RAND EUROPE

Strategic Impact Evaluation Fund Mid Term Review

Final Report

Joachim Krapels, Jirka Taylor, Marie-Louise Henham, Ben Baruch, Tom Ling

For more information on this publication, visit www.rand.org/t/RR1137

Published by the RAND Corporation, Santa Monica, Calif., and Cambridge, UK

RAND Europe is an independent, not-for-profit policy research organisation that aims to improve policy and decisionmaking in the public interest through research and analysis. RAND's publications do not necessarily reflect the opinions of its research clients and sponsors.

www.rand.org
www.rand.org/randeurope

Preface

The Strategic Impact Evaluation Fund (SIEF) was created with £29.5 million funding from the Department for International Development (DFID) over the period 2011/2012 to 2016/2017.[1] SIEF provides funding for impact evaluations focusing on human development outcomes, for building capacity on impact evaluations and for synthesising and disseminating the findings of impact evaluations along with wider evidence in its thematic subject areas. The fund's objective is to 'strengthen the effectiveness of DFID's and others' global aid policies, programmes and partnerships.' In the long run, SIEF is expected to contribute towards an improved delivery of DFID's and the World Bank's programmes with human development outcomes and towards improved human development outcomes for poor populations. To achieve these goals, the spending of SIEF funds is distributed across three pillars:

1) Impact evaluations of interventions with human development outcomes in four key areas:

   a) Maternal and child undernutrition

   b) Basic education service delivery

   c) Health systems and health related behaviours

   d) Water supply, sanitation, and hygiene for sustainable human development

2) Capacity building on impact evaluation methods

3) Dissemination and knowledge sharing of impact evaluation findings and results

This report presents the results of a mid term review (MTR) of SIEF, commissioned by DFID and undertaken by RAND Europe. The MTR had three main objectives:

1) To assess whether the design, management, implementation and governance of SIEF are appropriate for the achievement of the outputs and outcomes
2) To assess whether SIEF represents value for money for DFID to date and can be expected to do so across the life of the programme
3) To make recommendations for improvements in the design, management, governance and implementation of SIEF for the remaining period of implementation

RAND Europe is an independent not-for-profit research institute whose mission is to help improve policy- and decision making through research and analysis. We realise our mission by undertaking objective, balanced and relevant research and analysis; communicating our findings to a wide audience,

[1] A recent Memorandum of Understanding has extended the collaboration to 2020/2021. SIEF is also currently supported by a US$3 million grant from the Children's Investment Fund Foundation.

often through publications, many of which are available on our website; working in partnership with our clients; and working collaboratively with others. RAND Europe's work lies on the continuum between that of universities and consultancies, combining the academic rigour of universities and the professional, task-oriented approach of consultancies.

In order to declare any potential conflict of interest, we note that three researchers affiliated with RAND Corporation have been either recipients of or applicants for SIEF funding. As such, they were among the invitees to the survey conducted as part of this mid term review. None of these individuals were among the authors of this report.

For more information about RAND Europe or this document, please contact Professor Tom Ling:

RAND Europe
Westbrook Centre, Milton Road
Cambridge, CB4 1YG, United Kingdom
tling@rand.org

Abstract

The Strategic Impact Evaluation Fund (SIEF) was created with £29.5 million funding from the Department for International Development (DFID) over the period 2011/2012 to 2016/2017.[2] SIEF provides funding for impact evaluations focusing on human development outcomes, for building capacity on impact evaluations and for synthesising and disseminating the findings of impact evaluations along with wider evidence in its thematic subject areas. In the long run, SIEF is expected to contribute towards an improved delivery of DFID's and the World Bank's programmes with human development outcomes and towards improved human development outcomes for poor populations.

This report presents the results of a mid term review (MTR) of SIEF, commissioned by DFID and undertaken by RAND Europe. The MTR had three main objectives:

1) To assess whether the design, management, implementation and governance of SIEF are appropriate for the achievement of the outputs and outcomes
2) To assess whether SIEF represents value for money for DFID to date and can be expected to do so across the life of the programme
3) To make recommendations for improvements in the design, management, governance and implementation of SIEF for the remaining period of implementation

The basis for our conclusions and recommendations related to these research questions is evidence collected and synthesised through: a document and data review; 2 inception meetings with 6 participants; in-depth interviews (n=14); a survey including respondents who were unsuccessful applicants (n=89), respondents who were successful candidates (n=37), and Call 3 seed-funded applicants (n=2); and an online focus group with 10 DFID staff.

[2] A recent Memorandum of Understanding has extended the collaboration to 2020/2021. SIEF is also currently supported by a US$3 million grant from the Children's Investment Fund Foundation.

Table of Contents

Figures

Tables

Summary

Introduction

This report presents the results of a mid term review (MTR) of the Strategic Impact Evaluation Fund (SIEF), commissioned by the Department for International Development (DFID) and undertaken by RAND Europe. The MTR had three main objectives:

1) To assess whether the design, management, implementation and governance of SIEF are appropriate for the achievement of the outputs and outcomes
2) To assess whether SIEF represents value for money for DFID to date and can be expected to do so across the life of the programme
3) To make recommendations for improvements in the design, management, governance and implementation of SIEF for the remaining period of implementation

Research questions

The objectives of this MTR (as outlined above) revolve around six core research questions, which are outlined in the box below and which have been taken from the original terms of reference.

Box 1-1. Six core research questions

RQ1: Is the design of SIEF appropriate to achieve the intended outputs and outcomes? Specifically, how do SIEF systems and processes respond to the needs outlined in the business case? (Organisation for Economic Co-operation and Development Development Assistance Committee [OECD DAC] criteria: efficiency, effectiveness)

RQ2: To what extent are the funds being used appropriately and generating value for money (where VfM pertains to the extent to which issues of economy, efficiency and cost-effectiveness are embedded or not in working practices)? What are approx. (expected) unit costs of SIEF evaluations so far, and how do these compare to similar IEs? (OECD DAC criteria: efficiency, effectiveness)

RQ3: Pillar 1: Have the calls for proposals been managed effectively, and are the quality assurance mechanisms SIEF has in place for impact evaluation products adequate? (OECD DAC criteria: efficiency)

RQ4: Pillar 2: Are the impact evaluation trainings and clinics delivered so far relevant to beneficiaries and of high quality? (OECD DAC criteria: efficiency, effectiveness)

RQ5: Pillar 3: Are the initial evidence-to-policy notes of high quality, relevant and used? (OECD DAC criteria: effectiveness, relevance)

RQ6: Is gender being adequately considered in SIEF impact evaluations proposals and capacity development activities to date?

The basis for our conclusions and recommendations related to these research questions is evidence collected and synthesised through: a document and data review; 2 inception meetings with 6 participants; in-depth interviews (n=14); a survey including respondents who were unsuccessful applicants (n=89), respondents who were successful candidates (n=37), and Call 3 seed-funded applicants (n=2); and an online focus group with 10 DFID staff.

Result by research question

We structured the MTR in such a way that the first research question, which covers the entire programme, was addressed last and thereby built on the other five research questions. Thus, we came to the first research question last.

RQ2 – Value for money: The collected evidence suggests that SIEF unit costs for impact evaluations, while somewhat higher, do not differ substantially from those observed in other, broadly comparable settings. This applies both to the size of individual grants provided by SIEF and to the overall costs of the impact evaluation in question. In addition, costs stemming from the embedding of SIEF within the World Bank (WB) (i.e. administration fee and management costs) compare favourably to costs faced by donors in other impact evaluation (IE) funding programmes, and it is not possible to quantify the benefits resulting from this arrangement. Value for money is also driven by the effective management of Pillar 1, the value added from the workshops in Pillar 2 and the impacts of the evidence-to-policy (E2P) notes (all discussed below).

RQ3 – Pillar 1: In terms of timeliness and objectivity, the calls appear to have been managed effectively, yet the MTR raises some questions over the transparency and inclusivity of the calls. The perceived lack of transparency identified through the survey and in one interview is likely to be the result of the limited resources SIEF has available to communicate back on unsuccessful applications. With regards to inclusivity, it is clear that for the moment the inclusion of researchers from the Global South is not extensive. The quality assurance mechanisms in place appear to be adequate. Data from the interviews and the focus group confirm that SIEF is perceived as being good at selecting high-quality impact evaluations.

RQ4 – Pillar 2: On the basis of the available data it appears that the workshops that have been organised have been of high quality and have been relevant to the participants. One element that may require further attention is learning. Beyond the strict confines of research question 4, there is also a wider question as to what the best approach to capacity building might be. Interviews with ‘peers’ show that there are different views about ‘what works’ in capacity building, and that other programmes tend to work more closely and directly with researchers in, for example, the Global South to increase capacity. This raises a question, however, as to what the remit of SIEF is within the wider landscape of strengthening development. As highlighted under the first research question, this is not a question solely for the SIEF core team, but also one for the funders.

RQ5 – Pillar 3: The E2P notes are generally deemed to be of high quality by interviewees. Several interviewees made suggestions, which have been summarised in this chapter. Most important was perhaps the suggestion to include more information on the actual implementation of the interventions discussed.

E2P notes were generally considered to be relevant, yet several interviewees did comment that notes on their own will not be sufficient to bring evidence to policymakers. As such, the effectiveness of the E2P notes to inform policymakers was questioned. This opens up the realm of alternative activities of communication and dissemination, a realm that is already being explored by SIEF.

RQ6 – Gender: Gender is considered to different degrees in the activities of SIEF, but on the basis of the available data it cannot be concluded that gender has been considered adequately. First, while data collected under Pillar 1 is disaggregated by gender, more is needed to move towards 'gender-informed research designs', i.e. research designs that consider how gender might play a foundational role and how data might be collected on it. Second, female participation at the workshops could improve, especially in specific locations. The overall average of 41% against a backdrop of low women's participation in development research in the South is good, but to redress existing imbalances a higher target might be appropriate. It would also be helpful to know what the background of these women was (e.g. researcher or government official). Third, the E2P notes have reported on gender, but this has not always been done systematically or elaborately. For all three pillars gender therefore has been considered, but the dimension of gender could be addressed more systematically.

RQ1 – Overall assessment: The majority of interviewees agree that SIEF is on track to achieve its outputs, and the output figures also show that substantial progress has already been made to achieve the targets. In terms of achieving outputs, the design of SIEF therefore seems to have been appropriate. It has been able so far to rigorously select and subsequently support impact evaluations; to organise high-quality and relevant workshops and to publish a considerable number of generally well-received E2P notes. The results from the interviews, the focus group and the survey indicate that applicants and grantees view the quality and support from the SIEF core team as vital to outputs and to progress. Through the pillars, SIEF systems and processes have largely been able to respond to the needs of the business case, at least in terms of outputs.

Acknowledgements

The authors would like to thank David Rinnert, Anna Henttinen, Andrew Shaw, Chris Berry and Stephen Porter from DFID for continued project support. In addition, the authors would like to thank, and acknowledge the cooperation and support received from, Joost de Laat, Julieta Trias, Zaure Schwade and Aliza Marcus from the SIEF core team.

We would also like to thank Dr Christian van Stolk and Alexandra Pollitt (RAND Europe) for their review of the document and constructive comments in their role providing Quality Assurance on this project.

Abbreviations

CIFF	Children's Investment Fund Foundation
DFID	Department for International Development
DIME	Development Impact Evaluation
E2P	evidence to policy
ECD	early childhood development
IE	impact evaluation
LIC	low-income country
MTR	mid term review
NGO	non-governmental organisation
OECD DAC	Organisation for Economic Co-operation and Development Development Assistance Committee
PI	principal investigator
RQ	research question
Spanish IEF	Spanish Impact Evaluation Fund
SIEF	Strategic Impact Evaluation Fund
TTL	task team leader
VfM	value for money
WB	World Bank

1. Introduction and background

1.1. Background: the need for evidence

Impact evaluations play a critical role in the human development agenda. By focusing on the relationship between cause and effect and by including a counterfactual, they can provide authoritative guidance and evidence on the effectiveness of various policy approaches. As such, they not only represent accountability and monitoring tools, but also form the basis for decisions on whether individual programmes and actions should be continued, terminated or scaled up. What is more, impact evaluations are frequently able to comment on causal links between sectoral policies and cross-sectoral outcomes, thereby informing the understanding of and solutions to complex developmental challenges (Legovini 2010).[3]

The World Bank has established itself as a prominent centre for funding and conducting impact evaluations, a process that has manifested itself through the creation of Development Impact Evaluation (DIME) in 2005 and its re-launch in 2009, and the launch of the Spanish Impact Evaluation Fund (henceforth referred to as Spanish IEF) in 2007, which was hitherto the largest World Bank–run trust fund with an explicit focus on impact evaluation.[4] Following the closure of the Spanish IEF, the current Strategic Impact Evaluation Fund (SIEF) was established in 2012. This report presents the results of a mid term review (MTR) of SIEF, commissioned by the Department for International Development (DFID) and undertaken by RAND Europe. The MTR had three main objectives:

1) To assess whether the design, management, implementation and governance of SIEF are appropriate for the achievement of the outputs and outcomes
2) To assess whether SIEF represents value for money for DFID to date and can be expected to do so across the life of the programme
3) To make recommendations for improvements in the design, management, governance and implementation of SIEF for the remaining period of implementation

[3] Legovini, A. (2010) Development Impact Evaluation Initiative: A World Bank–Wide Strategic Approach to Enhance Developmental Effectiveness. Washington, DC: World Bank.

[4] Feinstein, ON (2012) Final Program Assessment Report: Spanish Impact Evaluation Fund (SIEF). As of 1 July 2015: http://siteresources.worldbank.org/INTHDOFFICE/Resources/5485648-1332253705502/SIEF_MG_120603_Final_Program_Assessment_Report_Osvaldo_Feinstein.docx

1.2. About the Strategic Impact Evaluation Fund

The Strategic Impact Evaluation Fund provides funding for impact evaluations focusing on human development outcomes, for building capacity on impact evaluations, and for synthesising and disseminating the findings of impact evaluations along with wider evidence in its thematic subject areas. Its objective is to 'strengthen the effectiveness of DFID's and others' global aid policies, programmes and partnerships.'[5] In the long run, SIEF is expected to contribute towards an improved delivery of DFID's and the World Bank's programmes with human development outcomes and towards improved human development outcomes for poor populations. To achieve these goals, the spending of SIEF funds is distributed across three pillars:

1) Impact evaluations of interventions with human development outcomes in four key areas:

a) Early childhood development and nutrition

b) Basic education service delivery

c) Health systems and health related behaviours

d) Water supply, sanitation, and hygiene for sustainable human development

At the time of writing, none of the impact evaluations has been completed, and the focus of the MTR was therefore on the organisation of the calls for proposals and the management of grants, rather than on any outcomes or impacts achieved by the funded impact evaluations.

2) Capacity building on impact evaluation methods. Pillar 2 currently includes a range of capacity-building activities, but, in accordance with the terms of reference, the focus of the MTR was on the workshops previously organised.

3) Dissemination and knowledge sharing of impact evaluation findings and results. Pillar 3 currently also includes a range of dissemination and communication activities, but, in accordance with the terms of reference, the focus of the MTR was on the evidence-to-policy (E2P) notes.

1.3. Research Questions for this MTR

The objectives of this MTR (as outlined above) revolve around six core research questions, which are outlined in the box below and which have been taken from the original terms of reference.

Box 1. Six core research questions

RQ1: Is the design of SIEF appropriate to achieve the intended outputs and outcomes? Specifically, how do SIEF systems and processes respond to the needs outlined in the business case? (Organisation for Economic Co-operation and Development Development Assistance Committee [OECD DAC] criteria: efficiency, effectiveness)

RQ2: To what extent are the funds being used appropriately and generating value for money (where VfM pertains to the extent to which issues of economy, efficiency and cost-effectiveness are embedded or not in

[5] SIEF is also currently supported by a US$3 million grant from the Children's Investment Fund Foundation.

working practices)? What are approx. (expected) unit costs of SIEF evaluations so far, and how do these compare to similar IEs? (OECD DAC criteria: efficiency, effectiveness)

RQ3: Pillar 1: Have the calls for proposals been managed effectively, and are the quality assurance mechanisms SIEF has in place for impact evaluation products adequate? (OECD DAC criteria: efficiency)

RQ4: Pillar 2: Are the impact evaluation trainings and clinics delivered so far relevant to beneficiaries and of high quality? (OECD DAC criteria: efficiency, effectiveness)

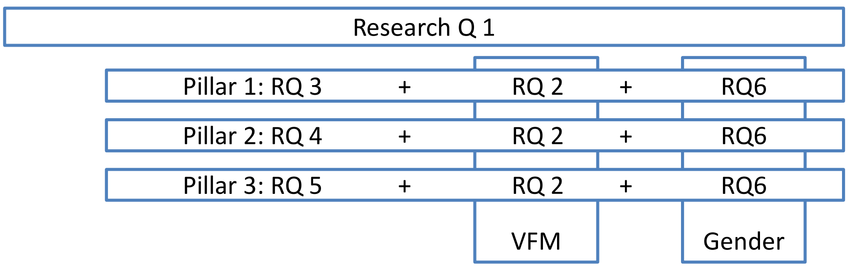
RQ5: Pillar 3: Are the initial evidence-to-policy notes of high quality, relevant and used? (OECD DAC criteria: effectiveness, relevance)

RQ6: Is gender being adequately considered in SIEF impact evaluations proposals and capacity development activities to date?

To thoroughly review the research questions, it was necessary for us to approach them from both an internal and an external viewpoint. The internal viewpoint requires us to examine how SIEF itself has been designed, while the external viewpoint places SIEF in the wider environment of impact evaluations to examine how SIEF is similar to or different from related initiatives. This interplay between the internal and the external viewpoint will be emphasized for a number of research questions.

Figure 1 below captures our understanding of the relationship between the basic research questions and their mutual interconnectedness. Q1 can be conceptualised as an overarching question to this MTR; as will be discussed in greater detail in Chapter 8, any answer to this question will to a large extent be informed by an exploration of the five remaining research questions. Q3, Q4 and Q5 correspond to the individual pillars that SIEF funds are organised in, i.e. impact evaluations of interventions (Q3), capacity building (Q4) and dissemination and knowledge sharing (Q5). Cutting across these three pillars are two underlying themes: value for money (Q2) and the gender dimension (Q6). These two themes are explored in all three pillars. This approach has enabled us to capture learning by taking into account the linkages between individual questions and by avoiding possible duplication of work that could result from considering each question in isolation.

Figure 1: Interpretation of the interrelations between the research questions



In addition to the six primary research questions, there are additional, secondary research questions of interest that have been explored in the course of the review and that have also been taken from the original terms of reference. Data collection that was intended to answer the basic research questions has also yielded information that has enabled the research team to formulate findings and recommendations with respect to these secondary areas. Table 1 below outlines under which primary research question the secondary research questions are addressed. At the end of each chapter, under the 'Summary and Answer' heading, the secondary research questions will be answered. For only one research question did we find that insufficient information was available to address it. This is the second-to-last research question in Table 2, on a funding cap. Unfortunately, this question therefore remains unanswered at this point.

Table 1: Overview of secondary research questions

Secondary question	Primary question under which it is addressed
Do the four key priority areas remain highly relevant (due to lack of robust evidence)? Are additions and amendments appropriate?	RQ1
Have the five key risks outlined in the SIEF business case so far been mitigated effectively?	RQ1
Has the collaboration between those receiving grant funding and SIEF so far been effective?	RQ3
Are the Secretariat's and the Board's role and configuration appropriate to the needs?	RQ1
Is the coordination between SIEF and other impact evaluation funds and initiatives effective?	RQ1
What progress has been made towards securing funding for SIEF from other donors?	RQ1
What kinds of beneficiaries participate in capacity development activities, and how are they being selected?	RQ4
Is the funding cap for individual studies having the desired effect? Are there unintended effects, for instance, on the ambition, scope or quality of proposals?	RQ3
Do SIEF and the individual IEs have uptake strategies and plans that give confidence that desired outcomes will be achieved?	RQ5

1.4. Understanding of evaluation criteria: what is relevance, what is effectiveness and what is efficiency?

As we indicated above, the research questions for this study are linked to several of the OECD's Development Assistance Committee (DAC) criteria for evaluating development assistance, as presented in the DAC Principles for Evaluation of Development Assistance,[6] namely, relevance, effectiveness and efficiency. The OECD defines these criteria as follows:

Relevance: The extent to which the aid activity is suited to the priorities and policies of the target group, recipient and donor. In evaluating the relevance of a programme or a project, it is useful to consider the following questions:

[6] OECD (n.d.) DAC Criteria for Evaluating Development Assistance. As of 25 June 2015: http://www.oecd.org/development/evaluation/49756382.pdf
OECD (1991) Development Assistance Committee Principles for Evaluation of Development Assistance. Paris: OECD. As of 25 June 2015: http://www.oecd.org/dac/evaluation/50584880.pdf

- To what extent are the objectives of the programme still valid?
- Are the activities and outputs of the programme consistent with the overall goal and the attainment of its objectives?
- Are the activities and outputs of the programme consistent with the intended impacts and effects?

Effectiveness: A measure of the extent to which an aid activity attains its objectives. In evaluating the effectiveness of a programme or a project, it is useful to consider the following questions:

- To what extent were the objectives achieved/are they likely to be achieved?
- What were the major factors influencing the achievement or non-achievement of the objectives?

Efficiency: Efficiency measures the outputs – qualitative and quantitative – in relation to the inputs. It is an economic term which signifies that the aid uses the least costly resources possible in order to achieve the desired results. This generally requires comparing alternative approaches to achieving the same outputs, to see whether the most efficient process has been adopted. When we evaluate the efficiency of a programme or a project, it is useful to consider the following questions:

- Were activities cost efficient?
- Were objectives achieved on time?
- Was the programme or project implemented in the most efficient way compared with alternatives?

In addition, the research questions introduce the terms appropriateness and adequacy. For the purposes of this review, *appropriateness* and *adequacy* both address the issue of whether a given aspect or component of SIEF and its operationalization is compatible with and suitable for the attainment of its objectives.

The diversity of the activities undertaken by SIEF (funding of impact evaluations, organisation of workshops) means that these evaluation concepts are operationalised in slightly different ways depending on the research question and SIEF activity. At the start of each chapter, we will outline how the criteria relevant to the respective research question are operationalised and what the benchmark will be for its assessment.

1.5. Structure of the report

This report is structured as follows: The remainder of this report discusses the objectives of and research questions for this review. The subsequent six chapters are structured around the six core research questions. The first research question will be treated last, however, as it brings together the entire MTR. After providing an outline of the methodology, we will therefore start with the research question on value for money.

For each question, we outline our interpretation of the question and include an account of evidence collected through the methods used to address the pertinent question, such as semi-structured interviews or a survey. Finally, the last chapter provides the research team's assessments and recommendations for DFID and the SIEF management team.

2. Methodology

2.1. Overview of approach

To quickly collect a substantial amount of data to inform the MTR, we chose a multi-method approach to allow different strands of data collection to occur simultaneously. Data gathered through the different methods, in turn, informed the responses to various research questions. Table 2 provides an overview of the methods applied and the research questions to which they provide input. We reiterate that the six core research questions are interrelated, which is reflected in the methodology. For example, interviews conducted to generate data primarily to answer research question 2 might also serve to generate data to answer research question 3.

Table 2. Overview of methods used in this review

Method	Participation/completion	Related research questions
Document and data review	n.a.	RQ1, RQ2, RQ3, RQ4, RQ5, RQ6
Inception interviews	2 inception meetings (DFID and SIEF); 6 participants	RQ1, RQ3, RQ6
Interviews in-depth	14 participants	RQ1, RQ2, RQ3, RQ5, RQ6
Survey	Unsuccessful applicants: 89 Successful applicants: 37 Call 3 seed funding: 2	RQ1, RQ3
Online focus group	10 participants (DFID staff)	RQ1, RQ3, RQ4, RQ5, RQ6

2.1.1. Document and data review

Different types of SIEF data documentation have been reviewed as part of the MTR, and their selection was driven by the six core research questions. The approach to the review was informed by the research questions; thus data and documentation were reviewed with the aim to provide answers. This targeted approach improved the efficiency with which the documents were reviewed. Table 3 shows the different types of data and documentation that have been reviewed and how they have informed the responses to the research questions.

Table 3. Overview of reviewed data and documentation

Data or documentation	Related RQ	Usage
SIEF annual reports	All RQs	The annual reports were used to inform all RQ, as they cover assessments of all Pillars of SIEF.
Annual reviews (by DFID)	All RQs	The annual reviews were used to inform all RQ, as they cover assessments of all Pillars of SIEF.
SIEF business case	All RQs	The business case was used to inform all RQ, as it details the considerations behind the design of SIEF.
Data on successful and unsuccessful applicants	RQ3	The data were used to chart changes in the regions in which successful IE will be conducted and to chart changes in the background of successful and unsuccessful principal investigators (PIs) (e.g. their location).
Workshop assessments and workshop pre- and post-tests	RQ4, RQ6	Data on training attendance has been used to understand the background of participants (e.g. the percentage female).
Evidence-to-policy notes	RQ5, RQ6	E2P notes were used in interviews with relevant stakeholders and have been examined to RQ5.
Updated Pillar 2 and 3 strategy	RQ5	Information on the updated strategy was used to align recommendations to changes already in progress within SIEF.

2.1.2. Interviews: inception and in-depth

Two types of interviews were conducted. First, open discussions during the inception phase informed the framing of the MTR and provided initial thoughts on the direction and progress of SIEF. Second, over the course of the project, in-depth interviews were conducted with a number of different stakeholders. These interviews followed a semi-structured interview protocol (see Appendix B). Not all interviewees were asked all questions in the interview protocol. Depending on the interviewee's relationship to and knowledge of SIEF, some questions were omitted in some interviews.[7] Interviewees were for the majority identified in collaboration with DFID during the inception phase, and they cover a range of stakeholders. The interviews combined served to collect data for all research questions, as Table 4 shows. Because interviews were conducted by multiple researchers, the team held a synthesis workshop to discuss the emerging findings and how stakeholder interviews inform answers to individual research questions.

[7] As such, any discrepancies in the frequency of individual references may be attributable to differences in the ability of stakeholders to speak to various topics of this review.

Table 4. Overview of conducted interviews

Interviewee group	Number of interviews	Related RQ	Usage
SIEF	7	All	SIEF staff were interviewed to generate a thorough understanding of the running and management of SIEF.
NGO	3	All	Non-governmental organisations (NGOs) are potential applicants as well as users of SIEF-generated evidence. Depending on their knowledge of and relationship with SIEF, they were able to provide their perspectives on all research questions.
Peers (e.g. International Initiative for Impact Evaluation [3ie], DIME)	3	RQ2, RQ3, RQ4	Peers are organisations relatively similar to SIEF, even though substantial differences exist. Data on their finances as well as their organisation of IE application and selection and training have been used to compare SIEF with other organisations.
Donors (Children's Investment Fund Foundation [CIFF])	2	All	During the inception phase, in-depth discussion were conducted with DFID. Over the course of the project, in-depth interviews were conducted with CIFF.

Finally, all interviews were numbered and assigned an interviewee group. Interviews are referenced in this report according to their number, as outlined in Table 5.

Table 5: Assigned interview and meeting numbers

Interviewee group	Assigned number
Donor	Interview 1
Donor/Peer	Interview 2
NGO	Interview 3
NGO	Interview 4
NGO	Interview 5
Peer	Interview 6
Peer	Interview 7
SIEF	Interview 8
SIEF	Interview 9
SIEF	Interview 10
SIEF	Interview 11
SIEF	Interview 12
SIEF	Interview 13
Donor	Meeting 1
SIEF	Meeting 2

2.1.3. Survey

To gather data from a large number of successful and unsuccessful applicants to SIEF, a survey was designed to capture views on SIEF as a funder, on the process of application and selection and on the running of the grants. The full list of survey questions, both for the successful and for the unsuccessful applicants to SIEF, are attached in Appendix C. The full survey methodology is described in Appendix D.

The survey of unsuccessful applicants is a shortened version of the survey of applicants that covers the following sections:

- Background
- SIEF and other funders: includes questions to identify what the unique aspects of SIEF are and why researchers apply to SIEF (i.e. anticipated benefits)
- Application process
- Running of the grant: questions on the interactions with SIEF over the course of the grant so far
- Conducting the IE: questions on the challenges encountered in conducting the IE
- Added value: questions on the value that participation in the application and selection process of SIEF may have generated

The survey instrument was developed by the research team based on their previous work evaluating research funding instruments and their understanding of SIEF stemming from document review and inception interviews with SIEF staff. It was further modified to reflect suggestions from SIEF and DFID staff.

