

# MAINSTREAMING SCALING INITIATIVE CASE STUDIES

*Deutsche Gesellschaft für Internationale  
Zusammenarbeit (GIZ)*

Dr. Richard Kohl

18 August 2023



# MAINSTREAMING SCALING

## A Case study of GIZ

by

Dr. Richard Kohl  
Strategy and Scale LLC  
18 August 2023

A Case Study for the Initiative on  
Mainstreaming Scaling in Funder Organizations

For the Scaling Community of Practice  
[www.scalingcommunityofpractice.com](http://www.scalingcommunityofpractice.com)

# Table of Contents

Acronyms

Background

Methodology

Findings

- Status of Scaling at GIZ

- Support for Scaling from the Sectoral Departments Unit that Advises on Scaling

Recommendations

- Mainstreaming Scaling

- Scaling Support

Results from the COP Webinar on GIZ Mainstreaming

Lessons learned



## Preface

The **Scaling Community of Practice (CoP)** launched an action research initiative on mainstreaming scaling in funder organizations in January 2023. This initiative has three purposes: to inform the CoP members and the wider development community of the current state of support for and operationalization of scaling in a broad range of development funding agencies; to draw lessons for future efforts to mainstream the scaling agenda in the development funding community; and to promote more effective funder support for scaling by stakeholders in developing countries. (For further details about the Mainstreaming Initiative, see the **Concept Note on the COP website**).

The Mainstreaming Initiative is jointly supported by Agence Française de Développement (AFD) and the Scaling Community of Practice (CoP). The study team consists of Richard Kohl (Lead Consultant and Project Co-Leader), Johannes Linn (Co-Chair of the Scaling CoP and Project Co-Leader), Larry Cooley (Co-Chair of the Scaling CoP), and Ezgi Yilmaz (Junior Consultant). MSI staff provide administrative and communications support, in particular Leah Sly and Gaby Montalvo.

The principal component of this research is a set of case studies of the efforts to mainstream scaling by selected funder organizations. These studies explore the extent and manner in which scaling has been mainstreamed, and the major drivers and obstacles. The case studies also aim to derive lessons to be learned from each donor's experience, and, where they exist, their plans and/or recommendations for further strengthening the scaling focus.

The present case study focuses on Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), the German bilateral development implementation agency. It was prepared by Richard Kohl under a consultancy financed by GIZ.

## Acronyms

BMZ	Federal Ministry for Economic Cooperation and Development
CC	Competence Center for Innovation, Learning, Knowledge Management within Methods, Digital Transformation, and Innovations unit of FMB
CGIAR	Consultative Group for International Agricultural Research
CoP	Scaling Community of Practice
FMB	Sectoral division Methods, Digital Transformation, and Innovations unit of the Sectoral (FMB) division of GIZ
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH
KI	Key Informant
KII	Key Informant Interview
VHF	vertical, horizontal, and functional

## Background

This case study is based on an assessment of scaling up at GIZ, the German bilateral international development implementation agency. It was commissioned by the Competence Center "Innovation, Learning, Knowledge Management" (hereafter CC for Competence Center) within the Methods, Digital Transformation and Innovations unit of the Sectoral (FMB) division of GIZ. The FMB division is the part of GIZ responsible for providing technical assistance and advice for the implementing projects of GIZ, particularly the Sector and Global Programs and bilateral projects. Another part of FMB is planning



advisors who play the lead role in GIZ in the designing of and providing ongoing support for GIZ programs.

GIZ has been working explicitly on mainstreaming for over a decade. In 2013 it commissioned an assessment of the status of scaling up within GIZ, which was conducted by Rambøll Management Consulting and completed in 2014 (hereafter referred to as the 2014 study).<sup>1</sup> The key findings of this study were:

- There was no common understanding of the definition of either “broad impact” -- how GIZ describes the expected impact or outcomes achieved by its projects -- nor of scaling up.
- Using a nomenclature of vertical, horizontal, and functional (VHF) scaling up,<sup>2</sup> the study found that most GIZ projects did engage in processes that could be characterized as vertical and horizontal scaling, even though they were not necessarily conceptualized or labeled as scaling. By contrast, very few projects were designed to or actually included functional scaling up.
- GIZ projects are characterized by a multi-level approach that works with partner systems at all the relevant levels, i.e., the macro, meso and micro levels, by engaging with the system at all levels supports scaling (a finding consistent with the Scaling CoP principles<sup>3</sup>).
- Scaling tends to occur through staff initiative, i.e., bottom up, especially when staff are rotated to new contexts and bring to bear their experience with interventions in other countries.
- GIZ’s administrative and bureaucratic structures did not actively support, encourage, or incentivize scaling up. FMB planning advisors “are not yet being used comprehensively as the bearers of knowledge on scaling up”.<sup>4</sup> Little use was being made of project progress reviews or the results of monitoring and evaluation generally to support or encourage scaling up.
- GIZ’s Broad Impact approach correlated with scaling in some areas but not others. The three components of Broad Impact that support scaling were: (a) close attention to aligning project design with the expressed interests of the national government and other local stakeholders (relevance); (b) the role of achieving quantitative results in terms of reaching a (large) number of beneficiaries (quantity); and (c) sustainability -- financial, institutional and political. By contrast, its emphasis on a different aspect of relevance, customization to the local context, had little correlation with impact at scale.
- Of the seven case studies reviewed, only a few actually achieved broad impact. GIZ projects were most successful in achieving relevance, and least successful in achieving quality, quantity, and sustainability.

