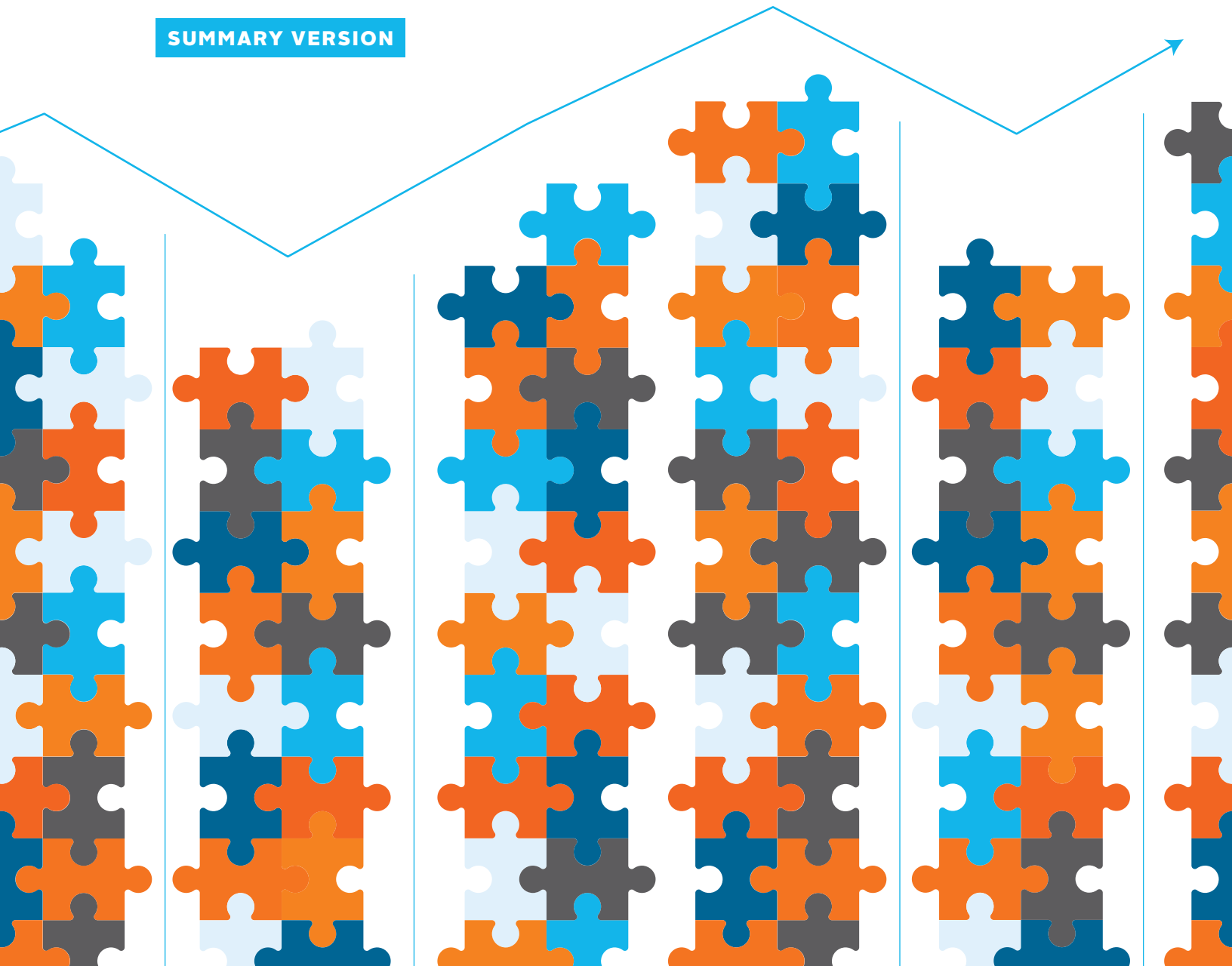
fhi360
THE SCIENCE OF IMPROVING LIVES

Guidance for EVALUATING Integrated Global Development Programs

SUMMARY VERSION



ACKNOWLEDGMENTS

From full-length document

This guidance document is the product of many thoughtful and knowledgeable contributors. We are incredibly grateful for every contribution, large and small. First, we thank the primary authors, who spent many hours thinking creatively on this topic:

Tessa Ahner-McHaffie
Aurélie Brunie, PhD, MS
Mario Chen, PhD, MS
Gina Etheredge, PhD, MA
Greg Guest, PhD, MA
Theresa Hoke, PhD, MPH
Rick Homan, PhD
Christine Kim, MSPH
Tricia Petruney, MA
Merywen Wigley, MSc

Next, we thank the editors: Kerry Aradhya and Natasha Mack.