Participants

Invitees to the survey included all task team leaders (TTLs) and PIs named on applications to the first three SIEF calls for funding, as provided by the SIEF core team.

Response rates

Response rates to each survey are reported in the table below, both combined and disaggregated by category of respondent.

Table 6. Number of invitees and survey response rates

	Grantees	Applicants	Call 3 participants
Number of unique contacts	79	357	10
Number of excused respondents*	0	12	2
Effective number of invitees	79	345	8
Complete responses	35 (44%)	81 (23%)	2 (25%)
Incomplete responses	21 (27%)	79 (23%)	0 (0%)
Responses included in analysis	**37 (47%)**	**89 (26%)**	**2 (25%)**
Of whom TTLs	*17*	*37*	*1*
Of whom PIs	*20*	*52*	*1*

Note: *Refers to invited participants who told/wrote the research team that they felt they were not in a position to answer the survey questions, for instance due to their limited involvement with SIEF.

2.1.4. Online focus group of DFID staff

In the interest of time, it was decided over the course of the project and in consultation with DFID to launch an online focus group to swiftly collect the perspectives of different stakeholders within DFID on SIEF. In contrast to in-depth interviews, the online focus group allows participants to respond to polling and open-text questions in their own time and at their convenience. Involvement in the online focus group can be spread out over days and includes the possibility for participants to respond to other stakeholders' comments. Based on the protocol of the semi-structured interviews, the RAND team, with support from an external supplier of the online platform, developed the protocol for the online focus group, attached in Appendix E.

The focus group protocol is a many ways similar to the interview protocol, yet a few changes were made to enhance the ease with which it could be completed. First, several questions were transformed into simple polls (i.e. yes/no), followed by open-text questions to elicit further elaboration of the answers chosen. Second, the total number of questions was reduced by the RAND team to focus only on the questions that seemed directly relevant to the participants from DFID.

Participants were identified by DFID. They included, in total, 10 DFID staff, both in-country experts and DFID sector advisors. All participants were first invited by DFID, after which the RAND team continued communication. Participation in the online focus group was voluntary, and answers were displayed anonymously on-screen. To increase engagement with the focus group, questions were spaced out. A first set of questions was launched on Wednesday, 22 April 2015, at 10 a.m. GMT, and a second set of questions was launched on Friday, 24 April 2015, at 10 a.m. GMT. Participants had until Sunday, 26 April 2015, to complete the focus group. Because the participants completed the focus group questions to different degrees, exact response rates are less straightforward, yet when the focus group closed the response rate to the different questions was as follows:

- 8 respondents had answered over 50% of the questions (4 of whom had answered over 90%)
- 2 respondents had answered less than 50% of the questions

The RAND team acted as moderator of the focus group to check participation rates and, in specific cases, to request further clarifications from participants on their answers. As the duration of the focus group was relatively short, follow-up questions were kept to a minimum to first allow participants to complete all the questions of the protocol.

2.2. Data limitations

It is important to recognise that each data collection method has its inherent limitations, which are briefly discussed below. In addition, the MTR prioritised issues identified in the terms of reference and was constrained by the resources allocated to the study. With respect to the online focus group, it is important to recognise that it included a strong element of self-selection from a group of DFID stakeholders and that it had uneven rates of participation. While these stakeholders represent individuals who are in the best position to comment on SIEF and its design, findings derived from this exercise build on a relatively small number of testimonies. Similarly, the number of interviews conducted with SIEF representatives, peers and NGOs was constrained by the scope of this study. It may be that further points and observations could have been elicited had more stakeholders been consulted; however, every effort was made to identify interviewees with the best knowledge of SIEF or in the best position to comment on its design and activities.

The survey results also need to be read with potential limitations and biases in mind. For instance, although guarantees of confidentiality were communicated by the research team, desirability bias is a possibility with respondents focusing on positive comments, particularly if they intended to participate in any future WB- or DFID-sponsored funding calls. Also, classification and analysis of responses is somewhat complicated by the fact that some respondents submitted multiple applications to SIEF, in a few instances in different clusters and in a different role (TTL vs PI). In those instances, we provided instructions to respondents as to what application they should refer to, but we do not have any measure of the degree to which this guidance was followed.[8] Respondent characteristics suggest, however, that the survey data can be understood as being broadly representative, because no particular position/cluster/call for proposal appears to be substantially overrepresented.[9] That said, caution is required with the disaggregation of survey data in questions pertinent only to successful applicants, due to small sample sizes. In summary, in our judgement, and based on later data collection confirming earlier findings, further resources might have strengthened the detail of the conclusions presented here, but not the overall balance.

[8] Though we note that only two survey respondents expressed confusion about what role/application they should use as the basis for their responses.

[9] Two types of irregularities can be observed from participant characteristics; however, both are attributable to the underlying contact database. First, the survey elicited more responses from the EDU and EDC clusters, which corresponds to the fact that the third call for proposal focused exclusively on these two areas. Second, there were comparatively fewer responses from Call 1 applicants, which is attributable to the fact that contact information for unsuccessful PI applicants from this call was not available.

3. Research question 2: SIEF value for money

RQ2: To what extent are the funds being used appropriately and generating value-for money (where VfM pertains to the extent to which issues of economy, efficiency and cost-effectiveness are embedded or not in working practices)? What are approx. (expected) unit costs of SIEF evaluations so far, and how do these compare to similar IEs?

(OECD DAC criteria: efficiency, effectiveness)

The institutional form of SIEF is in line with DFID's business plan commitment to make UK aid more effective by providing greater value for money. DFID understands value for money as an effort to 'maximise the impact of each pound spent to improve poor people's lives.'[10] In order to do so, the 3e's of economy, efficiency and effectiveness need to be assessed and maximized. In this context, *economy* refers to the cost of inputs to the intervention in question given their quality, *efficiency* captures the conversion of inputs into outputs, and *effectiveness* looks at whether the intervention's outputs lead to its desired outcomes. Importantly, VfM is not synonymous to preferring the least costly intervention; rather, it refers to the intervention that delivers the best outcomes in relation to its inputs.

As stated in the business case for SIEF, partnering with the World Bank and making use of its technical and commercial capacity will enable DFID to save on management costs associated with the management of the SIEF[11] while at the same time harnessing the organisational networks and profile of the World Bank. The World Bank has been awarded the highest Multilateral Aid Review (MAR) score of very good value for money and is in this aspect among the highest-ranking recipients of UK aid.[12]

---

[10] Department for International Development (2011) DFID's Approach to Value for Money (VfM). As of 25 June 2015: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67479/DFID-approach-value-money.pdf

[11] Department for International Development (n.d.) Business Case and Intervention Summary: Strategic Impact Evaluation Fund (SIEF) 2011/12–2016/17. As of 25 June 2015: http://iati.dfid.gov.uk/iati_documents/3744993.docx

[12] The Multilateral Aid Review (MAR) is an assessment of the performance of multilateral organisations that receive UK funding for development conducted systematically by the UK government. Department for International Development (2011) Multilateral Aid Review: Ensuring Maximum Value for Money for UK Aid Through Multilateral Organisations. As of 25 June 2015: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/67583/multilateral_aid_review.pdf. Department for International Development (2013) Multilateral Aid Review Update: Driving Reform to Achieve

Since no impact evaluations had been completed in time for the mid term review and thus no outcomes could be assessed, our examination of the use of SIEF funds and their generation of VfM is largely limited to an analysis of economy and efficiency. In practical terms this translates into an assessment of unit costs of funded outputs, of other costs incurred by SIEF and of existing processes guiding the allocation of SIEF funds. Due to unavailability of SIEF and comparator cost data on Pillar 2 and 3 activities, the discussion below focuses primarily on SIEF-funded impact evaluations. However, in our wider discussion of Pillars 2 and 3, we engage with the value (as the first part of the 'value for money' couplet) that they may add.

3.1. Comparing unit costs is methodologically imprecise but suggests SIEF is broadly in line with costs elsewhere

A comparative analysis of IE unit costs is fraught with uncertainties for several reasons. Most fundamentally, IE costs are generally hard to trace because they often use multiple funding sources (IEG 2012),[13] which is the case for SIEF evaluations as well. Other funding sources may include other World Bank trust funds, other external donor agencies, governments etc. Second, it may not always be easy to separate the costs of an impact evaluation from the costs of the very program the impact of which the evaluation is supposed to assess. What is more, the two may influence each other over time. For instance, Legovini (2010) observed that one of the biggest drivers of IE costs, data collection, usually serves not only the impact evaluation in question but also general programme monitoring, which may overstate the costs of the evaluation. At the same time, it may be argued that IE places greater demand on data collection in the first place, which could increase overall project costs when compared with a similar project without an IE component. This is not to say that data collection per se is an issue; quite on the contrary, it makes programmes and measurement and evaluation activities more efficient. However, it may make it harder to capture specific costs in an accurate manner.

Because of the difficulties in tracing IE costs, two types of unit costs need to be distinguished: a) the average size of a grant from each individual funder and b) the overall costs of an impact evaluation. According to available SIEF administrative documents, the average size of a SIEF grant in Call 1[14] was US$582,861, while the average overall cost of a SIEF-funded impact evaluation was US$1,027,136. However, it needs to be noted that when we look at averages, a great deal of variability across individual projects remains hidden. For instance, the smallest SIEF grant was US$180,000, while the largest was US$849,625. In terms of the relative size of SIEF's contribution, the share of SIEF funding as a proportion of overall IE costs ranged from 7% to 100%.

Multilateral Effectiveness. As of 25 June 2015:
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/297523/MAR-review-dec13.pdf

[13] Independent Evaluation Group (2012) World Bank Group Impact Evaluations: Relevance and Effectiveness. Washington, DC: World Bank. DOI: 10.1596/978-0-8213-9717-6. License:

Creative Commons Attribution CC BY 3.0

[14] Data on one funded IE in Call 1 not available. Financial data for Call 2 were not made available to the review team by SIEF and DFID. It is likely that with the introduction of the funding cap for Call 2, the overall average unit cost will have decreased.

To compare SIEF's costs with those of other 'peer' funders of impact evaluations, we attempted to collect data on projects funded by the following entities: other funding sources within the World Bank, 3ie and the Abdul Latif Jameel Poverty Action Lab (J-PAL). It is important to acknowledge that this approach can arrive at best at an approximate comparison, because the contexts in which these organisations operate differ. For instance, mechanisms such as the World Bank's DIME explicitly anticipate that successful applicants will be able to obtain co-funding from other sources, while 3ie made a conscious decision not to introduce such a requirement with a view to not disadvantage researchers from lower-income countries (LIC). Similarly, J-PAL explicitly states that its objective is to conduct randomized evaluations, which may have slightly different cost implications from the quasi-experimental approaches funded by other organisations. Another difference is that the policy areas in which other organisations fund IEs are not always limited to the four SIEF clusters.

As such, the main conclusions from this exercise need to be confined to conducting a high-level assessment of whether SIEF costs are broadly in line with those of peer entities.

Starting with comparator data on the size of funding grants, the final report of the Spanish IEF indicates that the mean unit cost of Spanish IEF grants was between US$150,000 and US$330,000, with the overall average size of IE grants amounting to US$188,807. A detailed breakdown by category is presented in Table 7 below.

Table 7. Spanish IEF unit costs

Activities	Budget (US$)	Number	Mean unit cost (US$)
1. Impact evaluation studies	$9,817,968	52	$188,807
1.1. Quick win IE studies*	$879,851	5	$175,970
1.2. Innovation fund IE studies**	$2,009,223	11	$182,657
1.3. Cluster IE studies	$6,928,894	36	$192,469
1.3.1 Health contracting/pay for performance reforms	$1,250,539	6	$208,423
1.3.2 Conditional cash transfers	$895,086	5	$179,017
1.3.3 Malaria control	$762,790	5	$152,558
1.3.4 Active labor market/youth employment	$1,276,350	8	$159,544
1.3.5 Basic education accountability	$1,022,601	5	$204,520
1.3.6 HIV/AIDS prevention	$981,561	3	$327,187
1.3.7 Early childhood development	$739,966	4	$184,992

Source: Authors' calculations, based on Feinstein (2012).[15]

*First call of proposals from existing Human Development Network impact evaluation clusters. They were 'expected to capitalize on work on Human Development issues that had already been going, searching for proposals of high quality, prospective impact evaluations that had immediate funding needs.'

**IEs not under a specific cluster

[15] Feinstein, ON (2012) Final Program Assessment Report: Spanish Impact Evaluation Fund (SIEF). As of 1 July 2015: http://siteresources.worldbank.org/INTHDOFFICE/Resources/5485648-1332253705502/SIEF_MG_120603_Final_Program_Assessment_Report_Osvaldo_Feinstein.docx

J-PAL grant sizes, as far as available data enable a comparison, appear to be in the same region as those of the Spanish IEF. According to the 2012 annual report (J-PAL 2012),[16] FY2012 awards in the Agricultural Technology Adoption Initiative averaged US$170,000, while FY2012 awards in the Governance Initiative averaged US$136,000.

As for 3ie, average grant size data are available for three open-window calls and are summarized in the table below.

Table 8. 3ie unit costs

Call	Number of grants	Average grant size (US$)
Open Window 1 – 2009	17	$249,000
Open Window 2 – 2010	30	$504,000
Open Window 3 – 2010	21	$423,000

Source: Morton et al. (2012)[17]

With respect to overall costs of impact evaluation projects, it is possible to compare project cost of SIEF-funded projects with those funded under programmes coordinated by DIME. Below is an overview of various programmes in DIME's portfolio as reported in Legovini (2010), which puts a total average cost at slightly over half a million US dollars. It is worth mentioning that the average contribution from donors to each was US$273,785.[18]

Table 9. DIME secretariat–coordinated programmes unit costs

Programme	Average unit cost (US$)
Africa Program for Education Impact Evaluation	$410,919
MIEP Malaria Impact Evaluation Program	$1,111,807
AADAPT-AFR Agricultural Adaptations Impact Evaluation Program in Africa	$604,071
DIME HIV/AIDS Impact Evaluation Initiative	$240,700
DIME FPD Impact Evaluation in Finance and Private Sector Development	$437,143
Average	$526,198

Source: Legovini (2010)

16 Abdul Latif Jameel Poverty Action Lab [J-PAL] (2012) 2012 Annual Report. As of 25 June 2015: http://www.povertyactionlab.org/doc/j-pal-annual-report-2012

17 Morton, J, Shaxson, L & Greenland, J (2012) Process Evaluation of the International Initiative for Impact Evaluation (2008-11). London: Triple Line Consulting/Overseas Development Institute. As of 25 June 2015: http://www.3ieimpact.org/media/filer_public/2013/01/07/3ie_proces_evaluation_2012_full_report_1.pdf

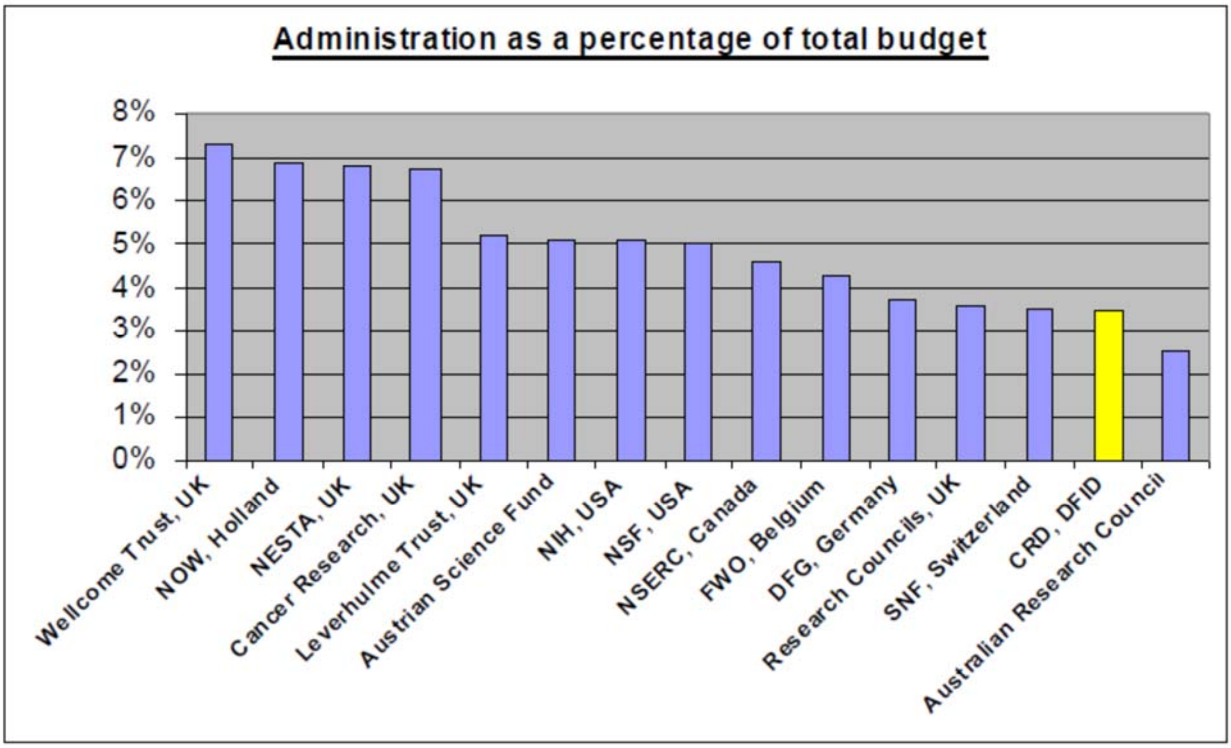
18 The rest of project cost was borne by governments ($225,224) and the World Bank's internal funds ($25,638) (Legovini 2010).

3.2. Other costs in addition to direct project costs should be considered

In order to assess the efficiency of SIEF, on top of direct project costs, it is also important to consider indirect costs stemming from running the programme and its daily operations. As stated in SIEF's administrative documents, the World Bank charges DFID a 5% fee to administer the trust fund, and management costs amount to approximately 5% of its overall budget. These two cost items can also be subject to an imperfect comparison with other funding programmes.

Under the Spanish IEF, programme management costs amounted to 4% of total budget and cluster coordination costs to an additional 2% of total budget (Feinstein 2012). Similarly, 3ie management costs total 5% of total budget. All these figures are broadly comparable to those observed in other funders, according to an overview produced by DFID (see Figure 2 below), as reported by Morton et al. (2012).

Figure 2: Overview of administration costs



Source: Morton et al. 2012, p.72

However, we reiterate the limitations of any conclusions drawn from the presented data. The overview above serves only as a demonstration of the range of research funders' management costs. Due to differences in their contexts and other aspects, such as size, they represent only illustrative comparators to SIEF. In addition, it appears that a perfect separation of costs borne by SIEF funds and other WB-originated sources may not be feasible. For instance, a fraction of the administrative support provided to SIEF is not charged to SIEF but, rather, absorbed by wider World Bank costs. To some extent, this may be an inevitable consequence of the decision to embed SIEF within the World Bank and may very well contribute to the overall efficiency of SIEF. However, for the present purposes, this may obfuscate the true SIEF cost breakdown.

With respect to the 5% trust administration fee charged to DFID, interviews with WB stakeholders confirmed that this is on the lower end of fees charged by the World Bank. As for external comparators, 3ie has a somewhat similar arrangement, in which it paid a hosting fee to Global Development Network (GDN).[19] Morton et al. (2012) calculated this fee to be 7%.

3.3. Processes are in place to strengthen cost-effectiveness

Finally, it is worth emphasizing that the SIEF structure has processes in place that are intended to further ensure an efficient use of existing resources. A budget cap of US$500,000–600,000 (depending on the cluster in question) was introduced during the second call for proposals, with the aim to further boost cost-effectiveness.[20] Each IE proposal is evaluated for 'budget realism' and assessed on that criterion, which helps ensure that proposed evaluations are costed appropriately from the outset.[21] To that end, budget proposals are required to use a unified application template, which streamlines the budgetary review process and facilitates a comparison across proposed projects. The importance of financial management is also evident in the survey of successful applicants, who indicated finances as the most frequent topic covered during their consultations with SIEF. Finances were also the area were SIEF grantees were most likely to be satisfied with support provided by the SIEF core team. This is particularly important since financial challenges were the most frequently reported type of arising issues.

It is worth noting that considerations surrounding management of available resources are not limited to Pillar 1 only. As exemplified by the recently approved modifications to Pillar 2 and 3 activities,[22] SIEF management is continually committed to maximizing the impact of training and policy activities within the budget available. It is too early to evaluate the outcome and impact of these modifications, but, based on projected reach figures, it is likely they will result in a reduction in training unit costs in terms of outputs.

3.4. Summary and answer

The collected evidence suggests that SIEF's unit costs, while somewhat higher, do not differ substantially from those observed in other, broadly comparable settings. This applies both to the size of individual grants provided by SIEF and to the overall costs of the impact evaluation in question. In addition, costs stemming from the embedding of SIEF within the World Bank (i.e. administration fee and management

---

[19] The fee includes a base payment (covering rent, IT, travel office and other services) and a percentage management charge on against the volume of 3ie grants.

[20] Considerations of cost-effectiveness during the management of call for proposals is also discussed in Chapter 5.

[21] Department for International Development (2013) Strategic Impact Evaluation Fund (SIEF) 2011/12–2016/17: Annual Review 2013. As of 25 June 2015: http://iati.dfid.gov.uk/iati_documents/4022758.docx.

[22] SIEF (2015) Cleared Concept Note: Pillar 2: IE Capacity Building (P152175) and Pillar 3: IE Dissemination and Policy Engagement (P152178).

costs) compare favourably to costs faced by donors in other IE funding programmes. It is not possible, however, to quantify the benefits resulting from this arrangement. In this context, and from the evidence available, it appears that SIEF funds are used appropriately and can be considered good value for money.

Of course, it is necessary to view these comparisons as broadly illustrative at best, since differences in contexts across individual funders puts considerable limits on direct comparability. Nevertheless, it is also important to note that the institutional setup of SIEF and its processes, along with SIEF management's continuing commitment to increase the effectiveness of its spending across all three pillars, lends further support to the conclusion presented above.

4. Research question 3: Management of calls for proposal

RQ3: Pillar 1 – Have the calls for proposals been managed effectively, and are the quality assurance mechanisms SIEF has in place for impact evaluation products adequate?
(OECD DAC criteria: efficiency)

There are two elements to the research question targeted at the first pillar: the first relates to the administration of calls for proposals, the second to quality assurance of funded impact evaluations. The effective management of calls for proposals touches on the administrative management of SIEF and has overlaps with the VfM study. It is an aspect of performance management that is also encountered relatively frequently among research funders. The second element touches on the adequacy of the mechanisms that are currently in place to ensure the quality of impact evaluations.

Effectiveness in the context of Pillar 1 refers to the management of the calls for proposals. Because effectiveness refers to the degree to which objectives are met, effectiveness means here that we review if and how SIEF has been able to attract and fund high-quality proposals. To further operationalise effectiveness in the context of Pillar 1, we identify four criteria:

- Timeliness: Are applications processed in a timely manner?
- Objectivity: Is the selection process perceived to be free of bias, and are checks in place to ensure selection is not biased?
- Transparency: Is the application and selection process perceived to be transparent?
- Inclusivity: Is SIEF able to attract and include applications from developing regions?

The review of adequacy includes a consideration of whether SIEF's quality assurance mechanisms are suitable and conducive to the achievement of SIEF's aims, which in this context are the identification and subsequent support of rigorous impact evaluation proposals and grants. As such we can divide the quality assurance mechanisms into two parts: a) the application and selection process of impact evaluation proposals and b) the management of grants.

We assessed the adequacy of the application and selection process by examining the following elements:

- internal organisation of the processes, where possible in comparison to other funders
- perception of stakeholders on the rigour of the processes
- the quality of the researchers funded (i.e. the preliminary 'output' of the process)

The adequacy of the grant management processes will be assessed through an examination of:

- internal organisation of the grant management processes, where possible in comparison to other funders
- successful and unsuccessful applicants' perceptions

In addition, we have collected information about the added value of participation in the application process to the applicants, both successful and unsuccessful. Should researchers indicate that the application to SIEF in itself was valuable, then that may provide indications that the application process was of high quality.

4.1. Effectiveness of management of calls

Assessing the effectiveness of the management of the calls, as explained under research question 2, can be approached from two angles: the financial angle and the process angle. The financial angle is discussed as part of the value-for-money assessment of research question 2; therefore the focus here is on the effectiveness of the process of application and selection.

An inseparable dimension of SIEF's call for proposals is their ability to collectively generate a robust body of evidence across the four thematic clusters and, more specifically, in selected focus areas within each cluster. As interviews with SIEF staff highlighted,[23] too vague a definition of a call for proposal can hinder subsequent effective formulation of evidence-based policies, and SIEF has progressively firmed up the parameters of each of its application rounds. While this observation does not directly impact on the management of each call for proposal, this effort has its implications further downstream and contributes to a more effective eventual use and leveraging of SIEF-funded IEs. We revisit this point in our discussion of Pillar 3 in Chapter 5.

With respect to the actual call management, its effectiveness is assessed in this MTR through the following criteria:

- Timeliness: Are applications processed in a timely manner?
- Objectivity: Is the selection process perceived to be free of bias, and are checks in place to ensure selection is not biased?
- Transparency: Is the application and selection process perceived to be transparent?
- Inclusivity: Is SIEF able to attract and include applications from developing regions?

4.1.1. Timeliness and usefulness

Of the participants in the focus group, five felt that the current application and selection is efficient by being quick and thorough,[24] while four responded that they did not know. In addition, two respondents said that they had noticed improved efficiency between Call 1 and Call 2 for proposals 'with SIEF providing clear guidance to teams developing study proposals'.[25] One respondent, however, felt that the

[23] Interview 8

[24] Focus group question: Do you perceive the current application and selection to be efficient? (i.e. is it quick and thorough without requiring unnecessary effort?)

[25] Focus group participant 3

volume of proposals received meant that the process of assessment of proposals was very time-consuming and could be done more efficiently.[26]

With regard to the survey respondents, 82% noted that they were informed about the outcome in a timely manner, and 87% noted that their participation in the process was useful. However, it should be noted that two survey respondents suggested that the dual process of concept note review and application review was unnecessarily burdensome in terms of applicants' preparation costs. In addition, two respondents felt that the requirement to have a World Bank TTL was too restrictive and could lead to delays in the application process in the event of unforeseen changes, though one respondent highlighted the TTL role as a good way to make sure that the goals of researchers and funders are better aligned.

Table 10. Applicants' experience with the application process (I)

Statement: The outcome of the application was communicated in a timely fashion.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
Successful	14 (38%)	17 (46%)	4 (11%)	2 (5%)	31 (84%)	6 (16%)
Unsuccessful	25 (32%)	38 (49%)	11 (14%)	3 (4%)	63 (82%)	14 (18%)
Total	39 (34%)	55 (48%)	15 (13%)	5 (4%)	94 (82%)	20 (18%)
Statement: The application was useful in our preparation for the project.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
Successful	16 (43%)	16 (43%)	4 (11%)	1 (3%)	32 (86%)	5 (14%)
Unsuccessful	24 (31%)	44 (56%)	10 (13%)	0 (0%)	68 (87%)	10 (13%)
Total	40 (35%)	60 (52%)	14 (12%)	1 (1%)	100 (87%)	15 (13%)

4.1.2. Objectivity

The interviews show that over the course of the evolution of SIEF changes have been made to the process of selection to reduce any possible conflict of interest.[27] This may have been in response to a risk of 'capture' by certain subgroups, i.e. the risk that specific groups are somehow better placed to win grants due to biases in the selection and application process. This risk was mentioned both by people in the World Bank[28] and by people in DFID.[29] The latter, however, noted that the comment was based on experiences two years ago and therefore may not apply to the process today. It seems from the interviews that measures have indeed been taken to address potential bias in selection by explicitly excluding from the selection panel parties who submitted the proposals. The new donor who only joined SIEF recently

[26] Focus group participant 6

[27] Interview 8, Interview 9

[28] Interview 9

[29] Focus group participant 10

observed that they experienced the selection process to be efficient and to provide a good balance between parties and evaluation criteria (e.g. technical rigour, policy relevance).[30]

The majority of the participants in the focus group (n=7) felt that the current application and selection is able to distinguish stronger applications from those which are weaker. However, one participant felt that the current application and selection is not capable of supporting such distinction.[31] In addition, more than half of the participants (n=5) felt that the current application and selection procedure is objective and does not prefer some candidates over others irrespective of the quality of the proposal. However, three respondents disagreed with this.[32] Not all participants were able to elaborate on their answer, but it was noted by two participants that SIEF is not always an objective funder, because there was 'was a "capture" by some partners so that their favoured areas were funded'[33] and World Bank researchers may have 'preferential access to funds over external institutions.'[34]

The survey shows that 80% of respondents thought that that the process itself was fair. Perhaps unsurprisingly, 97% of successful applicants deemed the process fair, but 71% of unsuccessful applicants also deemed the process fair. Combined, only five respondents strongly disagreed with the statement that the process was fair.

