implementation of education initiatives – both solutions and associated PD - at scale.

Evaluation purpose and objectives

The purpose of this Mid-Term Review is to understand what is required to institutionalise and sustain the education initiatives developed, piloted, and implemented in collaboration with stakeholders and to bring about real systemic change and impact at scale. The evaluation was conducted to answer four key questions:

1. What constitutes successful “scaling” for VVOB?
2. What efforts have BETTER programmes made towards institutionalisation and scaling?
3. To what extent are these efforts likely to lead to institutionalisation and scaling of education initiatives?
4. What factors have enabled and hindered programme towards institutionalisation and scaling?

Specifically, it aims to:

- Understand requirements for systemic change.
- Evaluate progress against goals.
- Generate learning for future programming.
- Enhance alignment with global and local goals.

Methodology

The overarching evaluation question is: **What is required to institutionalise and sustain VVOB's education initiatives to bring about systemic change and impact at scale?**

This was explored through four data collection methods:

- Desk-based review of programme documents.
- Remote interviews with global and country teams.
- In-depth case studies in Zambia, Uganda, and Vietnam.
- Sensemaking activities with country teams and a Community of Practice (COP) on scaling.

Key questions (KQs) were further broken down into 13 sub-questions. They explored the definition of scaling, progress made, contextual factors, effectiveness of activities, and future strategies.

The Review employed a Realist Evaluation approach, which examines how outcomes are triggered by mechanisms within specific contexts. This approach used Context-Mechanism-Outcome (CMO) hypotheses to understand the causal links between actions and results.

Data Analysis

Data was coded using the Framework approach, creating matrices for within-case and cross-case analysis. Insights were triangulated to understand programme effectiveness, challenges, and implications for institutionalisation and scaling. CMOs were tested throughout the evaluation by (1) documenting mechanisms from ToCs, (2) using case study data to assess how mechanisms influenced outcomes in various contexts; and (3) comparing hypotheses across cases.

Four sensemaking collaborative workshops were undertaken to verify findings, three of which took place as sensemaking workshops within-case analysis and two of which took place as sensemaking and validation workshops after across-case analysis. The global sensemaking workshop also explored broader implications and recommendations after cross-case analysis.

Ethics

The evaluation adhered to ethical guidelines set by UNICEF, UNEG, and other international standards. These include:

- Transparency and informed consent during data collection.
- Ethical handling of data and adherence to privacy standards.

This methodology ensured a robust, context-sensitive evaluation to understand and enhance VVOB's efforts in scaling and institutionalising education initiatives.

Findings 1: Progress along the ToC

Programme Theories of Change

As part of each of the three country case studies, country-level ToCs were mapped to the global ToC for the BETTER programme, focussing in particular on explicating the efforts that country teams are making to pursue scaling and institutionalisation, and unpacking how these are intended to contribute to the results and outcomes in the ToC.

Contexts

The following contextual factors identified as important:

- **Government commitment to education.** Significant government commitment to education is in place across almost all countries, however in some contexts, even where commitment is high, there are often systemic constraints to the effective functioning of the education system.
- **A specific need or gap in the country.** Across the eight countries, the interventions supported by VVOB meet a need or gap that is recognised by national stakeholders.
- **A broader agenda, movement or plan that can nest the initiative.** In several countries, there are broader agendas or movements that can nest the initiatives supported as part of the BETTER programme.
- **Previous success working in the sub-sector.** Across most countries, VVOB's efforts are supported by previous successes working with government.
- **Capacity in the public education system.** In several countries the availability and capabilities of staff at different levels of the public education system acts as a constraint to high-quality and impactful delivery of initiatives at scale.

Planned activities

A range of activities are being carried out to promote effective scaling and institutionalisation of the initiatives supported by the BETTER Programme, including:

- **Establishing formal engagement and approval structures with partners:** In several countries VVOB has helped set up formal structures to support planning and delivery of initiatives.
- **Use of informal engagement strategies:** Formal interaction is complemented by informal engagement in several countries. In Rwanda a member of the VVOB team is based at the Rwanda Basic Education Board (REB). In Zambia VVOB has an office at the Ministry of Education. In Uganda there are provisions for two VVOB team members to work at the MoES, which happens on a regular basis.
- **Engagement at the policy level:** In Rwanda, VVOB is working to ensure that the CPD programme is made mandatory and is enshrined in policy. In Uganda, the LEAP programme has engaged through formal and informal engagement structures to align with national policies, including the curriculum and teacher frameworks. In Vietnam, VVOB needs to tread carefully around mandate and ensure they are not seen by government to be attempting to adapt the policy environment.
- **Developing resources to support delivery at scale:** This activity is seen across most countries. Examples include the development of a CPD App and Learning Management System in Kenya, design of content for a diploma programme for school leaders in Rwanda, and creation of materials on early-grade reading in South Africa.
- **Building capacity through Technical Assistance and joint planning or delivery:** This activity is seen across many countries.
- **Establishing the scaling model:** A number of country teams are working on developing the mechanisms through which their initiatives will be scaled.
- **Monitoring implementation:** All eight countries carry out basic monitoring of delivery of the initiative using indicators such as the number of teachers, school leaders or master trainers trained, or the number of schools affected by delivery of the initiative.
- **Building quality assurance (QA) systems:** Several countries have some processes in place to assess the quality of delivery of their education initiatives, and the outcomes of these initiatives.
- **Activities to generate evidence that the initiative is needed:** Several countries conducted activities to generate evidence that the initiative is needed, while in others, such as Zambia and South Africa, such activities were not needed as VVOB is responding to a need that was already evident to stakeholders based on pre-existing data from global and regional learning assessments.
- **Activities to generate evidence of best practice for the initiative:** Several countries are working to generate evidence on what form the initiative should take.
- **Activities to generate evidence that the initiative is working:** In all countries activities are underway to generate evidence that the initiatives are working. However, in most cases the findings of these activities will only be available in the outer years of the programme (2025-26).
- **Aligning scaling ambitions and developing a compelling, sustainably-funded scaling strategy:** Several countries are engaging in activities that sit under the umbrella of developing scaling strategies.

Progress against planned activities

Across the eight countries, progress in implementing planned activities is largely on track as indicated by the completion of milestones in programme operational plans. However, for complex processes related to institutionalisation and scaling, completion of activity milestones tends to provide only a limited indication of likely progress along the ToC. The harder-to-measure qualitative aspects of these activities are critical. For instance, meetings may be held with key stakeholders, but it is the quality and nature of interaction at these meetings, and the level of engagement and interest from key actors that critically shape the extent to which the completion of activities will translate into results and outcomes further along the ToC.

Progress along the ToC is also likely to be shaped by the selection and timing of key activities. This varies across countries. Most country teams are using a mix of complementary strategies to pursue progress towards the outcomes in the ToC. Progress across several activities is noteworthy:

- **Generating evidence that the initiative is working:** In most countries, activities are taking place to generate evidence on the initiative's impact on students' learning outcomes. However, this evidence will only be available in the outer years of the programme (2025-26). There is a related risk that initiatives will be scaled before evidence is available on whether they are effective in improving outcomes for students.
- **Generating evidence of best practice for the initiative:** All country teams have reviewed existing evidence on the initiatives and types and modalities of PD that they promote, and many have done pilots or small research studies to generate evidence to inform the design of the initiatives. One key reflection question is whether such evidence is sufficient to shed light on the Minimum Viable Initiative: the core components that must be delivered well in order for the solution and associated PD to have impact at scale.
- **Building compelling, sustainably-funded scaling strategies:** As discussed above, most countries are making some efforts under this broad umbrella, but in most cases these are not comprehensive, and often not shared and agreed with partners.
- **Building QA Systems:** While most countries are making some efforts to monitor the quality of delivery of their initiatives, comprehensive QA systems have not been widely implemented from the beginning of programme implementation.

Results areas

The global ToC refers to four key results related to scaling and institutionalisation:

1. Strengthening the evidence base to make initiatives more credible.
2. Establishing coalitions that support the scaling and institutionalisation of initiatives.
3. Increased coherence in the policy/regulatory environment to support the uptake of initiatives.
4. Formulating clear, compelling scaling strategies with sustainable funding.

In Zambia, we see two further targeted results related to scaling and institutionalisation:

5. Integration of the initiative into the MoE's existing systems for pre-service training, INSET, school monitoring and data management. This is linked to point 4 above, in that such integration would lower the discrete costs of implementing the initiative, which in turn would support cost-effectiveness and sustainability.
6. Support to government partners to build structures and capabilities to sustain high-quality planning and delivery of the initiative.

A distinct feature of the Zambia programme is that Catch Up is being implemented across large parts of the country with funding from multiple donors (not just DGD). This has created some additional scope for the team to pursue a more ambitious institutionalisation agenda.