We also express our extreme gratitude to the reviewers for their thoughtful and constructive comments: Sohail Agha, Laura Ahearn, Joy Noel Baumgartner, Aurelie Brunie, Cindy Clapp-Wincek, Elizabeth Larson, Kate MacQueen, Mike Merrigan, Emily Namey, Hallie Preskill, Diana Rutherford, Pamela Velez-Vega, and Lucy Wilson.

The participants of a workshop on evaluating integration merit our thanks as well: Sohail Agha, Laura Ahearn, Joy Noel Baumgartner, Mario Chen, Cindy Clapp-Wincek, Bob Engelman, Samuel Field, Nathanael Goldberg, Paul Hutchinson, Daniel Kidder, Elizabeth Larson, Katherine Lew, Kristin Lindell, Kate MacQueen, Shira Mitchell, Veronica Olazabal, Hallie Preskill, Kyle Rearick, Laura Reichenbach, and Joseph Sany. Their keen insight and discussion laid the foundation for this document.

Finally, we thank the Bill & Melinda Gates Foundation for providing the guidance and funding to develop this document.

SEPTEMBER 2016

PREFACE

This document includes the executive summary of the **Guidance for Evaluating Integrated Global Development Programs**. The full document[1] includes detailed, integration-specific guidance throughout the evaluation lifecycle, with examples of methodologies and approaches that are best suited for integrated, multisector programs. The guidance can serve as a roadmap for those conducting evaluations of integrated programs and can be used in-full or in-part through the individual modules. This executive summary provides overviews of each of the seven modules. The summaries of each module included here have key questions and key considerations that illustrate some of the issues addressed in the guidance, some approaches that are best suited to evaluate integrated programs, and important questions to consider when planning an evaluation. Although this summary is not comprehensive, it is illustrative and can help guide evaluators to plan thoughtful evaluations of integrated programs.

CONTENTS

Evaluation in a Program Lifecycle 2
Executive Summary 4
Formative Research 6
Developing Performance Indicators 7
Program Monitoring 8
Process Evaluation 9
Cost Analyses 10
Evaluating Impact 11
Evaluating Scale-up 12
Conclusion 13

VIEW THE FULL DOCUMENT HERE:

https://www.fhi360.org/resource/guidance-evaluating-integrated-global-development-programs

1 Ahner-McHaffie, T., Brunie, A., Chen, M., Etheredge, G., Guest, G., Hoke, T., Homan, R., Kim, C., Petruney, T., & Wigley, M. (2016). Guidance for Evaluating Integrated Global Development Programs. Retrieved from https://www.fhi360.org/resource/guidance-evaluating-integrated-global-development-programs

EVALUATION IN A PROGRAM LIFECYCLE

AFTER IMPLEMENTATION

The modules in this guidance document reflect various types of evaluative activities. For the most part, they reflect a logical progression of evaluative activities and processes within a program's lifecycle using a before, during, and after implementation structure. Nevertheless, many of the activities discussed can happen simultaneously and be interdependent. Although these activities have a logical progression (and phases wherein they are most important), time and thought are ideally devoted to the planning and results of each activity throughout implementation.

Executive Summary

INTRODUCTION

Integrated development is an intentional approach that links the design, delivery, and evaluation of programs across sectors to produce an amplified, lasting impact on people's lives. Whether explicitly or implicitly, integrated approaches are based on the premise that the interaction between interventions from two or more sectors will generate benefits beyond a vertical intervention, such as improved outcomes or operational benefits. Evaluation is a valuable tool in making evidence-based judgements about the comparative value of integrated versus vertical programming and therefore about whether and how to implement integrated development programs. Importantly, the desired goal of integration varies greatly depending on the perspective, priorities, and ultimate aim of a given decision maker. Funders may emphasize cost efficiencies or enhanced sustainability, whereas program implementers may prioritize time savings, improving user satisfaction, or reducing inequality. Therefore, the evidence they require for informed decision making will vary in nature.

Yet relative to vertical development programs, integrated approaches to development are more complex in design; more complicated in implementation (with a greater number and more diverse range of actors involved); more diverse in terms of inputs, outputs, and outcomes; and innovative, and therefore possibly requiring more adaptive or emergent thinking. They also bring together different cultures of research and evaluation among sectors. These characteristics have implications for how integrated models are assessed and evaluated with respect to the questions being asked and the methods and designs used to answer those questions.