Table 11. Applicants' experience with the application process (II)

Statement: The application assessment and decision were fair.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
Successful	14 (40%)	20 (57%)	0 (0%)	1 (3%)	34 (97%)	1 (3%)
Unsuccessful	12 (16%)	40 (55%)	17 (23%)	4 (5%)	52 (71%)	21 (29%)
Total	26 (24%)	60 (56%)	17 (16%)	5 (5%)	86 (80%)	22 (20%)

4.1.3. Transparency

There appear to be reports from several data sources that the application and selection process of SIEF is not perceived to be fully transparent. First, an interviewee from an NGO[35] noted that the feedback received from SIEF after an unsuccessful application was very minimal compared with feedback received from 3ie. The latter apparently were more elaborate in providing feedback, including working with applicants to come up with a way forward for an unsuccessful proposal. One focus group participant also

30 Interview 1

31 Another participant (11%) said that they did not know.

32 Another participant (11%) said that they did not know. On a slightly related note, one interviewee highlighted that the credentials of the proposed research team is one of the criteria for proposal selection (Interview 9). As a result, well-established researchers may be afforded greater benefit of the doubt in the event of uncertainties about a proposed project in earlier stages of the process. This may be another factor contributing to the regional disparity of funded calls discussed in the section on inclusivity below.

33 Focus group participant 10

34 Focus group participant 2

35 Interview 5

noted that transparency could be improved regarding the process of the identification of evaluations at the country level by the TTL.[36]

The perceived lack of transparency coupled with minimal feedback to substantiate applications outcomes also appeared in the survey. While two thirds (66%) of respondents agreed that the decision and reasons underlying the outcome were transparent, this sentiment varied substantially depending on the outcome of the application – 89% of successful applicants agreed the process was transparent, while only 55% of unsuccessful ones did so. A review of the open-text comment boxes at the end of the survey further shows that eight of them related their dissatisfaction to the perceived lack of feedback. In addition, three respondents explicitly mentioned it would be beneficial to publish selection criteria along with the list of selected projects. Also, a small number of comments suggested increased transparency may help dispel any perceived bias in the selection process (by region, policy area, etc.). The importance of transparency was further underscored by one survey respondent, who felt that due to a perceived lack of transparency in the results announcement, the partnerships they established at the proposal stage were strained as partners went on to rethink any collaboration on potential future projects.

Table 12. Applicants' experience with the application process (III)

Statement: The decision and reasons underlying the outcome were transparent.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
Successful	14 (38%)	19 (51%)	3 (8%)	1 (3%)	33 (89%)	4 (11%)
Unsuccessful	13 (17%)	30 (38%)	26 (33%)	9 (12%)	43 (55%)	35 (45%)
Total	27 (23%)	49 (43%)	29 (25%)	10 (9%)	76 (66%)	39 (34%)

4.1.4. Inclusivity

The fourth criterion of effectiveness identified relates to the inclusivity of the applications. Already in the inception phase, it was noted by SIEF that it struggles to include researchers and research institutions from the Global South in the funding of impact evaluations.[37] This observation was also made in the interviews and in the focus group. Interviewees from NGOs[38] and donors[39] observed that the inclusion of researchers from the Global South is often hard to achieve, but that without their inclusion, it is difficult to build capacity in these countries. Three focus group participants noted that SIEF is successful at reaching researchers globally, beyond the existing recognised cluster of researchers, yet five participants answered they did not know and one participant stated that SEIF is not successful at reaching researchers across the globe.

Administrative data from the calls for proposals of SIEF also illustrate that, while applications do have PIs from all over the world, successful PIs tend to be from either North America or Europe. Data from Call 1

[36] Focus group participant 7

[37] Meeting 2

[38] Interview 4, Interview 5, Interview 6

[39] Interview 1

is less complete than data from Call 2 and Call 3, as in Call 1 often the only researcher mentioned is the World Bank TTL. These proposals are listed under the column 'WB TTL'. Hence the data appear to be less accurate for Call 1.

For all calls, the majority of successful PIs came from Europe or North America, with only two successful PIs from Asia in Call 1 and Call 2. It is very likely that the successful PIs from North America and Europe have in-country partners with which they work, but that information is not readily available from the administrative data. The administrative data confirm the general concern that is difficult to actually get PIs from the Global South.

Table 13. Geographical spread of SIEF PIs

	Africa	Asia	Europe	Middle and South America	North America	Oceania	WB TTL
Call 1							
Unsuccessful							157
Seed funded only				1	6		
Fully funded		1	4		18		
Call 2							
Unsuccessful	16	25	13	11	50	3	17
Seed funded only							2
Fully funded		1	2		24		
Call 3							
Unsuccessful	22	19	9	4	22	1	
Seed funded so far		1	2	1	13		

Note: WB TTL refers to proposals which in the data only have a task team leader of the World Bank listed on the application and no other PIs.

While the collaboration with in-country partners cannot be inferred from the administrative data, it is possible to show the regions in which successful impact evaluations are being conducted. Using a slightly different coding from the PI regions, Table 14 shows that the majority of impact evaluations are based in African countries, and that in fact the success rate of applications for impact evaluations in Africa has increased. There has been a decrease in the number of regions in which successful impact evaluations are taking place. While this need not be a direct problem, it could in the future hamper the generalizability of the evidence if certain regions are underrepresented. This is a point to which we will return in the recommendations.

Table 14. Geographical spread of funded IEs

	Africa	EAP	ECA	LAC	MENA	SAR	Multiple
Call 1							
Unsuccessful	67	15	2	40	4	27	2
Seed funded	3	0	1	2	0	1	0
Fully funded	10	2	1	2	1	7	0
Call 2	Africa	EAP	ECA	LAC	MENA	SAR	
Unsuccessful	47	9	2	33	4	40	
Seed funded	1	1				1	
Fully funded	14			4		8	
Call 3	Africa	EAP	ECA	LAC	MENA	SAR	
Unsuccessful	32	8	6	9	1	21	
Seed funded	10	3		2		2	
Fully funded							

Note: EAP refers to East Asia and Pacific; ECA refers to Europe and Central Asia; LAC refers to Latin America and Caribbean; MENA refers to Middle East and North Africa; SAR refers to South Asia Region.

Survey results: examining inclusivity through collaborations

Table 15 summarises how frequently SIEF proposals collaborated with external parties. Through such collaborations, inclusiveness can potentially be increased. Academic institutions were the most frequent collaborators on bids submitted for SIEF consideration (65%), followed by NGOs (47%) and non-affiliated experts (16%). Collaboration with academic organisations was more frequent among successful applicants (85%) compared with unsuccessful applicants (47%), but the opposite was observed for non-governmental organisations. Other collaborators mentioned by individual respondents include local governments and specialized technical firms (e.g. survey experts).

Table 15. Collaborations on SIEF applications

	Academic institutions	NGOs	Non-affiliated experts	Other
Successful	33 (85%)	15 (38%)	6 (15%)	5 (13%)
Unsuccessful	50 (56%)	45 (51%)	15 (17%)	7 (8%)
Total	83 (65%)	60 (47%)	21 (16%)	12 (9%)

4.1.5. Survey results: indicators of effectiveness broken down by cluster

When these data are broken down by clusters (see Table 38 in Appendix I), we see that health respondents reported the least positive experience with the application process in the majority of areas. Water supply, sanitation, and hygiene (WASH) respondents were least likely to view their participation in the application process as useful, and ECD respondents were least likely to agree that the application process was fair.

4.2. The adequacy of quality assurance mechanisms

The review of the adequacy of SIEF's quality assurance mechanisms includes a consideration of whether SIEF's processes and practices are suitable and conducive to the achievement of SIEF's aims. Quality assurance occurs at both the application and the selection stages, as well as during the management of actual grants. In the following section we will first address the adequacy of the application and selection process. Then we will address the adequacy of grant management. Finally, we will outline what both successful and unsuccessful applicants highlighted as being the added value of applying to SIEF.

4.2.1. Adequacy of the application and selection processes

The adequacy of the quality assurance mechanisms was considered by examining the following elements:

- internal organisation of the processes, where possible in comparison with other funders
- perception of stakeholders on the rigour of the processes
- the quality of the researchers funded (i.e. the preliminary 'output' of the process)

Internal organisation of the application and selection process

The first assessment of the adequacy of the application and selection process can be made by reviewing how the process has been organised, and by verifying how this organisation compares with that of other funders of impact evaluations. The process of selection is outlined in full in Appendix K. It consists of several steps in which multiple stakeholders, representing both donor and World Bank staff, assess the proposals on technical merit and relevance to the call. This process is not dissimilar from the selection process at peers, such as 3ie,[40] in which multiple rounds of proposals are scored on technical rigour and (policy) relevance.[41]

Once proposals make it through the first round, they receive seed funding to prepare a full proposal. The full proposals are then reviewed by an internal World Bank specialist and an external specialist. Upon clearance by SIEF and World Bank Regional/Global Practice, projects will move to full funding.

The predetermined technical criteria, and the fact that through multiple rounds multiple stakeholders are involved, give confidence that the internal organisation of the application and selection process is adequate to select proposals of a high quality.

Perception of stakeholders on the rigour of the processes

The second way to assess the adequacy of the application and selection process is through an examination of the perceptions of the stakeholders involved. On the basis of the interviews and focus group, it appears that the application and selection processes in place at SIEF are recognised as adequate to ensure high-quality impact evaluation products. Five interviewees from different backgrounds emphasised that the selection procedure of SIEF is able to identify high-quality impact evaluations.[42] Asked what the main

---

[40] See: http://www.3ieimpact.org/funding/open-window/ow-faq/#27

[41] Interview 7

[42] Interview 1, Interview 2, Interview 3, Interview 6, Interview 9

objectives of SIEF are, seven focus group participants highlighted SIEF's key objectives as funding and supporting rigorous impact evaluations. On a related note, four survey respondents explicitly highlighted the high degree of technical capacity of the SIEF team, noting that this differs from other funders and contributes to an overall emphasis on the robustness of project methodological designs. In addition, the survey shows that, overall, at least 80% of respondents agreed or strongly agreed that the requirements of the application process are clear (89%) and reasonable (83%). As such, the application and selection process was viewed largely positively by the survey participants.

Table 16. Applicants' experience with the application process (IV)

Statement: The steps and requirement of the application process were clear from the start.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
Successful	17 (46%)	15 (41%)	4 (11%)	1 (3%)	32 (86%)	5 (14%)
Unsuccessful	26 (33%)	45 (57%)	7 (9%)	1 (1%)	71 (90%)	8 (10%)
Total	43 (37%)	60 (52%)	11 (9%)	2 (2%)	103 (89%)	13 (11%)

Statement: The cost and burden associated with meeting application requirements were reasonable.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
Successful	9 (24%)	18 (49%)	7 (19%)	3 (8%)	27 (73%)	10 (27%)
Unsuccessful	15 (19%)	54 (69%)	9 (12%)	0 (0%)	69 (88%)	9 (12%)
Total	24 (21%)	72 (63%)	16 (14%)	3 (3%)	96 (83%)	19 (17%)

A concern was raised by several interviewees that the strong emphasis on the technical rigour of proposals by SIEF may have negative unintended consequences for capacity building. It was mentioned in two interviews that there is a risk that the current selection criteria will always exclude 'weaker' teams from regions in the world where capacity building is an urgent need. Applications from traditional Western universities are likely to still provide the strongest applications, yet this may mean that SIEF will not fund impact evaluations at institutions at which capacity building is a real need. In addition, it was mentioned in two interviews that the emphasis on high-quality designs may result in studies on smaller projects and pilots rather than on existing government programmes, as the latter are more difficult to incorporate into a study design. Still, the potential scalability of existing government programmes is often deemed to be greater by policy makers, which would make data collection on them more important. We will return to the potential tension between capacity building and high-quality impact evaluations in the final chapter.

The quality of the researchers funded (i.e. the preliminary 'output' of the process)

The third way in which we assess the adequacy of the application and selection process is by examining the researchers who have actually been funded. If the process is adequate, it will deliver high-quality impact evaluations. Because none of the impact evaluations are finished yet and therefore cannot be reviewed, we are limited here to taking the funded researchers as preliminary proxy.

As was mentioned above under inclusivity, the PIs tend to come from Western countries. Within those countries, the researchers funded tend to be based at well-known universities and research institutes, such as Harvard University; University of California, Berkeley; and the London School of Hygiene and

Tropical Medicine. On the basis of the reputations of the institutions to which the researchers are attached, it further appears that high-quality researchers have been funded. This in turn seems to confirm that the application and selection processes have been adequate.

4.2.2. Adequacy of grant management

The adequacy of the grant management processes can be assessed in a similar way to that of the application and selection process, through an examination of:

- internal organisation of the grant management processes, where possible in comparison to other funders
- perception of successful and unsuccessful applicants

Internal organisation of the grant management processes

When projects have been cleared for full funding, payment is phased in tranches that are dependent on the achievement of milestones. The milestones have been predetermined, are measureable and are standardised across funded projects (though there is room for flexibility as necessary).[43] The perception of the SIEF core team has been that the payment in tranches has been important to get researchers to meet milestones and for SIEF to be able to monitor milestones. The payment tranches are set as follows for Call 1 and Call 2:

- 40% after final SIEF and WB Regional/Global Practice clearance of the full IE technical proposal
- 5% after uploading baseline data and documentation to the MMT
- 45% upon submission and SIEF approval of the baseline report and report validating the evaluation design
- 10% after uploading the endline data and endline documentation to the MMT

Perception of successful and unsuccessful applicants

The second part of the assessment of the adequacy of the grant management processes is to look at how responsive it is to the funded researchers – because ultimately, SIEF's objectives can be met only in so far as the funded IEs deliver on their own objectives. Hence it was important for us to collect data on applicants' and grantees' satisfaction with their interaction with the SIEF core team and with the level of support they received.

Two thirds of successful respondents indicated that they interacted with SIEF at least every six months over the course of the duration of the funded project, and 6% of respondents consulted with SIEF more than once a month.

[43] Interview 8

Table 17. Grantees' frequency of interaction with SIEF

Frequency	Count/ proportion
More than once a month	2 (6%)
Every 1–3 months	12 (35%)
Every 3–6 months	9 (26%)
Every 6–12 months	6 (18%)
Less than once a year	5 (15%)

Topics most frequently covered by these consultations were, perhaps unsurprisingly, finances (71%) and project milestones (60%). Other subjects include communications and dissemination (37%) and any arisen problems (34%). Access to in-country partners and policymakers were considerably less frequently discussed (9% each). This is in line with observations regarding the perceived benefits of submitting an application to SIEF, in which in-country access also featured less prominently than other topics. Other issues discussed with SIEF include discussion of wider cluster activities and concept note review.

A corresponding picture emerged when respondents were asked how satisfied they were with SIEF core team support during regular consultations. As summarized in Table 18 below, the proportion of satisfied grantees notably outweighed that of dissatisfied grantees with respect to support in the areas of finances, milestones, problems and 'other'. A much more mixed picture was observed in the areas of access to in-country partners and policymakers.

Table 18. Grantees' satisfaction with SIEF core team support, by activity

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
Finances of the project	12 (39%)	14 (45%)	0 (0%)	1 (3%)	26 (84%)	1 (3%)	4 (13%)
Milestones of the project	7 (23%)	13 (42%)	3 (10%)	1 (3%)	20 (65%)	4 (13%)	7 (23%)
Technical/research problems experienced	8 (29%)	6 (21%)	4 (14%)	0 (0%)	14 (50%)	4 (14%)	10 (36%)
Access to in-country contacts and partners	3 (12%)	2 (8%)	7 (28%)	0 (0%)	5 (20%)	7 (28%)	13 (52%)
Access to (in-country) policymakers	3 (12%)	3 (12%)	6 (24%)	0 (0%)	6 (24%)	6 (24%)	13 (52%)
Other [as specified in text above]	2 (12%)	6 (35%)	1 (6%)	0 (0%)	8 (47%)	1 (6%)	8 (47%)

Successful applicants were invited to comment on how SIEF compares with other, non–World Bank funders in a series of categories. On the whole, large majorities agreed that SIEF's reporting requirements are reasonable, that SIEF is open to consultation about any arising issues and that its capacity-building activities are a valuable resource to the conduct of the funded projects. Opinions were more mixed with respect to whether SIEF provides greater access to in-country partners and policymakers, and whether it makes use of country offices in support of funded IEs. In addition, seven survey respondents volunteered that their communications with the SIEF team were easier/more open than with those with other funders, further testifying to the level of support provided by the core team. By contrast, however, one respondent found the level of communication very poor.

Table 19: Answers to the question 'Compared with other non–World Bank funders where applicable, how strongly do you agree with the following statements about the project process?'

	SA	A	D	SD	AC	DC	NA
The cost and burden associated with reporting requirements are reasonable when compared with the potential benefits.	9 (26%)	17 (49%)	4 (11%)	2 (6%)	26 (74%)	6 (17%)	3 (9%)
SIEF is open to consultation about any problems encountered in the project.	13 (37%)	13 (37%)	0 (0%)	0 (0%)	26 (74%)	0 (0%)	9 (26%)
SIEF provides greater access to in-country contacts or partners.	6 (18%)	8 (24%)	8 (24%)	1 (3%)	14 (41%)	9 (26%)	11 (32%)
SIEF provides greater access to (in-country) policymakers.	6 (18%)	5 (15%)	9 (27%)	1 (3%)	11 (33%)	10 (30%)	12 (36%)
SIEF makes use of country offices and contacts to support the impact evaluation.	9 (26%)	3 (9%)	9 (26%)	1 (3%)	12 (35%)	10 (29%)	12 (35%)
SIEF offers capacity-building activities (e.g. workshops, measurement support seminars) that are a valuable resource to running the impact evaluation.	12 (34%)	13 (37%)	3 (9%)	1 (3%)	25 (71%)	4 (11%)	6 (17%)

Note: SA = Strongly Agree; A = Agree; D = Disagree; SD = Strongly Disagree; AC = Agree Combined; DC = Disagree Combined; NA = Not Applicable

Challenges encountered thus far (survey results)

Approximately one third of respondents (32%) shared that they had not encountered any challenges so far in conducting their impact evaluations. The most frequent types of challenges mentioned were financial (51%), data collection (24%) and engagement of local partners (24%). Research design (14%) and recruitment of participants (3%) were notably less frequently mentioned. Other types of challenges also indicated by individual respondents include both external factors (timely implementation of government policy under evaluation, bureaucratic delays, outbreak of Ebola) and internal factors (delays in access to funding).

Table 20. Challenges encountered so far by grantees

Challenge	Count/ proportion
Financial	19 (51%)
Design	5 (14%)
Data	9 (24%)
Recruitment	1 (3%)
Engagement	9 (24%)
Dissemination	0 (0%)
None so far	12 (32%)

Some respondents indicated the various forms of support the SIEF core team was able to provide in order to overcome these challenges. Most often, these include being open to consultation (54%), providing flexibility around project implementation (35%) and providing technical advice (34%). Help with accessing in-country partners, policymakers and World Bank country offices was mentioned considerably less frequently. It should be noted that a lower response for some of the options need not imply that SIEF is 'failing' in this respect. It may simply mean that these are not areas in which grantees were in need of support. To explore these matters in more depth was, however, beyond the scope of the survey.

Table 21. How does SIEF help overcome arisen challenges?

Area	Count
Open consultation	20 (54%)
Flexibility	13 (35%)
Technical advice	11 (30%)
In-country contacts	1 (3%)
Access to policymakers	2 (5%)
Access to WB country offices	4 (11%)

When asked in which areas SIEF core support could be improved, respondents suggested the following areas: peer-review comments with a view to publishing in a journal (30%), provision of access to in-country partners (30%), access to policymakers (22%) and technical support (16%). However, the generally high proportions of respondents who did not feel that support can be improved in these areas suggest an already high level of satisfaction with the existing degree of support.

Other potential areas of support mentioned by individual respondents are greater flexibility of top-up funding and help with administering contracts and subcontracts with research institutions.

Table 22. Areas of SIEF core team support that could be improved

Area	Count
Peer review	11 (30%)
Tech support	6 (16%)
Access partners	11 (30%)
Access policymakers	8 (22%)
Other	6 (15%)

4.2.3. Added value of the application process

Finally, we have collected information about the added value of participation in the application process to the applicants, both successful and unsuccessful. Should researchers indicate that the application to SIEF in itself was valuable, then that may provide indications that the application process was of high quality. Survey respondents were asked whether any added value was generated through their participation in the application process. None was reported in 21% of cases (by respondents explicitly selecting the 'none' option). Where some added value was reported, this took the form of improved knowledge about proposal preparation (41%), new contacts and connections (34%) and improved knowledge about the design of impact evaluations (33%). Other forms of added value reported by individual respondents include an opportunity to obtain feedback on proposed materials and ideas and repurposing those for other contexts.

There is not a large difference between successful and unsuccessful applicants with respect to this question.

Table 23. Added value generated through the application process

	None	Improved knowledge/change in approach about preparing funding applications	Improved knowledge/change in approach about designing research projects	New contacts/collabora-tions made through the application	Other
Successful	10 (27%)	12 (32%)	12 (32%)	14 (38%)	3 (8%)
Unsuccessful	16 (18%)	39 (44%)	29 (33%)	29 (33%)	7 (8%)
Total	26 (21%)	51 (41%)	41 (33%)	43 (34%)	10 (8%)

Broken down by application clusters (see Table 51 in Appendix I), early childhood development (ECD) and education applicants were more likely to agree that some added value has been generated than were their health and WASH counterparts. In particular, new contacts and connections were much less frequently mentioned by respondents from the latter two clusters.

4.3. Summary and answer

In terms of timeliness and objectivity, the calls appear to have been managed effectively, yet the MTR raises some questions over the transparency and inclusivity of the calls. The perceived lack of transparency identified through the survey and in an interview is likely to be the result of a limit to the resources SIEF has available to communicate back on unsuccessful applications. With regards to inclusivity, it is clear for the moment that the inclusion of researchers from the Global South is not extensive.

The quality assurance mechanisms in place appear to be adequate. Data from the interviews and focus group confirm that SIEF is perceived as being good at selecting high-quality impact evaluations. In fact, the point was raised that a strict focus on quality may risk missing opportunities for capacity building. The possible tension to these two aims is further addressed in the final chapter.

4.3.1. Secondary research questions

Has the collaboration between those receiving grant funding and SIEF so far been effective?

The word effective here needs to be redefined slightly from the way it has been operationalised in this chapter so far. Ideally, the effectiveness of the collaboration is evaluated by the outcome of the collaboration, namely, the impact evaluations. Because these are not yet available, our measure of effective collaboration between grantees and SIEF for the MTR relates to the frequency and usefulness of the interactions between SIEF and the grantees. If the interactions are sufficiently frequent and informative, then the collaboration can be seen to be effective. Overall, the survey results of successful applicants reported in this chapter show that interactions with SIEF have largely been perceived as useful. SIEF is perceived as being open to consultation and as offering support with capacity building. Areas of improvement were also mentioned, such as access to partners and the introduction of peer review comments that can help with publication. Overall, the collaboration so far seems to have been effective, yet a full judgement can really only be given once the impact evaluations are finished, as the effectiveness of the collaboration is best judged by the product of the collaboration.

5. Research question 4: impact evaluation trainings and clinics

RQ4: Pillar 2 – Are the impact evaluation trainings and clinics delivered so far relevant to beneficiaries and of high quality? (OECD DAC criteria: efficiency, effectiveness)

The data used to assess research question 4 on Pillar 2 largely stems from surveys conducted among participants. These surveys are of two types:

- A pre- and post-test of participants to assess their knowledge of impact evaluations and to assess the learning gain achieved by the workshops
- An overall workshop assessment to understand participants' opinions of the workshops

The data cover a number of regional (8) and country-specific workshops (6). Out of these 14 workshops, survey data are available for 9 countries. The kind of data collected does not differ between the regional and the country-specific workshops.

The research question can be divided into two parts. First, the quality of the workshops will be assessed using data from both surveys. Second, the relevance of the workshops will be assessed using data from the overall workshop assessment. The overall workshop assessment therefore feeds into both parts, as it covers a range of different questions. Table 24 shows the questions included in the overall workshop assessment that will be used in this chapter and notes the parts for which they are relevant.

Table 24. Overview of workshop assessment questions

Question	Part
How would you rate the overall quality of the workshop? [Quality]	Quality
How would you rate the overall usefulness of the workshop? [Usefulness]	Relevance
Relevance to the Bank's mission/to the fight against poverty [Relevance to Bank Mission]	Relevance
My knowledge and/or skills increased as a result of this workshop. [Knowledge/skills increased]	Relevance
The knowledge and/or skills gained through this workshop are directly applicable to my work. [Knowledge/skills apply to work]	Relevance
To what extent did the workshop achieve its announced objectives? [Objective achieved]	Quality
To what extent did the workshop fulfil your learning needs? [Learning needs fulfilled]	Relevance
How would you rate the workshop content? [Content/subject matter]	Quality
How would you rate the order in which the content was presented? [Order content was presented]	Quality
How would you rate the materials used during the workshop? [Materials]	Quality

Questions are answered on a scale from 1–5 (1=very poor; 2=poor; 3=neither; 4=good; 5=very good). The data in the tables that follow show the average scores of all respondents.

5.1. Quality of training

The questions relating to quality of workshop training (Table 25) consistently show high scores, generally above 4, with the exception of a few workshops, such as Bangladesh (2014). The quality, content and materials are highly rated, and together the scores suggest that the quality of the workshops was good. However, on a more cautious note, the responses to the question 'Objective achieved' are consistently lower than the responses to all other questions, raising questions about relative attitudes to this most important question. The data can be disaggregated by gender, and this shows that there are not noticeable gender differences in scores (see Appendix H).

Table 25. Quality of workshop training

	Ethiopia (2012)	Korea (2012)	India (2013)	Senegal (2013)	Bangladesh (2014)	Ethiopia (2014)	Korea (2014)	Peru (2015)	Total
Quality	4.4	4.4	4.3	4.3	3.9	4.4	4.1	4.5	4.3
Objective achieved	4.3	4.2	3.9	4.3	3.7	4.1	4.0	4.0	4.1
Content/subject matter	4.5	4.5	4.2	4.4	4.2	4.6	4.4	4.5	4.4
Order content was presented	4.3	4.5	4.3	4.2	3.9	4.3	4.2	4.4	4.2
Materials	4.3	4.5	4.2	4.3	4.1	4.5	4.3	4.3	4.3
N	69	59	49	71	89	52	105	51	545

Note: Chile not included as data was not available

Participants' opinions are not the only data source that can used to assess quality of training, the results of a test that SIEF administers among participants to assess their knowledge gain as a result of the workshops can also be used. The 10 questions of the test that are used to estimated knowledge gain are attached in Appendix G. Table 26 shows the average percentage of correct answers on the pre- and post-test at the different workshops. It is difficult to assess the 'performance' for these figures because there are no benchmarks to which they can be compared. A 'good' score is therefore difficult to determine. It is encouraging that the two largest gains in knowledge were made recently, in 2014. However, the overall average gain in knowledge, of around 20 percentage points,[44] implies that out of a total of 10 questions, participants, on average, answered 2 additional questions correctly on the post-test than they did on the pre-test.[45]

[44] This is the non-weighted average of all the '% gain in knowledge'.

[45] The scores of Korea (2012) are based on 7 questions. After that workshop this list of questions was extended to 10.

Table 26. Workshop knowledge gains

	Korea (2012)	India (2013)	Chile (2013)	Senegal (2013)	Bangladesh (2014)	Ethiopia (2014)	Korea (2014)	Peru (2015)
% correct answers on pre- test	48%	58%	69%	43%	53%	48%	38%	45%
% correct answers on post-test	64%	73%	90%	59%	68%	72%	72%	60%
% gain in knowledge	16%	15%	21%	16%	15%	24%	34%	15%

5.2. Relevance to beneficiaries

The second part of the research question on Pillar 2 relates to the relevance of the workshop to beneficiaries. Relevance is measured through a number of questions included in the workshop assessment, complemented with demographic data of the participants. Five questions that relate to relevance are included in Table 27. These show that, overall, the relevance of the workshops is rated highly by the participants.

