

MAY 2022

CEDIL Methods Brief 8

When time is of the essence: Timely evaluations in international development



Centre of
Excellence
for Development
Impact and Learning

1. What are timely evaluations and what can they be used for?

Timely evaluations produce results when they are needed to inform decision making while using appropriate, rigorous designs. While traditionally evaluations are carried out at the end of the programme, the methods explored in this brief highlight evaluations opportunities during the course of a programme and are important for informing decisions on project design for which information is needed in a timely manner. Timely evaluations can be used when trying out new intervention approaches, when working in a new context, or when the context is rapidly changing.

Timely evaluations can answer a range of questions that are relevant at different stages of an intervention.

- *Using formative research and evaluation to design feasible and useful interventions:* What are the priority outcomes in the target population and what are the constraints to achieving them? What barriers deter the intended participants from taking part? Is the intervention being implemented as planned? Are planned technologies appropriate and are they functioning as intended under field conditions?
- *Different types of timely evaluation approaches for different needs:* Which programme design approaches are best to ensure programme take-up? How can monitoring data be used to test and improve programme performance? Can multicomponent projects be unpacked to identify the contribution of single project activities to specific outcomes, as well as the optimal combination of components?
- *Using qualitative and participatory approaches to assess impact:* How can qualitative data be used for getting early estimates of impact? What can be learned from evaluation findings about the effectiveness and cost effectiveness of different project design choices in different contexts? How can knowledge accumulated throughout the lifetime of a project about how and when causal processes occur inform and update the programme's theory of change?

The above questions are time-sensitive. Answering these questions helps ensure that sufficient time is built in for adapting the focus and scope of an intervention as it is implemented, as well as for learning lessons from these changes. These changes can in turn

improve the effectiveness of interventions.

This methods brief reviews the range of available methods for conducting timely evaluations in international development.

2. Using formative research and evaluation to design feasible and useful interventions

Formative research and formative evaluation are often neglected parts of the project planning and design process. Formative research identifies priority issues and constraints to achieving development goals, thus informing what a programme ought to do. Formative evaluation is an assessment of the feasibility of a proposed project in its early stages.

Formative research and evaluation may draw on the full range of evaluation methods, both quantitative and qualitative. Participatory approaches, such as rapid rural appraisal and participatory rural appraisal, are an important subset of qualitative data.

Formative research

Most formative research uses focus groups, small samples of respondents, and direct observation. For example, a study of antenatal care (ANC) and iron folic acid (IFA) supplementation in Niger conducted semi-structured interviews with pregnant women, observed ANC sessions, conducted exit interviews from those sessions, and held focus groups. The results from this research were interpreted in a two-day workshop with key stakeholders. The intervention based on this formative research included behaviour change communication in communities, quality improvements in health centres, and the provision of key supplies, including IFA supplements. The study did not give a timeline but is likely to have taken three to four months from inception to completion.

An example of formative research using quantitative analysis is the use of qualitative comparative analysis (QCA) to explore pathways from disasters to post-flood disorder. Many combinations of intervening variables may lead to social unrest, especially when incoming migrants are forced into urban areas that are already overcrowded, hazardous environments. QCA was used to test three models of post-disaster unrest in 26 cases across sub-Saharan



Africa.ⁱ The models of contributory factors were rapid urbanisation, youth bulge, and high inequality. The study found that under-delivery with respect to what had been promised was a consistent driver of disorder. People did not protest against the consequences of the disaster itself, but against the policy that followed: a prompt response with promises that were not kept or promises that were unrealistic from the beginning. The use of QCA produced policy-relevant findings based on a relatively small dataset of urban cities. The implication from the findings was clear: decision makers should have a realistic policy response to disasters that is actually implemented, facilitated by good disaster preparedness policy and arrangements.

Formative evaluation

Once a programme has begun, formative evaluation can be used to test its feasibility. There are usually three key aspects to the feasibility of a social intervention: its acceptability by the target population, whether it can be implemented as planned, and whether any proposed technology functions as intended under field conditions. All of these can be assessed by a range of evaluation methods.

Acceptability can be assessed quantitatively through surveys on participation or adoption among the target population. Qualitative data collection, such as focus groups and facilitated community exercises (e.g. rapid rural appraisal

methods), can be used to assess barriers to taking part, for example a lack of creches being a barrier to women with young children being able to participate or the need to ensure that influential people in the community speak out in favour of the project.