The purpose of this document is to provide evaluators, funders, and development practitioners with guidance on evaluating integrated development programs. Although general "good evaluation practices" are woven throughout its contents, the document focuses on the unique characteristics of evaluating the complexities associated with integrated, multi-sector program implementation. These unique aspects — centering on the concept of interaction across sectors and activities — provide the common thread that ties this guidance document together.

BEFORE IMPLEMENTATION

1 | FORMATIVE RESEARCH

Formative research informs program content, design, and operation. The primary task is to determine the key problem or nexus of problems, and opportunities and ideas for addressing them prior to implementation.

KEY QUESTIONS

- *How can the interaction among sectors be used to enhance efficiencies, outcomes, or sustainability?*
- *How do stakeholders from different sectors perceive the goals of the integrated intervention and the theory of change that underlies it? What are their expectations and how do they differ?*
- *What challenges might program staff and beneficiaries face with integration?*
- *How can logistical efficiencies be enhanced? Conversely, what are potential challenges and negative effects that might arise due to the interactive nature of the design?*
- *What are potential effects of the integrated program on the community and the larger socioeconomic system?*

KEY CONSIDERATIONS

Mapping the stakeholder environment and the broader system can help develop specific evaluation questions and measures for a complex project, and allows focus on areas that may have the greatest impact.

Early development of an integrated theory of change and/or a logic model, or alignment of different models, can help stakeholders and evaluators identify emergent outcomes, relationships between activities, and best practices.

Inclusive participatory techniques are effective ways to build or revise a theory of change/logic model; enhance cross-sector cooperation and communication; and identify potential areas of convergence, incompatibility, or unintended consequences associated with integration.

Through various methods, stakeholders can provide insights into how integration may affect: services within their sector of operation, the social structure of the larger community, local and regional governance, access to and cost of services, and dynamics within the household (such as gender or age).

2 | DEVELOPING PERFORMANCE INDICATORS

Performance indicators are the precise measures used to assess success by a program or activity. Indicators for integrated programs can include **sector-specific indicators** and **value-added indicators**. Collection of sector-specific indicators is either required or recommended for distinct programs, and can sometimes be standardized by sector or funder. Value-added indicators measure amplified effects or synergy beyond what would have occurred in a vertical program.

KEY QUESTIONS

- *What indicators can be used to measure the expected integrated outcomes (amplified effects or synergies) of the program?*
- *Do program designers or evaluators need to harmonize indicators across sectors?*
- *What indicators can be used to measure activities in two or more sectors?*

KEY CONSIDERATIONS

Careful planning is needed when developing indicators for integrated programs, as the potential for greater number and complexity of indicators is high. Harmonizing indicators between activities and/or stakeholders is necessary and can be achieved by using proxy indicators, adapting data collection to streamline as much as possible, and convening stakeholders to prioritize indicators.

Meetings between experts in particular sectors, or meetings with stakeholders on specific projects, can help identify and harmonize key indicators, and can reinforce the common understanding of the intervention's outcomes and impact and how they will be achieved.

In order to choose or develop value-added indicators, program designers need to identify what "integration" means within the context of the program, and the pathways through which the program is intending to have an effect on its goals. An integrated program can use traditionally sector-specific indicators to track integration if the program anticipates that value-added will be measured through those indicators.

The most thoughtfully chosen indicators need to be paired with a monitoring and evaluation system that is designed to show relationships between the outcomes produced (the instances in which a sector's outcome indicator is the result of multiple aspects of integrated programming).

DURING IMPLEMENTATION

3 | PROGRAM MONITORING

Program monitoring is the routine, systematic observation and recording of program implementation and problems using the performance indicators developed and other monitoring processes, including analysis and feedback about the progress of the program to the donors, implementers, and beneficiaries (for example, through site visits and periodic stakeholder meetings).

KEY QUESTIONS

- *What data collection processes or forms can be adapted to serve more than one sector's purpose?*
- *Who should have access to the data to best facilitate integrated monitoring, and who should meet to discuss the data?*
- *Can progress, or lack thereof, in one sector inadvertently have consequences in another sector (and how can that be monitored)?*
- *How will the monitoring system track beneficiary or household access to or usage of multiple services within different sectors?*

KEY CONSIDERATIONS

Responsibility for the collection, housing, analysis, and reporting of data needs to be clearly documented, and may cross sectors or organizations depending on how the program is designed. Different levels of integration will require different solutions to the problems of integrating monitoring systems.