Table 27. Workshop relevance

	Ethiopia (2012)	Korea (2012)	India (2013)	Senegal (2013)	Bangladesh (2014)	Ethiopia (2014)	Korea (2014)	Peru (2015)	Total
Usefulness	4.6	4.4	4.1	4.5	4.0	4.6	4.4	4.4	4.3
Knowledge /skills increased	4.4	4.5	4.2	4.1	3.6	4.2	4.1	4.3	4.1
Knowledge /skills apply to work	4.3	4.3	4.1	4.2	3.9	4.3	4.0	4.1	4.1
Learning needs fulfilled	4.2	4.3	3.8	4.0	3.4	3.9	3.8	3.8	3.8
Relevance to Bank Mission	4.3	4.7	4.3	3.9	3.7	N/A	4.2	N/A	4.1

The only area that scores below 4 on average (non-weighted) is the fulfilment of learning needs; the responses to this question are consistently are lower than for other questions. However, the average score of 3.8 on this item is still considered a high score. To get a better understanding of what might be improved in terms of learning needs, we have analysed the general comments of participants in the Bangladesh (2014) and Korea (2014) workshops. Specifically, we looked at the general comments of the participants who scored the fulfilment of learning needs as 3 or below to see if there are any general lessons that can be learned. Most frequently mentioned in the open comment boxes are suggestions to include more practical examples and case studies. The comments suggest that a more explicit link to actual cases and practices could be made in the workshops.

Finally, to assess the relevance of the workshops, the background of the participants can be examined to understand what audiences the workshops are reaching. A first observation can be made on the number of women attending the workshops. With an average participation rate of 41%, it cannot be said that women have been hugely underrepresented, but in specific contexts (Senegal, India, Ethiopia), the aim could be to increase the proportion of female participants in the future. Furthermore, it cannot be determined from the data what the background of these women was – for example, if they were researchers or government officials. Finally, it needs to be noted that the dimension of gender is broader than just counting numbers; this is a topic to which we will return under the sixth research question.

A second observation relates to the professional background of the participants. As Table 28 shows, the majority of participants in the workshops have been classified as government officials. It was observed in one of the interviews that the inclusion of government officials is necessary to create a policy environment receptive of research evidence, and to stimulate demand for impact evaluations.[46] Within the capacity of the workshops, the 'demand' side of impact evaluations therefore seems to be covered. The 'supply' side, however, seems to have been less covered. Only four workshops had participants from an academic environment (graduate student/academic institution). Combined with representatives from a research institution, there were fewer than 200 participants who could be classified as representing the supply side.

[46] Interview 1

Table 28. Overview of workshop participants

	Ethiopia	Korea	India	Chile	Senegal	India	Bangladesh	Korea	Ethiopia	Total
Number of participants	127	110	95	119	139	62	134	122	57	965
Number of male participants	88	48	56	61	101	47	87	65	41	569
Number of female participants	39	62	39	58	37	15	47	57	16	395
Proportion of female participants	31%	56%	41%	49%	27%	24%	35%	47%	28%	41%
Number of moderators	20	10	11	32	27	22	31	12	7	172
By affiliation										
Government official	77	36	20	57	93	12	36	55	22	408
Representative of an NGO organization	9	7	13		18		53	7	5	112
Graduate student/member of an academic institution	5	45		19				36		105
World Bank staff	23	6	14	6	15	10	27	16	14	131
Representative of a research institution		12	34		3		15	8	4	76
Representative of an international donor organization	13	4	14	17	10		3		1	62
Other				20		40			11	71

5.3. Survey results: expectation of capacity building

The survey results show that among successful applicants of SIEF, 71% agreed with the statement that 'SIEF offers capacity building activities (e.g. workshops, measurement support seminars) that are a valuable resource to running the impact evaluation.' By contrast, when asked about the expected benefits of SIEF as a funder, only 11% of successful applicants and 34% of unsuccessful applicants listed capacity building. Although the data are limited on this topic, they seems to suggest that while the capacity-building activities organised by SIEF are valued, there is not a great expectation that SIEF will help with capacity building.

5.4. Additional comments on workshop monitoring and materials

5.4.1. Monitoring

Overall, it should be noted that SIEF maintains a robust workshop monitoring system. Data collected through questionnaires cover a range of important aspects pertaining to participants' satisfaction, experience and knowledge gains through a mix of qualitative and quantitative indicators. It may be desirable to collect additional information on participants themselves, in particular with respect to government officials and NGO representatives. One potential indicator of interest may be the reason for, and expectations of, participation on the part of particular individuals, with a possible classification and taxonomy of factors, such as institutional and political affiliation, or level of seniority. One reason for doing so would be to further enhance the existing process and output workshop indicators. Second, coupled with follow-up surveys over time (as discussed in section 9.2.2), this may help inform a typology of which type of individual is the most appropriate and suitable workshop attendee.

5.4.2. Materials

A brief review of the workshop materials listed on the SIEF website,[47] conducted by the research team, found the materials to be of high quality and generally suitable for their intended purpose. This is perhaps not surprising since they are based on the World Bank's handbook *Impact Evaluation in Practice*.[48]

Several elements are in particular noteworthy. Workshop materials appear to follow to a large extent a standardised agenda, which logically progresses from introductory topics to more specialised and technical lessons. A distinction is made between materials used in sessions with policymakers and those used in more technical sessions with researchers. Accordingly, policy-track materials are generally made more accessible and focus on high-level messages. Some materials strive to incorporate interactive elements, for example, by placing summary test questions throughout individual courses.[49] Each workshop also incorporates a series of case studies, which serve as practical examples. However, particularly in light of reviewed participant feedback, we note that there is room for increasing the number of practical examples and better integrating them with core technical presentations. Also of benefit is the fact that a large number of materials are available in French and Spanish, as is the underlying IE handbook, which should further improve their relevance and accessibility.

Ultimately, however, an inherent limitation of the workshop materials is the relatively short time devoted to each workshop and the resulting constraints on knowledge gain and retention. In light of this challenge, the planned development of the online course, along with an update of the IE handbook, can be seen as promising ways to provide continuous support to participants after the conclusion of individual workshops.

5.5. Summary and answer

On the basis of the available data, it appears that the workshops that have been organised have been of high quality and have been relevant to the participants. One element that may require further attention is learning. While still good, the scores for the fulfilment of learning needs have been somewhat lower, and open-text comments seem to suggest that practical examples could further improve the quality of the workshops.

---

47 http://www.worldbank.org/en/programs/sief-trust-fund#7

48 Gertler, PJ. Martinez, S, Premand, P et al. (2011) Impact Evaluation in Practice. World Bank. As of 25 June 2015: https://openknowledge.worldbank.org/handle/10986/2550 License: CC BY 3.0 IGO.

49 In this context, one interviewee mentioned the possibility of distributing wireless voting machines to participants for use when a multiple-choice question comes up.

Beyond the strict confines of research question 4, there is also a wider question as to what the best approach to capacity building might be. The SIEF management team have been working to address this question and to refine the design of Pillar 2 activities, culminating in the decision to undertake the following additional activities (in addition to the existing offer of training through workshops and clinics): a) update the existing handbook *Impact Evaluation in Practice* and b) develop an online course on impact evaluation methods (SIEF, 2015).[50]

This review has identified additional, wider questions. Interviews with 'peers' show that there are different approaches to capacity building and that other programmes tend to work more closely and directly with researchers in, for example, the Global South, to increase capacity.[51] This raises a question, however, as to what the remit is of SIEF. As highlighted under the first research question, this is a question not solely for the SIEF core team, but also for the funders. Related to that is the observation that it is not always clear what the relationship between the workshops of Pillar 2 and the impact evaluations funded under Pillar 1 is. In the final chapter we will return to these questions.

5.5.1. Secondary research question

What kinds of beneficiaries participate in capacity-development activities, and how are they being selected?

The main beneficiaries participating in the workshops for which there is data available are government officials (see Table 28 above). Of all participants, 41% have been women. Participants are primarily identified by SIEF's 'customers,' i.e. country teams or TTLs. If there are external partners involved in the preparation of a workshop, additional nominations for attendees can be made. This arrangement places the SIEF team in a reactive position, in which it helps to provide training according to the specifications and preferences of those originating the request for training, and is reflective of SIEF's Pillar 2 partial function as an 'on-demand' facility for interested parties. As such, SIEF is largely only in a position to comment on the proposed selection of participants (e.g. by calling for a greater participation of female researchers).[52]

[50] SIEF (2015) Cleared Concept Note: Pillar 2: IE Capacity Building (P152175) and Pillar 3: IE Dissemination and Policy Engagement (P152178). 25 March 2015.

[51] Interview 6, Interview 7

[52] Interview 10

6. Research question 5: evidence-to-policy notes

RQ5: Pillar 3: Are the initial evidence-to-policy notes of high quality, relevant and used?
(OECD DAC criteria: effectiveness, relevance)

While the fifth research question on Pillar 3 is focused on the evidence-to-policy (E2P) notes, the data collected under this research question allow us to address more broadly the topic of communication of evidence to policymakers. This is important, as a large majority of successful applicants expected that SIEF would assist them in communicating and disseminating the results of their projects to various groups of stakeholders. The expected audiences include practitioners and NGOs in the study country (74%), policymakers in the study country (74%), policymakers in donor countries (82%) and international authorities (76%) (see Table 29 for details). Assistance with dissemination was less expected by academic audiences. Given this expectation, and given the objective of SIEF to inform policymakers of research evidence, the third pillar is already evolving to be much broader than the E2P notes. Still, because the research question asks that we focus on the E2P notes, this chapter will pay dedicated attention both to the E2P notes and to communication to policymakers more broadly.

The research question and the corresponding evaluation criteria cover a number of concepts. We will use the evaluation criteria of relevance and effectiveness to structure this chapter. Relevance is the first part of the research question we will review. In the context of the E2P notes, relevance can be defined as the suitability of the E2P notes to the objectives and activities of SIEF. Under relevance we also include the quality of the E2P notes, as high quality of the E2P notes seems to be a pre-requisite for their relevance. Relevance is assessed though the perceptions of stakeholders.

Effectiveness in the context of the E2P notes relates to the extent to which the E2P notes attain their objectives. Thus, we need to review if the E2P notes are used and are informing policy making. Under *effectiveness*, we therefore include usage, and we will rely on both the perceptions of stakeholders and administrative data from SIEF to review it.

Table 29. Grantees' expectations of communication assistance by SIEF

Target audience	Count/ proportion
Academia	15 (48%)
NGOs in study countries	26 (74%)
Policymakers in study countries	26 (74%)
Policymakers in donor countries	27 (82%)
International authorities	26 (76%)

6.1. Relevance and quality

Rather than provide a simple summary of the data collected, we have tried to provide practical guidance based on the feedback we received from interviewees on the E2P notes. Spread over Appendix L, we have copied one E2P note as an example. In the comment boxes we summarise the feedback we have received through the interviews. Overall, the E2P notes were rated highly in the interviews,[53] and the comments on the selected E2P notes are simply points mentioned for further improvement.

6.2. Effectiveness and usage of E2P notes

Usage is very difficult to establish without conducting a very elaborate survey of recipients. Therefore, given the confines of this MTR, distribution itself is taken as a first proxy of usage. The E2P notes are posted online and distributed through the SIEF quarterly newsletter, which has a distribution of 3000+:

'Our newsletter goes to more than 3000 non World Bank people worldwide. The "open rate" averages 40 percent, which is more than twice average open rate for World Bank newsletters. The newsletter's list is more than 60 percent in developing countries, including government officials, local university researchers and local NGOs. The newsletter highlights new publications, interviews, policy notes and updates on the SIEF Trust Fund.'[54]

While this is quite an extensive email list, one interview with a large NGO shows that the interviewee was not aware of the E2P notes, yet did rate them highly.[55] This can suggest that apart from changes to the content listed above, SIEF can continue to focus on broadening the distribution. Expanding the email list is one of the ways to do that, but interviewees also mentioned dissemination through other online channels. Suggestions were the use of new media (e.g. a Tweet on a new E2P note) and 'advertising' E2P notes on well-known websites related to human development, such as Devex.

[53] Interview 1, Interview 3, Interview 4

[54] SIEF Annual Report 2013–2014

[55] Interviewee was presented with a copy of the E2P note by a member of the research team. Interview 4

6.3. Other Pillar 3 activities and wider thoughts on dissemination

It is important to note that evidence-to-policy notes represent only one of an array of activities undertaken within the framework of Pillar 3 (SIEF 2015).[56] Other multi-media products have included feature stories, short videos and a slide show (all available online). In addition, the SIEF team plans to expand the portfolio of available instruments to include on-demand support to teams interested in publishing in peer-reviewed journals and in developing policy materials around existing evidence in selected focus areas within the four policy clusters covered by SIEF. It is envisaged by SIEF that these materials will be of use during future policy workshops, organised in cooperation with the Pillar 2 programme.[57] Given the expectations of the successful applicants, there is likely to be greater demand for these activities once the impact evaluations near completion.

The data collected provides some insights on the new activities. Where relevant we have tried to link comments from the data collected to the new initiatives proposed by SIEF.

New media

The use of new media (e.g. Twitter, etc.) was mentioned in interviews as one of the ways in which the online reach of SIEF could be enhanced.[58] It was noted in an interview that while the E2P notes are of good quality, they in themselves can only go so far in communicating results. To reach policymakers, multiple communication lines will be required.

A clear strategy in the use of new media is advisable, however; for example, being consistent in the way that new E2P notes are promoted. The use of videos has been suggested by SIEF; opinions on their effectiveness have been mixed in the interviews.[59]

Summary of evidence

SIEF will start reviewing and mapping IE evidence in each of the four clusters to communicate summaries to policymakers. This approach resonates well with comments in the focus group, in which it was observed that policymakers are unlikely to read individual impact evaluations to draw lessons from them. We were told that what policymakers need are clear synthesis documents providing key lessons that are based on a multitude of evidence.[60]

6.4. Summary and answer

The E2P notes are generally deemed to be of high quality by interviewees. Several interviewees made suggestions, which have been summarised in this chapter. Most important, perhaps, was the suggestion to

---

[56] SIEF (2015) Cleared Concept Note: Pillar 2: IE Capacity Building (P152175) and Pillar 3: IE Dissemination and Policy Engagement (P152178). 25 March 2015.

[57] SIEF Pillar 2 and Pillar 3 Concept Note

[58] Interview 1, Interview 3

[59] Interview 3, Interview 4

[60] Focus group participant 10, Interview 3

include more information on the actual implementation of the interventions discussed. E2P notes were generally considered to be relevant, yet several interviewees did comment that notes on their own will not be sufficient to bring evidence to policymakers. As such, the effectiveness of the E2P notes to inform policymakers was questioned. This opens up the realm of alternative activities of communication and dissemination, which are already being explored by SIEF. The additional comments we were able to distil from the data on such alternatives will hopefully be useful in the design of such activities. Finally, the usage of the notes is difficult to establish, yet it appears that SIEF could explore ways to increase the distribution of the E2P notes to thereby increase the chances of them being used, and to ultimately increase their effectiveness.

6.4.1. Secondary research question

Do SIEF and the individual IEs have uptake strategies and plans that give confidence that desired outcomes will be achieved?

Already at the proposal stage, submissions are reviewed to see if dissemination and impact are taken into consideration appropriately from the outset of the project. At this stage, any potential risks (e.g. reputational, opposition to the project) are also considered, and proposal teams receive feedback with additional suggestions (if appropriate) from the selection committee. Once the project is funded and underway, the SIEF core team agree on a strategy for communication and post-completion dissemination for each impact evaluation, which is tracked over time through regular communication between SIEF and project teams. As part of this process, based on lessons from the Spanish IEF, the SIEF team also encourage project teams to engage in an early dialogue with relevant policymakers and to meet with them on a regular basis, because such strategies are in place to help ensure efforts will be made to achieve outcomes. Overall, while at this stage it is difficult to estimate to what degree they will contribute to any outcomes, the current arrangements lend confidence that outcomes will be documented.

7. Research question 6: consideration of gender

RQ6: Is gender being adequately considered in SIEF impact evaluations proposals and capacity development activities to date?

The last of the main research questions examined relates to gender. Similar to value for money, the dimension of gender underpins all of the pillars. As such, the extent to which gender has been adequately considered for all pillars has been a core part of our assessment. It is not easy, however, to specify upfront what constitutes 'adequate' and what does not. There are no predetermined benchmarks. Our approach has therefore been to first specify for each pillar what we perceive as the minimum consideration of gender and then to assess whether or not SIEF has aimed to move beyond the minimum. The consideration of gender can then only be considered adequate when more has been done than just the minimum.

The minimum considerations of gender that serve as the benchmark for the three pillars are:

- Pillar 1: data disaggregated by gender
- Pillar 2: workshop participants' scores disaggregated by gender; the objective is to include women, but no concrete actions are undertaken to achieve this
- Pillar 3: separate reporting on outcomes for women

7.1. Gender in proposals

Several observations can be made with respect to the gender dimension in the SIEF application process. Overall, no SIEF calls have been designed with an explicit objective to address gender issues, though it is conceivable that at least some of the projects evaluated by the funded IEs will have differential impact by gender. Nonetheless, gender considerations are manifested in several ways. At the most fundamental level, it is a SIEF requirement that data are collected and disaggregated by gender as part of funded projects. As three interviewees[61] (both SIEF and external) stressed, this is a fairly standard requirement that is common in other research environments as well. Furthermore, applicants are invited to give special consideration to gender as part of their application package. This feature on the application form is expanded for the third call for proposals, but it falls short of a requirement to have a special gender-specific component in the proposed project. Three participants in the focus group out of four who answered the question further confirmed that they thought gender was an important part of the work of

[61] Interview 6, Interview 7, Interview 8

SIEF. As one participant noted, 'most of the proposed evaluations that I have reviewed have a significant gender component'.[62]

Based on conducted interviews, it appears that this arrangement is not dissimilar from other funders. For instance, 3ie generally include subgroup analysis in their products, which needs to meaningfully reflect the context of the evaluated programme – in some instances, disaggregation by gender is appropriate, in other instances, disaggregation by other groupings, such as youth or other particular age groups, may be more suitable.[63]

While SIEF therefore appears to fulfil the minimum requirement, there are no direct indications it has substantially moved beyond this. As a way forward, SIEF can therefore consider to move from what one interviewee referred to as a 'gender informed analysis' towards a 'gender informed design'.[64] The implication is that gender is considered to be an important aspect of research from the start. This is vital, as another interviewee observed that in terms of data collection, the gender of the researcher can matter. For example, in certain circumstances, women might be more willing to provide information to a female researcher than to a male researcher.[65] In addition, SIEF currently does not require impact evaluations to be sufficiently powered to allow for the detection of gender-separate effects. In circumstances where gender is deemed an important factor, this could be introduced as a requirement.

7.2. Gender in capacity building

As was observed in Chapter 5, the share of women participating in the workshops is still below that of men, yet with an average participation of 41% women cannot be said to have been absent. The assessment of the quality and relevance of the workshops shows no differences between men and women either, as both groups rate the workshops highly. In this context it is worth reiterating the limitations faced by SIEF, in that the primary responsibility for identifying workshop participants rests outside of the SIEF team. And as one interviewee observed,[66] even where the SIEF team makes a recommendation for more female participants to be involved, this may not be feasible due to practical realities on the ground. The minimum requirement seems therefore to have been reached, but there are questions surrounding the capacity of SIEF to move beyond that. To further consider the dimension of gender in the workshops, it may therefore be necessary for SIEF and the donors to review how the organisation of workshops is taking place and what their exact purpose is. This question ties in to the larger question surrounding the workshops that was raised in Chapter 5.

[62] Focus group participant 1

[63] Interview 7

[64] Interview 13

[65] Interview 4

[66] Interview 8

7.3. Gender in E2P notes

To understand how gender has been reported on in the E2P notes, we reviewed all the 35 E2P notes that have been produced to date for gender. All E2P notes were scanned for the words 'women', 'girls' and 'boys'. The results per E2P note are listed in Appendix J. The scan shows that while most of the impact evaluations had relevance for gender, these contributions were only made on an ad-hoc basis. Specific subheadings on gender are used in only some of the E2P notes, and there seems to be some inconsistency in the level of detail provided. The minimum requirement seems, therefore, not to have been reached by the E2P notes.

7.4. Summary and answer

Gender is considered to different degrees in the activities of SIEF, but on the basis of the available data it cannot be concluded that gender has been considered adequately. First, although data collected under Pillar 1 is disaggregated by gender, what may be needed in addition is a move towards 'gender informed research designs', i.e. research designs that consider how gender might play a role and how data might be collected on this role. Second, while overall the average of 41% female participants is not bad, female participation at the workshops could still improve, especially in specific locations, as could data collection on the female participants' background. It would be helpful to know whether the background of these women was in research or in government, for example. Third, although the E2P notes have reported on gender, this has not always been done systematically or elaborately. For all three pillars, gender therefore has been considered, but the dimension of gender could be addressed more systematically.

8. Research question 1: overall design of SIEF

RQ1: Is the design of SIEF appropriate to achieve the intended outputs and outcomes? Specifically, how do SIEF systems and processes respond to the needs outlined in the business case?
(OECD DAC criteria: efficiency, effectiveness)

The first research question serves as both the start and the end of the MTR. As the overarching question, it provides an overview of the MTR and simultaneously summarises the evidence from all the other research questions. As such it also acts as the conclusion.

To address this final, overarching research question, we have divided it into two parts. First, we will assess the appropriateness of the design of SIEF to achieve the intended outputs and outcomes. Appropriateness in this respect is defined as the extent to which the activities of SIEF allow for and are conducive to the attainment of its objectives. We will review the appropriateness in two ways:

- Review of the progress made towards planned outputs and outcomes
- Review of stakeholder perceptions on the broader design of SIEF

Second, we will assess the extent to which SIEF systems and processes respond to the needs outlined in the business case. The primary need outlined in the business case is the need for more and better evidence of the (cost-)effectiveness of interventions. To assess whether SIEF is responding to this need, we will rely on the perceptions of stakeholders collected through the interviews, focus group and survey.

Finally, the evaluation criteria highlighted under this research question are effectiveness and efficiency. As this chapter brings together the other five research questions, the operationalization of these evaluation criteria will depend on the pillar. The previous chapters have each specified how the evaluation criteria have been operationalized in the context of the different pillars, and in this chapter we bring together and summarize the evidence arising from these previous research questions.

8.1. Appropriateness of SIEF design

The review of the appropriateness of the design of SIEF to achieve the intended outputs and outcomes has been approached from two angles: first, the progress in achieving outputs and outcomes has been assessed (i.e. the delivery of the Theory of Change), and, second, the design of the Theory of Change has been assessed.

8.1.1. Progress is achieving outputs and outcomes

The Theory of Change provides a very clear outline of the long-term outcomes and impacts that SIEF aims to achieve and of the activities that SIEF will undertake in order to reach these outcomes and impacts (see Appendix A). As such, the Theory of Change represents a reference frame for this research question.

As described previously, SIEF activities are organised within 3 pillars that together cover all the activities undertaken within the programme. In brief, for the MTR, the pillars and their related output and outcome targets[67] were the following:

- Pillar 1: funding of impact evaluations (IE) of interventions with a human development outcome
  - Output: 45 high-quality IE are expected from SIEF, of which (at least) 40 will be published
  - Outcome: it is expected that '50% of SIEF funded impact evaluations trigger changes in the design/implementation of evaluated programs, including decisions to kill/continue/scale up of programs'
  - Outcome: it is expected that 'in 2020/2021 40% of SIEF funded impact evaluations inform program or policy changes beyond the evaluated program'
- Pillar 2: organising and running capacity-building activities on IE methods and implementation techniques, including training and workshops
  - Output: by 2017 at least 1500 participants are trained in impact evaluation techniques
- Pillar 3: strategic dissemination of IE findings through evidence-to-policy notes
  - Output: by 2017, 66 evidence-to-policy notes produced
- For all pillars combined, the following long-term impacts are expected:
  - 'Improved delivery of DFID and World Bank programmes with human development outcomes through better use of evidence and research.'
  - 'Improved development outcomes for the poorest.'

Additional actions and targets are included in the revised Logframe, but these were outside the scope of the MTR, which commenced before the revision was completed. Given that the outputs and outcomes cover all three pillars, the other research questions follow from and, in turn, have served to answer the first research question. The first research question has therefore been partly, but not wholly, dependent on the answers generated to the other five research questions.

The available data, ranging from early conversations and the annual reviews to in-depth interviews with stakeholders, seem to indicate that generally SIEF is on track to achieve the intended outputs (see Table 30). For each pillar, we can further illustrate the progress made, based on the previous chapters.

[67] Output and outcome targets were taken from the revised SIEF Logframe, dated 5 March 2015.

Table 30: Status of outputs by pillar

Pillar	Output	Target (by 2017/18)	Current status
Pillar 1	IE	45 (54 in business case)	50 in progress
Pillar 2	Workshops	20 to be organised	10 organised
Pillar 2	Participants	1500 to be reached	900 reached
Pillar 2	IE designs	–	156 produced
Pillar 3	Evidence-to-policy notes	66 to be published	25 produced*

*Since no SIEF-funded impact evaluations have been completed yet, the heretofore produced evidence-to-policy notes cover other IE work, most notably findings from studies funded under the previous Spanish IEF.

Pillar 1 (Chapter 4)

Effectiveness has been assessed with respect to timeliness, objectivity, transparency and inclusivity. In terms of timeliness and objectivity, the calls appear to have been managed effectively, yet the MTR raises some questions over the transparency and inclusivity of the calls. Furthermore, the quality assurance mechanisms in place appear to be adequate. Data from the interviews and focus group confirm that SIEF is perceived as being good at selecting high-quality impact evaluations. In fact, the point was raised that a strict focus on quality may risk missing opportunities for capacity building.

Pillar 2 (Chapter 5)

On the basis of the available data, it appears that the workshops organised so far have been of high quality and have been relevant to their participants. One element that may require further attention is learning. While still good, the scores for the fulfilment of learning needs were somewhat lower, and open-text comments seem to suggest that practical examples could further improve the quality of the workshops.

Pillar 3 (Chapter 6)

The E2P notes are generally deemed to be of high quality by interviewees. Several interviewees made suggestions, which have been summarised in this report. Most important, perhaps, is the suggestion to include more information on the actual implementation of the interventions discussed. E2P notes were also considered to be relevant, yet several interviewees did comment that notes on their own will not be sufficient to bring evidence to policymakers.

Given the fact that no SIEF-funded impact evaluations have been completed yet (most of the IE from Call 1 and Call 2 are scheduled to complete in 2016/17), it is not yet possible to comment on the extent to which SIEF is on track to meet its outcome targets; indeed, this is also not the aim of the MTR. However, the data collected do provide insights into what might be effective ways to reach outcomes. For example, what kind of evidence will be useful from the perspective of the policymaker? These comments have been included under the relevant research questions. In the final chapter, on recommendations, we will return to a few of these insights, because they can serve as guidance for the future.

8.1.2. Design of the Theory of Change

Two aspects related to the design of SIEF and the appropriateness to achieve the outcomes were explored in the interviews with stakeholders. The first relates to the three pillars of SIEF; the second, to the positioning of SIEF within the World Bank.

Three pillars of SIEF

SIEF is structured along three pillars, each performing a different function. It was observed, however, that the linkages between the three pillars are not always clear and straightforward.[68] For example, while it is clear that the E2P notes will be based on funded impact evaluations, the link between the capacity-building activities and the funded impact evaluations is less straightforward. Not all funded researchers will attend the workshops, and, in turn, not all workshop participants who work in research will submit a proposal to SIEF. The survey shows that from the perspective of successful applicants this is in line with their expectations, as they expect to see more benefits from SIEF in the area of communication and dissemination than in capacity building. It does raise a larger question about the exact purpose and identity of the SIEF workshops, a question also raised in the annual review of 2014.[69] This does not seem to be a matter that can be resolved by the SIEF core team alone, as they are not the only stakeholder – it concerns the expectations of all stakeholders of the purposes of different programmes, such as SIEF, DIME and 3ie.[70] To avoid unnecessary duplication and to ensure an effective 'division of labour', some coordination might be required from the donors and the World Bank to ensure that it is clear to programmes such as SIEF what they are expected to deliver.

SIEF situated within the World Bank

A specific trait of SIEF is its positioning within the World Bank and, in particular, its position within Global Practices within the World Bank (as opposed to its research arm). The interviews, focus group and survey illustrate that there are both benefits and drawbacks to this arrangement. Some of the benefits mentioned are as follows:

- The survey shows that when asked about the benefits applicants expected from an application to SIEF, the most frequently mentioned benefit is post-project communication and dissemination.
- Compared with other funders, the survey also shows that SIEF is perceived as open to consultation and having reasonable reporting requirements and strong technical skills to provide support to ongoing projects.
- Several interviewees mentioned that the benefit of situating SIEF within the Global Practices of the World Bank means that the research evidence is much closer to project design and implementation and can therefore be much more directly communicated.[71]

---

[68] Interview 6

[69] Annual Review 2014: 'To overcome any confusion between the various WB capacity building mechanisms for evaluation, there should be clearer coordination between SIEF, DIME and CLEAR and a clear articulation of the differences between them.'

