



Call for Proposals 4: Nimble Evaluations

Since 2012, the World Bank's [Strategic Impact Evaluation Fund](#) (SIEF) has funded three open, competitive calls for proposals for impact evaluations in low- and middle-income countries focused on at least one of four issues: early childhood development and nutrition, basic education, health systems, and water and sanitation. The current portfolio has around 60 evaluations spread across 31 countries. Many of these evaluations are led by world-class researchers; many are closely linked to government programs; and many are answering questions that bilateral and multilateral aid organizations find relevant.

SIEF is now announcing its fourth call for proposals. This call will focus on rapid, low-cost evaluations – or nimble evaluations.¹ A follow-on call will provide incentives for iterative policy experimentation after results of Call 4 evaluations are submitted. In call 4, SIEF will for the first time seek to match impact evaluation researchers with operational teams in the World Bank and the United Kingdom's Department for International Development (DFID) to encourage the use of nimble evaluations for better program design and scale-up. Research teams will be able to propose evaluation designs for specific information needs articulated by World Bank and DFID client governments and non-government partners ([see appendix of World Bank and DFID projects](#)). In these cases, funds for implementing the interventions being evaluated will be covered by whichever organization is sponsoring the program. It is expected these organizations will also be able to provide easier access to administrative data.

In addition, SIEF will also accept applications from researchers who have identified planned or existing interventions needing evaluation that are neither on the list nor connected with World Bank or DFID projects. However, as in former calls for proposals, evaluation teams in these situations will need to arrange for separate funding for intervention implementation activities.

This note provides details on the objectives of the fourth call for proposals and explains the process SIEF will use for selecting evaluations and matching researchers to operational teams.

1 Objectives of the call

SIEF will use this fourth call for proposals to meet five objectives that advance SIEF's core mandate of supporting evidence-based policy design in low and middle-income countries.

OBJECTIVE 1 Generating rigorous evidence on implementation

¹ For some examples, please see the following talk by Dean Karlan: <https://vimeo.com/218836475> and his corresponding slides: <http://pubdocs.worldbank.org/en/626921495727495321/Nimble-RCTs-WorldBankMay2017-v4.pdf>

Ensuring implementation fidelity is a major challenge in many middle- and low-income countries, and this affects the inference we can make in the context of an impact evaluation. When measured impacts of a program suggest limited or even no effect, it is typically not clear if the program was the wrong one for the policy problem or if the program was just not implemented correctly.

To remedy this, SIEF's call will make implementation the research focus or, as economist Esther Duflo has argued, encourage researchers to become "plumbers" by using field experiments to learn *how* to do things, rather than just learn *what* to do.² For example, before testing the impact of an early childhood development intervention on child outcomes - a long and expensive evaluation - an evaluation could test who is more likely to deliver the intervention on time, volunteers minimally incentivized with something like a T-shirt or small stipend, or individuals who are regularly remunerated for their work. It could also test where targeted parents are most likely to take-up the intervention - in health clinics or in their homes during a door-to-door campaign? These would be more nimble evaluations that could likely be completed in a matter of months, and if neither variant successfully delivered the intervention (or generated sufficiently high take-up, in the second case), there would be no need to track children for a much longer period and administer an expensive data collection instrument to measure their development.

Because of this focus on implementation, evaluations need not be designed prior to the start of a program to estimate whether or not it is working; rather they can be designed around interventions already under implementation, where problems have already been identified and require troubleshooting. We expect nimble evaluations to provide empirically rigorous troubleshooting.

OBJECTIVE 2 Creating a portfolio of low-cost, rapid evaluations

Evaluations focused on take-up, service delivery, and program adherence can provide much needed evidence on implementation in low and middle-income countries and also help create a portfolio of rigorous low-cost, rapid evaluations. Currently many operational teams shy away from evaluating components of their projects because they think that experiments are very expensive and must take at least three years to complete. This is not the case. For example, in an experiment in Indonesia that was completed in less than six months, researchers tested how small modifications to identity cards could affect the amount of subsidized rice purchased by poor households and the level of unofficial price mark-ups.³ The innovation in identity card design was then scaled up to more than 15 million individuals.