Focus groups with staff and beneficiaries, direct observation of project activities, and surveys of project staff knowledge of implementation practice can be used to evaluate whether the programme can be implemented as planned.

Before any of this, the planned technology should be tested under field conditions by getting representatives of the target population to try it out.

All of the above can be done quickly and cheaply during the early months of a project.



3. Different types of timely evaluation approach for different needs

Rapid impact evaluation to improve programme impact (adaptive trials)

Rapid impact evaluations, also called adaptive trials, use data collected early in the life of a project to identify desirable changes in programme design based on real-time evidence. These studies are often randomised evaluations in which different programme design options are tested through different treatment arms. These approaches are common A/B testing that compare two treatment arms. They may also be factorial designs, which have three arms—A, B, and A+B—to which a fourth, untreated, control arm may be added.

For example, indoor residual spraying in Mali is promoted by door-to-door mobilisation in villages. A USAID project considered the use of text messaging as an alternative means of mobilisation. A rapid randomised controlled trial (RCT) was able to show quickly that an approach based on the use of mobile phones was both less effective and more costly than door-to-door visits.ⁱⁱ Based on this finding, the existing approach was retained.

A second example of an adaptive trial design was an experiment to test three employment interventions among Syrian refugees in Jordan.ⁱⁱⁱ The three interventions were a small, unconditional cash transfer; the provision of information to increase the ability to signal skills to employers; and a behavioural nudge to strengthen job search motivation. Data were collected through ‘rapid follow-up’ phone interviews six weeks into the programme, followed by two detailed follow-up surveys two and four months later. There was no effect from any intervention at six weeks, but all interventions showed an effect after two months, and cash transfers had the largest effect after four months.

Rapid impact evaluations are characterised by three features. First, they measure the impact on a ‘low-level outcome’ such as adoption, which is both easily observed and is expected to occur quickly. Second, the unit of assignment is at the individual level, so that a simple RCT (or other impact evaluation design) can be used rather than a cluster RCT, which is far more costly and time-consuming to administer. Third, for any impact evaluation with primary data collection, a rapid RCT has to be planned in close collaboration with the implementing agency, with their buy-in to features such as variations in treatment arms.



Adaptive learning using monitoring data: rapid cycle testing

Adaptive learning is simply the idea that we learn from the evidence and adapt programme design or implementation approaches to ensure we achieve the desired outcomes. All impact evaluations have a learning function: does an intervention work, for whom, and at what cost? However, most impact evaluations have too long a timeframe to inform decisions during the life of the project. One way of ensuring more timely findings to inform decisions is to use monitoring data for the evaluation.

Monitoring data alone can tell us what happened but not why or how to fix any problems we see. For example, the data may show that enrolments of programme participants are falling well below the intended target, but those data do not tell us why or what to do to increase them. However, evaluative uses of these data can be planned for this purpose.

An example of adaptive learning that can use monitoring data is rapid cycle testing. Rapid cycle testing makes small changes in programme design that are randomly applied to a subsample of the target population. The rest of the target population receive the programme as planned ('treatment as usual'). This is an A/B design, comparing two treatment arms within the treatment group. No untreated control is needed, so the study can be conducted using monitoring data. Compared to rapid impact

evaluations, which are conducted by the research team with additional data collection, rapid cycle testing is built into the project and uses monitoring data.

The approach requires the outcome being analysed to be included in the monitoring data, and also that it can be expected to respond to the intervention within a short period. Generally, this means that the outcome needs to be an indicator lower down the causal chain, such as participation or adoption of the behaviour or technology being promoted.

Assessing impact by looking at patterns in the monitoring data: statistical process control and interrupted time series

Monitoring data allow project management to see if the programme is on track and to identify problem areas that require attention. This process is formalised in statistical process control, which uses continuous or real-time data either to identify possible problems or to assess the effect of changes in practice. The approach may be applied to look at patterns in the data over time or across geographical areas.

For example, a programme employs agents to facilitate the formation of women's groups. Project managers receive monthly data on the number of groups formed in each district. National project management might identify numbers starting to drop in a particular district and take remedial action before the problem



worsens. They may also note that a particular district is performing much better than others. This case warrants further identification. It may be unintentional misreporting or fraud, or it may be a genuine success as a result of some innovative approaches being used. If the latter is the case, the implementing agency may consider learning through applying a positive deviance approach. Positive deviance—first applied in Bangladesh and Vietnam to address child health and nutrition—identifies good performers (households with the same resources as other households but with healthier children) and then employs a participatory learning approach to identify what good performers are doing differently, before spreading these approaches more generally.^{iv}

Statistical process control is more commonly applied to monitoring a time trend. In this case, managers look for unexpected deviations or measure the response to a change in project practice. The impact of such changes may be formally tested using interrupted time series analysis, a regression-based approach to determine whether the change in practice coincided with an improvement in outcomes.