[70] This point also resonates in Interview 6

[71] Interview 7, Interview 9, Interview 13

- One interviewee mentioned that the practical implementation experience of staff at the World Bank helps with the assessment of the future potential of proposals.[72]

Most of the benefits identified through the survey pay tribute to the work performed by the SIEF core team, and to acknowledge these benefits is crucial. For the moment, however, these benefits do not yet appear to touch on the added value of the location of SIEF within the World Bank. It should be remembered, though, that most survey respondents expect the added value of SIEF to show only towards the end of the impact evaluation, and that the IEs were not yet at that stage when we conducted the survey and focus group. Furthermore, the interviews and one focus group participant[73] did explicitly refer to the added value of situating SIEF within the World Bank.

Some of the drawbacks of situating SIEF within the World Bank that were mentioned are:

- The difficulty of getting a partner within the World Bank (i.e. a TTL) if you do not already have one to prepare a proposal[74]
- Risk of 'capture' and the perception of preferential treatment of World Bank researchers[75]

Finally, one of the 'selling points' of SIEF is the access to World Bank contacts and (in-country) policymakers and partners.

8.2. Response of SIEF processes to needs

The primary need outlined in the business case is the need for better evidence to improve interventions with human development outcomes. To assess whether SIEF is responding to this need, we collected data from various stakeholders on their perceptions of SIEF. If SIEF is seen as a funder addressing a gap in knowledge by multiple stakeholders, then there are indications that SIEF is addressing the need. In the absence of the actual impact evaluations, which will only be completed in the future, this is the best proxy we have to assess whether or not SIEF is responding to the need.

Interviews and focus group

While it is known that SIEF is not the only funder of impact evaluations on interventions with a human development outcome, the data from the interviews and the focus group show that there is a gap in knowledge and that the impact evaluations funded by SIEF are generating evidence that can be expected to fill this gap. Explicit mention of this was made by both NGOs[76] who work on human development and by the multiple donors of SIEF.[77]

---

[72] Interview 1

[73] Focus group participant 3

[74] Interview 5

[75] Interview 9, Focus group participant 6, Focus group participant 10

[76] Interview 3, Interview 4

[77] Interview 1, Meeting 1

Survey results: reasons for applying to SIEF

The survey shows that the majority of respondents did not apply for any other source of funding in parallel with their SIEF application. Among unsuccessful applicants, almost three quarters of applicants (74%) submitted an application only to SIEF, while the split among successful applicants was even.

Of those who did submit a parallel funding application, this application was successful in nearly three quarters of cases. This proportion was even higher among those whose SIEF application was ultimately successful. This is not very surprising, as it suggests SIEF quality criteria are aligned with those of other funders.[78]

Table 31. Applications for complementary sources of funding

	Applied successfully	Applied unsuccessfully	Did not apply	Prefer not to say
Successful	16 (41%)	3 (8%)	19 (49%)	1 (3%)
Unsuccessful	9 (10%)	6 (7%)	65 (74%)	8 (9%)
Total	25 (20%)	9 (7%)	84 (66%)	9 (7%)

Of all offered explanations, the most frequently stated reason (46%) for applying to SIEF funding only was that SIEF was the only identified potential funder for the proposed project in question. The next most frequently mentioned reasons were high perceived chance of winning (23%) and acceptability of preparation costs (17%). Interestingly, a consideration of the chance of winning was notably more prevalent among successful applicants.

Eleven respondents (four successful and seven unsuccessful applicants) explicitly stated that they perceived SIEF to be the best match for their application given either the design of the proposed project, its topic or its timing. The offer of seed funding and SIEF's acceptance of parallel funding applications were also noted as reasons for applying to SIEF.

Table 32. Reasons for applying to SIEF only

	SIEF was the only identified potential funder that was receiving applications	SIEF application was seen as most acceptable in terms of preparation costs	The chances of receiving SIEF funding were perceived to be the highest of all potential funders	Lack of capacity to prepare multiple applications and/or manage multiple grants	Prefer not to say
Successful	10 (53%)	3 (16%)	7 (37%)	1 (5%)	0 (0%)
Unsuccessful	32 (44%)	13 (18%)	14 (19%)	6 (8%)	3 (4%)
Total	42 (46%)	16 (17%)	21 (23%)	7 (8%)	3 (3%)

[78] Other funders approached by SIEF applicants included other WB trust funds, Children's Investment Fund Foundation, ESRC, Gates Foundation, Johns Hopkins University, University of Maryland, National Institutes of Health, International Growth Centre, Oxford University, KCP, Korean Development Institute, Wellcome Trust, Templeton Foundation, and USAID.

Among successful applicants, SIEF funding appeared to be of critical importance in approximately half of all cases. In 23% of instances, respondents suggested that their project would have been abandoned if they had not received SIEF funding. In another 23% of cases, respondents indicated they would have applied for the next round of funding. In 43% of instances, respondents suggested they would have applied for another source of funding.[79]

Table 33. Answers to the question 'What would have happened in the absence of SIEF funding?' (successful applicants only)

Outcome	Count/ proportion
Would have abandoned project	8 (23%)
Would have applied for funding from elsewhere	15 (43%)
Would have applied for the next round of SIEF funding	8 (23%)
Other	4 (11%)

Among unsuccessful applicants, the failure to obtain SIEF funding led to the abandonment of the proposed project (or at least of the component of the project that relied on SIEF funding) in 31% of cases, and another 17% respondents indicated that they reused or intended to reuse their application in the next SIEF call for applications. About two fifths of respondents (37%) applied for funding elsewhere or were considering doing so.[80]

Table 34. Answers to the question 'What happened in the absence of SIEF funding?' (unsuccessful applicants only)

Outcome	Count/ proportion
Applied for funding from elsewhere	30 (37%)
Applied for the next round of SIEF funding	14 (17%)
Abandoned project	25 (31%)
Other	12 (15%)

Survey results: perceived benefits from SIEF funding application

Survey participants were invited to indicate what benefits, in addition to funding support, they were considering when applying for SIEF funding. The wording of this question reflects an underlying assumption that the primary reason for any SIEF application is funding. Respondents were able to confirm this assumption by explicitly selecting 'none' among the offered options, which 21% of them did (26% of successful and 19% of unsuccessful applicants). The most frequently sought benefit was support

[79] These included 3ie, other World Bank funds, ESRC, IGC, Hewlett, NIH, Wellcome Trust, Gates Foundation, and WPP.

[80] These other potential funding sources included ESSPIN, IMMANA, J-PAL, NIH, other World Bank Trust Funds, Parents Teachers Association (PTA), and DFID-ESRC.

with post-project communications and dissemination (42% of all respondents). All the remaining options were selected by 20–27% of respondents. With the exception of communications and dissemination, unsuccessful applicants were generally more likely to expect additional benefits than were their counterparts who were ultimately selected for SIEF funding.

Other benefits explicitly mentioned by individual respondents include the rigorousness of the proposal stage, quality control provided by SIEF and the reputation and connections of SIEF and WB more widely, which would help with final dissemination.

Table 35. Expected benefits of SIEF application

	Technical support with proposal development	Technical support with project execution	Capacity building	Access to in-country contacts and partners	Access to (in-country) policymakers	Post-project communication and dissemination	None	Other
Successful	8 (21%)	9 (24%)	4 (11%)	6 (16%)	7 (18%)	21 (55%)	10 (26%)	3 (8%)
Unsuccessful	21 (24%)	23 (26%)	30 (34%)	24 (27%)	18 (20%)	32 (36%)	17 (19%)	8 (9%)
Total	29 (23%)	32 (25%)	34 (27%)	30 (24%)	25 (20%)	53 (42%)	27 (21%)	11 (9%)

When these data are disaggregated by the role of respondents (see Table 36), one notable difference emerges. Provision of access to in-country partners and policymakers was much more frequently mentioned by PIs (among whom it ranked the highest after communication and dissemination) than by TTLs. This is perhaps not very surprising given SIEF's design, because one of the purposes of the requirement to have a TTL is precisely to facilitate in-country contacts via World Bank networks. Nevertheless, two observations can be made. First, there still appear to be some TTLs who would like SIEF to assist with in-country contacts. And, second, though in-country access ranks higher as an expected benefit among PIs, its frequency does not differ very much from that of other benefits in this group. This suggests that while applicants appreciate and expect assistance with in-country access, they may not view this assistance as the primary non-financial contribution of SIEF.[81]

[81] We make this point explicitly to contrast with an expectation shared by a member of the SIEF core team that in-country access constitutes one of the biggest added values of SIEF.

Table 36: Expected benefits of SIEF application (disaggregated by respondents' position)

	PI	TTL	Total
Technical support with proposal development	17 (23%)	12 (23%)	29 (23%)
Technical support with project execution	18 (25%)	14 (26%)	32 (25%)
Capacity building	20 (27%)	14 (26%)	34 (27%)
Access to in-country contacts and partners	23 (32%)	7 (13%)	30 (24%)
Access to (in-country) policymakers	22 (30%)	3 (6%)	25 (20%)
Post-project communication and dissemination	30 (41%)	23 (43%)	53 (42%)
None	18 (25%)	9 (17%)	27 (21%)
Other	5 (7%)	6 (11%)	11 (9%)

Table 56 in Appendix I shows a breakdown of benefit expectations by applicant cluster. Applicants in the ECD cluster indicated the highest number of potential benefits, whereas WASH applicants reported the fewest. Consideration of capacity building and assistance with access to policymakers were particularly infrequent among health and WASH applicants.

8.3. Summary and answer

The majority of interviewees agree that SIEF is on track to achieve its outputs. The output figures also show that substantial progress has already been made to achieve the targets. In terms of achieving outputs, the design of SIEF therefore seems to have been appropriate. It has been able, so far, to rigorously select and subsequently support impact evaluations, to organise high-quality and relevant workshops and to publish a considerable number of E2P notes. To achieve these outputs, the quality and support from the SIEF core team has appeared as vital to progress in the interviews, the focus group and the survey. Through the pillars, SIEF systems and processes have largely been able to respond to the needs of the business case, at least in terms of outputs.

As this is a mid term review, it is not yet possible to assess the achievement of outcomes. However, questions raised over the course of this MTR on the different pillars do provide indications that the pillars as (narrowly) defined in this MTR are unlikely on their own to fully achieve the outcomes. To build capacity of researchers, more handholding support may be needed and a larger inclusion of researchers from the Global South seems to be required.[82] Furthermore, to reach policymakers, a reliance on E2P notes may be too confining.[83] These observations have also been made by SIEF; therefore, a new range of activities has been developed by SIEF within Pillar 2 and Pillar 3 to help achieve the outcomes. Under the relevant research questions, we have provided some feedback on these based on the data collected.

Finally, the different pillars and the range of activities currently developed by SIEF under Pillar 2 and Pillar 3 do raise a question as to what the expectations of SIEF are, also in comparison with other programmes funded, such as DIME and 3ie. Because of limited budgets, not every programme can do

[82] Interview 6

[83] Interview 1

everything, and some degree of prioritisation is inevitably required. There is a risk of overlap, as was also observed in the annual review, and there is a risk that efficiencies that could be achieved through a division of labour will be lost. This is not simply a matter for the SIEF core team to consider. How this risk is mitigated will also depend on the expectations of the donors and their vision of the functions of different programmes.

8.3.1. Secondary research questions

A large number of the secondary research questions have been squared under the first research question. Within the scope of the MTR, we cannot address these at the same length as the primary research questions. Still, based on the available evidence above, we can provide preliminary answers to these questions. To fully answer them would require additional analysis.

Box 8-1: Secondary research questions

Do the four key priority areas remain highly relevant (due to lack of robust evidence)? Are additions and amendments appropriate?

Based on evidence collected through the interviews and focus groups, there are no indications that the priorities are no longer relevant insofar as the priorities of SIEF's stakeholders are concerned.[84] SIEF is seen to be making a valuable contribution to filling the gap in knowledge in these areas. What is more, the potential for this contribution is maximised by increasingly targeted and coordinated calls for proposals and monitored by the SIEF team, as evidenced by undertaken portfolio reviews in all four clusters. With respect to additional clusters, it was mentioned in one of the interviews that new areas have been considered, such as labour and employment. However, so far these have not been pursued.[85]

Have the five key risks outlined in the SIEF business case so far been mitigated effectively?

The five key risks outlined in the business case, which are listed in Appendix F, are here briefly addressed in turn. It should be noted that a full assessment of these risks is beyond the scope of the MTR.

1. SIEF is used for non-strategic evaluations that offer limited value to DFID or globally: risk of not including low-income countries. The current overview of impact evaluations funded (see Chapter 4) shows that the majority (79%) are being conducted in Africa (49%) and South Asia (30%). As such, there does not seem to be a direct risk of missing low-income countries. Furthermore, it is worth reiterating that, as an explicit selection criterion, SIEF funds only those impact evaluations with potential for scalability and replicability in LIC settings.

2. The initiative is used to substantiate pre-determined positions rather than for rigorous, evidence-

---

[84] We reiterate that, as outlined in section 1.4, relevance refers to the suitability of the aid activity to the priorities of stakeholders, including donors, recipients and beneficiaries.

[85] Interview 8

based approaches. The MTR shows that, overall, the quality assurance processes in place are adequate and are rated highly by multiple stakeholders (see Chapter 4). This should guard against the substantiation of pre-determined positions. In addition, submitted IE proposals have generally been considered of high quality[86] and no cluster portfolio review found any risk of substantiating pre-determined positions. As such, SIEF impact evaluations can be expected to generate a robust evidence base.

3. There is limited uptake of evidence from impact evaluations. Because the current report is an MTR, the answer to this question can only really be determined at a later date. Still, actions can be taken to ensure there will be uptake. One of the actions is to move beyond just having evidence-to-policy notes. This is already being done by SIEF and is further discussed in Chapter 6.

4. Impact evaluations do not adopt a high level of rigour. Similar to risk 2, the quality assurance processes in place are likely to guard against this risk.

5. Impact evaluations substitute for rigorous evaluations that should otherwise be taken up by individual World Bank or DFID programmes. This is hard to assess on the basis of the MTR, as the MTR only looks at SIEF and not at the wider relationship between DFID and WB.

Are the secretariat's and the Board's role and configuration appropriate to the needs?

A detailed review of the appropriateness of the role of the Board and Secretariat is beyond the scope of this MTR. We can, however, comment on a related matter, which is the perception of stakeholders of the functioning of the Board and the Secretariat. The role and the work of the secretariat have generally been rated highly by both stakeholders and survey participants. As such, they largely seem to meet the needs. Still, in the various chapters indications are provided for the improvement of the pillars.

The governance of SIEF is reflective of its close working relationship with the relevant WB Global Practices, and evidence collected through interviews with senior staff[87] and a document review suggests that the arrangement is well received by all parties. In practical terms, senior management have been able to provide strategic steer as required, incorporating input from relevant stakeholders.

Is the coordination between SIEF and other impact evaluation funds and initiatives effective?

For the purpose of this secondary research question, we can define effective coordination as ensuring that the different initiatives make sure they do not duplicate efforts and that they collectively meet the requirements of funders. On the basis of interviews with other impact evaluation funds (i.e. 'peers'), it appears that there could be additional coordination, for example, in the area of training. However, this also raises a larger question as to who does what. The expectations of the different programmes will need to be clear to ensure an effective division of labour. This is not solely the responsibility of SIEF or the peers; it also requires input from donors, as listed in this chapter. We return to this topic in the recommendations.

[86] Based on the 2013/2014 SIEF Annual Report, which commented on Call 1 proposals.

[87] Interview 1, Interview 8, Interview 9, Interview 13

What progress has been made towards securing funding for SIEF from other donors?

On the basis of information available, SIEF has been able to secure additional funding from one new donor, the Children's Investment Fund Foundation.

9. Conclusions and recommendations

9.1. Observations and recommendations

On the basis of the collected evidence, we put forth the following observations and recommendations for consideration to both the SIEF core team and the donors of SIEF. In addition, we list several methods to be considered in the future as proxies to measure the outcomes of SIEF.

9.1.1. Overall design

The relationships among the pillars

The first observation relates the overall design of SIEF. It has been observed several times already in this MTR that there is a risk of tension between the various pillars. On the one hand, SIEF aims to fund only the highest-quality impact evaluations (Pillar 1), and this generally leads to a selection of researchers from well-established Western universities. On the other hand, SIEF is committed to global capacity building through workshops (Pillar 2). At the moment the result of these two aims seems to be that the 'audiences' for these two pillars are very different. The workshop participants (including government officials), even with training, are for the moment unlikely to win funding, while successful applicants are unlikely to be in need of the training provided at workshops. This may well have been anticipated by SIEF, but it does mean that there is often no immediate and direct link in the near future between the people being trained at workshops and the researchers winning funding.

Furthermore, while there is a clear 'gravitational pull' of all three pillars towards addressing the important questions in development, this could be strengthened. By more explicitly identifying 'key performance questions' (i.e. questions that, if answered, would support a significant change in meeting identified needs) and more systematically engaging with policymakers and practitioners about what these questions are, there could be an easier orientation of the three pillars both towards each other and towards global priorities. SIEF is in a strong position to articulate key performance questions for development that are amenable to analysis through the methodologies that underpin impact evaluations.

In theory, there is a role for TTLs to coordinate and mediate the achievement of outcomes across all three pillars. A detailed exploration of the extent to which TTLs take such an active role was beyond the scope of this report; however, several related comments can be made. First, there appears to be substantial variation in the degree of involvement of TTLs in their respective IE projects.[88] Second, although this is

[88] Interviews 8 and 9

circumstantial evidence at best, we note that approximately half of the surveyed funded TTLs (47%) reported consulting about dissemination with SIEF, and a slightly higher share (56%) saw communications as one of the benefits of having SIEF funding. At the same time, very few of the surveyed TTLs indicated a need or desire for help from SIEF to gain access to relevant policymakers, suggesting that there may be already existing avenues from TTLs to get their buy-in outside of SIEF's activities. Finally, SIEF training activities welcome, and are to a large extent driven by, input from relevant global practices and country teams. Unfortunately, this review is not able to comment on how much of a role TTLs themselves play in these processes.

Division of labour between programmes

Our first observation links into a second observation about the functions that different programmes perform. Apart from SIEF, there are other funders of impact evaluations who use different methods of funding, such as 3ie and DIME. In part, these programmes attract similar donors for funding. To avoid duplication of effort, it may be worthwhile for both donors and programmes to set clear expectations as to what the functions of the different programmes are. Alternatively, if the different programmes see themselves in open competition to drive up the standard of impact evaluations, then this should be made clear to prospective applicants. Such clarity might possibly help to resolve the potential tensions outlined above. Furthermore, it might help to resolve the current lack of clarity over, for example, training and capacity building, as identified in the annual review of 2014. Generating such clarity should not, however, be the sole responsibility of the individual programmes, but is best achieved through consultation with all stakeholders.

9.1.2. Pillar 1: calls for impact evaluations

Overall the calls have been effective, and at this relatively early stage, the evidence supports the conclusion that quality assurance mechanisms have been fit for purpose. Two areas, however, appear to warrant further attention: transparency and inclusivity.

Transparency could be increased through clearer communication of the reasons why proposals have not been successful. This is, of course, time- and resource-consuming, but in order to remove doubts about the process of selection, (some) feedback on the decisions seems necessary. It also would also help meet the aim of capacity building.

Inclusivity will be more difficult to achieve, given the inherent tensions outlined above. Researchers from the Global South currently do not feature prominently on SIEF-funded impact evaluations. We assume that this is likely the result of the strict quality criteria SIEF adheres to in the selection process. Building capacity in the Global South is not something that can easily be done in the short term, and thus it seems unlikely that there will be a substantial number of high-quality proposals from these regions in the near future. However, at least as important as building capacity in the Global South is building capacity in both North and South to collaborate in successful IEs. There are options for SIEF:

- SIEF could aim to fund only the highest-quality research by waiting for good responses to its ITTs. It is not likely that inclusivity would increase as a result, but such a funding strategy would continue to produce impact evaluations of the highest academic standing.

- SIEF could actively support submissions from the Global South.
- SIEF could promote greater meaningful collaboration between researchers from North and South.

The open-text boxes in the survey yielded a number of practical comments from researchers, in addition to the observations above, that the RAND team can make available to SIEF in anonymized format.

9.1.3. Pillar 2: workshops and capacity building

On the basis of the available data, the quality and relevance of the workshops are in line with targets of the Logframe. In terms of participants, however, the workshops have had a proportionally large presence of governmental officials and relatively fewer researchers and academics. Government officials are of course important. They can become 'smart consumers' of impact evaluations and help build effective demand for such products. While SIEF is likely to already actively target researchers, this is an effort that should be continued to contribute to capacity building in conducting, as well as using, impact evaluations.

The new activities proposed by SIEF in the strategy for Pillar 2 and Pillar 3, such as an online course, are also likely to contribute to such capacity building. In planning for these new activities, it is important that SIEF be explicit not only about the intended content of the learning, but also about the intended behaviour changes.

9.1.4. Pillar 3: E2P notes and dissemination

Overall, the E2P notes are seen to be of high quality, but suggestions have been made above on the layout and the content on the basis of the data collected. A wider point emerged in the focus group that relates to the type of information that is used by policymakers. Individual impact evaluations were deemed unlikely to change policy, but it was thought that guidelines and synthesis documents might. The synthesis documents proposed by SIEF therefore resonate quite well with these findings and may provide material on top of the E2P notes that can be used by policymakers.

In addition, alternative means of communication, such as social media, can be explored for the dissemination of the results. Plans for alternative means of communication are already in the updated strategy for Pillars 2 and 3. Considerations of who might promote the findings (who are the so-called trusted boundary spanners between the world of research and policymakers and practitioners), and how to recruit these individuals, would help support further impacts. Alongside this, DFID, CIFF and the World Bank all have strong networks that could also be mobilised.

9.1.5. Gender

While the dimension of gender has been addressed in all three pillars, it could still be integrated more systematically into SIEF's processes. It is worth emphasising that, in addition to using impact evaluations to produce gender-disaggregated data and gender mainstreaming, SIEF also has opportunities for using impact evaluations to look at transforming gender norms and behaviours associated with gender roles. Gender in this respect goes beyond counting women and includes men and boys. In all three pillars, the concept of gender could be present from the start to inform researcher designs,[89] workshop recruitment and E2P notes design, thus avoiding the risk of ad-hoc reporting. In this way, gender would start to become an integral part of the research and of the reporting.

9.2. Future measurement of outcomes and impacts

Finally, we conclude the MTR with a number of suggestions for the measurement of outcomes and impacts in the future. These suggestions are based on the experience of the RAND team with the measurement of research outcomes and impacts.

9.2.1. Pillar 1: calls for impact evaluations

The outcomes and impacts of impact evaluations can be measured in a number of ways, commonly using three main methods:

- Bibliometrics: to assess the academic utility and visibility of the impact evaluations, bibliometrics could be performed on the studies funded through SIEF. The (field-normalised) number of citations can provide an estimate of the use that other researchers have made of the work. In addition, citation levels are often used as a proxy of the quality of research; we note, however, that there are caveats around this assumption. Bibliometrics might also allow international collaborations to be identified and tracked.
- Survey of researchers: researchers can be anonymously surveyed to ask them what they think the impact has been of their work. Researches are likely to know if their work has had an impact on policy or practices, and a survey of all researchers funded through SIEF could provide an overview of SIEF impacts at a portfolio level. A survey could also be extended to consumers of research (e.g. DFID and WB).
- Case studies: to highlight and understand particular outcomes or impacts, in-depth case studies could be conducted. Such case studies could trace from beginning to end how the study was funded, how it was implemented and what the impact were.
- Together these methods can contribute to producing a version of the 'payback model', which articulates the domains of contribution and the routes to achieving these.[90] This model can then provide a basis for learning, ongoing review and adaptation.

9.2.2. Pillar 2: workshops and capacity building

Overall, to understand if the capacity-building workshops indeed had a lasting effect on the participants, follow-up surveys should be conducted. Depending on the availability of email addresses, a survey could

---

[89] In addition, SIEF currently does not require impact evaluations to be sufficiently powered for gender-separate effects, and in circumstances where gender is deemed an important factor, this could be introduced as a requirement.

[90] The payback model is built on work by Martin Buxton and Stephen Hanney in the Health Economics Research Group (HERG). Buxton, M, and Hanney, S (1996) How can payback from health services research be assessed? Journal of Health Services Research 1(1):35-43. For a discussion of its operationalisation in research evaluation, see Marjanovic, S (2009) Payback Framework. In: Ling, T and Villalba van Dijk, L (eds) Performance Audit Handbook: Routes to effective evaluation. Santa Monica, Calif.: RAND Corporation.

be distributed to former participants to understand whether and how they use the knowledge gained through the capacity-building workshops. These surveys would have to be tailored to the different types of participants, because the usage will vary between an academic and a governmental official. Academics can be surveyed about their ability to develop strong IE proposals and execute strong IE projects and to contribute to the emerging body of evidence in their respective areas.

However, 'capacity' should not only be conceived of as stemming from individuals. It is also a product of effective networks. Follow-up surveys could provide a good vehicle for also asking questions to support a social network analysis (SNA) focused on understanding how knowledge is generated and used within these research and policy networks.[91] Once central figures have been identified within these networks, more qualitative interviews might be conducted to understand how the networks function. In particular, it would be important to understand how 'bonding social capital'[92] is achieved to consolidate relationships within existing communities (e.g. South–South relationships) and to also address 'bridging social capital', creating new linkages across communities (e.g. research–policymakers or South–North). Over time, SIEF should consider whether developing an email-based SNA (tracking networks through email traffic) might provide a more automatic and scalable solution to collecting data than do surveys.

Keeping in mind the inherent difficulties in identifying suitable workshop participants as far as policymakers and NGO representatives are concerned, it may also be worth keeping track of the initial reasons for their inclusion among workshop participants and to take this into account in follow-up surveys. Relevant questions would probe post-workshop career trajectory and longevity to verify if individuals had been in a position to make an impact in the first place. Subsequent questions would focus on how knowledge and information gained through SIEF workshops were leveraged to effect change in policy and practice in their respective fields.

Similar follow-up surveys can be conducted in the future with participants in online courses.

9.2.3. Pillar 3: E2P notes and dissemination

Measuring the policy impact of communication and dissemination material is probably the most difficult of all three pillars. Monitoring downloads and online distribution can be used as a first proxy of usage. Including hyperlinks to the original article in E2P notes could provide data about how many people used the E2P notes as a gateway to the original article, and in the future Tweets might serve as such a gateway to E2P notes.

---

[91] For detailed discussion of social network analysis, see, for example, the following titles: Hanneman, R, and Riddle, M (2005) Introduction to Social Network Methods. Riverside, Calif.: University of California. Hawe, P, and Ghali L (2007) Use of social network analysis to map the social relationships of staff and teachers at school, Health Education Research, 23(1), pp 62-69. Krebs, V (2004) It's the Conversations, Stupid! The Link between Social Interaction and Political Choice. In: Ratcliffe, M, and Lebkowsky, J (eds) Extreme Democracy, pp 114-128. Laumann, E, Marsden, P, and Prensky, D (1989) The Boundary Specification Problem in Network Analysis. In: Freeman, L, White, D, and Romney, A (eds) Research Methods in Social Network Analysis. Fairfax, VA: George Mason Press. Scott, J (2000) Social Network Analysis: A Handbook. New York: Sage. Wasserman, S, and Faust, K (1994) Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press.

[92] For a detailed discussion of social capital, including its bridging and bonding functional subtypes, see Halpern, D (2005) Social Capital. Cambridge: Polity Press.

In addition, the surveys of both researchers and workshop participants can include questions on the E2P notes and other materials prepared by SIEF. Researchers may have heard of usage of 'their' E2P notes, and government officials attending the workshops may have started using the E2P notes after attending.

In addition to understanding dissemination as framed by 'SIEF to user' it would also help to use the network analysis described above to understand how new routes to impact are being created involving 'SIEF to many users/many users to each other/users to SIEF'. In addition to survey- or email-based analyses, social networks provide a basis not only for communicating, but also for researching impact and usage. It is not likely that there will be a single suitable measure of effective communication, but triangulating a variety of sources to understand the effectiveness of individual learning, bonding within existing communities and bridging to create new relationships would create a powerful conceptual basis for future learning.