At midline, progress towards a more credible evidence base is rated as at risk of not being achieved across most countries. Across all eight countries, MEL and research is being done to generate evidence that the initiatives are working. However, the findings of evaluations and other research will only be available in the later years of the programme (2025-26).

The initiative being supported by a strong, steady coalition for change is rated as on track. At midline, we see significant progress across the three case study countries.

The limited evidence available that initiatives are working combined with extremely strong coalitions for change interacted in a problematic manner in some countries, whereby in the absence of evidence partners filled the void with anecdotal perspectives that the initiatives work, including unproven confidence in the success of the programme.¹

At midline we see considerable progress towards increased coherence/adjustments in the policy/regulatory environment.

Ensuring clear, compelling and sustainably funded scaling strategies are in place is assessed as at risk, at this Mid-Term Review point. In Zambia, VVOB has begun engaging with the government to agree on a long-term vision for the initiative, and to develop a costed plan and Investment Case grounded in this vision. In Uganda and Vietnam, there is not yet an aligned vision for scaling across partners.

Impact and outcomes

Looking across the four results areas, we see that while robust evidence is not yet available that initiatives are working, strong coalitions in support of these initiatives have already been established. This creates a risk that scaling will (or is already) taking place without *evidence* that the initiative is working. This risk is likely to apply across the eight countries in light of the timing of planned MEL activities. The magnitude of this risk may vary across countries.

Findings 2: Findings against evaluation questions

What constitutes successful “scaling” for VVOB?

Global Understanding

VVOB defines scaling as achieving lasting systemic change beyond merely expanding programme reach. It focuses on altering structures and mechanisms to create sustainable education systems. The approach emphasises:

- Working within government systems to embed initiatives.
- Elevating teachers and school leaders as agents of change.
- Maintaining a long-term presence in focus countries.

Scaling is conceptualised as “scaling an initiative” (ensuring quality, impact, and sustainability) rather than just “scaling a programme.”

¹ For example, in Vietnam, stakeholders across government (at all levels) were extremely positive regarding TALK and its likely impact. A group of stakeholders stated that the initiative is definitely working, citing extremely unlikely results being achieved for all children in all implementation districts.

VVOB employs three strategies to operationalise scaling, including:

1. Addressing significant, persistent educational challenges based on evidence.
2. Focusing on policy-aligned issues to secure government buy-in.
3. Building coalitions for broad-based support, engaging multiple stakeholders.

Country-Level Interpretations include;

- Institutionalisation: Embedding programmes in government systems through policy alignment and transferring ownership to local institutions.
- Scaling: Expanding programme reach while maintaining quality, often via pilots and gradual rollouts.
- Systemic Change: Deep integration into frameworks, requiring cultural, structural, and policy shifts with decentralisation and community involvement.

The key elements for success identified across global and country definitions included sustained quality and funding, integration into government systems, technical capacity within government to sustain high-quality delivery of the initiative, building evidence for programme credibility and adaptable scaling models.

VVOB country programmes pursue diverse scaling outcomes tailored to local needs and capabilities. These include expanding geographically to new regions, building local capacity to independently sustain initiatives, supporting shifts in ownership and transferring programme leadership to local governments or partners, embedding innovations into government policies and standards for systemic adoption, prioritising underserved or marginalised communities to reduce disparities and changing mindsets and practices among educators and administrators for long-term innovation adoption.

VVOB BETTER programmes are relevant and aligned with local contexts. Uganda's LEAP Programme aligns with competency-based curriculum reforms and integrates agriculture and entrepreneurship, addressing national goals. Zambia's Catch Up Programme targets foundational learning gaps with strong government demand for national scaling. Vietnam's TALK programme engages national experts, aligns with policies, and leverages provincial networks for credibility and scalability.

What efforts have BETTER programmes made towards institutionalisation and scaling?

Institutionalisation and scaling progress depend on local contexts, government support, and evidence generation. Common challenges include resource constraints, bureaucratic delays, and insufficient evidence to validate interventions.

Effective practices, such as strong government collaboration and capacity building, were identified as critical for long-term sustainability and systemic integration.

Each country has a unique ToC developed based on local contexts, with all ToCs aligned to VVOB's global vision of institutionalisation and scaling for systemic change.