4. Using qualitative and participatory approaches to assess impact

There have been various participatory approaches to assessing programme impact. An early approach was beneficiary assessment, which collected beneficiary views of programme impact. Rapid rural appraisal and participatory evaluation are more comprehensive approaches to evaluation using participatory approaches. These approaches go beyond simply asking beneficiaries for their views by actively engaging them in the evaluation.

Participatory evaluation draws on techniques of participatory analysis such as social mapping, oral life histories, transects, and group activities to discuss theories of change. For example, a participatory impact assessment of a water, sanitation, and hygiene (WASH) project in Ghana and Burkina Faso engaged community members in identifying indicators to develop an Empowerment in WASH Index, so that programme effects were measured using indicators that the beneficiaries themselves identified as being important.^v

More sophisticated approaches to participatory evaluation have also been developed. One of these is outcome harvesting, which builds on outcome mapping. Rather than traditional, objective-based evaluation, which assesses interventions against their objectives, outcome harvesting engages stakeholders in identifying key outcomes, how they have changed, and how the intervention has contributed to those changes. By working ‘from the outcome backwards’, the interviews are more likely to identify other causal factors than when interviews focus on the role of the intervention.



The Qualitative Impact Protocol (QulP) is a similar approach, with additional methods adopted to reduce possible biases. It uses outcome harvesting and orients data collection towards what changes have taken place in the community and what the drivers of change have been.^{vi} The approach is reinforced by the fact that neither the respondents nor the data collectors know what intervention is being evaluated. The developers of QulP call this 'blindfolding' rather than 'blinding', as the 'blindfold' can be removed at any time.

These qualitative approaches to assessing impact can be conducted more cheaply and quickly than quantitative impact evaluations such as RCTs which require large sample size (also known as 'large n' impact evaluations). There may be doubts about the strength of their causal findings, but various approaches can be employed to strengthen confidence in them. These include framing the evaluation with a theory of change to establish plausible causal relationships between the intervention; ruling out other explanatory factors (an approach formalised in General Elimination Methodology); and triangulating findings from multiple data sources or using different approaches.

5. When should we use timely evaluations?

Timely evaluation methods offer scientifically sound and rigorous study designs that provide timely information to inform decisions on programme design. Different approaches may be appropriate at different stages of the programme cycle.

Table 1 shows the sorts of question that timely evaluations can be used to address, and the approaches to be used. It also shows the 'timely element' that features in each approach. The questions are divided into four categories depending on the approach: experimental, monitoring, management, and qualitative. A range of approaches are listed under each heading. A discussion of all of these approaches can be found in the accompanying CEDIL Methods Paper, 'Timely evaluation in international development'. A selection of these methods (marked with an asterisk) have been mentioned in this brief.

Table 1: Overview of questions and approaches

	TIMELY QUESTIONS	TIMELY METHODS	TIMELY FEATURE
EXPERIMENTAL	What (combination) of programme components are likely to generate the greatest impact?	Factorial design*	Multiple intervention arms trialled at design stage
	What programme features are likely to generate the greatest impact?	Adaptive trials*	Experiment on variations in programme design
	Is there early evidence of programme impact?	Stepped-wedge trial	Entire population gets treated sequentially, so there may be early evidence of impact
	How to choose between two variations of an intervention?	A/B testing*	Random assignment of individuals to competing alternatives (special case of adaptive trials)
MONITORING	Are the changes in time series random, or do they reflect impact?	Statistical control analysis*	Collect and plot monitoring data on time series graphs (can apply interrupted time series)
	What barriers constrain the impact of project activities?	Bottleneck analysis	Stepwise participatory approach to identify barriers to timely or successful implementation and to impact
	What contribution is the programme making as it is implemented?	Contribution analysis	Multistep approach to validate theory of change, which can be applied at any stage of the project
MANAGEMENT	How do different combinations of programme features and contextual conditions mediate impact?	Quality comparative analysis	Analysis using a small n sample with fewer data requirements
	How can local-context complex problems be identified and addressed?	Problem-driven iterative adaptation	Allows adaptation of the intervention through feedback loops of learning emerging from project activities
	Is the intervention acceptable, deliverable, and adaptable?	Rapid cycle design and testing	As soon as the intervention begins, early learning from qualitative data rapidly collected informs any necessary adaptation to the project
	Can rapid and responsive change be developed in complex environments?	Development evaluation	A multistep approach designed to infer the contribution the programme is making to particular outcomes
	How can large amounts of complex information in large-scale, open-ended projects be collected and monitored?	Most significant change	A participatory monitoring and evaluation qualitative method using stories of change from the field on intermediate outcomes
QUALITATIVE	What are the perceived drivers of change of the project over a predefined recall period?	QuIP*	Use narrative causal statements; lower cost and more rapid than large n impact evaluation
	How can context-specific knowledge be acquired from project participants to inform decision makers?	Rapid rural appraisal*	A participatory method designed for rapid learning of relevant, timely, and accurate information about the challenges on the ground to better understand the context of the intervention
	How can rapid learning take place during the formative research phase?	Behaviour-centred design	Focuses on the short-term behavioural response to project activities for rapid learning of impact during the design phase