The 2014 study of scaling within GIZ made a number of recommendations to increase scaling within GIZ’s project activities. First, create a common understanding and language around the concepts of broad impact, scaling, and the relationship between the two. Second, improve knowledge management so that the design of new projects can benefit from and leverage learning from past and concurrent projects. Third, leverage management’s role in setting goals for individual staff and also staff rotation to

<sup>1</sup> Lennart Raetzell and Jan Oliver Scheller (2014) Corporate Strategy Evaluation Broad impact / Scaling up. Bonn, Germany: Rambøll Management Consulting GmbH for GIZ. Only the executive summary of the study, published in 2014, is available in English.

<sup>2</sup> Loosely speaking, vertical scaling up involves either affecting systems changes and capacity building at a “higher” e.g., national level, horizontal scaling up involves replicating an intervention in additional geographic areas, and functional scaling up involves adding additional components or greater dosage or intensity of existing components of an intervention.

<sup>3</sup> Scaling Community of Practice. “Scaling Principles and Lessons”. 2012.

[https://www.scalingcommunityofpractice.com/wp-content/uploads/2022/03/Scaling-Principles-and-Lessons\\_v3.pdf](https://www.scalingcommunityofpractice.com/wp-content/uploads/2022/03/Scaling-Principles-and-Lessons_v3.pdf)

<sup>4</sup> Anonymous interview with GIZ staff member



incentivize scaling. Fourth, strengthen administrative structures, particularly those supporting co-funding, to encourage scaling. Finally, integrate scaling into project design from the outset and into monitoring systems, applying scalability criteria available.

The current review to some extent can be considered an update of the 2014 study. It was designed to answer the following four questions, which overlap but are not identical with those of the 2014 evaluation. These questions were:

1. What is the status of scale and scaling at GIZ, including challenges to increasing scaling at GIZ?
2. What are the options and recommendations for addressing increasing scaling at GIZ?
3. What is the need and demand from GIZ staff for support for scaling up?
4. What are the recommendations for a scaling support strategy for the GIZ team that advises on scaling?

To foreshadow our findings, we did find that many of the findings, gaps, and recommendations of the 2014 report were still relevant. Further work is still required on a widely understood and utilized definition of scale, knowledge management for scaling needs to be further improved, staff needs incentives to implement and prioritize scaling, and monitoring efforts can be strengthened to support scaling.

Before turning to our methodology, it is important to make one other point. In the German system bilateral development cooperation agreements and their translation and implementation into projects are done by two separate agencies, BMZ – the Federal Ministry for Economic Cooperation and Development, and GIZ. BMZ negotiates with developing country governments, and partner ministries in particular, the overall goals for a project in broad terms, as well as the contributions in cash and in-kind from the German and host country government or ministry. GIZ designs and implements projects to achieve those goals but does not set them. As such, the extent to which scaling occurs in GIZ projects, or projects are designed with future scaling in mind, is constrained by the agreements negotiated by BMZ.

Finally, the views presented in this assessment are those of the author alone, and do not necessarily represent the views of GIZ as an organization nor that of the unit that advises on scaling, i.e., the CC unit. That said, this report could not have been written without the support, cooperation and feedback from Arlett Stojanovic and Mara Holzenthal of GIZ. It also benefited from comments from other GIZ staff, Johannes Linn of the Scaling Community of Practice (CoP), and observations and questions posed during three webinar presentations.

The rest of this study is organized as follows. The second section briefly describes the methodology used, including some caveats regarding the findings. The third section lists the findings of the scaling review conducted in 2022, covering questions 1 and 3 above. The fourth section presents the recommendations for GIZ if it wishes to increase its scaling efforts, and what types of support would facilitate that given the current situation, i.e., questions 2 and 4.

## Methodology

It is important in any study such as these to define what we mean by scale – the noun and objective -- and scaling or scaling up – the verb or process of achieving scale. As used in this study, scale is defined as sustainable impact at a significant share (coverage or reach) as compared to the size of the need, demand, or problem. In other words, we define scale as large scale; otherwise, any reach, coverage or numbers of beneficiaries can be considered scale. The size relative to the denominator is key; reaching



two million people with incomes under \$2 per day in a nutrition program in India is not large or significant scale as it is equivalent to 1/10<sup>th</sup> of a percent of the population; in Mali it would be scale as equivalent to 10 percent of the population.

Scaling or scaling up is the process of reaching scale. It is important to note that just because a project or activity includes activities that could be considered scaling, does not mean it will reach large or significant scale. Similarly, the time frame is important. A project may achieve scale within its own time frame of, in the GIZ case, commonly three years, but given that scaling usually takes more than five years, this is unlikely. Or it may create the foundations for and catalyze a scaling process that will continue after the project ends, which could include a follow-up donor project. In either case, scaling continues, and large scale will be reached after the end of the initial project.

This scaling assessment faced some constraints that are worth noting. Ideally to answer these questions we would have conducted a relatively comprehensive inventory or at least sample of GIZ projects. This would have allowed us to assess what percentage had gone to scale (and their results), could have gone to scale and did not, and those that likely had little scaling potential. Complemented by extensive staff interviews, especially of those involved in scaling projects, as well as senior management, this would have allowed to get an accurate sense of how much scaling is actually happening as well as what the drivers are when scaling does occur, and obstacles or lacunae when it doesn't. Such information would have been useful in determining what types of scaling support services would be most helpful in maintaining and strengthening scaling at GIZ. However, currently there is no data available on what scale, impact, and sustainability GIZ projects achieve, let alone how many achieve or catalyze large scale.