In Uganda the ToC evolved from prior programmes like Teaching Agriculture Practically (TAP) and integrates practical skills in agriculture and entrepreneurship to address employability. Institutionalisation is emphasised through close collaboration with the Ministry of Education and Sport (MoES) and its emerging specialised agencies, focusing on affordability and sustainability of CPD programmes.

In Zambia the ToC focuses on addressing foundational learning gaps in primary education through the Catch Up programme. Integration into government systems is central, with

efforts to build capacity at all levels and align monitoring and evaluation processes with government structures.

In Vietnam, lessons from the previous BAMl programme informed the ToC, which added a focus on evidence generation as a result area. The programme engages national experts and champions to align interventions with government policies, particularly in early childhood education and language development.

The progress on scaling milestones across countries varies, with each country facing unique challenges and utilising different strengths. Delays in research and administrative aspects, as seen in Uganda and Vietnam, highlight the need for robust risk management and contingency planning to mitigate potential setbacks.

Common challenges in implementing institutionalisation and scaling activities include:

1. **Research and Evidence Generation:** Both Uganda and Vietnam faced delays related to research, underscoring the challenges of managing external partnerships finding qualified researchers, and the importance of timely data collection.
2. **Communication and Visibility:** Kenya encountered delays in activities aimed at visibility and dissemination (such as success stories and newsletters), highlighting the potential need for stronger communication planning.

While programme activities are mostly being delivered as intended across Uganda, Zambia, and Vietnam, there are notable limitations in evidence generation. In addition, resource constraints (both internal and within partner organisations) limit the teams' ability to fully achieve their scaling ambitions.

To what extent are these efforts likely to lead to institutionalisation and scaling of education initiatives?

Programmes show significant progress in terms of **partnership building**, **policy alignment**, and **capacity development**, but achieving **improved teaching** and **school leadership** at scale will depend on further **evaluation**, **monitoring**, and the ability to address existing challenges in **quality control** and **government alignment and support**.

Across case study countries, the evidence for the program's **relevance** and **need** has been successfully generated, but the long-term impact on teaching and leadership remains to be fully assessed.

We do not yet know if the context, and action is triggering a mechanism that *will* or *has* led to more effective school leadership and teaching at scale. This isn't just because we don't have the evidence yet but is because we are not yet exploiting all of the opportunities available, to increase the chances of triggering all of the mechanisms. With this said, we can consider *likely outcomes*.

We observe many exploited CMOs within the ToC that are holding true, however there are also unexploited CMOs that can be built on in order to increase the likelihood of achieving impact at scale. The more exploited CMOs within the ToC that hold, the higher the likelihood of BETTER programme outcomes being realised.

The broader agenda, movement or plan within a context and previous success working in the sub-sector through previous programmes can be leveraged to enact activities to monitor the quality of implementation and generate evidence that the initiative is working. These contextual factors can also be leveraged to enact activities to engage partners on the scaling strategy, aligning ambition and documenting a clear,

compelling and sustainably funded scaling strategy. These actions would feed into the results areas related to the initiative being more credible and having a clear, compelling and sustainably funded scaling strategy. Achieving these results areas would increase the likelihood of BETTER programmes achieving more effective leadership and teaching at scale.

When reviewing the exploited and unexploited mechanisms, Longer Term Results Areas that are currently implicit in many ToCs and explicit in Zambia's ToC might present other unexploited mechanisms. These include Longer Term Result Areas: (1) there is capacity at all levels of the system to implement the initiative and (2) the initiative is integrated into the system.

What factors have enabled and hindered progress towards institutionalisation and scaling?

Across case study countries, activities that have been delivered have been effective. What is more pressing is **understanding the activities that have not been done**, either for logistical reasons (procurement for research studies) or because they were not seen as so important or prioritised in busy schedules (scaling strategies). In both cases, this is linked to the ambition of programmes within short timeframes (perhaps with the exception of Zambia), and has led teams to consider (in sensemaking workshops) if timelines need to be extended or ambitions reduced within current timeframes.

The **ESC** and other tools have been used in various ways across countries, reflecting different stages of programme development, institutionalisation, and scaling.

Overall, the **ESC tool** has been an important tool for **internal reflection and early planning**, but its **application with external partners** has been **challenging** due to complexity and the need for **contextual adaptation**. Its use internally has been predominantly limited to initial reflections.

In most countries, the ESC tool is used primarily for **internal planning** and identifying gaps in scaling strategies. M&E and research have also played crucial roles in informing efforts toward institutionalisation and scaling. In all three case study countries, **M&E and research** have played a central role in informing decisions about programme design, scaling, and institutionalisation. However, the effectiveness of these efforts has varied.