Endnotes

- ⁱ E. Plänitz (2020) 'Natural disasters and political disorder: Why urban flooding turns violent. Applying a fuzzy-set qualitative comparative analysis', *Peace Economics, Peace Science and Public Policy*, 26(2), p. 20190046. <https://doi.org/10.1515/peps-2019-0046>
- ⁱⁱ K. Mangam, E. Fiekowsky, M. Bagayoko, L. Norris, A. Belemvire, R. Longhany, C. Fornadel and K. George (2016) 'Feasibility and effectiveness of mHealth for mobilizing households for indoor residual spraying to prevent malaria: A case study in Mali', *Global Health: Science and Practice*, 4(2), pp. 222–237. <https://doi.org/10.9745/GHSP-D-15-00381>
- ⁱⁱⁱ S. Caria, G. Gordon, M. Kasy, S. Quinn, S. Shami and A. Teytelboym (2020) *An adaptive targeted field experiment: Job search assistance for refugees in Jordan*. CESifo Working Paper Series 8535.
- ^{iv} D.R. Marsh, D.G. Schroeder, K.A. Dearden, J. Sternin and M. Sternin (2004) 'The power of positive deviance', *BMJ*, 329, p. 1177. doi:10.1136/bmj.329.7475.1177
- ^v E. Bisung and S. Dickin (2019) 'Concept mapping: Engaging stakeholders to identify factors that contribute to empowerment in the water and sanitation sector in West Africa', *SSM—Population Health*, 20(9), p. 100490. doi: 10.1016/j.ssmph.2019.100490; PMID: 31993485; PMCID: PMC6978480
- ^{vi} J. Copestake, M. Morsink and F. Remant (2019) *Attributing development impact: The Qualitative Impact Protocol (QuIP) case book*. Rugby: Practical Action Publishing.

About this brief

This brief was written by Marcella Vigneri and Howard White. It is partly based on Vigneri, M. (2022). *Timely evaluation in international development. CEDIL Methods Working Paper 7*. London and Oxford: Centre of Excellence for Development Impact and Learning. <https://doi.org/10.51744/CMWP7>

Suggested citation

Vigneri, M., and White, H. (2022). *When time is of the essence: Timely evaluations in international development evaluation*. CEDIL Methods Brief 8. London and Oxford: Centre of Excellence for Development Impact and Learning. <https://doi.org/10.51744/CMB8>

Photo credits

p.3: Alex Tilley, WorldFish, P.4: © European Union 2016 – European Parliament, p.5: UN Women Asia and the Pacific, p.6: ILO Asia-Pacific, p.7: Dominique Mercier, Water Alternatives



@CEDILProgramme



Centre of Excellence for Development Impact



Centre of
Excellence
for Development
Impact and Learning



About CEDIL

The Centre of Excellence for Development Impact and Learning (CEDIL) is an academic consortium supported by the UK Government through UKaid. The mission of the centre is to test innovative methodologies in evaluation and evidence synthesis and to promote evidence-informed development. CEDIL-supported projects fall into three programmes of work: evaluating complex interventions, enhancing evidence transferability, and increasing evidence use.

For more information on CEDIL, contact us at cedil@opml.co.uk or visit our website www.cedilprogramme.org