Instead, to answer these questions, we conducted a review of documents relevant to scaling, though not project documents or evaluations, supplemented by a small number of key informant interviews (KIIs). The documents reviewed included GIZ's scaling guidance framework and case studies, tools like the scaling canvas, and external consultant reports. Over a dozen GIZ staff were interviewed, both based in Germany as well as several working on the ground in developing countries. It included staff and middle management working on global projects, individual sectors like social sector, fiscal strengthening, and nutrition, and on projects in both Africa and Asia. It is important to note that these KIIs were neither a random nor representative sample but were drawn from the informal network of contacts of the unit that advises on scaling. With the important exceptions of the heads of GIZ's two monitoring and evaluation divisions, Central Projects and Strategic, we did not interview any directors of the various departments, senior management, or management board.

As such, the findings and recommendations shared in this report are based on qualitative data and should be understood as such. Should the resources be available in the future, we do recommend that a more thorough census of GIZ's aggregate impact at scale, the number of projects that do and do not achieve large scale (and where, why and under what circumstances) would be useful.

## Findings

### Status of Scaling at GIZ

A major finding is that there appears to be a fairly broad consensus among GIZ staff on the importance of scaling and interest in doing more. Many GIZ staff interviewed felt scaling is important and an organizational priority but not clear what that would mean in practice. By contrast, we were not able to identify any formal mandate, directive, or goals from GIZ management for scaling, such as explicit mention of scaling in GIZ's Mission, Vision or current or recent strategy documents.

Most people believed that many GIZ projects are engaged in some form of scaling; in perceived order of frequency, vertical scaling (or systems change) was the most common, horizontal scaling (geographical



or replication) less common, and functional scaling (transfer of successful solutions to another context or service) occurs infrequently.<sup>5</sup> This is not surprising, because the majority of GIZ projects target either strengthening the capacity of government partners, often ministries and usually at the national level, or improving the policy enabling environment, or both.

This belief, consistent with the 2014 scaling assessment, appears to derive from conflating scaling as a process and achieving large scale. Most of the KIs interviewed believe that because they were doing vertical scaling (capacity building or systems strengthening), they were achieving some degree of scale. However often that did not appear to be the case empirically if we use the definition of large scale; these systems were not often and not necessarily systems with large scale reach and impact. If all systems change is seen as a form of vertical scaling, then almost all GIZ initial projects are scaling and, according to this logic, achieving scale. The same was true of horizontal scaling; project scale was often at a small fraction of the overall population size or number of administrative districts, departments, regions, etc. Yet in fact, and as is widely acknowledged, few if any GIZ projects achieve scale in the initial three years; given the short project duration and limited financial resources, large scale is nearly impossible to achieve in that time frame.

Concomitantly, there continues to be a confusion between replication and scaling. (In this case, we are using replication to mean horizontal scaling with no changes or adaptation of either the program or model to the new environment or context, or systems changes or capacity building/technical assistance of the new context or actors in that context). In fact, in most cases the program must be modified or adapted to take into account differences in context, or resources available in that context. GIZ staff point out that because projects are not designed for scaling, their unit costs<sup>6</sup> and other resources and capacity requirements usually exceed the resources and capacity available at large scale. (This is also true in terms of GIZ's own resources to support scaling.) Because of this, project modifications and redesign are very frequently required to increase efficiency, reduce unit costs, and where possible, achieve economies of scale. Some of the KIs explicitly stated that many of their colleagues understand scaling to either assume replication with no changes, or, alternatively, see scaling as if X number of people or places were reached with Y euros, then 10xX people or places can be reached with 10xY euros. In other words, they design projects for scale assuming a proportionate increase in resources. Yet many sources define this is a growth, not scale, where scale is an exponential or non-linear increase in impact with an incremental increase in resources.<sup>7</sup>

As noted in the 2014 scaling evaluation, GIZ's concept of broad impact – relevance, quality, quantity, and sustainability – is substantially aligned with scaling principles but presents some challenges in terms of its actual implementation. Relevance is a positive factor for scaling as it means that local ownership and buy-in is more likely, and therefore likely to translate into domestic mobilization of resources and political will needed for scaling, such as in a follow-up project or after project end. The GIZ project design process invests substantial time and resources in interviewing all relevant stakeholders, identifying their needs, aligning project goals with those needs as well as doing analyses of the context at the target scale of the project. In fact, relevance appears to be the greatest priority for GIZ projects, along with quality or impact.

---

<sup>5</sup> Digital innovations or components of a project design are now a major focus of GIZ, and this emphasis may lead to greater functional scaling in the future.

<sup>6</sup> By unit cost we mean either the initial costs, as part of the project, of delivering the program born by domestic actors or ongoing costs after project end. The relevant denominator is either a given number of beneficiaries, a geographic unit (e.g., village, municipality, district, region), or an administrative unit (e.g., national ministry, local or district offices of a national ministry, bureau, or division within a national ministry). GIZ projects often also face a similar challenge in terms of their own unit costs, such as the costs of capacity building per any of the units listed above – beneficiaries, geographic or administrative units.

<sup>7</sup> See, for example, Rizwan Tayabali, (2014) PATRI Framework for Scaling Social Impact,

<https://www.inclusivebusiness.net/sites/default/files/wp/PATRI-Framework.pdf>, p.3



However, relevance also implies that each design needs to do something new and be customized to the context at what is almost always small scale in initial projects. Moreover, because the agreements that BMZ negotiates with partner ministries often include numerous goals, projects often contain a large number of interventions which does not facilitate scaling. Both features tend to be inimical to scaling.

Relevance is prioritized without either the recognition, let alone explicit consideration, of potential tradeoffs between relevance and the other three goals. Quantity is generally a low priority and the residual factor after GIZ considers how to achieve relevance and impact sustainably, given a project's financial constraints.