The evaluations funded under Call 4 will also serve as important case studies that SIEF itself can use to learn about how to best design nimble RCTs and the contexts in which they can be most effective.

OBJECTIVE 3 Spurring innovations in the use of administrative data and other survey alternatives

SIEF is particularly interested in learning about cheaper data capture methods in low and middle-income countries. For example, to what extent can existing administrative data, once accessed, be used in evaluations? To what extent is additional investment in data quality needed? Moreover, are there other

² For more detail, see Duflo, Esther (2017), "Richard T. Ely Lecture: The Economist as Plumber," *American Economic Review*, 107(5): 1-26. Ungated version available here: <https://economics.mit.edu/files/1256>

³ For more details on this experiment, please see here: <https://www.povertyactionlab.org/node/9407>

alternatives to sending in enumerators for household, school, and clinic surveys to measure an evaluation's main outcome(s) of interest?

OBJECTIVE 4 Incentivizing iterative experimentation

Both researchers and international aid organizations recognize that the path from evidence to policy is not straightforward. It is unusual for a policy to be immediately scaled up after a single evaluation. More often, evaluation results may lead to new questions and new ideas for interventions that themselves need to be evaluated, which in turn may lead to new ideas and new designs that need to be evaluated.

To incentivize this kind of experimentation, SIEF will split the funds set aside for nimble evaluations into two separate calls for proposals – Call 4 (the current call for proposals) and Call 5, which would only be for those teams who have completed a nimble evaluation in Call 4 and are proposing a follow-up experiment to what was completed during Call 4. Funds in Call 5 would be available on a rolling basis, conditional on submission of all SIEF deliverables in Call 4, and would not be sufficient to cover all teams awarded funding in Call 4.

OBJECTIVE 5 Matching first class researchers to operational teams

SIEF requires that all funded evaluations be overseen by a World Bank task team leader (TTL). This requirement allows the World Bank to provide necessary fiduciary oversight of SIEF funds and to set up a conduit for dissemination to decision-makers in governments that World Bank TTLs regularly communicate with. In the past, some researchers interested in SIEF funding have found it difficult to match with a World Bank TTL.

In this upcoming call for proposals, SIEF will aid the matching process by allowing researchers to design evaluations for specific policy problems that are the topic of current World Bank and DFID projects that have already expressed an interest in incorporating nimble evaluations into their project design ([see list](#)). Successful applications will then be matched to project TTLs interested in the proposed designs.

2 How will this all work?

2.1 Eligibility

SIEF solicits proposals through competitive calls for proposals open to both internal and external research teams. Ultimately, all SIEF funded evaluations will require an evaluation task team leader (TTL) from the World Bank who will provide fiduciary oversight and serve as a conduit of results to key policymakers and other operational colleagues working in same region. SIEF funds are restricted to evaluation-related activities (evaluation design, data collection, quality assurance for data collection, analysis and writing, and in-country dissemination) and cannot be used for an intervention's implementation activities. SIEF will provide a maximum of \$150,000 for each evaluation. Principal investigators must have completed an evaluation in the past.

Funding for the evaluation will be disbursed upon submission of deliverables, which include:

- (i) Approval from the relevant World Bank country management unit
- (ii) Human subjects certification of all investigators
- (iii) Approval of an Internal Review Board or other research ethics entity
- (iv) Baseline data and end line data (if applicable)

- (v) Baseline report on balance, statistical power, cost data, and implementation updates
- (vi) End line report on evaluation results, intervention costs, and implementation fidelity
- (vii) Brief note on learning about the process of implementing a nimble evaluation. For example, insights on challenges and facilitators in data capture and/or iterative experimentation.
- (viii) Timely quarterly reporting on evaluation progress

2.2 Screening

To select evaluations for funding in Call 4, there will be two rounds of screening: Screening 1, a quick application to check for the quality of the evaluation design and its alignment with the objectives of the call for proposals and to facilitate matches between operational and research teams; and Screening 2, which will be based on a joint proposal submitted by the research teams and a World Bank TTL or lead operational adviser from DFID, depending on the sponsor of the project. Details on this can be found below.