Internal and contextual factors that have supported the delivery of institutionalisation and scaling activities include government support and alignment, experienced teams and scale-up demand. Internal capabilities and contextual factors have created a supportive foundation for both the delivery and scaling of education initiatives, demonstrating the importance of **strong relationships with governments, contextual awareness, and strategic internal collaboration**.

Additionally in most countries, **flexibility** and **adaptability** have allowed teams to respond to changing contexts, creating opportunities for **growth and expansion** where new relationships and **partnerships** can be established, which is crucial in settings where **policy shifts** or **external factors** may require adjustments.

The programme teams have noted that they were confronted by various challenges in their work. These challenges span a wide range of areas, from data integrity issues and misaligned objectives to natural disasters and government policy changes. **These challenges have required adaptations, and these have tended to delay, rather than derail, the programmes.**

In addition to the barriers identified by teams, a large barrier for institutionalising and scaling is clearly articulating *what* is being institutionalised and scaled. It is inevitable in complex, resource constrained systems for initiatives to be “reduced” at the point of scaling. This might include a reduction of training days for CPD (like in Vietnam) or reductions in the components of the training material. Across case study countries the Minimum Viable Initiative was rarely agreed with partners.

Unanswered questions around scaling, funding and capacity exist across several countries and quality monitoring has also presented challenges in several countries.

How can the BETTER programmes and future scaling initiatives be more effective?

Experimental innovation is based on two key concepts: (1) design thinking and (2) evidence-based decision-making.² These two concepts do not always sit comfortably together. Design thinking is about doing something truly innovative (and so unlikely to have rigorous evidence related to it and a high chance it fails), while evidence-based approaches are about doing something that is firmly grounded in evidence and low risk (and so unlikely to be truly new). Experimentation brings the possibility of radical initiatives while evidence-based approaches bring rigour to the innovation process. With experimentation comes risk. The challenge is to articulate a level of risk of failure that is acceptable to everyone involved.

Stage gates are decision moments, where in advance, a set of criteria is **established to evidence that the innovation is ready** to move to the next stage of the innovation pathway. This is essential to reduce risk. If the criteria aren’t met, new strategies or approaches need to be worked out *before* moving the innovation to the next stage on the innovation pathway.

The BETTER programmes have certain advantages for institutionalisation and scaling, but also face unique challenges as VVOB has chosen to innovate and pilot *with and through government* -- and approach that is outside of the usual model where innovation and piloting takes place outside of government and the initiative is later passed over to government to scale up.

Institutionalising within government from the point of design and piloting has many important advantages over other approaches. Bold et al’s seminal study (Scaling-up What Works: Experimental Evidence on External Validity in Kenyan Education, 2013) on external validity in scaling education programmes evidenced that the short-term gains of NGO implemented initiatives in controlled experimental settings are lost in weak public institutions. Within the VVOB approach the evidence of impact from a pilot is more likely to reflect the realities that will be encountered at the time of scaling up if the piloting organisation is the same organisation as the one that will be responsible for scaling up.

However, along with these advantages come challenges. Current models for innovation do not take account of institutionalising within government during the design and piloting process. While elements of the “traditional” innovation pathway and standards of evidence can be borrowed, this evaluation also provides an opportunity to **build a new model of innovation** which accounts for institutionalising from the outset.

² Mcannan, Blomkamp and Lewis, 2018. *The rise of public sector innovation labs: experiments in design thinking for policy*. Available from: <https://www.researchgate.net/publication/323664854> The rise of public sector innovation labs experiments in design thinking for policy (accessed 19 November 2018).

Across the BETTER programme, each country programme is at a different point in the pathway and the standards of evidence, however there is little alignment between progress and standards of evidence.

It was recognised within sensemaking workshops that building a strong coalition and generating evidence on the initiative (and then potentially needing to pivot) **can work against each other.** Building a strong coalition has some aspects of advocacy and this can make for difficult conversations if the evidence that is generated during piloting does not yield positive results.

This potential “working at cross purposes” can be addressed by including government partners in the innovation process from the beginning. Introducing the concepts of experimental design, evidence generation and potential pivots from the outset and embedding them in the workplan can ready partners for these conversations and build capacity in experimental innovation.