As noted above in our discussion of definitions, because relevance aligns with context at small (project) scale rather than potential large scale, GIZ projects that eventually do scale through follow-up project have to both undertake additional analysis of large-scale systems and then be redesigned to take into account large-scale constraints or spaces, particularly financial, fiscal and implementation constraints. It can take the GIZ staff responsible for designing a follow-up project 1-2 years to redesign a project to align with large-scale constraints. As noted above, GIZ resources do not usually expand proportionately with the increase in scale (i.e., they don't receive from BMZ five times the funding to cover five times the number of locations), requiring them to figure out either how to achieve the same impact with fewer resources, or to implement a more modest program that has less impact.

The GIZ staff we talked to stated that this means that in most cases scaling is not built in from the beginning, even though they believe that much of the staff understands that building scaling in from the beginning is a core principle of scaling. Two implications of this are that scale, or even the foundations for scale, are not achieved in initial projects. As discussed above, when scaling does occur, principally through follow-on projects, it requires significant additional time for a redesign. Nonetheless, judging from the interviews, achieving large scale does occur with some frequency, mostly through follow-on projects and to a lesser extent through global projects.<sup>8</sup> (Again, it is important to note that this is an impression based on a limited number of interviews with a potentially biased sample, and not a systematic review of what happens to GIZ individual or global projects.)

This suggests that many of GIZ's initial, small-scale projects do support scaling in at least two ways. First, they provide some form of proof of concept. Second, because of the heavy emphasis on local ownership and participation, initial projects create a demonstration effect and generate requests from local government partners, interest from other donors in follow-up projects go to scale; in fact, often both interest from the government and other funders. However, as there is no systematic tracking of scaling, it is hard to know precisely how much scaling occurs through follow-up projects or through other mechanisms. Thus scaling appears to be largely ad hoc and opportunistic in response to the interests of host governments and other donors rather than strategic.

The need for project redesign also reflects the fact that monitoring and evaluation are not set up to produce learning and documentation for scaling, but primarily for accountability and for maximizing efficiency. GIZ's project monitoring and evaluation framework adheres closely to the OECD DAC guidelines, though some of the indicators included could contribute to a scaling assessment if it were to be required in the future, e.g., questions of sustainability. Both the GIZ and OECD guidelines do not have an explicit focus on scaling. For example, the GIZ matrix does not assess the sensitivity of specific activities and their impact to local context, what implementation capacity or resource requirements are, nor relevant financial information like unit costs or potential for economies of scale. Cost and resources

---

<sup>8</sup> GIZ global projects refer to projects in the same sector and/or a common issue e.g., Private Business Action for Biodiversity, which are designed from the beginning to be implemented in multiple countries. These usually combine a set of knowledge products, tools and approaches combined with individual country projects and interventions customized to the specific needs, context, and priorities of each country. Individual country projects may or may not make use of the common knowledge projects, tools, and approaches.



issues are looked at through the lens of efficiency, which is measured in terms of how well the project used resources to achieve its goals, not unit costs or scalability. That said, there is a general question about what the future prospects of the project are, though there is little guidance on how to answer that question; it is left to consultant evaluation teams to make a qualitative, ad hoc judgment about future scaling potential unless the project had a predecessor project. Since in almost all cases creating the foundations for scaling, let alone impact at large scale, is not a project goal, scaling is not really an evaluation category or goal.<sup>9</sup>

The same problem arises in terms of leveraging prior experience in the design of new projects in other countries. While there has been some effort to develop short, accessible summaries of projects to serve that purpose, they do not currently contain enough information for someone designing a new project to know whether an intervention or component would be relevant or appropriate to a new project given its goals and the local context. To the extent that this information is available at all, a project manager would have to read an entire project evaluation report which is quite lengthy. Moreover, even if evaluations did contain most of the information needed for scaling, currently only forty percent of projects are evaluated by the Central Project Evaluation unit. (Cross cutting evaluations are conducted by the Strategic Evaluation unit to help fill the learning gap.) In other words, as noted in the 2014 evaluation, knowledge management for scaling up remains a challenge. As a result, according to our KIIs most new project designs do not make use of elements from past projects unless the FMB planning advisor has personal knowledge of those projects.

GIZ management is acutely aware of the knowledge management issues and has a strong organizational commitment to leveraging project experience in future projects and other countries, what it calls standardization. There is currently a major interest in designing and implementing a system to document project interventions and develop an accessible database of project information and other information relevant for scaling. GIZ does not see any tension between standardization and the emphasis by project designers on ensuring that designs are relevant and customized to the local context. While the tension between relevance and standardization probably doesn't exist in terms of aligning with the needs and interests of host country stakeholders, it is probably less true regarding customization to local context. Standardization and, for that matter, horizontal scaling within countries, entails that the role of context be made explicit, so that future applications can consider whether an intervention is appropriate to a new context, or what adaptations might be necessary. GIZ is now working to better document projects and specific interventions so that there will exist a library or catalogue to support standardization; hopefully, that will include documenting both the important contextual factors as well as unit costs and resources needed. However, until a knowledge management system that covers these factors is in place, an FMB project advisor's ability to leverage or 'scale' prior GIZ experience in the design of a new project is chiefly dependent on their own knowledge, experience and awareness of previous interventions and their network of contacts within GIZ.