Integrated programs require integrated program monitoring processes and teams. These teams would ideally comprise both monitoring and evaluation staff with specific expertise in individual sectors and monitoring and evaluation staff experienced in program integration.

It is important to cross-reference changes and identify synergistic interrelationships in order to capture cross-sector changes.

Reporting on an integrated program may be challenging because of a larger number of, and greater variation in, activities being implemented and a wider variety of stakeholders.

4 | PROCESS EVALUATION

Process evaluation is a method of assessing and understanding how a program is being implemented, focusing on the program's operations and service delivery.

KEY QUESTIONS

- *What level of integration is occurring across sectors?*
- *What is the quality of the program components?*
- *Can data on the implementation experience explain how any observed amplified or synergistic effects were achieved? If none were achieved, can the data explain why they were not?*
- *What strategies are working for or inhibiting the cross-sector coordination or collaboration required by the program?*
- *What implementation experiences may be unique to cross-sector programs?*
- *Are the target beneficiary population(s) being reached, and with which activities?*
- *Are households or individuals accessing more than one part of the intervention (and if so, how many)?*

KEY CONSIDERATIONS

Evaluators should maintain a bird's-eye view when seeking and generating evidence, and be aware of what has been achieved by monitoring and what evidence still needs to be explored. Having a good understanding of how integration was supposed to be achieved, and the potential constraints and value added of integration is key for the process evaluator.

Feedback for the process evaluation should be sought from a multi-sectoral group of stakeholders; pains should be taken to make sure no sector dominates the evaluation. Familiarity with systems methods can be an additional asset here to identify parts of the system that need to be evaluated and explored.

5 | COST ANALYSES

Cost analysis is a technique for documenting the extent to which any operational benefits occur in integrated programs and the size of these gains.

KEY QUESTIONS

- *What are the costs of integrated versus vertical programming (including short- and long-term costs, cost efficiencies, financial costs or savings associated with negative outcomes or missed opportunities, etc.)?*
- *What are the cost implications of additional inputs and processes necessary for the successful management, coordination, and delivery of an integrated program that are unique in comparison to the implementation of vertical programs?*

KEY CONSIDERATIONS

Costs of an integrated program will often be frontloaded as compared to standard vertical programs. The requisite expertise and time needed increases as the complexity of program design increases, although this will vary based on the type and extent of integration.

When thinking about costs, it is often useful to distinguish between different phases of a program, including design, preparation, and operational phases. Any operational benefits may not be realized until the operational phase, suggesting higher start-up costs for integrated programs.

Many cost analysis methods require allocating and separating all costs and measures of an outcome to particular sectors or activities. As the degree of integration increases, it is increasingly difficult and subjective to allocate funds to distinct sectors, especially for resources like labor. For this reason, cost-effectiveness analysis of integrated programs is particularly challenging unless a single effectiveness metric can be used.

In most cases, a comparator will be required so that the costs of an integrated program can be assessed with respect to a non-integrated approach.

AFTER IMPLEMENTATION

6 | EVALUATING IMPACT

An **impact evaluation** assesses the outcomes and impact that can be attributed to a particular intervention by comparing outcomes between intervention group(s) and a counterfactual (control group). However, evaluating impact for complex integrated programs is sometimes best served by employing a methodological perspective that extends beyond the traditional experimental and quasi-experimental design. This document adopts a definition of impact that includes a broader range of methods than included in traditional impact evaluations.

KEY QUESTIONS

- *Were the planned amplified effects or operational benefits from integration realized?*
- *To what extent were operational benefits and/or amplified or synergistic outcomes due to integration (if they were observed)?*
- *Can we demonstrate that integration led to these effects?*
- *How and why did integration produce effects beyond those observed in a vertical program? If no change was measured, why did integration not produce effects?*
- *Did integration result in unanticipated effects?*

KEY CONSIDERATIONS

Impact evaluations for integrated models should be undertaken thoughtfully, and only if there is a reasonable consensus on what models need to be tested, and what can be tested, rigorously. Experimental or quasi-experimental designs are not always appropriate for dynamic or complex program models.

The most effective way to measure the degree to which amplification occurs — if at all — is through a full factorial experimental design. Depending on the priorities of the study, however, utilizing a fractional factorial design or testing an integrated program against a counterfactual can be enough to demonstrate success for some aims.

Challenges for impact evaluation include the larger numbers of data points necessary for the counterfactual and the compounding problem of uneven exposure when multiple interventions are involved.

Case studies are appropriate to: identify the how and why of the added value of integration, document non-linear pathways, illustrate context, and describe the process behind observed changes. Systems approaches can also be used to augment the evaluation of integrated programs, and better understand documented amplified and synergistic effects.