A new model for institutionalisation and scaling can add important elements based on lessons learnt during this evaluation. We elaborate on this in the Conclusion and Recommendations Sections in this report.

The promising practices emerging from VVOB’s educational programmes across different countries reflect a combination of collaborative design, evidence-based decision-making, relationship building, and adaptive implementation. These practices are not only effective in promoting successful implementation but also contribute to long-term sustainability and scalability. By focusing on the active involvement of local stakeholders, iterative development, and alignment with national priorities, VVOB has created a framework for success that can be replicated and adapted across diverse educational contexts.

The ESC, M&E systems, and research best practices are key tools that can significantly enhance the design, implementation, and scaling of educational initiatives. By integrating these tools more systematically into scaling processes, organisations like VVOB can improve the effectiveness, efficiency, and sustainability of their interventions. Below are ways these tools can better inform scaling initiatives:

- Frequent and Structured Use of the ESC Tool
- Integrating M&E Systems for Real-Time or “light touch” Data Collection and Analysis – balancing urgency and evidence
- Use of Evidence to Inform Adaptations in Programme Design
- Strengthening the Evidence Base to Reduce Scaling Risks

There are examples within BETTER countries of ways to overcome barriers. However, the biggest barriers to institutionalisation and scaling, that bring the biggest risks, sit within two core areas (1) system capacity and (2) systems integration. In the Conclusions and Recommendations sections below, we propose approaches that increase the likelihood of overcoming these barriers.

Conclusions

VVOB’s BETTER Programme recognises that in order to move the needle on learning, real system change is required, through the scaling of education initiatives that work to increase teaching and learning quality at scale. The eight BETTER country programmes engage in a range of activities to promote effective scaling and

institutionalisation of the initiatives that they support. A lot of progress has taken place for BETTER programmes in each country in a very short period of time (2022-2024).

At this Mid-Term Review reflection point, progress towards Result Areas is mixed yet mainly consistent across countries. This Mid-Term Review provides significant evidence that BETTER programmes are relevant and aligned to partner objectives. However, overall, **the success of BETTER programmes hinges on generating evidence along the causal chain, and aligning scaling strategies**, with this evidence in mind. BETTER programmes are entering a pivotal phase in 2025, moving to the second phase of programming where final institutionalisation activities and ambitions for scaling need to be aligned and agreed, with haste.

Institutionalizing and sustaining educational initiatives at scale, particularly those developed in collaboration with partners, requires a strategic approach that focuses on systems change, government ownership, long-term sustainability, capacity building, and integration in systems.

While the BETTER programme draws on institutional and sector tools and approaches,³ VVOB is currently missing an overall model that underpins their approach to institutionalisation and scaling.

The key lessons that emerge from the evaluation include:

1. Clarify and unpack ToCs and use them regularly
2. Ensure a clear pathway for innovation and evidence integration
3. Explicit risk acknowledgment and explore risk appetite
4. Expand partner involvement in collaborative decision-making
5. Institutionalise learning and evidence use
6. Continue building relationships and explicate building capacity
7. Add a Result Area to the ToC that refers to the integration of the initiative into government structures.

Based on the lessons learnt within this Mid-Term Review, we propose a starting point for VVOB to develop an institutionalisation and scaling model (and tools) to bring about real systemic change and impact at scale.

Moving in a sequential chronological order, a new model might include:

Developing initiatives that leverage established relationships, built on previous VVOB programmes that are nested in policy or specific initiatives.

Iterating through prototype development, testing and adaptation, to find the Minimum Viable Initiative to solve the problem. VVOB is already undertaking this work across their BETTER programmes. However, light touch M&E moments focusing on the causal chain can be added to the current approach to this work.

Once it is believed that the Minimum Viable Initiative is in place, pilot the Initiative to demonstrate that the initiative works. A successful pilot will provide evidence that the innovation had demonstrable impact and, ideally, that it offers a comparative improvement in terms of effectiveness, efficiency and/or quality over existing approaches. It will also generate new learning on what works, what doesn't and why.

³ Including the ESC, VVOB scaling materials, VVOB ToC templates etc.