GIZ is also heavily committed to leveraging the potential of digital technologies for development impact and scaling them. This is still in its early stages, so it is hard to draw any conclusions. However, in practice the strength of GIZ's close partnerships with local governments often means that the

---

<sup>9</sup> GIZ project evaluations are done by teams of external consultants using a common framework based on the OECD DAC guidelines. In that framework, there is one question about systems change and one about scaling: "To what extent has the intervention achieved structural or institutional changes (e.g., for organizations, systems and regulations)?" and "To what extent did the intervention serve as a model and/or achieve broad-based impact?" If a follow-up project is anticipated or already exists by the time the evaluation is conducted, then the evaluators are asked to answer the question: "Extent to which developed concepts, approaches, instruments and lessons learnt have been or will be used or are further developed by the follow-on project." Footnote is based on the 2020-11 English version of the Central Project Evaluation Unit's Evaluation Matrix, which was kindly supplied by CPE to the author.



institutionalization of digital and other information technologies is done based on relationships and local political economy considerations rather than what makes the most sense in terms of scaling and sustainability. For example, almost all GIZ projects have as their primary partner an individual government ministry. While it may make sense to embed new digital and information technologies in a single information technology ministry devoted to supporting the rest of the government, GIZ's partner ministry is often able to get such units put into their own sectoral ministry where they are less likely to have the capacity and resources to maintain them, let alone scale them. This issue is even more problematic when it makes the most sense for the digital or IT innovations to be housed in an independent third party such as a university or private sector actor.

In sum:

- Scaling at GIZ is heavily constrained by the fact that BMZ negotiates partnership agreements with host country governments that specify a project's broad goals and objectives; GIZ is left to design and implement those objectives through its projects. To the extent that scaling, or potential for future scaling, is present or absent from such agreement constrains what GIZ can do. This constraint is reinforced by short project durations of three years, and the fact that BMZ agreements tend to include a large number of components.
- There is general awareness of scaling and its importance, and of the VHF framework. However, inclusion of elements that could be considered forms of VHF appears to be interpreted as scaling is occurring, regardless of whether these lead to large scale being achieved or the foundations for future scaling being established.
- There remains a lack of a common understanding what scale and scaling mean and requires, especially when economies of scale and scope or alignment with constraints at large scale are needed for vertical and horizontal scaling to be viable.
- Scaling is rarely integrated into initial project designs. While these designs do rely heavily on stakeholder and systems analyses, these are often conducted at a scale relevant to the project level. As such, they do not analyze or take into account unit costs and potential for economies of scale, financial or fiscal constraints in terms of either domestic or GIZ resources, and implementation or other constraints at large scale. Therefore, project designs, implementation, and evaluations, are not set up to support scaling should projects engender demand for scaling in the future.
- GIZ's broad impact approach in many ways strongly aligns with commonly accepted principles to support scaling, especially in terms of relevance to local stakeholders and sustainability. At the same time, the fact that there is a strong emphasis on customization to local context (and stakeholders) means that there is little incentive to replicate or incorporate learning from previous projects, or to explicitly consider the relative merits and tradeoffs between relevance and the other broad impact pillars, especially quantity or scale. As a result, scale is often the residual size possible after relevance, quality/impact, and sustainability are taken into account. Given limited resources it is often quite small compared to the size of the problem.
- Scaling of projects does appear to occur relatively often, though it is hard to know precisely what that means in practice without any tracking data. Most commonly scaling takes the form of follow-up projects initiated at the behest of partner governments, leveraging co-funding from other donors, additional funding from BMZ, or both.
- Follow-up projects usually have to conduct new analyses of large-scale systems and retrofit, adapt and simplify initial project designs to align them with large scale constraints. Similar resource constraints also apply to GIZ's own project support resources which usually do not increase proportionately with target scale. This reassessment and redesign process takes



significant additional time and resources that could be largely avoided if scaling was built into project designs from the beginning, a point made both in the 2014 evaluation and in GIZ's own scaling guidance framework that has not been translated into practice. Global projects appear to have more success in sharing information across countries, but this appears to be highly variable.

- In terms of project documentation to support standardization and scaling, current approaches to monitoring and evaluation and knowledge management generally provide only some of the necessary information to make decisions about future scaling let alone design follow-on projects. The information that does exist is not easily accessible to project designers.
- GIZ is aware of its knowledge management challenges and particularly the lack of a knowledge base for standardization, and addressing this issue is at present a major focus. GIZ believes that greater use of standardization can be achieved without rebalancing its priorities between relevance, quality, and quantity; whether this will be true in practice remains to be seen.
- Digital innovations and using digital as a tool for scaling, as well as scaling digital itself, is also a major focus in GIZ; a framework and tools to that end are currently under development.

### Support for Scaling from the Sectoral Departments Unit that Advises on Scaling

As noted above, following the publication of the 2014 scaling evaluation, GIZ developed a guidance framework based on the horizontal, vertical, and functional (VHF) conceptual approach along with tools to support scaling up. Most of the tools were existing tools that could be equally well applied to support scaling. The GIZ toolkit Capacity Works includes a number of highly relevant tools that are commonly used in project design such as the Actor Analysis, Societal Currents and Trends, and Process Analysis. In addition to the Capacity Works tools, the unit that advises on scaling also began development of other tools such a scaling business canvas, began to make use of the Scaling Scan originated by one of the scaling advisors that GIZ has embedded in CGIAR centers, and piloted all of this in a trial training workshop.

However by and large, it appears that while the VHF concept is widely understood and applied, this application commonly takes the form of renaming activities that would take place anyway rather than as a means of achieving scale as a goal unto itself. For example, most GIZ projects include strengthening the capacity of partner ministries, but it seems likely that this would occur anyway; program designers are not targeting vertical scaling as a goal unto itself. In general, the overall guidance document and the scaling specific tools are rarely used outside of a small number of staff who went through a few pilot scaling training workshops that occurred in the last few years.