Screening 1

The Screening 1 application form is available [here](#).

All applications received by May 30, 2018 will be considered.

At the time of application for Screening 1, applicants do not need to be matched to a World Bank TTL.

As in the past, applicants can propose an evaluation design of any intervention affecting outcomes related to SIEF's core themes: basic education; health and health systems; early childhood development and nutrition; and water, sanitation, and hygiene. These interventions can be implemented by governments, non-governmental organizations, or by researchers themselves, although SIEF will not provide any funding for implementation activities.

In Call 4, applicants will also have the opportunity to design an evaluation for an existing World Bank or DFID project. An [online appendix](#) lists all projects expressing an interest in building a nimble evaluation into a project design and includes a short description of the policy problem and any data access that World Bank or DFID teams could facilitate. All projects listed would have guaranteed funding for evaluated interventions, freeing researchers of the need to separately fundraise for implementation. On the application form for Screening 1, applicants should indicate which project they are proposing a design for (using the IDs listed in the appendix). Their application information will be shared with relevant World Bank TTLs/DFID staff only if they make it past Screening 1. The review process will take into consideration that the interventions and evaluation designs (including data collection plans) proposed in these cases will be hypothetical. Applicants proposing an intervention and evaluation for these articulated information needs can consult the websites of the World Bank and DFID to find more information about current projects and general policy challenges in these countries.⁴ These documents may also describe existing information systems in the sector in question.

⁴ DFID: <https://devtracker.dfid.gov.uk/>

World Bank: <http://www.worldbank.org/en/what-we-do>

Screening 1 applications will be reviewed only by World Bank and external evaluation experts and subject experts who are not submitting applications for Call 4 and who are not operational TTLs. The SIEF team will aim to inform all teams of their Screening 1 status by June, 30, 2018. There is no target for the number of evaluations that successfully pass Screening 1.

The following criteria will be used for screening:

Criteria that all applications must meet for Screening 1. If any answer is "no," SIEF funding cannot be awarded.

Will the design of the evaluation help uncover causal impacts of an intervention?

Would the evaluation build evidence on SIEF's core themes?

Would SIEF funding account for at least 75 percent of total evaluation costs?

Is it feasible for an end line report to be submitted within 18 months of funding?

Is there a principal investigator who has completed an impact evaluation in the past?

Does the study design include cost data collection?

Does the study design include data collection related to implementation fidelity of any interventions under evaluation?

Is the country of study a World Bank client country?⁵

Criteria that will be used in the case of tie-breaks. Affirmative answers to these questions should give applications higher priority for getting past Screening 1. More affirmative answers should give applications more priority.

Will the evaluation use or improve upon existing administrative data?

Is the country of study a DFID-priority country?⁶

Is there an identified policy or program related decision that will be affected by the results of the evaluation?

Will the evaluation generate evidence on how to improve implementation of a program?

Does the proposed evaluation design include multiple, iterative experiments?

Is there potential for iterative experimentation in general, even if only with additional funding?

Have funds for implementation of the intervention under evaluation already been secured?

Does the evaluation build evidence for identified World Bank or DFID projects?

Are application responses succinct and clear?

⁵ For a list of World Bank client countries, see here: <http://www.worldbank.org/en/where-we-work>

⁶ DFID-priority countries include: Afghanistan, Bangladesh, Myanmar, Democratic Republic of Congo, Ethiopia, Ghana, India, Iraq, Jordan, Kenya, Kyrgyzstan, Lebanon, Liberia, Malawi, Mozambique, Nepal, Nigeria, West Bank and Gaza, Pakistan, Rwanda, Sierra Leone, Somalia, South Africa, Sudan, South Sudan, Syria, Tajikistan, Tanzania, Uganda, Yemen, Zambia, and Zimbabwe.

Screening 2

Teams that pass Screening 1 will be invited to submit a full proposal by August 15, 2018.