Appendices

The following appendices are included:

Appendix A: Theory of Change for SIEF

Appendix B: Interview protocol

Appendix C: Survey questions

Appendix D: Detailed survey methodology

Appendix E: Focus group protocol

Appendix F: Key risks from the business case

Appendix G: 10 questions to assess knowledge gain

Appendix H: Workshop data gender-disaggregated

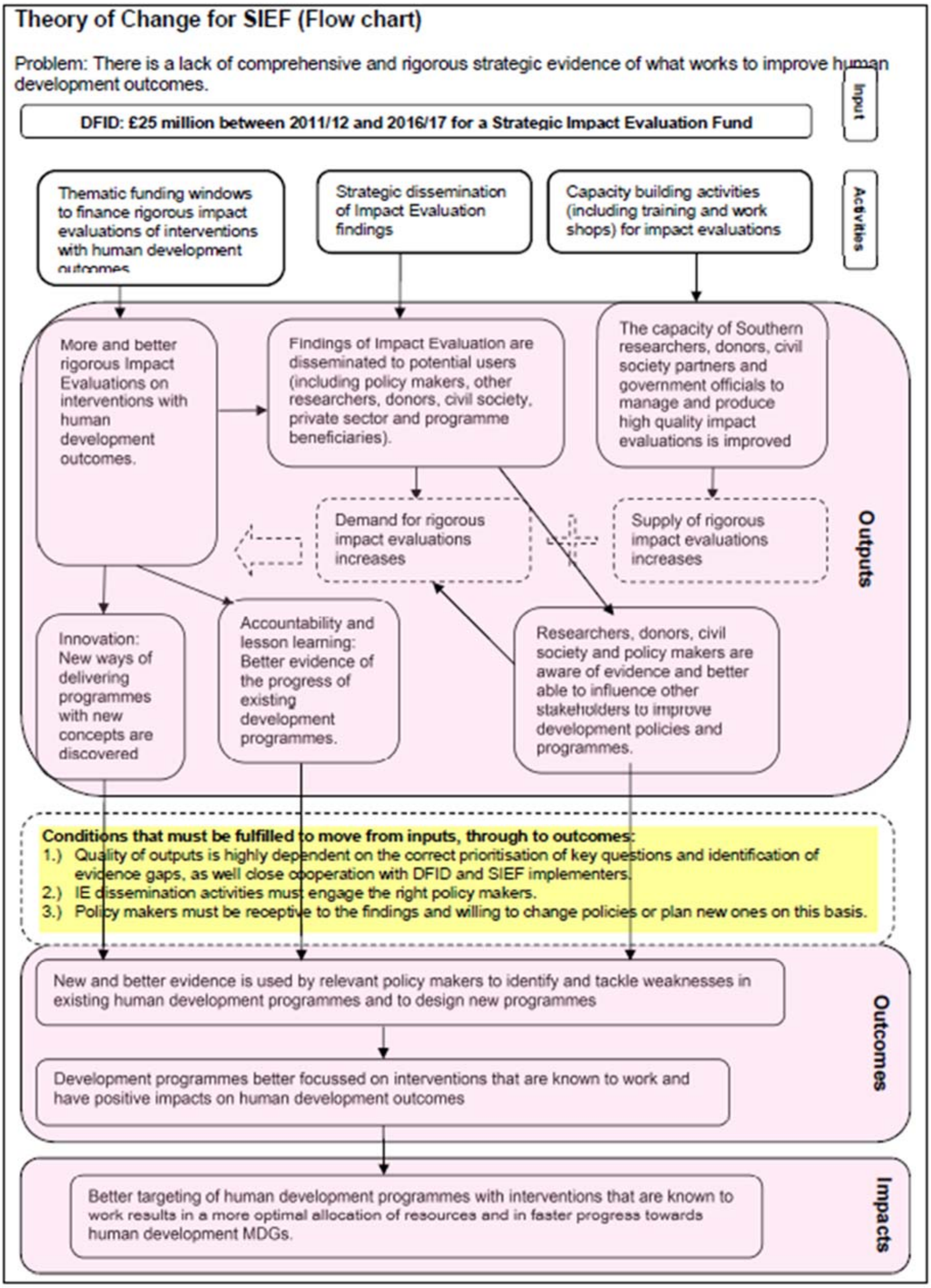
Appendix I: Additional survey results

Appendix J: Results E2P notes gender check

Appendix K: SIEF proposal selection process

Appendix L: Comments on E2P notes

Appendix A: Theory of Change for SIEF



Appendix B: Interview protocol

Role in relation to SIEF

1. What is your role in relation to SIEF? (e.g. donor, peer)

SIEF in general (may not be relevant for peers)

2. What do you think are the main objectives of SIEF and is it on track to achieve them?
3. Do you think SIEF will be able to reach policy makers and practitioners to generate lasting change?
4. Why do you think that?

Pillar 1 – Application and Selection

(for those involved in the design of SIEF)

5. What was the original rationale behind the design of the application and selection process?
    a. [possible follow-up: why is peer-review only included once seed funding has already been given?]

(For all)

6. Do you perceive the current application and selection to be efficient? (i.e. is it quick and thorough enough?)
7. Do you feel the current application and selection is able to distinguish strong from weaker applications?
8. What criteria are used to determine the quality of the proposal?
    a. Which criteria weigh most heavily?
9. Do you think the current application and selection is objective (i.e. that it does not prefer some candidates over others irrespective of the quality of the proposal)?
10. Is SIEF able to attract good quality proposals from good quality research teams which are individually of high quality and collectively coherently address priority issues?
11. Are you aware of any existing mechanisms or indicators in place to measure the proposals on their gender focus and sensitivity?
    a. If yes – could you detail them?
12. Do evaluations which fall into some particular clusters address gender more adequately than other clusters (e.g. Education/health?)?

Pillar 2 – Training

13. Do you feel the workshops are reaching the intended audiences?
    a. If not, what groups are being missed (e.g. developing countries)
14. Do you think the content and form of the workshops is suitable? Why/why not?
15. Do you think the way workshops are organised (i.e. very much as an on-demand facility) is effective? Why/why not?
16. Do you feel that there is a balance in gender of participants involved in workshops?
    a. And the right balance in gender of workshop facilitators?

  b. To what extent do you feel that there are sufficient efforts to achieve a balance in participants'/facilitators' gender?
  c. Do you feel that it is important that these efforts are made? Why?
  d. To what extent do you feel that workshop materials provided are tailored to address the needs of both males and females participants?
  e. To what extent do capacity building workshops prioritise/pay attention to gender in their content?

Pillar 3 –Evidence-to-policy

(where relevant)

17. Have you read E2P notes before? And do you find them useful?
    a. What would make an E2P note useful to you?
18. What criteria should be used to determine its quality and utility?
19. Are you aware of any comparable publications that seek to disseminate impact evaluation findings?
20. Is there another kind of output that you think would better communicate the research findings?
21. Do you think the E2P notes reach policy makers? If so, how? If not, why not?
22. To what extent do you think the evidence-to-policy notes show sensitivities and awareness of gender?
    a. Prompt: data disaggregated by gender
23. Do you have any thoughts on how you could improve the policy notes to consider gender more fully?

VfM

(where relevant)

24. What is the unit cost of impact evaluations funded under your programme? Would you be able to share funding data with us?
25. Could you comment on the following cost categories:
    a. Management fees (i.e. fee for managing and disbursing donor funds)
    b. Administrative costs
26. Are there other non-monetary cost categories you would like to highlight?
    c. E.g. in-kind contributions from within your organisation/partner organisations…

Appendix C: Survey questions (successful and unsuccessful)

Survey of successful applicants

SIEF Mid Term Review: Survey of fundees

SIEF Mid Term Review: Survey of fundees

Page 1

Department for International Development

UKaid from the British people

RAND EUROPE

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Welcome to the SIEF Mid Term Review Survey

Thank you very much for taking the time to complete our survey. This questionnaire aims to inform the ongoing Mid-Term Review of the Strategic Impact Evaluation Fund. The objective of the review is to determine whether SIEF is on course to achieve its objectives and whether the programme represents good value.

All responses will be treated confidentially and stored by RAND Europe. The identity of responders will be known only to the members of the independent RAND Europe research team and will not be shared with SIEF/DFID. No quotes or comments will be attributed to anybody unless expressly permitted by respondents.

If you have submitted multiple applications to SIEF, please refer to your earliest successful application when answering the survey questions.

SIEF Mid Term Review: Survey of fundees

Page 2

Background

1. Did any of the following collaborate on this successful bid?

Please select all that apply.

- ☐ Academic institutions
- ☐ Non-governmental organisations
- ☐ Non-affiliated experts
- ☐ Other, please specify

2. How did you hear about SIEF funding?

https://smapp2.rand.org/surv4/Print.aspx?SurveyID=8lKH4641&Title=Y&Breaks=Y&AllPages=Y&Pages=[13/05/2015 12:04:27]

- ○ Through a colleague
- ○ Through a previous grant
- ○ General web search
- ○ Online announcement (web or email)
- ○ Other, please specify

3. Did you apply for any other funding in parallel with SIEF or in the course of your SIEF grant?*
   - ○ Yes, successfully
   - ○ Yes, unsuccessfully
   - ○ No
   - ○ Prefer not to say

4. How likely do you think that you would be able to finance your impact evaluation from other funding sources?
   - ○ Very Likely
   - ○ Likely
   - ○ Unlikely
   - ○ Very Unlikely
   - ○ Don't know

SIEF Mid Term Review: Survey of fundees

Page 3

Why SIEF only

5. What were your reasons for applying for SIEF funding only?*

   Please select all that apply
   - ☐ SIEF was the only identified potential funder that was receiving applications
   - ☐ SIEF's application was seen as most acceptable in terms of preparation costs
   - ☐ Perceived chances of receiving SIEF funding were the highest of all potential funders
   - ☐ Lack of capacity to prepare multiple funding applications and/or manage multiple grants
   - ☐ Prefer not to say
   - ☐ Other, please specify

SIEF Mid Term Review: Survey of fundees

Page 4

Other funding and applications

6. Approximately what percentage of your project does SIEF funding support?

- ◯ Less than 20%
- ◯ 20-39%
- ◯ 40-59%
- ◯ 60-79%
- ◯ 80-100%

7. Who did you receive complementary funding from?

8. Where did you apply for complementary funding?

SIEF Mid Term Review: Survey of fundees

Page 5

Anticipated benefits of SIEF

9. When you applied to SIEF, were there any potential benefits that you considered beyond receiving financing from SIEF?

Please select up to three most important benefits.

- ☐ Technical support with proposal development
- ☐ Technical support with project execution
- ☐ Capacity building
- ☐ Access to in-country contacts and partners
- ☐ Access to (in-country) policy makers
- ☐ Post-project communication and dissemination
- ☐ None
- ☐ Other, please specify

10. What do you think would have happened to your project if you had not been granted SIEF funding?

SIEF Mid Term Review: Survey of fundees

13. When you interact or consult with the core SIEF team, what topics or issues are addressed?
Please select all that apply.
- ☐ Finances of the project
- ☐ Milestones of the project
- ☐ Technical/research problems experienced
- ☐ Access to in-country contacts or partners
- ☐ Access to (in-country) policy makers
- ☐ Communication and dissemination
- ☐ Other, please specify

14. For the options selected in the question above, how would you rate the support provided by SIEF?

	Very useful	Useful	Not very useful	Not useful at all	Not applicable
Finances of the project	○	○	○	○	○
Milestones of the project	○	○	○	○	○
Technical/research problems experienced	○	○	○	○	○
Access to in-country contacts or partners	○	○	○	○	○
Access to (in-country) policy makers	○	○	○	○	○
Other [as specified above]	○	○	○	○	○

15. Compared with other non-World Bank funders where applicable, how strongly do you agree with the following statements about the project process?

	Strongly Agree	Agree	Disagree	Strongly Disagree	Not applicable
The cost and burden associated with reporting requirements are reasonable when compared with the potential benefits	○	○	○	○	○
SIEF is open to consultation about any problems encountered in the project	○	○	○	○	○
SIEF provides greater access to in-country contacts or partner	○	○	○	○	○
SIEF provides greater access to (in-country) policy makers	○	○	○	○	○
SIEF makes use country offices and contacts to support the impact evaluation	○	○	○	○	○
SIEF offers capacity building activities (e.g. workshops, measurement support seminars) that are a valuable resource to running the impact evaluation	○	○	○	○	○

https://smapp2.rand.org/surv4/Print.aspx?SurveyID=8lKH4641&Title=Y&Breaks=Y&AllPages=Y&Pages=[13/05/2015 12:04:27]

16. Do you expect SIEF to enable you to communicate and disseminate the results of the project to the following audiences?

	Yes	No
Academia	○	○
Practitioners and NGOs in the study-country	○	○
Policy makers in the study-country	○	○
Policy makers in donor countries	○	○
International authorities (e.g. UN, WHO)	○	○

17. Do your interactions with SIEF differ from interactions with other funders? If so, how?

SIEF Mid Term Review: Survey of fundees

Page 8

Conducting the impact evaluation

18. What, if any, are the major challenges encountered in conducting the impact evaluation (i.e. challenges that, if unresolved, would compromise the validity of the entire impact evaluation)?

Please select all that apply.

- ☐ Financial challenges
- ☐ Research design
- ☐ Data collection
- ☐ Recruitment of participants
- ☐ Engagement of local partners
- ☐ Dissemination and communication of results
- ☐ None to date
- ☐ Other, please specify

19. How does SIEF help you at the moment to address the challenges?

Please tick all that apply.

- ☐ Open to consultation
- ☐ Providing flexibility around project implementation
- ☐ Providing technical advice
- ☐ Providing in-country contacts and partners
- ☐ Providing access to (in-country) policymakers

- ☐ Providing access to World Bank offices in study countries
- ☐ Other, please specify

20. In what areas can SIEF core team support activities be improved?

Please tick up to three most important items.

- ☐ Peer-review comments to help prepare for publication in a journal
- ☐ Technical support on IE design and implementation
- ☐ Providing access to in-country contacts or partners
- ☐ Providing access to (in-country) policy makers
- ☐ Other, please specify

SIEF Mid Term Review: Survey of fundees

Page 9

Added value

21. What added value, if any, has been generated through your participation in the application process?

Please select all that apply.

- ☐ None
- ☐ Improved knowledge/change in approach about preparing funding applications
- ☐ Improved knowledge/change in approach about designing research projects
- ☐ New contacts/collaborations made through the application
- ☐ Other, please specify

22. If any added value has been generated through your participation in the application process, could you elaborate on if and how you can use it in the future?

23. What do you expect the impact of your evaluation to be? Who do you think will consider its findings and for what purpose?

SIEF Mid Term Review: Survey of fundees

Page 10

Final comments

24. Do you have any other comments you would like to share?

Survey of unsuccessful applicants

SIEF Mid Term Review: Survey of applicants

Page 1

Department for International Development | UKaid from the British people | RAND EUROPE

THE WORLD BANK IBRD • IDA | WORLD BANK GROUP

Welcome to the SIEF Mid Term Review Survey

Thank you very much for taking the time to complete our survey. This questionnaire aims to inform the ongoing Mid-Term Review of the Strategic Impact Evaluation Fund. The objective of the review is to determine whether SIEF is on course to achieve its objectives and whether the programme represents good value.

All responses will be treated confidentially and stored by RAND Europe. The identity of responders will be known only to the members of the independent RAND Europe research team and will not be shared with SIEF/DFID. No quotes or comments will be attributed to anybody unless expressly permitted by respondents.

If you have submitted multiple applications for SIEF funding, please refer to your earliest application when answering the survey questions.

SIEF Mid Term Review: Survey of applicants

Page 2

Background

1. Did any of the following collaborate on your bid?

Please select all that apply.

- ☐ Academic institutions
- ☐ Non-governmental organisations
- ☐ Non-affiliated experts
- ☐ Other, please specify

2. How did you hear about SIEF funding?

- ○ Through a colleague
- ○ Through a previous grant
- ○ General web search
- ○ Online announcement (web or email)
- ○ Other, please specify

3. Did you apply for any other funding in parallel with SIEF?*

- ○ Yes, successfully
- ○ Yes, unsuccessfully
- ○ No
- ○ Prefer not to say

SIEF Mid Term Review: Survey of applicants

Page 3

Why SIEF only

4. What were your reasons for applying for SIEF funding only?*

Please select all that apply

- ☐ SIEF was the only identified potential funder that was receiving applications
- ☐ SIEF's application was seen as most acceptable in terms of preparation costs
- ☐ Perceived chances of receiving SIEF funding were the highest of all potential funders
- ☐ Lack of capacity to prepare multiple funding applications and/or manage multiple grants
- ☐ Prefer not to say
- ☐ Other, please specify

SIEF Mid Term Review: Survey of applicants

Page 4

Other funding and applications

5. Where did you apply for complementary funding?

SIEF Mid Term Review: Survey of applicants

Page 5

Anticipated benefits of SIEF

6. When you applied to SIEF, were there any potential benefits that you considered beyond receiving financing from SIEF?

Please select up to three most important benefits.

- ☐ Technical support with proposal development
- ☐ Technical support with project execution
- ☐ Capacity building
- ☐ Access to in-country contacts and partners
- ☐ Access to (in-country) policy makers
- ☐ Post-project communication and dissemination
- ☐ None
- ☐ Other, please specify

7. What happened to your project after your SIEF application?

- ○ Applied for funding from elsewhere
- ○ Applied for the next round of SIEF funding
- ○ The project was abandoned
- ○ Other, please specify

SIEF Mid Term Review: Survey of applicants

Page 6

Application process

8. How strongly do you agree with the following statements about the application process?

	Strongly agree	Agree	Disagree	Strongly disagree
The steps and requirements of the application process were clear from the start	○	○	○	○
The application was useful in my preparation for the project	○	○	○	○
The outcome of the application was communicated in a timely fashion	○	○	○	○

The decision and reasons underlying the outcome were transparent	○	○	○	○
The cost and burden associated with meeting application requirements were reasonable	○	○	○	○
The application assessment and decision were fair	○	○	○	○

SIEF Mid Term Review: Survey of applicants

Page 7

Added value

9. What added value, if any, has been generated through your participation in the application process?

Please select all that apply.

- ☐ None
- ☐ Improved knowledge/change in approach about preparing funding applications
- ☐ Improved knowledge/change in approach about designing research projects
- ☐ New contacts/collaborations made through the application
- ☐ Other, please specify

10. If any added value has been generated through your participation in the application process, could you elaborate on if and how you can use it in the future?

SIEF Mid Term Review: Survey of applicants

Page 8

Final comments

11. Do you have any other comments you would like to share?

Appendix D: Detailed survey methodology

The survey was designed to capture views on SIEF as a funder, on the process of application and selection and on the running of the grants. The survey of unsuccessful applicants involved a shortened version of the survey of applicants, which covers the following sections:

- Background
- SIEF and other funders: includes questions to identify what the unique aspects of SIEF are and why researchers apply to SIEF (i.e. anticipated benefits)
- Application process
- Running of the grant: questions on the interactions with SIEF over the course of the grant so far
- Conducting the IE: questions on the challenges encountered in conducting the IE
- Added value: questions on the value that participation in the application and selection process of SIEF may have generated

The survey instrument was developed by the research team based on their previous work evaluating research funding instruments and their understanding of SIEF, stemming from document review and inception interviews with SIEF staff. It was further modified to reflect suggestions from SIEF and DFID staff.

Participants

Invitees to the survey included all task team leaders (TTLs) and principal investigators (PIs) named on applications to the first three SIEF calls for funding, as provided by the SIEF core team. Given the fact that different sections of the survey were going to be relevant to different respondents depending on the outcome of their application and given that only the selection for seed funding in Call 3 was known at the time of the survey's launch, three different surveys were prepared:

- Survey of grantees: for successful applicants in Call 1 and 2
- Survey of applicants: for unsuccessful applicants in Calls 1, 2 and 3
- Survey of Call 3 participants: for seed-funded applicants in Call 3

Since some individuals submitted multiple applications (both within and across calls), duplicates were identified and the following approach was taken:

- All duplicates were removed so that each individual received only one invitation to participate.
- Those who had at least one successful application among their submissions were invited to participate in the *Survey of grantees* and were asked to refer to their earliest successful application (with a view to maximize the duration of the reference period).
- Those who had multiple unsuccessful applications among their submissions were invited to participate in the *Survey of applicants* and were asked to refer to their earliest application (with a view to maximize the duration of the reference period).
- Therefore, *Survey of Call 3 participants* invited only seed-funded individuals who had not submitted a successful application to any previous SIEF call for funding.

There was no difference between TTLs and PIs in the survey instruments and invitations.

Implementation

All three surveys were launched on April 1, 2015. At one-week intervals, two reminders were sent to those who had not responded yet and who had not contacted the research team to excuse themselves. A final appeal was sent on April 23, 2015, and the survey was closed on April 27, 2015.

Response rates

Response rates to each survey are reported in the table below, both combined and disaggregated by category of respondent.

Table 37. Survey response rates

	Grantees	Applicants	Call 3 participants
Number of unique contacts	79	357	10
Number of excused respondents*	0	12	2
Effective number of invitees	79	345	8
Complete responses	35 (44%)	81 (23%)	2 (25%)
Incomplete responses	21 (27%)	79 (23%)	0 (0%)
Responses included in analysis	**37 (47%)**	**89 (26%)**	**2 (25%)**
Of whom TTLs	*17*	*37*	*1*
Of whom PIs	*20*	*52*	*1*

Note: *Refers to invited participants who contacted the research team feeling they were not in a position to answer the survey questions, for instance, due to their limited involvement with SIEF.

Appendix E: Focus group protocol

Section 1: Wednesday

Opening questions

1. [Poll] Have you had many interactions with SIEF?
    a. Yes
    b. Only occasionally
    c. No
2. [Open] What do you think are the main objectives of SIEF? (If you do not know, please answer 'don't know')
    a. [Open text]
3. [Poll] Is SIEF on track to achieve these objectives?
    a. Yes
    b. No
    c. Don't know
4. [Open] Why do you think that?
    a. [Open text]

Pillar 1 – Application and Selection

5. [Poll] We are interested in your perspective on SIEF, irrespective of your knowledge of SIEF. It would be helpful, however, if you could indicate how confident you feel that you know and understand SIEF?
    a. Very confident
    b. Confident
    c. Not confident
    d. Not at all confident
6. [Poll] Do you perceive the current application and selection to be efficient? (i.e. is it quick and thorough without requiring unnecessary effort?)
    a. Yes
    b. No
    c. Don't know
7. [Poll] Do you feel the current application and selection is able to distinguish strong from weaker applications?
    a. Yes
    b. No
    c. Don't know
8. [Poll] Do you think the current application and selection is objective (i.e. that it does not prefer some candidates over others irrespective of the quality of the proposal)?
    a. Yes
    b. No
    c. Don't know

9. [Open] Overall, could you indicate why you think SIEF is an efficient and objective funder of impact evaluations or not? And do you think differences exist by sector (e.g. nutrition versus education)?
    a. [Open text]

Pillar 1 – SIEF's reach

10. [Poll] Is SIEF successful at reaching researchers globally, beyond the existing recognized clusters of researchers?
    a. Yes
    b. No
    c. Don't know
11. [Poll] Is SIEF recognised as a valued partner for funding and training in impact evaluation?
    a. Yes
    b. No
    c. Don't know
12. [Poll] Does SIEF contribute to capacity building in impact evaluation in the Global South?
    a. Yes
    b. No
    c. Don't know
13. [Open] Can you clarify and illustrate your answers?
    a. [open text]

Section 2: Friday

Pillar 2 – Training

14. [Poll] Are you familiar with training provided by SIEF?
    a. Yes
    b. No
15. [Poll] Do you feel the workshops are reaching the intended audiences?
    b. Yes
    c. No
    d. Don't know
16. [Poll] Do you think the content and form of the workshops is suitable?
    a. Yes
    b. No
    c. Don't know
17. [Poll] Do you think the way workshops are organised (i.e. very much as an on-demand facility) is effective?
    a. Yes
    b. No
    c. Don't know
18. [Open] Can you clarify and illustrate your answers?
    a. [open text]

Pillar 3 – Evidence-to-policy

19. [Poll] Do you find the evidence developed through SIEF grants to be useful in the design and implementation of programmes and intervention?
    a. Yes
    b. No
    c. Don't know
20. [Poll] Do you find the communication material developed by SIEF, such as the evidence-to-policy notes, to be useful in the design and implementation of programmes and intervention?
    a. Yes
    b. No
    c. Don't know
21. [Open] Can you clarify and illustrate your answers?
    a. [open text]

Gender

22. [Poll] To what extent is gender an important part of the work that SIEF does?
    a. To a great extent
    b. To a small extent
23. [Open] Can you clarify and illustrate your answer?
    a. [open text]

Reaching policy makers

24. [Poll] Do you think SIEF will be able to reach policy makers and practitioners to generate lasting change?
    a. Yes
    b. No
25. Why do you think that?
    a. [open text]

Appendix F: Key risks from the business case

Risk	Grading	Mitigation
SIEF is used for non-strategic evaluations that offer limited value to DFID or globally. The World Bank has had a track record of focussing on Latin America and there are significantly more success stories and evidence coming from Latin America than LICs. There is a risk that the impact evaluation financed from SIEF will not focus on low income countries or the relevant MDGs. Out of the 50 impact evaluations under the Spanish Trust Fund, 76% were in DFID countries and 60% were in LICs	Medium likelihood and high impact.	DFID policy leads have worked with the Bank to define the priority clusters and evidence gaps, as well as the criteria that determines which impact evaluations will be funded. Furthermore, SIEF steering committee (which includes DFID participants) will approve the work programme and monitor that it is being implemented appropriately. It has been agreed that SIEF will only finance impact evaluations that can demonstrate replicability or scalability in a developing country (LIC setting).
The Initiative is used to substantiate pre-determined positions rather than for rigorous evidence-based approaches.	Low likelihood and high impact.	This is mitigated by having rigorous and substantial quality assurance processes in place at entry, during and exit of the impact evaluation so that the impact evaluations are genuinely independent research.
There is limited uptake of evidence from impact evaluations. It is more difficult to show evidence of how the findings of an impact evaluation have fed into policy development and changed existing practice.	Medium likelihood and high impact	Firstly, SIEF will develop a dissemination and communication strategy and undertake a number of dissemination activities to ensure that findings are well known and taken on board. DFID's involvement in the key governance and technical scrutiny structures for the SIEF and a more strategic focus on key policy relevant gaps in the evidence will also mitigate this risk.
Impact evaluations do not adopt a high level of rigour.	Low likelihood and high impact.	The Spanish Impact Evaluation Fund has a track record of delivering high quality products. In addition, World Bank will use their procurement systems and pre-determined criteria to select high quality bids for Impact Evaluations.
Impact Evaluations substitute for rigorous evaluations that should otherwise be taken up by individual World Bank or DFID programmes, or the capacity building components of SIEF are used to finance activities where benefits mainly accrue to World Bank or DFID.	Medium likelihood and low impact	The World Bank has committed through IDA16 to undertake 51 impact evaluations. Of these only 25% are in human development (and therefore eligible for SIEF funding). Additionally, the SIEF timeline implies that only 15% could in practice apply for SIEF funding. This risk can be mitigated by ensuring that the World Bank must meet its IDA commitment *in addition* to SIEF impact evaluations.