In addition, explicit discussions and workplans with regards to:

- Scaling
- M&E, Research and Learning
- Capacity
- Systems integration

Scaling actions and workplans could include:

- Agreeing and communicating exactly what the Minimum Viable Initiative is. In other words, what are the core components of the initiative that are considered essential for the initiative to work?
- Simplifying the ESC and use it with partners
- Discussing institutionalisation and scaling options and costs, and documenting options and agreements
- Explicating and assessing each partner's appetite for risk
- Agreeing on stage gates for what evidence is needed that the initiative is working before moving on to the next stage, based on agreements regarding risk
- Agreeing the ambition for scaling the initiative and the plan to achieve scale
- Identifying and planning for key moments of reflection during the scale-up process

M&E, Research and Learning actions and workplans needed could include:

- Planning and designing evidence generation activities based on the stage gates plan
- Designing M&E for quality monitoring, not only implementation monitoring

Capacity actions and workplans could include:

- Explicitly assessing the capacity within the system to scale up
- Explicating the capacity gaps and planning with partners to address them or find new implementation strategies
- Linking quality monitoring plans to an explicit QA Framework that predicts potential weaknesses as scaling takes place

Systems integration actions and workplans could include:

Considering all of the functions of programme implementation and documenting how the QA Framework, quality monitoring, MEL functions, updates to materials and any technology (including maintainance) will be integrated into government systems as scaling takes place. Estimate the cost implications of integrating the programme into government systems. identify how this could be funded.

At key moments of reflections, during the scale up process, meeting with partners to access if programme strategies need to change, based on:

- Quality implementation monitoring results
- Capacity challenges
- Success of system integration plans
- Emerging evidence
- If the modelled costing holds
- Model changes in the ESC

Recommendations

Based on the conclusions of this evaluation we propose a number of recommendations to be considered for the next period of BETTER.

Global recommendations

Recommendation 1: Clarify VVOB's vision of what constitutes successful scaling. Develop, in close collaboration with country teams, an overarching model on innovating, piloting, implementing and scaling through government institutions for VVOB.

Consider starting with the proposed model provided in this report, adapting where needed to suit VVOB's needs.

Recommendation 2: Ensure a clear pathway for innovation and evidence integration. By acknowledging where evidence is lacking, partners can make informed decisions about risks and manage expectations. This will ensure that scaling efforts are not based solely on the desire to "get things done", but on a balanced approach that integrates evidence, evaluation, and risk management.

This pathway might be integrated within a VVOB scaling model (Recommendation 1) or be developed in parallel and/or in the absence of a full model.

Recommendation 3: Consider VVOB's risk appetite and balancing VVOB's long history of supporting partners to institutionalise and scale up, while also being an evidence-based organisation.

Provide guidance to country teams on how to navigate risks and scaling at the country level.

Recommendation 4: Explicitly add capacity to implement and integration into the system as Longer-Term Result Areas. This includes embedding QA systems, MEL functions, updates to materials, professional training, integrated technology and long-term maintenance plans.

Recommendation 5: Change the second Result Area to "increased coherence in the policy environment" rather than "adjustments in the policy environment".

Recommendation 6: Within the context of the findings of this Review and Recommendations 1, 2 and 3, consider the scaling ambition that can be reasonably achieved in the programme period.

Country-level recommendations

Recommendation 1: Articulate the Minimum Viable Initiative as soon as possible to partners. Produce explicit guidance on what aspects of the model can be adapted, and what shouldn't, taking into account legitimate cases where the model may need to be tailored to contextual factors during scaling.

Recommendation 2: Clarify and unpack ToCs and use them regularly to adapt strategies if needed.

Recommendation 3: Document scaling plans for the remaining 2 years of the BETTER programme in each country. Check understanding with government partners and align on the vision for scaling. Considerations in this process should include:

The scaling ambition that can be reasonably achieved in the programme period and align the ambition to what stage of scaling can be achieved *with evidence that the initiative is working*. (i.e. What happens next if evidence does not find that the initiative works?)

How the programme will be funded at the end of VVOB programme support

An assessment of capacity to sustain quality delivery of the programme at scale

Recommendation 4: Establish quality monitoring and QA systems at the beginning of programming. Consider the risks during the scaling process, linked to capacity assessments and quality monitoring and adjust programming when needed.

Recommendation 5: Consider contracting Learning Partners for the duration of BETTER programmes, rather than contracting specific evaluation deliverables. Learning Partners can work with programme teams to clarify and unpack ToCs, design and implement nimble evaluation activities, reflect of ToC in light of emerging evidence and supporting a step-wise approach to evidence generation and scaling.

Recommendation 6: Expand collaborative decision-making with partners by bringing them into the innovation pathway process. Involve partners in considering adaptations that might need to be made to the initiative and build support for the process, rather than the output.