Combining qualitative and quantitative methods is important particularly for integrated programs, as combining methods has the potential to reveal not only what occurred, but why.

7 | EVALUATING SCALE-UP

Scale-up is the process of expanding the reach of a successfully tested practice in order to benefit more people and to develop sustainable and institutionalized programs and policies long-term.

KEY QUESTIONS

- *Does the pooling of resources by multiple sectors lead to more robust development solutions that can be sustained over time when implemented on a broad scale, under real-world conditions?*
- *Do integrated models that show promise as proof-of-concept pilots retain their added value and feasibility when scaled up?*
- *Are there priorities or perspectives that supported integration at the pilot level that may not be present everywhere (and what can be done to mitigate these differences)?*

KEY CONSIDERATIONS

Evaluators must take into account that the coordination of different actors may differ from pilot to scale-up, as the feasibility of decision makers from multiple sectors conducting joint budget exercises or sharing supervisory responsibility will be influenced by different factors in different areas. Examining the success or failure of scale-up should include examination of these integrated management and coordination activities.

Organizations or partners supporting an integrated pilot intervention may not represent the motivations of their broader communities; preparing for scale-up requires examination of support for cross-sector collaboration beyond the pilot area, to ensure that it is not merely an idiosyncrasy of a small group of actors.

Expectations should be managed for any project. Ideally, for the scale-up of an integrated project, managers should be able to assign interventions to sites that promise the most success, and should communicate that similar comparison sites may be difficult to find and that variation in implementation is expected. This variation is not necessarily a burden, but is rather an opportunity to better understand the factors favoring and impeding effective implementation of an integrated intervention.

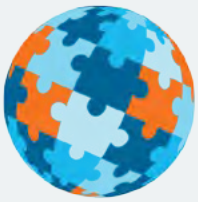
Conclusion

Integrated programming offers unique opportunities for time savings, value for money, enhanced human resources, strengthened capacity of local community or government, provider or user satisfaction, reach, equity, sustainability, amplified effects, and systems change. The integrated nature of the programming also has potential challenges and costs. Indeed, costs for M&E are likely to be greater than for vertical programs.

Evaluation is a valuable tool in making evidence-based judgements about the comparative value of integrated versus vertical programming and therefore whether and how to implement integrated development programs. Importantly, the desired goal of integration varies greatly depending on the perspective, priorities, and ultimate aim of a given decision maker. Funders may emphasize cost efficiencies or enhanced sustainability, whereas program implementers may prioritize time savings, improving user satisfaction, or reducing inequality. Therefore, the evidence they require for informed decision making will vary in nature.

Formative research can help reveal the scenarios and problem sets that are best suited for integrated approaches, and where the need is greatest. Process and impact evaluations can offer proof-of-concept findings to test the feasibility and results of innovative integrated interventions. Implementation science can identify best practices for the replication or scale-up of proven multi-sector models. High-quality monitoring and evaluation within non-research settings can help assess progress and guide subsequent adaptations and improvements. Cost analyses can help to identify the components of a multifaceted program that offer the best return on investment, and the most efficient means to a desired outcome. Generally speaking, greater complexity and diversity — coupled with a focus on interaction or amplification effects — have implications for how integrated development programs are monitored and evaluated. This affects all components of an evaluation — from developing logic models and costing approaches to choosing indicators and an evaluation design. Additional research and evaluation objectives and questions, specific to integration, also need to be considered.

This guidance document raises issues for strategic consideration and provides suggestions derived from experience to help support evaluators, funders, and development practitioners in evaluating the unique aspects of integrated development approaches. Our hope is that this document will be used to improve the evaluation of integrated programs — so that we take the opportunity to learn from what is being done now, and so that the evidence base around integration continues to grow.



THE POWER OF INTEGRATION

fhi360.org/integrated-development

This is a summary version, the full-length version is at **https://www.fhi360.org/resource/guidance-evaluating-integrated-global-development-programs**

fhi360
THE SCIENCE OF IMPROVING LIVES

SEPTEMBER 2016

FHI 360 is a nonprofit human development organization dedicated to improving lives in lasting ways by advancing integrated, locally driven solutions. Our staff includes experts in health, education, nutrition, environment, economic, development, civil society, gender, youth, research, technology, communication and social marketing — creating a unique mix of capabilities to address today's interrelated development challenges. FHI 360 serves more than 70 countries and all U.S. states and territories.

For more information:

GREGORY BECK
Integrated Development Director
askid@fhi360.org