Appendix G: 10 questions to assess knowledge gain

Source: Answer Key English (SIEF)

4- When is the best time to design an impact evaluation?

a) Just after the program starts **1**
a) Just before the end of the program **2**
b) During program design **3**
c) Anytime if you have good experts **4**

5-The outcomes of a program are the products and services it delivers?

a) True **1**
b) False **2**

6- An impact evaluation measures the causal impact of a program by simply comparing the outcomes of its beneficiaries before and after receiving the program

a) True **1**
b) False **2**

7- In the context of an impact evaluation, what is the definition of a ‘counterfactual’?

a) The comparison group **1**
b) The impact of the program **2**
c) The situation of the beneficiaries in absence of the program **3**
d) The control group **4**

8- Which of the following methods does NOT provide an accurate measure of the causal effect of a program?

a) Randomized assignment **1**
b) Before and After comparison of beneficiaries **2**
c) Regression discontinuity design **3**
d) None of the above **4**

9- Which of the following represents ‘random assignment’?

a) Randomly selecting 10 households from a village to be surveyed **1**
b) A program manager selects 3 districts for the evaluation and does not reveal the criteria for selection **2**
c) A program that assigns benefits to beneficiaries based on a poverty index **3**
d) A program that uses a lottery to determine the beneficiaries **4**

10- Assuming you have budget for 2000 interviews, which option would you choose to increase maximise power?

a) Interview 20 households per village in 100 villages **1**

b) Interview 50 households per village in 40 villages **2**
c) Interview 500 households per village in 4 villages **3**

11- Which of the following is the main assumption associated with Difference in Difference estimates?

a) In the absence of the program, the treatment and control groups will experience the same trend in outcome indicators over time **1**
b) In the absence of the program, the treatment and control groups will experience different trends in outcome indicators over time **2**
c) Treatment and control groups experience different shocks that affect outcome indicators (rainfall, drought, etc.) **3**

12- Which of the following is a limitation associated with Regression Discontinuity Design method?

a) Identifies the impact of the intervention only on the beneficiaries near the discontinuity and not for the entire population **1**
b) Identifies the impact of the intervention only on the beneficiaries far from the discontinuity and not for the entire population **2**
c) Identifies the impact of the intervention far from the discontinuity **3**
d) Estimation requires a large sample size at the far end of the distribution **4**

13- Which of the following is an effective indicator for primary school enrolment?

a) Number of students enrolled in primary school **1**
b) Percentage of students attending the school **2**
c) Percentage of school age children **3**
d) Percentage of primary school age children that are enrolled in school **4**

Appendix H: Workshop data gender-disaggregated

Female	Ethiopia	Korea	India	Senegal	Bangladesh	Ethiopia	Korea	Peru	Total
Quality	4.5	4.4	4.3	4.3	3.9	4.4	4.1	4.4	4.3
Usefulness	4.6	4.4	4.1	4.5	4.0	4.6	4.4	4.4	4.4
Relevance to World Bank mission	4.3	4.7	4.3	3.9	3.7		4.2		4.2
Knowledge/skills increased	4.4	4.5	4.2	4.1	3.6	4.2	4.1	4.3	4.2
Knowledge/skills apply to work	4.3	4.3	4.1	4.2	3.9	4.3	4.0	4.1	4.1
Objective achieved	4.3	4.1	3.9	4.3	3.7	4.1	4.0	4.0	4.1
Learning needs fulfilled	4.2	4.2	3.8	4.0	3.4	3.9	3.8	3.8	3.9
Content/subject matter	4.6	4.5	4.2	4.3	4.2	4.6	4.4	4.6	4.4
Order content was presented	4.2	4.5	4.3	4.2	3.9	4.3	4.2	4.4	4.2
Materials	4.3	4.5	4.2	4.2	4.1	4.5	4.3	4.3	4.3
Participants	22	34	21	20	32	13	51	15	208

Male	Ethiopia	Korea	India	Senegal	Bangladesh	Ethiopia	Korea	Peru	Total
Quality	4.4	4.4	4.3	4.3	3.9	4.4	4.1	4.5	4.3
Usefulness	4.6	4.4	4.1	4.5	3.9	4.6	4.3	4.4	4.4
Relevance to World Bank mission	4.3	4.7	4.3	3.9	3.7		4.2		4.2
Knowledge/skills increased	4.4	4.4	4.2	4.1	3.6	4.2	4.1		4.2
Knowledge/skills apply to work	4.3	4.2	4.1	4.2	3.9	4.3	4.0	4.3	4.2
Objective achieved	4.3	4.1	4.0	4.3	3.7	4.1	4.0	4.1	4.1
Learning needs fulfilled	4.2	4.3	3.8	4.0	3.4	3.9	3.8	3.8	3.9
Content/subject matter	4.5	4.5	4.2	4.4	4.1	4.6	4.4	4.5	4.4
Order content was presented	4.2	4.5	4.3	4.2	3.9	4.3	4.2	4.4	4.2
Materials	4.3	4.5	4.2	4.3	4.0	4.5	4.3	4.3	4.3
Participants	47	25	28	51	57	39	54	36	337

Appendix I: Additional survey results

The following are additional tables for Chapter 4.

Table 38: Indicators of effectiveness (broken down by cluster)

Statement: The steps and requirement of the application process were clear from the start.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
ECD	16 (35%)	24 (52%)	6 (13%)	0 (0%)	40 (87%)	6 (13%)
EDU	22 (51%)	20 (47%)	1 (2%)	0 (0%)	42 (98%)	1 (2%)
Health	5 (36%)	6 (43%)	1 (7%)	2 (14%)	11 (79%)	3 (21%)
WASH	2 (13%)	11 (73%)	2 (13%)	0 (0%)	13 (87%)	2 (13%)
Total	43 (37%)	60 (52%)	11 (9%)	2 (2%)	103 (89%)	13 (11%)
Statement: The application was useful in our preparation for the project.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
ECD	16 (36%)	26 (58%)	3 (7%)	0 (0%)	42 (93%)	3 (7%)
EDU	19 (44%)	19 (44%)	5 (12%)	0 (0%)	38 (88%)	5 (12%)
Health	4 (29%)	7 (50%)	3 (21%)	0 (0%)	11 (79%)	3 (21%)
WASH	1 (7%)	9 (64%)	3 (21%)	1 (7%)	10 (71%)	4 (29%)
Total	40 (35%)	60 (52%)	14 (12%)	1 (1%)	100 (87%)	15 (13%)
Statement: The outcome of the application was communicated in a timely fashion.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
ECD	14 (31%)	25 (56%)	5 (11%)	1 (2%)	39 (87%)	6 (13%)
EDU	21 (49%)	18 (42%)	3 (7%)	1 (2%)	39 (91%)	4 (9%)
Health	4 (31%)	5 (38%)	2 (15%)	2 (15%)	9 (69%)	4 (31%)
WASH	1 (7%)	11 (79%)	2 (14%)	0 (0%)	12 (86%)	2 (14%)
Total	39 (34%)	55 (48%)	15 (13%)	5 (4%)	94 (82%)	20 (18%)
Statement: The decision and reasons underlying the outcome were transparent.						
	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
ECD	8 (18%)	22 (50%)	13 (30%)	2 (5%)	30 (68%)	15 (34%)
EDU	16 (37%)	12 (28%)	10 (23%)	5 (12%)	28 (65%)	15 (35%)
Health	2 (14%)	7 (50%)	3 (21%)	2 (14%)	9 (64%)	5 (36%)
WASH	1 (7%)	9 (64%)	3 (21%)	1 (7%)	10 (71%)	4 (29%)

Total	27 (23%)	49 (43%)	29 (25%)	10 (9%)	76 (66%)	39 (34%)

Statement: The cost and burden associated with meeting application requirements were reasonable.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
ECD	9 (20%)	32 (70%)	4 (9%)	1 (2%)	41 (89%)	5 (11%)
EDU	13 (30%)	21 (49%)	9 (21%)	0 (0%)	34 (79%)	9 (21%)
Health	1 (8%)	8 (62%)	2 (15%)	2 (15%)	9 (69%)	4 (31%)
WASH	0 (0%)	12 (80%)	2 (13%)	1 (7%)	12 (80%)	3 (20%)
Total	24 (21%)	72 (63%)	16 (14%)	3 (3%)	96 (83%)	19 (17%)

Statement: The application assessment and decision were fair.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
ECD	7 (17%)	25 (60%)	10 (24%)	0 (0%)	32 (76%)	10 (24%)
EDU	16 (38%)	17 (40%)	7 (17%)	2 (5%)	33 (79%)	9 (21%)
Health	2 (15%)	9 (69%)	0 (0%)	2 (15%)	11 (85%)	2 (15%)
WASH	1 (8%)	10 (83%)	0 (0%)	1 (8%)	11 (92%)	1 (8%)
Total	26 (24%)	60 (56%)	17 (16%)	5 (5%)	86 (80%)	22 (20%)

Table 39. Indicators of effectiveness (broken down by position)

Statement: The steps and requirement of the application process were clear from the start.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
PI	24 (35%)	35 (51%)	7 (10%)	2 (3%)	59 (87%)	9 (13%)
TTL	19 (40%)	25 (52%)	4 (8%)	0 (0%)	44 (92%)	4 (8%)
Total	43 (37%)	60 (52%)	11 (9%)	2 (2%)	103 (89%)	13 (11%)

Statement: The application was useful in our preparation for the project.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
PI	25 (37%)	34 (51%)	7 (10%)	1 (1%)	59 (88%)	8 (12%)
TTL	15 (31%)	26 (54%)	7 (15%)	0 (0%)	41 (85%)	7 (15%)
Total	40 (35%)	60 (52%)	14 (12%)	1 (1%)	100 (87%)	15 (13%)

Statement: The outcome of the application was communicated in a timely fashion.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
PI	22 (33%)	31 (47%)	9 (14%)	4 (6%)	53 (80%)	13 (20%)
TTL	17 (35%)	24 (50%)	6 (13%)	1 (2%)	41 (85%)	7 (15%)
Total	39 (34%)	55 (48%)	15 (13%)	5 (4%)	94 (82%)	20 (18%)

Statement: The decision and reasons underlying the outcome were transparent.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
PI	14 (21%)	24 (36%)	19 (28%)	10 (15%)	38 (57%)	29 (43%)
TTL	13 (27%)	25 (52%)	10 (21%)	0 (0%)	38 (79%)	10 (21%)

Total	27 (23%)	49 (43%)	29 (25%)	10 (9%)	76 (66%)	39 (34%)

Statement: The cost and burden associated with meeting application requirements were reasonable.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
PI	13 (19%)	44 (65%)	8 (12%)	3 (3%)	57 (84%)	11 (16%)
TTL	11 (23%)	28 (60%)	8 (17%)	0 (0%)	39 (83%)	8 (17%)
Total	24 (21%)	72 (63%)	16 (14%)	3 (3%)	96 (83%)	19 (17%)

Statement: The application assessment and decision were fair.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined
PI	13 (21%)	33 (53%)	11 (18%)	5 (8%)	46 (74%)	16 (26%)
TTL	13 (28%)	27 (59%)	6 (13%)	0 (0%)	40 (87%)	6 (13%)
Total	26 (24%)	60 (56%)	17 (16%)	5 (5%)	86 (80%)	22 (20%)

Table 40. Grantees' frequency of interaction with SIEF (broken down by position)

	Every 1–3 months	Every 3–6 months	Every 6–12 months	Less than once a year	More than once a month
PI	6 (33%)	3 (17%)	4 (22%)	4 (22%)	1 (6%)
TTL	6 (38%)	6 (38%)	2 (13%)	1 (6%)	1 (6%)
Total	12 (35%)	9 (26%)	6 (18%)	5 (15%)	2 (6%)

Table 41. Grantees' frequency of interaction with SIEF (broken down by cluster)

	Every 1–3 months	Every 3–6 months	Every 6–12 months	Less than once a year	More than once a month
ECD	3 (27%)	5 (45%)	1 (9%)	2 (18%)	0 (0%)
EDU	4 (36%)	2 (18%)	2 (18%)	2 (18%)	1 (9%)
Health	2 (29%)	3 (43%)	1 (14%)	0 (0%)	1 (14%)
WASH	3 (50%)	0 (0%)	2 (33%)	1 (17%)	0 (0%)
Total	12 (35%)	9 (26%)	6 (18%)	5 (15%)	2 (6%)

Table 42. Grantees' satisfaction with SIEF core team support by activity (broken down by position)

Finances of the project

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicabl e
PI	5 (26%)	9 (47%)	1 (5%)	1 (5%)	14 (74%)	2 (11%)	3 (16%)
TTL	4 (25%)	8 (50%)	3 (19%)	1 (6%)	12 (75%)	4 (25%)	0 (0%)
Total	9 (26%)	17 (49%)	4 (11%)	2 (6%)	26 (74%)	6 (17%)	3 (9%)

Milestones of the project

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicabl e
PI	5 (26%)	7 (37%)	0 (0%)	0 (0%)	12 (63%)	0 (0%)	7 (37%)
TTL	8 (50%)	6 (38%)	0 (0%)	0 (0%)	14 (88%)	0 (0%)	2 (13%)

Total	13 (37%)	13 (37%)	0 (0%)	0 (0%)	26 (74%)	0 (0%)	9 (26%)

Technical/research problems experienced

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
PI	5 (26%)	5 (26%)	1 (5%)	0 (0%)	10 (53%)	1 (5%)	8 (42%)
TTL	1 (7%)	3 (20%)	7 (47%)	1 (7%)	4 (27%)	8 (53%)	3 (20%)
Total	6 (18%)	8 (24%)	8 (24%)	1 (3%)	14 (41%)	9 (26%)	11 (32%)

Access to in-country contacts and partners

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
PI	6 (33%)	2 (11%)	2 (11%)	0 (0%)	8 (44%)	2 (11%)	8 (44%)
TTL	0 (0%)	3 (20%)	7 (47%)	1 (7%)	3 (20%)	8 (53%)	4 (27%)
Total	6 (18%)	5 (15%)	9 (27%)	1 (3%)	11 (33%)	10 (30%)	12 (36%)

Access to (in-country) policymakers

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
PI	7 (37%)	2 (11%)	1 (5%)	0 (0%)	9 (47%)	1 (5%)	9 (47%)
TTL	2 (13%)	1 (7%)	8 (53%)	1 (7%)	3 (20%)	9 (60%)	3 (20%)
Total	9 (26%)	3 (9%)	9 (26%)	1 (3%)	12 (35%)	10 (29%)	12 (35%)

Other

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
PI	6 (26%)	6 (26%)	1 (4%)	1 (4%)	12 (52%)	2 (9%)	9 (39%)
TTL	6 (38%)	7 (44%)	2 (13%)	0 (0%)	13 (81%)	2 (13%)	1 (6%)
Total	12 (31%)	13 (33%)	3 (8%)	1 (3%)	25 (64%)	4 (10%)	10 (26%)

Table 43. Grantees' satisfaction with SIEF core team support by activity (broken down by cluster)

Finances of the project

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
ECD	4 (36%)	3 (27%)	0 (0%)	1 (0%)	7 (64%)	1 (9%)	3 (27%)
EDU	3 (33%)	6 (67%)	0 (0%)	0 (0%)	9 (100%)	0 (0%)	0 (0%)
Health	3 (50%)	3 (50%)	0 (0%)	0 (0%)	6 (100%)	0 (0%)	0 (0%)
WASH	2 (40%)	2 (40%)	0 (0%)	0 (0%)	4 (80%)	0 (0%)	1 (20%)
Total	12 (39%)	14 (45%)	0 (0%)	1 (3%)	26 (84%)	1 (3%)	4 (13%)

Milestones of the project

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
ECD	2 (18%)	2 (18%)	2 (18%)	0 (0%)	4 (36%)	2 (18%)	5 (45%)

EDU	1 (13%)	5 (63%)	1 (13%)	1 (13%)	6 (75%)	2 (25%)	0 (0%)
Health	2 (29%)	4 (57%)	0 (0%)	0 (0%)	6 (86%)	0 (0%)	1 (14%)
WASH	2 (33%)	3 (50%)	0 (0%)	0 (0%)	5 (83%)	0 (0%)	1 (17%)
Total	7 (23%)	13 (42%)	3 (10%)	1 (3%)	20 (65%)	4 (13%)	7 (23%)

Technical/research problems experienced

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
ECD	3 (33%)	0 (0%)	0 (0%)	0 (0%)	3 (33%)	0 (0%)	6 (67%)
EDU	1 (10%)	5 (50%)	2 (20%)	0 (0%)	6 (60%)	2 (20%)	2 (20%)
Health	2 (40%)	1 (20%)	1 (20%)	0 (0%)	3 (60%)	1 (20%)	1 (20%)
WASH	2 (50%)	0 (0%)	1 (25%)	0 (0%)	2 (50%)	1 (25%)	1 (25%)
Total	8 (29%)	6 (21%)	4 (14%)	0 (0%)	14 (50%)	4 (14%)	10 (36%)

Access to in-country contacts and partners

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
ECD	2 (22%)	0 (0%)	1 (11%)	0 (0%)	2 (22%)	1 (11%)	6 (67%)
EDU	0 (0%)	1 (13%)	4 (50%)	0 (0%)	1 (13%)	4 (50%)	3 (38%)
Health	1 (25%)	1 (25%)	1 (25%)	0 (0%)	2 (50%)	1 (25%)	1 (25%)
WASH	0 (0%)	0 (0%)	1 (25%)	0 ()%)	0 (0%)	1 (25%)	3 (75%)
Total	3 (12%)	2 (8%)	7 (28%)	0 (0%)	5 (20%)	7 (28%)	13 (52%)

Access to (in-country) policymakers

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
ECD	2 (22%)	0 (0%)	1 (11%)	0 (0%)	2 (22%)	1 (11%)	6 (67%)
EDU	0 (0%)	2 (25%)	3 (38%)	0 (0%)	2 (25%)	3 (38%)	3 (38%)
Health	1 (25%)	1 (25%)	1 (25%)	0 (0%)	2 (50%)	1 (25%)	1 (25%)
WASH	0 (0%)	0 (0%)	1 (25%)	0 (0%)	0 (0%)	1 (25%)	3 (75%)
Total	3 (12%)	3 (12%)	6 (24%)	0 (0%)	6 (24%)	6 (24%)	13 (52%)

Other

	Very useful	Useful	Not very useful	Not useful at all	Agree combined	Disagree combined	Not applicable
ECD	1 (14%)	3 (43%)	0 (0%)	0 (0%)	4 (57%)	0 (0%)	3 (43%)
EDU	1 (14%)	1 (14%)	1 (14%)	0 (0%)	2 (29%)	1 (14%)	4 (57%)
Health	0 (0%)	1 (100%)	0 (0%)	0 (0%)	1 (100%)	0 (0%)	0 (0%)
WASH	0 (0%)	2 (67%)	0 (0%)	0 (0%)	2 (67%)	0 (0%)	1 (33%)
Total	2 (12%)	6 (35%)	1 (6%)	0 (0%)	8 (47%)	1 (6%)	8 (47%)

Table 44: Compared with other non–World Bank funders where applicable, how strongly do you agree or disagree with the following statements about the project process? (broken down by position)

The cost and burden associated with reporting requirements are reasonable when compared with the potential benefits.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
PI	5 (26%)	9 (47%)	1 (5%)	1 (5%)	14 (74%)	2 (11%)	3 (16%)
TTL	4 (25%)	8 (50%)	3 (19%)	1 (6%)	12 (75%)	4 (25%)	0 (0%)
Total	9 (26%)	17 (49%)	4 (11%)	2 (6%)	26 (74%)	6 (17%)	3 (9%)

SIEF is open to consultation about any problems encountered in the project.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
PI	5 (26%)	7 (37%)	0 (0%)	0 (0%)	12 (63%)	0 (0%)	7 (37%)
TTL	8 (50%)	6 (38%)	0 (0%)	0 (0%)	14 (88%)	0 (0%)	2 (13%)
Total	13 (37%)	13 (37%)	0 (0%)	0 (0%)	26 (74%)	0 (0%)	9 (26%)

SIEF provides greater access to in-country contacts or partners.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
PI	5 (26%)	5 (26%)	1 (5%)	0 (0%)	10 (53%)	1 (5%)	8 (42%)
TTL	1 (7%)	3 (20%)	7 (47%)	1 (7%)	4 (27%)	8 (53%)	3 (20%)
Total	6 (18%)	8 (24%)	8 (24%)	1 (3%)	14 (41%)	9 (26%)	11 (32%)

SIEF provides greater access to (in-country) policymakers.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
PI	6 (33%)	2 (11%)	2 (11%)	0 (0%)	8 (44%)	2 (11%)	8 (44%)
TTL	0 (0%)	3 (20%)	7 (47%)	1 (7%)	3 (20%)	8 (53%)	4 (27%)
Total	6 (18%)	5 (15%)	9 (27%)	1 (3%)	11 (33%)	10 (30%)	12 (36%)

SIEF makes use of country offices and contacts to support the impact evaluation.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
PI	7 (37%)	2 (11%)	1 (5%)	0 (0%)	9 (47%)	1 (5%)	9 (47%)
TTL	2 (13%)	1 (7%)	8 (53%)	1 (7%)	3 (20%)	9 (60%)	3 (20%)
Total	9 (26%)	3 (9%)	9 (26%)	1 (3%)	12 (35%)	10 (29%)	12 (35%)

SIEF offers capacity-building activities (e.g. workshops, measurement support seminars) that are a valuable resource to running the impact evaluation.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
PI	6 (26%)	6 (26%)	1 (4%)	1 (4%)	12 (52%)	2 (9%)	9 (39%)
TTL	6 (38%)	7 (44%)	2 (13%)	0 (0%)	13 (81%)	2 (13%)	1 (6%)
Total	12 (31%)	13 (33%)	3 (8%)	1 (3%)	25 (64%)	4 (10%)	10 (26%)

Table 45: Compared with other non–World Bank funders where applicable, how strongly do you agree with the following statements about the project process? (broken down by cluster)

The cost and burden associated with reporting requirements are reasonable when compared with the potential benefits.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable

ECD	3 (25%)	4 (33%)	2 (17%)	0 (0%)	7 (58%)	2 (17%)	3 (25%)
EDU	4 (36%)	7 (64%)	0 (0%)	0 (0%)	11 (100%)	0 (0%)	0 (0%)
Health	2 (29%)	2 (29%)	1 (14%)	2 (29%)	4 (57%)	3 (43%)	0 (0%)
WASH	0 (0%)	5 (83%)	1 (17%)	0 (0%)	5 (83%)	1 (17%)	0 (0%)
Total	9 (26%)	17 (49%)	4 (11%)	2 (6%)	26 (74%)	6 (17%)	3 (9%)

SIEF is open to consultation about any problems encountered in the project.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
ECD	5 (42%)	2 (17%)	0 (0%)	0 (0%)	7 (58%)	0 (0%)	5 (42%)
EDU	4 (36%)	5 (45%)	0 (0%)	0 (0%)	9 (82%)	0 (0%)	2 (18%)
Health	2 (29%)	4 (57%)	0 (0%)	0 (0%)	6 (86%)	0 (0%)	1 (14%)
WASH	2 (33%)	3 (50%)	0 (0%)	0 (0%)	5 (83%)	0 (0%)	1 (17%)
Total	13 (37%)	13 (37%)	0 (0%)	0 (0%)	26 (74%)	0 (0%)	9 (26%)

SIEF provides greater access to in-country contacts or partners.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
ECD	3 (27%)	2 (18%)	1 (9%)	0 (0%)	5 (45%)	1 (9%)	5 (45%)
EDU	1 (9%)	4 (36%)	2 (18%)	0 (0%)	5 (45%)	2 (18%)	4 (36%)
Health	2 (29%)	1 (14%)	2 (29%)	0 (0%)	3 (43%)	2 (29%)	2 (29%)
WASH	0 (0%)	1 (17%)	3 (50%)	1 (17%)	1 (17%)	4 (67%)	1 (17%)
Total	6 (18%)	8 (24%)	8 (24%)	1 (3%)	14 (41%)	9 (26%)	11 (32%)

SIEF provides greater access to (in-country) policymakers.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
ECD	2 (18%)	2 (18%)	1 (9%)	0 (0%)	4 (36%)	1 (9%)	6 (55%)
EDU	2 (20%)	1 (10%)	3 (30%)	0 (0%)	3 (30%)	3 (30%)	4 (40%)
Health	2 (29%)	1 (14%)	2 (29%)	0 (0%)	3 (43%)	2 (29%)	2 (29%)
WASH	0 (0%)	1 (17%)	3 (50%)	1 (17%)	1 (17%)	4 (67%)	1 (17%)
Total	6 (18%)	5 (15%)	9 (27%)	1 (3%)	11 (33%)	10 (30%)	12 (36%)

SIEF makes use of country offices and contacts to support the impact evaluation.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
ECD	4 (36%)	0 (0%)	2 (18%)	0 (0%)	4 (36%)	2 (18%)	5 (45%)
EDU	3 (27%)	1 (9%)	2 (18%)	0 (0%)	4 (36%)	2 (18%)	5 (45%)
Health	2 (29%)	1 (14%)	2 (29%)	0 (0%)	3 (43%)	2 (29%)	2 (29%)
WASH	0 (0%)	1 (17%)	3 (50%)	1 (17%)	1 (17%)	4 (67%)	1 (17%)
Total	9 (26%)	3 (9%)	9 (26%)	1 (3%)	12 (35%)	10 (29%)	12 (35%)

SIEF offers capacity-building activities (e.g. workshops, measurement support seminars) that are a valuable resource to running the impact evaluation.

	Strongly agree	Agree	Disagree	Strongly disagree	Agree combined	Disagree combined	Not applicable
ECD	6 (50%)	3 (25%)	0 (0%)	0 (0%)	9 (75%)	0 (0%)	3 (25%)
EDU	4 (36%)	3 (27%)	1 (9%)	1 (9%)	7 (64%)	2 (18%)	2 (18%)
Health	2 (29%)	4 (57%)	1 (14%)	0 (0%)	6 (86%)	1 (14%)	0 (0%)
WASH	0 (0%)	4 (67%)	1 (17%)	0 (0%)	4 (67%)	1 (17%)	1 (17%)

Total	12 (34%)	13 (37%)	3 (9%)	1 (3%)	25 (71%)	4 (11%)	6 (17%)

Table 46. How does SIEF help overcome arisen challenges? (broken down by position)

	Open consultation	Flexibility	Technical advice	In-country contacts	Access to policymakers	Access to WB country offices
PI	10 (50%)	7 (35%)	6 (30%)	0 (0%)	1 (5%)	3 (15%)
TTL	10 (59%)	6 (35%)	5 (29%)	1 (6%)	1 (6%)	1 (6%)
Total	20 (54%)	13 (35%)	11 (55%)	1 (3%)	2 (5%)	4 (11%)

Table 47. How does SIEF help overcome arisen challenges? (broken down by cluster)

	Open consultation	Flexibility	Technical advice	In-country contacts	Access to policymakers	Access to WB country offices
ECD	8 (62%)	4 (31%)	2 (15%)	0 (0%)	0 (0%)	1 (8%)
EDU	5 (42%)	5 (42%)	5 (42%)	1 (8%)	2 (17%)	3 (25%)
Health	4 (57%)	3 (43%)	1 (14%)	0 (0%)	0 (0%)	0 (0%)
WASH	4 (67%)	2 (33%)	3 (50%)	0 (0%)	0 (0%)	0 (0%)
Total	20 (54%)	13 (35%)	11 (30%)	1 (3%)	2 (5%)	4 (11%)

Table 48. Areas of support that could be improved (broken down by position)

	Peer review	Tech support	Access to partners	Access to policymakers
PI	5 (25%)	1 (5%)	5 (25%)	5 (25%)
TTL	6 (35%)	5 (29%)	6 (35%)	3 (18%)
Total	11 (30%)	6 (16%)	11 (55%)	8 (22%)

Table 49. Areas of support that could be improved (broken down by cluster)

	Peer review	Tech support	Access to partners	Access to policymakers
ECD	3 (23%)	1 (8%)	3 (23%)	3 (23%)
EDU	4 (33%)	1 (8%)	6 (50%)	4 (33%)
Health	2 (29%)	1 (14%)	2 (29%)	2 (29%)
WASH	2 (33%)	3 (50%)	1 (17%)	0 (0%)
Total	11 (30%)	6 (16%)	11 (30%)	8 (22%)

Table 50. Leverage broken down by position

	None	Improved knowledge/change in approach about preparing funding applications	Improved knowledge/change in approach about designing research projects	New contacts/collaborations made through the application	Other
PI	18 (25%)	30 (41%)	21 (29%)	28 (38%)	4 (5%)
TTL	8 (15%)	21 (38%)	20 (36%)	15 (27%)	6 (11%)

Total	26 (20%)	51 (40%)	41 (32%)	43 (34%)	10 (8%)

Table 51: Leverage broken down by cluster

	None	Improved knowledge/change in approach about preparing funding applications	Improved knowledge/change in approach about designing research projects	New contacts/collaborations made through the application	Other
ECD	10 (20%)	22 (45%)	16 (33%)	18 (37%)	6 (12%)
EDU	8 (17%)	22 (48%)	16 (35%)	20 (43%)	2 (4%)
Health	5 (28%)	5 (28%)	3 (17%)	4 (22%)	1 (6%)
WASH	4 (25%)	3 (19%)	6 (38%)	2 (13%)	1 (6%)
Total	28 (22%)	51 (41%)	41 (33%)	43 (34%)	10 (8%)

Table 52. Topics covered during consultations with SIEF (broken down by position)

	Finances	Milestones	Problems	Partners	Policymakers	Dissemination	Other
PI	13 (65%)	10 (50%)	4 (20%)	1 (5%)	2 (10%)	5 (25%)	3 (15%)
TTL	12 (71%)	11 (65%)	8 (47%)	2 (12%)	1 (6%)	8 (47%)	2 (12%)
Total	25 (68%)	21 (57%)	12 (32%)	3 (8%)	3 (8%)	13 (35%)	5 (14%)

Table 53. Topics covered during consultations with SIEF (broken down by cluster)

	Finances	Milestones	Problems	Partners	Policymakers	Dissemination	Other
ECD	8 (67%)	6 (50%)	2 (17%)	2 (17%)	1 (8%)	7 (58%)	3 (25%)
EDU	8 (73%)	6 (55%)	6 (55%)	1 (9%)	1 (9%))	3 (27%)	0 (0%)
Health	5 (71%)	6 (86%)	3 (43%)	0 (0%)	1 (14%)	2 (29%)	1 (14%)
WASH	4 (67%)	4 (67%)	1 (17%)	0 (0%)	0 (0%)	2 (33%)	1 (17%)
Total	25 (71%)	21 (60%)	12 (34%)	3 (9%)	3 (9%)	13 (37%)	5 (14%)

The following are additional tables for Chapter 6.