The lack of attention to scaling, and use of the scaling framework and tools, appears to be due to the fact that scaling has not been explicitly made an organizational priority for GIZ in general, nor for FMB planning advisors who lead the project design effort (and provide ongoing support to project implementation). While there appears to be substantial interest in scaling, the planning advisors already face a large number of mandatory cross-cutting themes and accompanying tools or guidance documents that they are required to integrate into project designs, such as gender and climate change. Project implementation teams in turn are responsible for implementing those designs, which, as referred to earlier, often contain a large number of goals, key performance indicators, and components. Almost all KIs interviewed, including planning advisors, report being overwhelmed by the number of required issues and frameworks that have to be considered, and as a result *de facto* scaling is crowded out.

While demand may be limited, there is a clear need for scaling tools and support. The probable exception to this is GIZ's global and follow-up projects that target scaling. The extent to which global projects do attempt (let alone succeed) to generate cross-country learning and adoption of interventions



from other countries appears to vary widely and to be primarily a function of global project leadership. Those projects that emphasize cross-country learning, such as a current global nutrition project, have made some use of the guidance framework, tools, and support from the scaling unit staff. The designers of follow-up projects need support, especially when there is a substantial need to redesign projects to facilitate scaling. However, in the grand scheme of GIZ projects, whether country specific or global, use of GIZ's scaling tools and support from the scaling unit is well below actual need or potential.

The KIs interviewed indicated that GIZ staff would like more support for scaling but are wary of this being a mandatory addition to their already substantial load. In other words, they do not want integration of scaling to be compulsory, nor to be required to use scaling guidance frameworks, templates, and tools. Their preference would be for additional human resources to help integrate scaling into project designs and support project implementation, or at project end or at the beginning of follow up projects if that is when scaling considerations become relevant.

## Recommendations

### Mainstreaming Scaling

GIZ leadership should consider making explicit the role and goals of scaling and impact at scale. This could include setting explicit targets for the organization's overall impact at scale and the role of each part of the portfolio in achieving that objective. In other words, setting targets for the number of initial projects that target large scale or that are expected to lead to scaling up through follow-up projects. Another way to promote scaling would be to require program advisors to make explicit the tradeoffs between relevance, sustainability, quality, and quantity, and potentially give greater weight to quantity.

To support a greater emphasis on scaling, the existing work on knowledge management and standardization needs to be taken to its conclusion. Monitoring and evaluation can also provide greater support by expanding the number of indicators relevant to scaling and providing greater guidance on how to assess future scalability, taking into consideration the substantial international literature on scalability criteria and measurement. Strategic evaluations can be more targeted towards supporting standardization and scaling by identifying best practices for particular goals and their sensitivity to context. Project evaluations could be done earlier in cases of projects that have been identified as having scaling potential.

### Scaling Support

Providing scaling support should begin with existing demand and build on that. That means focusing initially on providing training and support to FMB program advisors who lead the project design on the one hand, and to global and follow-up projects on the other. While program advisors can integrate scaling into initial designs, they do not have the time to fully accompany projects on a scaling journey, or to support developing scaling opportunities as these arise. Training of program advisors needs to be supplemented by a roster of experienced scaling experts or even embedded scaling advisors to work with global projects (similar to what GIZ has done with CGIAR centers). These could be either existing GIZ staff with scaling experience, external consultants, or both.

Support for scaling needs to be supplemented by some additional investment in tools. The VHF framework is fine, but it needs to be made clear to GIZ staff that large scale is the goal, and that vertical, horizontal, or functional scaling pathways are means to that end, and not ends in themselves. A common reference on what large scale is and explicitly how it relates to Broad Impact and relevance, quantity, quality, and sustainability would be helpful. The initial steps towards developing or completing additional tools, like the scaling business canvas, should be completed. In the case of the canvas in



particular, the issues it covers are excellent, but it needs some gaps to be filled in. In particular, clear definitions of what is target scale and identifying and aligning with the systems constraints at scale is essential.

## Results from the COP Webinar on GIZ Mainstreaming

The results of the GIZ assessment were presented as part of a Webinar sponsored by the Scaling CoP on 28 March 2023. In addition to the author of the assessment, Richard Kohl, the webinar included remarks from:

- Arlett Stojanovic, Head of CC Innovation, Learning, Knowledge Management at GIZ
- Bishnu Datta Gautam, former Joint Secretary for Planning and Development Assistance in the Ministry of Federal Affairs and General Administration, Nepal
- Benjamin Kumpf, Head of Innovation for Development Facility at the OECD Development Assistance Committee, and
- Larry Cooley, Co-Chair, Scaling Community of Practice.

In her remarks, Ms. Stojanovic noted that despite the fact that GIZ works in 120 countries with an annual budget of €4 billion, i.e., has a substantial footprint, GIZ isn't leveraging its potential to achieve the impact at scale that it could. She noted that GIZ does have a unit that advises on scaling, located inside its technical advisory division, FMB (referred to in the main part of this note above). She and her colleague, Mara Holzenthal, decided in early 2022, to initiate a new assessment of GIZ's scaling work to get a fresh look at scaling within GIZ from an unbiased, external view. The goal of the assessment was to look at what was happening, what was working or not working, what were the impediments, and how could GIZ and this unit in particular improve the whole discussion and effort around scaling.

Ms. Stojanovic confirmed Dr. Kohl's finding that the plethora of topics that GIZ projects have to address has overwhelmed project design and implementation, leaving little time, attention or prioritization to scaling. She confirmed that short-term projects, in the GIZ case of three years duration, focus first and foremost on project goals that are achievable within the project time frame. She also agreed that getting program advisors to be explicit in their tradeoffs between relevance, sustainability, quality, and quantity is important, and that in many cases this could lead to a rebalancing of relevance which currently is usually the highest priority. This is particularly because the more relevant or customized an intervention is at a local level the more difficult it can often be to scale.