If during Screening 1 an applicant has proposed an evaluation design for a project that is not on the list, then a World Bank TTL will need to agree to jointly apply for Screening 2 with the researcher and serve as the task team leader of the evaluation only, without any explicit ties between the evaluation and ongoing operations. Researchers can directly contact TTLs that they know and want to work with, but SIEF can also facilitate an introduction to a TTL working on the topic in the country of study.

If during Screening 1 the applicant has proposed an evaluation design for a World Bank or DFID project on the list, then the relevant World Bank or DFID operational team must also agree to jointly apply for Screening 2. Again, SIEF will make the introduction between the research team and the World Bank TTL, after which the research team can communicate directly with the TTL and find out more detail about the policy problem, data access, and viability of their proposed evaluation design. In the case of DFID projects, a World Bank TTL must also agree to serve as the task team leader of the evaluation only, again without any explicit ties between the evaluation and ongoing operations.

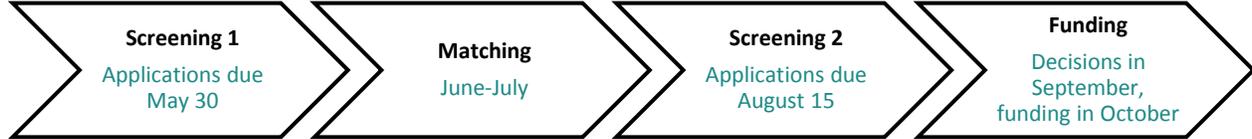
SIEF will actively try to match researchers to operational teams, but teams that pass Screening 1 that cannot generate the interest of any World Bank TTLs will not be able to submit a full application during Screening 2. Applicants that have made it past Screening 1 will also be invited to come to the World Bank (at their own expense) to participate in a networking event (sometime in June or July) that will be designed to facilitate matches and discussions between researchers and operational teams.

During this stage of the application process, all Screening 1 criteria will still apply. Teams must also provide sufficient evidence that the evaluated intervention(s) have funding and can start soon. Teams will need to describe the data they plan to use in detail and include proof that they can access the data and that the proposed sample size will be sufficient to detect statistically meaningful effects. Teams will also need to fill out a budget template that will be provided and attest that SIEF funding would account for a minimum of 75 percent of total evaluation costs. Evaluations that employ quasi-experimental designs will also have to provide evidence for the assumptions underlying these methods (for example, evidence of equal pre-trends for differences-in-differences strategies or evidence of strong treatment prediction and common support between the treatment and control groups in propensity score matching strategies).

Once Screening 2 applications are received, SIEF will organize a second technical review, again by internal and external researchers and subject experts that are not applying for Call 4 funding. A regional screening committee will need to provide each application that makes it past Screening 2 a no-objection to ensure that the evaluation will not interfere with the World Bank's ongoing policy dialogue in the relevant country, and SIEF's Steering Committee will make the final assessments for funding for all applications that make it past Screening 2 and the regional committee. The total budget for nimble evaluations will limit the number of evaluations chosen for funding.

2.3 Timing

An approximate timing for the different stages of screening and matching is provided below.



3 Budgets

The overall budget for the call is USD 3 million, which will be separated into an initial round of \$2 million (Call 4) and follow-up round of \$1 million (Call 5). For evaluation teams indicating no interest in matching with a World Bank or DFID operational team, there will be a funding cap of \$100,000. For evaluations electing to participate in the matched pool or matching with another World Bank or DFID operation, there will be a cap of \$150,000. These teams will also have funding for the implementation of the evaluated interventions (as they will be project activities) and access to any available administrative data.

SIEF intends that these capped amounts should cover all or most of the expenses related to the evaluation – specifically at least 75 percent of all evaluation costs; SIEF funds in Call 4 should not be top-up funds for an ongoing, larger, not-so-nimble evaluation.

There will be no seed funding provided for Call 4.

4 Queries

Please direct all queries to siefimpact@worldbank.org. Prior to the completion of Screening 1, SIEF will not match researchers to TTLs. After Screening 1, SIEF will only match researchers to TTLs who have agreed to be matched.