Table 54. Grantees' expectations of communication assistance (broken down by position)

	Academia	NGOs in study countries	Policymakers in study countries	Policymakers in donor countries	International authorities
PI	6 (33%)	16 (84%)	17 (89%)	14 (82%)	15 (79%)
TTL	9 (69%)	10 (63%)	9 (56%)	13 (81%)	11 (73%)
Total	15 (48%)	26 (74%)	26 (74%)	27 (82%)	26 (76%)

Table 55. Grantees' expectations of communication assistance (broken down by cluster)

	Academia	NGOs in study countries	Policymakers in study countries	Policymakers in donor countries	International authorities
ECD	4 (36%)	10 (83%)	11 (92%)	10 (91%)	10 (83%)
EDU	4 (40%)	8 (73%)	8 (73%)	9 (82%)	9 (82%)
Health	4 (67%)	5 (71%)	5 (71%)	5 (71%)	5 (83%)
WASH	3 (60%)	4 (67%)	3 (50%)	4 (80%)	3 (50%)
Total	15 (48%)	26 (74%)	26 (74%)	27 (82%)	26 (76%)

The following are additional tables for Chapter 8.

Table 56: Expected benefits of SIEF application broken down by cluster

	Technical support with proposal development	Technical support with project execution	Capacity building	Access to in-country contacts and partners	Access to (in-country) policy-makers	Post-project communi-cation and dissemination	None	Other
ECD	10 (20%)	10 (20%)	19 (39%)	13 (27%)	14 (29%)	26 (53%)	8 (16%)	6 (12%)
EDU	12 (26%)	12 (26%)	13 (28%)	8 (17%)	7 (15%)	16 (25%)	11 (24%)	1 (2%)
Health	4 (22%)	4 (22%)	0 (0%)	7 (39%)	2 (11%)	8 (44%)	4 (22%)	3 (17%)
WASH	3 (18%)	6 (35%)	2 (12%)	2 (12%)	2 (12%)	4 (24%)	4 (24%)	2 (12%)
Total	29 (23%)	32 (25%)	34 (27%)	30 (24%)	25 (20%)	53 (42%)	27 (21%)	11 (9%)

Table 57. Expected benefits of SIEF application broken down by position

	Technical support with proposal development	Technical support with project execution	Capacity building	Access to in-country contacts and partners	Access to (in-country) policy-makers	Post-project communication and dissemination	None	Other
PI	17 (23%)	18 (25%)	20 (27%)	23 (32%)	22 (30%)	30 (41%)	18 (25%)	5 (7%)
TTL	12 (23%)	14 (26%)	14 (26%)	7 (13%)	3 (6%)	23 (43%)	9 (17%)	6 (11%)
Total	29 (23%)	32 (25%)	34 (27%)	30 (24%)	25 (20%)	53 (42%)	27 (21%)	11 (9%)

Table 58. How did you hear about SIEF (by position)?

	Colleague	Previous grant	Web search	Online announcement	Other
PI	35 (51%)	5 (7%)	2 (3%)	23 (34%)	7 (10%)

TTL	16 (33%)	8 (16%)	1 (2%)	21 (43%)	7 (14%)
Total	51 (44%)	13 (11%)	3 (3%)	44 (38%)	14 (12%)

Table 59. How did you hear about SIEF (by cluster)?

	Colleague	Previous grant	Web search	Online announcement	Other
ECD	19 (39%)	5 (10%)	2 (4%)	15 (31%)	8 (16%)
EDU	16 (36%)	5 (11%)	1 (2%)	20 (45%)	2 (5%)
Health	8 (42%)	2 (11%)	0 (0%)	7 (7%)	2 (11%)
WASH	8 (50%)	2 (13%)	0 (0%)	2 (13%)	4 (25%)
Total	51 (40%)	13 (10%)	3 (2%)	44 (35%)	14 (12%)

Appendix J: Results E2P notes gender check

E2P note (Note: where the title does not specify the country of focus, this is specified in brackets for convenience.)	Section on gender?	Other in-text coverage of issues specific to women
Tanzania: Can Conditional Cash Transfers Encourage Safer Sexual Behavior?	Yes – Men and women appeared to respond similarly to the payments.	Yes (as there is at least a section on it)
Rwanda: Will More People Be Tested for HIV if Clinics Are Paid Extra?		Yes – 'improving health care for pregnant women and children' is particularly important for Sub-Saharan Africa
Haiti: Can Non-Public Schools Fill the Gap for Poor Children?		
Argentina: Can Performance Payments Improve Newborn Health?		Yes – 'enrolling pregnant women and children in the program'
Bangladesh: Can Conditional Cash Transfers Improve Nutrition?		Yes – 'give women more say over family finances'
Liberia: Can Employment Opportunities Help Build Peace?		Yes – focus on men.
Turkey: Do job Training Programs Help People Find Jobs?		Yes – 'After being stratified by gender and age'
Philippines: Are Cash Transfer Programs Effective?		Yes – 'regular check-ups for pregnant women and children'
Can Disadvantaged Kids Ever Catch Up with Better-Off Peers? (Jamaica)		
Tanzania: Can Local Communities Successfully Run Cash Transfer Programs?	Yes – 'The program also had a positive effect on education, particularly for girls and the community's most vulnerable children.'	Yes (as there is at least a section on it)

Can Demand for Toilets Be Encouraged? Evidence from Indonesia		Yes – 'least one member who defecated in the open, with men and children doing it more often than women'
Cambodia: Challenges in Scaling Up Preschools		'run by trained women from the community'
Using Low-Cost Private Schools to Fill the Education Gap: An Impact Evaluation of a Program in Pakistan		'gender equality among children'
What Gets People to Wash Their Hands? Impact Evaluation Evidence from Peru and Vietnam		
What's the Long-Term Impact of Conditional Cash Transfers on Education? (Colombia)	'Girls benefited more than boys…'	Yes (as there is at least a section on it)
If You Give the Poor Cash, Does It Help? (Uganda)	'In particular, the cash grants gave women a real boost, underscoring that access for finance is critical to helping women escape poverty.'	Yes (as there is at least a section on it)
The Challenge of Ensuring Adequate Stocks of Essential Drugs in Rural Health Clinics (Zambia)		
Can Entrepreneurship Training Improve Work Opportunities for College Graduates? (Tunisia)		Yes – 'Randomization was conducted at the project level and stratified by gender and study subject' … 'Two-thirds of the applicants were women'
Can Scholarships Help Keep Kids in School? (Cambodia)	'The impact was the same for girls and boys.'	Yes (as there is at least a section on it)
Do Grants to Communities Lead to Better Health and Education? (Indonesia)		Yes – 'prenatal visits for pregnant women'… 'extremely poor households with children or pregnant women' … '27 percent for financial assistance for pregnant women to use health services'
Can Public Works Programs Help the Poor during Crises? (Latvia)		
Rwanda: Can Bonus Payments Improve the	Several, including: • 'The likelihood that a woman would	Yes (as there is at least a section on it)

Quality of Health Care?	give birth in a health center—instead of at home— increased among women served by facilities that could qualify for performance-based bonus payments.' • 'Women living in these areas also were more likely to bring in their babies for preventive care.'	
Do Wage Subsidies Help Young Women Get Jobs? (Jordan)	Several, including 'Vouchers boosted young women's employment – at least while the vouchers were active.'	Yes (as there is at least a section on it)
Pakistan: Can Low-Cost Private Schools Improve Learning?		Yes – '61.2 percent for girls'
Before Crisis Hits: Can Public Works Programs Increase Food Security? (Ethiopia)		
Can Cash Transfers Help Children Stay Healthy? (Burkina Faso)	'Nor was the gender of the child significant when it came to the effect of conditional cash transfers.'	
Can Small Farmers Protect Themselves Against Bad Weather? (Nicaragua)		Yes – 'The program targeted the main caregivers in the households, who almost always are women'
Is Preschool Good for Kids? (Mozambique)		
How Can We Make Schools Work Better? (Indonesia)		
Do School Feeding Programs Help Children? (Burkina Faso, Laos and Uganda)	'Educational gains were mixed: In Burkina Faso, girls enrolled in schools with a feeding program showed small increases in scores on math tests, but there was no significant impact for boys; in Uganda, boys were less likely to have to repeat a grade, but there was no noticeable effect on girls.'	Yes (as there is at least a section on it)
Do Vouchers for Job Training Programs Help? (Kenya)	'But women shown videos of women working in traditionally male jobs, such as auto repair, and told that wages were higher in such fields, were more likely to use their vouchers differently.'	Yes (as there is at least a section on it)
Do Food Supplements Help Children in Times of Economic Crisis? (Indonesia)	'The program targeted poor children from age 6 months to 60 months and continued nutritional assistance for pregnant women.'	Yes (as there is at least a section on it)
Can Targeted Assistance Help the Very Poor? (Chile)		Yes – 'Women, in particular, had little or no experience in the labor market, further hampering their chances of

	finding work.' Plus other references.
Can Computers Help Students Learn? (Colombia)	Yes – 'even distribution of students across grades and genders'
Does Linking Teacher Pay to Student Performance Improve Results? (India)	Yes – 'There was no difference in results based on student demographics, including gender'. Plus other references.

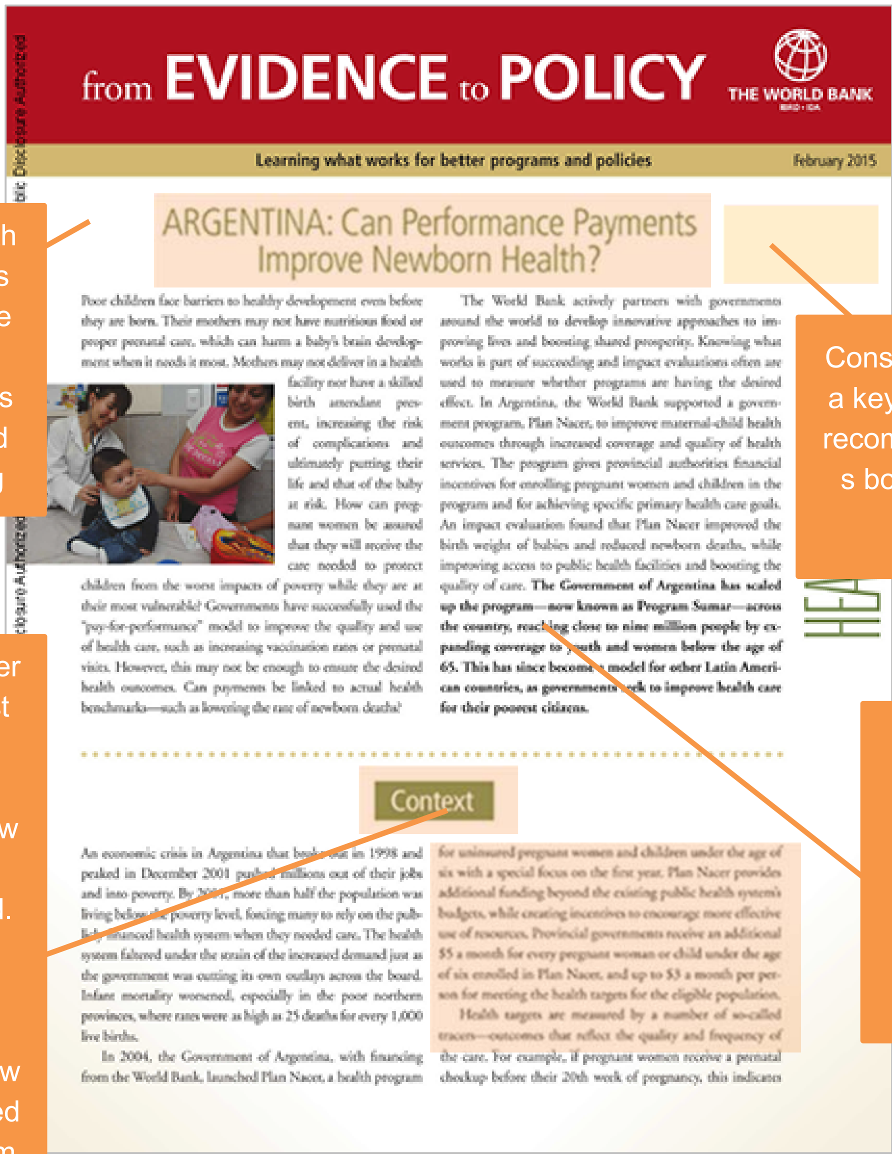
Appendix K: SIEF proposal selection process

Figure 3: Schematic of the SIEF selection process

Call for proposals	(Reviews of the entire calls are performed.)
Technical review (SIEF team and consultants)	Review of the proposals by the SIEF team and consultants. Check if proposals are essentially impact evaluations (e.g. is there a counterfactual; is it sufficiently powered). A traffic light system is used to indicate the quality of proposals.
Technical review committees (World Bank, DFID, CIFF)	Review of the proposals by the technical review committees and the World Bank, DFID and CIFF. Check the proposals for technical merit and relevance to the call. Responses are clustered by the SIEF team.
Evaluation review committee	Review of the proposals by the Evaluation Review Committee. They use the following criteria: - Policy/Cluster Relevance (weight 40%) - Technical quality (weight 40%) - Capacity for Delivery/Country Engagement (weight 10%) - Timeline for results (weight 5%) - Realism for Budget (weight 5%) Members include: DFID, CIFF, WB-global practices, WB-chief economist office
Steering committee review	The Steering Committee finally reviews the proposed funding allocation and upon approval seed funding is distributed.

Appendix L: Comments on E2P notes

Current length of four pages should be the maximum – policy makers do not spend long reading



from EVIDENCE to POLICY

THE WORLD BANK

Learning what works for better programs and policies

February 2015

ARGENTINA: Can Performance Payments Improve Newborn Health?

Poor children face barriers to healthy development even before they are born. Their mothers may not have nutritious food or proper prenatal care, which can harm a baby's brain development when it needs it most. Mothers may not deliver in a health facility nor have a skilled birth attendant present, increasing the risk of complications and ultimately putting their life and that of the baby at risk. How can pregnant women be assured that they will receive the care needed to protect children from the worst impacts of poverty while they are at their most vulnerable? Governments have successfully used the "pay-for-performance" model to improve the quality and use of health care, such as increasing vaccination rates or prenatal visits. However, this may not be enough to ensure the desired health outcomes. Can payments be linked to actual health benchmarks—such as lowering the rate of newborn deaths?

The World Bank actively partners with governments around the world to develop innovative approaches to improving lives and boosting shared prosperity. Knowing what works is part of succeeding and impact evaluations often are used to measure whether programs are having the desired effect. In Argentina, the World Bank supported a government program, Plan Nacer, to improve maternal-child health outcomes through increased coverage and quality of health services. The program gives provincial authorities financial incentives for enrolling pregnant women and children in the program and for achieving specific primary health care goals. An impact evaluation found that Plan Nacer improved the birth weight of babies and reduced newborn deaths, while improving access to public health facilities and boosting the quality of care. **The Government of Argentina has scaled up the program—now known as Program Sumar—across the country, reaching close to nine million people by expanding coverage to youth and women below the age of 65. This has since become a model for other Latin American countries, as governments seek to improve health care for their poorest citizens.**

Context

An economic crisis in Argentina that broke out in 1998 and peaked in December 2001 pushed millions out of their jobs and into poverty. By 2002, more than half the population was living below the poverty level, forcing many to rely on the publicly financed health system when they needed care. The health system faltered under the strain of the increased demand just as the government was cutting its own outlays across the board. Infant mortality worsened, especially in the poor northern provinces, where rates were as high as 25 deaths for every 1,000 live births.

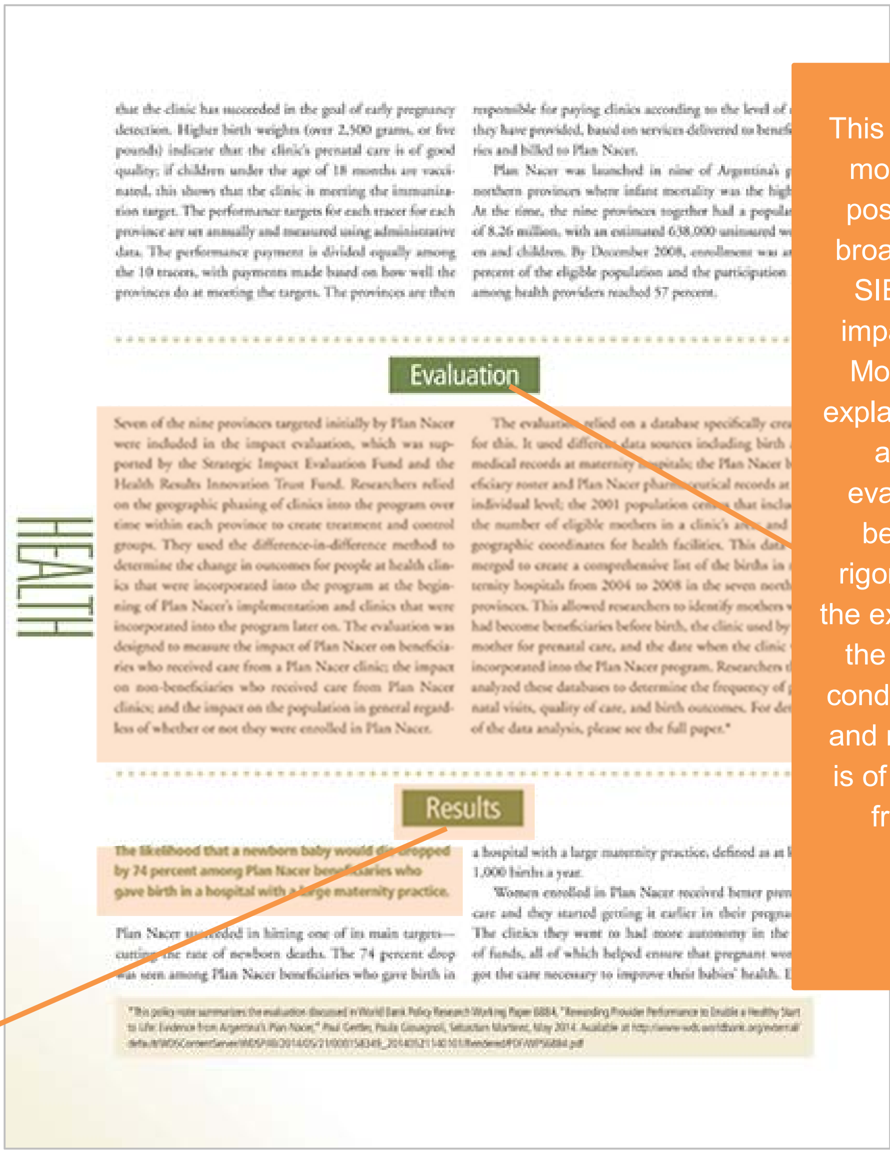
In 2004, the Government of Argentina, with financing from the World Bank, launched Plan Nacer, a health program for uninsured pregnant women and children under the age of six with a special focus on the first year. Plan Nacer provides additional funding beyond the existing public health system's budgets, while creating incentives to encourage more effective use of resources. Provincial governments receive an additional $5 a month for every pregnant woman or child under the age of six enrolled in Plan Nacer, and up to $3 a month per person for meeting the health targets for the eligible population.

Health targets are measured by a number of so-called tracers—outcomes that reflect the quality and frequency of the care. For example, if pregnant women receive a prenatal checkup before their 20th week of pregnancy, this indicates

Consider adding a key findings & recommendation s boxes at the start

As part of the wider context in the first page, consider adding in information on how the programme was implemented. While not all practical details can be provided, an overview of how it was implemented will be valued from the policy perspective.

Ensure study question is answered on first page – only some of the notes did this


that the clinic has succeeded in the goal of early pregnancy detection. Higher birth weights (over 2,500 grams, or five pounds) indicate that the clinic's prenatal care is of good quality; if children under the age of 18 months are vaccinated, this shows that the clinic is meeting the immunization target. The performance targets for each tracer for each province are set annually and measured using administrative data. The performance payment is divided equally among the 10 tracers, with payments made based on how well the provinces do at meeting the targets. The provinces are then responsible for paying clinics according to the level of they have provided, based on services delivered to benefi ries and billed to Plan Nacer.

Plan Nacer was launched in nine of Argentina's northern provinces where infant mortality was the high At the time, the nine provinces together had a popula of 8.26 million, with an estimated 638,000 uninsured w en and children. By December 2008, enrollment was a percent of the eligible population and the participation among health providers reached 57 percent.

HEALTH

Evaluation

Seven of the nine provinces targeted initially by Plan Nacer were included in the impact evaluation, which was supported by the Strategic Impact Evaluation Fund and the Health Results Innovation Trust Fund. Researchers relied on the geographic phasing of clinics into the program over time within each province to create treatment and control groups. They used the difference-in-difference method to determine the change in outcomes for people at health clinics that were incorporated into the program at the beginning of Plan Nacer's implementation and clinics that were incorporated into the program later on. The evaluation was designed to measure the impact of Plan Nacer on beneficiaries who received care from a Plan Nacer clinic; the impact on non-beneficiaries who received care from Plan Nacer clinics; and the impact on the population in general regardless of whether or not they were enrolled in Plan Nacer.

The evaluation relied on a database specifically cre for this. It used different data sources including birth medical records at maternity hospitals; the Plan Nacer b eficiary roster and Plan Nacer pharmaceutical records at individual level; the 2001 population census that inclu the number of eligible mothers in a clinic's area and geographic coordinates for health facilities. This data merged to create a comprehensive list of the births in ternity hospitals from 2004 to 2008 in the seven north provinces. This allowed researchers to identify mothers had become beneficiaries before birth, the clinic used by mother for prenatal care, and the date when the clinic incorporated into the Plan Nacer program. Researchers analyzed these databases to determine the frequency of natal visits, quality of care, and birth outcomes. For de of the data analysis, please see the full paper.*

Results

The likelihood that a newborn baby would die dropped by 74 percent among Plan Nacer beneficiaries who gave birth in a hospital with a large maternity practice.

Plan Nacer succeeded in hitting one of its main targets—cutting the rate of newborn deaths. The 74 percent drop was seen among Plan Nacer beneficiaries who gave birth in a hospital with a large maternity practice, defined as at l 1,000 births a year.

Women enrolled in Plan Nacer received better pre care and they started getting it earlier in their pregna The clinics they went to had more autonomy in the of funds, all of which helped ensure that pregnant wo got the care necessary to improve their babies' health.

*This policy note summarizes the evaluation discussed in World Bank Policy Research Working Paper 6884, "Rewarding Provider Performance to Enable a Healthy Start to Life: Evidence from Argentina's Plan Nacer," Paul Gertler, Paula Giovagnoli, Sebastian Martinez, May 2014. Available at

This section could be moved to the end, possibly even as a broader summary of SIEF's criteria for impact evaluations. Most respondents explained they trusted a SIEF impact evaluation to have been conducted rigorously. As such, the explanation of how the evaluation was conducted technically and methodologically is of less importance from the policy perspective

Results section could be condensed. The main results are likely to be communicated already through the headline findings in **bold**.

women who weren't enrolled in the program but who went to a Plan Nacer clinic for care benefitted: the rate of neonatal mortality among babies born to these women declined by 22 percent.

The mortality rate for newborns was reduced thanks to two things—a decline in the probability that a baby would be born with a low birth weight, and better care for low birth weight babies.

A baby's birth weight is a good indicator of child's future health and abilities. Low birth weight, defined as less than 2,500 grams—or five pounds—is associated with higher rates of infant death, more health problems and delays in learning later on in life. The probability of having a low birth weight baby was reduced by nine percent for all women who went to Plan Nacer clinics—regardless of whether the women were enrolled in the program—and for Plan Nacer beneficiaries, the probability decreased by 19 percent.

The probability of a low birth weight baby dying was cut by more than half if the mother was enrolled in Plan Nacer, compared with babies born to mothers who didn't go to a Plan Nacer clinic. Overall, researchers calculated that 54 percent of the drop in newborn mortality was due to the decline in low birth weight babies, and the remaining 46 percent was due to improved care for babies born under five pounds.

Women in Plan Nacer went for more prenatal visits and were less likely to deliver by cesarean section.

Prenatal visits are important because they allow health workers to identify problems early on and to provide the preventive care to protect women's health and the health of their babies. Women enrolled in the program were 5.6 percentage points more likely to receive the recommended tetanus vaccine (translating into almost a 25 percent drop in the percent of women who didn't get it). There also was a 21 percent drop in birth by cesarean sections—an indication of quality prenatal care. Overall, for every 100 women enrolled in Plan Nacer, there were 68 additional prenatal care visits.

The program encouraged accountability and results not just procedures.

When the program was designed, provinces received mo[…] control over health care because they had a say over ho[…] to distribute funds to clinics that enrolled in Plan Nac[…] At the same time, provinces were held accountable becau[…] they wouldn't receive funding from the national governme[…] unless certain health targets were reached, including enro[…]ing more eligible women and children in Plan Nacer. Th[…]

Examples of Plan Nacer Targ[…]

First prenatal checkup before week 20 of […]
APGAR score of over 6
Birth weight more than 2,500 grams
Children under 18 months given vaccines […]
Mothers receive reproductive counseling […] of delivery
Preventative checkups for children under 6
Staff trained to provide care for indigenous populations

two-pronged payment approach gave local governments an incentive to ensure that clinics enrolled potential beneficiaries and an incentive to ensure that clinics did a better job taking care of them.

The program has provided millions of services to the people who need it most, while proving cost effective in terms of helping people stay healthy.

In the nine provinces where Plan Nacer started, spending on health care rose 1.4 percentage points to 3.5 percent of the provincial budgets between 2005 and 2008, for a total of $107 million spent on the health program. Most of the money, about 71 percent in 2008, went to clinics for services provided to beneficiaries. The second largest category was investment in medical equipment and other facility necessary hardware. Most of this spending happened in the first year, when it made up 30 percent of spending, falling to eight percent in 2008. As the program has grown, overhead costs have decreased, from 16 percent of total provincial health expenditures in 2005 to just 11 percent by 2008. The program also has been cost-effective when compared with the number of years of life that could be lost because of illness, disability or early death.

Consider adding a sub-heading on 'Implementation' and/or 'how-to'. This can help readers navigate to essential implementation information.

Only some notes considered gender explicitly. Consider increasing use of Gender in Focus boxes

It was mentioned that a graphic or figure could help to illustrate the reporting of the results.

Conclusion

Plan Nacer's incentive-based model helped improve maternal and new born health, and ensure that children have the chance at a good start in life. The evidence from this evaluation will equip policy makers in low and middle income countries with additional information when designing health programs aimed at improving specific outcomes. As governments around the world look for ways to create effective programs to help their poorest citizens, the results from this impact evaluation provide an example of how health sector reforms can give children the right start in life.

HEALTH

The Strategic Impact Evaluation Fund, part of the World Bank Group, supports and disseminates research evaluating the impact of development projects to help alleviate poverty. **The goal is to collect and build empirical evidence that can help governments and development organizations design and implement the most appropriate and effective policies for better educational, health and job opportunities for people in developing countries.** For more information about who we are and what we do, go to: http://www.worldbank.org/sief.

The Evidence to Policy note series is produced by SIEF with generous support from the British government's Department for International Development.

THE WORLD BANK
IBRD • IDA

THE WORLD BANK, STRATEGIC IMPACT EVALUATION FUND
1818 H STREET, NW WASHINGTON, DC 20433

Produced by the Strategic Impact Evaluation Fund
Series Editor Alaka Marcus; Writer: Daphna Berman

As well as instructions on how to implement the approach taken, also consider incorporating input on how to scale it, if this is available. Scalability is on top of the agenda for policy makers