She sees an important opportunity to make major changes in GIZ work that will favor scaling. Starting in late Spring 2023 GIZ will be implementing a new strategy that calls for it to go beyond individual projects towards pursuing integrated solutions that address global challenges. Ms. Stojanovic also noted that capacity building in scaling for planning advisors and integrating scaling into monitoring and evaluation were issues that GIZ would consider.

Arlett emphasized that currently GIZ is pursuing scaling through what it calls standardization. Standardization is not that the same approach should be implemented and replicated everywhere, but rather having a portfolio of building blocks or specific interventions that have been proven in multiple contexts and that can be combined, modified, and adapted to new contexts so that not every project has to reinvent the wheel. In this view each project would have a combination of customized, standardized building blocks.

Bishnu Datta Gautam reflected on his experience on working with GIZ projects in Nepal. He highlighted GIZ's work in urban governance and its relationship to Nepal's recent federalized governance structure, particularly in the area of helping local governments (municipalities) improve resource generation and tax



collection. The initial project that worked with ten municipalities was highly effective and the Government of Nepal fast tracked scaling of improved revenue generation to all municipalities in the country. This revenue improvement plan was a high priority for the national government and obviously of strong interest to local governments. Not only were the fiscal interventions contained and scaled through a GIZ follow-on project and technical support, they were also mainstreamed into national government efforts and other donor projects working with municipalities. These became mutually reinforcing.

Ben Kumpf from the OECD DAC had several takeaways from this discussion for the bilateral community. He noted that despite the fact that bilateral assistance has steadily increased, it has been swamped by needs generated by ongoing crises like the pandemic and the related issue of food systems and security. Greater adoption of good practices and governance are necessary but not sufficient, scaling is required by which he meant not just more resources, but more impact from any given level of resources.

Looking at innovations specifically, Ben stated that a recent DAC meeting concluded that only small pockets of certain bilateral donor organizations innovate strategically and scale them over time. Most innovation activities reflect many of the same problems with scaling in development sector projects. While innovations often do address a substantial real-world problem, they are not designed with a clear vision and strategy for scale from the start should they prove to successfully address that problem. In particular, they explicitly identify an organization with the resources and incentives to take an innovation that 'graduates' from the innovation portfolio to full scale, i.e., an organization to play the intermediary role. Instead, there remains a widespread assumption that if innovators can demonstrate proof-of-concept, whether in the context of a project or pilot, someone else will pick it up and run with it. Who that someone is rarely if ever specified, let alone specifying the strategic efforts, resources and commitment to the advocacy and marketing necessary to affect scaling, hand-off, or technology transfer.

That said, Ben felt that there were two outliers that were positive examples of good scaling practices. One was work on market systems development that look systemically as to how we can use a portfolio of efforts to support small local markets. The other was that some innovations did have a strategic vision of scale through the public sector, usually in very stable governance contexts. In these cases, government partners were intimately involved in project design and implementation from the outset – actual participation and not just approval; in Ben's view this is very much what the current emphasis on localization should be about. In many cases this was about scaling local innovations, particularly in the health sector. He cited as an example the launch of and support for local innovation platforms and ecosystems in health services by Grand Challenges Canada and USAID's Global Health bureau. These have explicitly looked for or supported local innovations that could be scaled through the public sector.

Finally, Ben noted that the DAC monitoring and evaluation guidelines already provide a solid foundation for supporting scaling. Nonetheless the guidelines could benefit from further reinforcement in regard to scaling. To facilitate this, the DAC has created a task force to identify and recommend ways to strengthen not only evaluation guidelines but strategic approaches to innovation and scaling generally. One example of this is more flexible financing mechanisms that would be more aligned with the adaptation that usually characterizes the innovation and scaling process and has historically not characterized project approaches.

Larry Cooley drew several lessons from the presenters. First, in terms of sustainability, scaling ultimately must transfer interventions to either local private or public sector actors, or to a partnership of the two. If projects or innovators want to scale permanently, they need to identify which of those pathways they intend to pursue, and how to get there. Second, the predominant modality of donor projects, while it has its strengths, presents challenges for scaling. If projects and their evaluations are to support scale, they do need to take seriously the term "catalytic" and move from self-contained and delimited efforts to instruments of systemic change. Finally, scaling requires taking seriously host country leadership and localization as both Ben Kumpf and Bishnu Datta Gautam noted, be it markets or government. This



means identifying a local change where the forces pushing for change are almost strong enough to move from stasis to action. In other words, where there is an opportunity to shift the balance by facilitating a process or addressing existing constraints; particularly where an external actor might have an advantage or role, albeit temporary. By focusing on the local definition of the problem and what is feasible in terms of local institutions, external actors can truly play a catalytic role that affects large scale change rather than achieving a limited set of outcomes and facts on the ground.

## Lessons learned

There are a number of takeaways from the GIZ experience which may be useful to other funder organizations and the scaling in international development generally. These are:

1. It is important to have a clear internal definition of what achieving scale or impact at scale is, and to distinguish between processes that lead to scale versus scaling, i.e., the noun versus the verb. Scale should be targeting large scale, as defined as some percentage of the problem or the denominator of those in need. Otherwise, any and all levels of reach or coverage can be described as scale and the term loses its meaning. In the GIZ case, it appears that anything that looks like vertical, horizontal, or functional scaling therefore results in “scale” being achieved, even if in terms of the pure quantity dimension scale is actually quite small. Having a clear, common definition of large scale, impact at scale, and scaling creates common goals, language and understanding and helps an organization align internally. This definition needs to then be translated into a meaningful vision of large scale for individual projects or at least follow-on projects,
2. The strong emphasis on consultations with local stakeholders as part of relevance, systems analysis, adaptability, and sustainability in GIZ project designs and evaluations are strengths of GIZ and create some of the most important preconditions for scaling. While systems change, technical assistance and capacity building can and should be used to help support scaling, GIZ and donor resources in general are usually too limited to address the gap between the resource and implementation requirements and constraints at scale alone. If scaling is not built into project designs from the beginning, then subsequent scaling via standardization or horizontal scaling needs to anticipate complementing additional systems change with substantial modifications to the project design. These start with achieving cost reductions or economies of scale and scope and aligning with limited implementation capacity.
3. GIZ has done more than most bilateral funders in creating a scaling framework and tools and offering technical support; a scaling guidance document, tools, and examples of scaling. These have probably contributed to the general awareness of scaling amongst the staff. At the same time, it appears that these are not much used in project design and implementation generally. However, knowledge may lead to awareness and even a change in attitudes, but it often does not by itself lead to action or changes in behavior. In the case of scaling at GIZ, since neither project goals, time frame nor staff incentives encourage scaling in initial projects, the greater knowledge of GIZ staff about scaling has not led to much demand for scaling support services. In the face of competing demands on staff time, scaling will not be a priority unless that is clearly stated by leadership; demand for scaling support will need to show that it can save busy project designers and managers time and effort.
4. In the case of GIZ the proliferation of cross-cutting themes that have to be included in all projects, e.g., climate change, gender, extreme poverty, crowds out considerations of scaling. This is reinforced by the fact that in general, projects with numerous goals and broad mandates usually contain many individual components and interventions and are difficult to scale. Staff are overwhelmed by all of these, creating one of probably many organizational obstacles to scaling. This



appears representative of many donor organizations. There are certainly legitimate differences of opinion on how to balance these tradeoffs, thus the concept of optimal scale. However, in the GIZ case and no doubt others, decisions about tradeoffs between relevance and quality vs. scale, or cross-cutting objectives versus scale, are not being made explicitly. At least in the GIZ case, scale appears the residual category in terms of what quantity can be achieved after all other considerations or priorities have been taken into account. At a minimum, this implicit prioritization needs to be made explicit, if not reconsidered.

5. GIZ's responsiveness to local stakeholders means that when the latter are interested in scaling, usually through a follow-up project, that interest is likely to be acted upon. The same is true when co-funders are interested in supporting a follow-up scaling project. Previous studies of IFAD also show that a focus on local stakeholders and partnerships facilitates scaling; when IFAD has an in-country presence follow-up projects are more likely. However, the GIZ study confirms that if scaling is not built in from the beginning it makes subsequent scaling more difficult, consuming time and resources. GIZ projects that cover X multiple in terms of geography or population rarely have X times the resources (and that tends to be even more true in terms of local contributions, even if they are in kind). As a result, in the GIZ case, follow-up scaling projects often have to redesign interventions to take into account resource and capacity constraints at scale, both on the part of GIZ as well as domestic constraints for sustainability. This occurs despite the fact that GIZ apparently does an excellent job in mapping systems and systems constraints and aligning its projects with local constraints because this mapping and alignment is mostly done at (small) project scale not for the systems and constraints that exist at large scale. Thus, before scaling can proceed, mapping has to be repeated and alignment with large-scale constraints is needed.
6. As noted, several times, there are many positive factors in GIZ's project approach that could and eventually do support scaling. However, GIZ could increase its leveraging of this foundation if there were an explicit mandate for scale and scaling from GIZ's senior leadership i.e., contained in strategy documents, translated into objective measurable goals, and included in some percentage of initial project designs from the beginning. The same is true if the obstacles to scale listed above probably are to be addressed and overcome.
7. The GIZ case illustrates that incentives matter. Project designers respond to clear messages about what factors have primary importance, whether that be gender, climate-change, or relevance to local stakeholders. In GIZ, responsiveness to local actors, capacity building and taking into account local context are high priorities, and projects are designed and implemented that way. When local actors prioritize scale, and the resources are available, GIZ does scaling.
8. Explicit use of a portfolio approach that at least tracks (or sets targets for) either how many projects are attempting to achieve large scale, impact at scale (by country, region or sector), or the overall impact of scale of GIZ's efforts might be an important first step in prioritizing scale and scaling, or at least being explicit about tradeoffs. At a minimum, it would provide a clear view on how resources are being allocated and what impact is being achieved. This would allow GIZ to look at its overall impact in terms of multiple dimensions, starting with the components of Broad Impact – quality, quantity, relevance, and sustainability – and set explicit targets for quantity or scale. Even without building scaling in from the beginning, this could still take the form of targets for, e.g., the percentage of projects that will have follow-up projects, the quantity projects or sector efforts will achieve individually and, in the aggregate, or the number of projects that will achieve impact at national scale. A portfolio approach would help GIZ to measure their contribution to global goals like the SDGs and climate change. It would seem that the new strategy favoring integrated solutions with global impact would be facilitated by applying a portfolio approach.



9. The GIZ monitoring, evaluation, adaptation, and learning (MEAL) systems and frameworks already contain several elements that support scaling but require supplementation as well as an explicit focus on the topic. Like systems used by many international development organizations, they favor accountability. In the GIZ case a major focus is on the efficiency of inputs use, and whether project goals have been achieved, even if those goals are at small scale. As is generally true with the OECD DAC guidelines on which it is based, GIZ's M&E framework doesn't generate easily accessible documentation for internal or external actors to decide whether or how to scale a project or some of its components. Nor does it provide information relevant to assessing scalability or scaling strategies, e.g., evidence of impact, implementation capacity and resource requirements, sensitivity to local context, unit costs and potential for economies of scale. For GIZ and most organizations, revisions to MEAL frameworks may be a necessary condition for mainstreaming and scaling, and at a minimum would significantly facilitate it.

