EFFECTIVENESS REVIEW SERIES

'enhancing effectiveness through evidence-based learning'

HOW ARE EFFECTIVENESS REVIEWS CARRIED OUT?

Every year since 2011, Oxfam Great Britain (GB) has conducted rigorous evaluations on samples of randomly selected mature projects (projects that are a substantial way through implementation or have finished) to help the organisation understand and evidence whether its work is resulting in positive change in the lives of the women and men with whom and for whom it works.

These 'Effectiveness Reviews' consider Oxfam GB's work across six thematic areas:

- Humanitarian Response
- Accountability
- Livelihoods support
- Resilience
- Women's Empowerment
- Good Governance (which brings together "Citizen voice" and "Policy Influence")

Evaluations under the first two thematic areas - humanitarian and accountability - consider the degree to which interventions meet agreed standards. Accountability Reviews consider the extent to which projects have met Oxfam's standards for accountability to partners and communities. Evaluations of humanitarian responses consider the extent to which the response under review met internationally agreed humanitarian standards.

Evaluations of the final four thematic areas – Livelihoods, Resilience, Women's Empowerment and Good Governance – aim to assess the impact of our projects. They consider whether there is evidence of a cause-effect relationship between the intervention and observed outcomes/ impact, even where the intervention is only one factor contributing to that change.

Oxfam GB uses four distinct methodologies to evaluate projects selected as part of the Effectiveness Review series, as no one evaluation approach is appropriate for all programmes and contexts.

An enumerator in the Philippines carrying out an Effectiveness Review 2014/15 interview using a mobile phone. Photo: Emily Tomkys/Oxfam
Top: Climate change adaptation and advocacy project, Nepal, evaluated under resilience in 2012/13. Photo: Jisu Mok/Oxfam

HUMANITARIAN RESPONSE

WHAT WE ARE MEASURING

Evaluations conducted of our Humanitarian Responses consider the extent to which the intervention has met internationally agreed humanitarian standards. These are based on the hypothesis that a good-quality programme will lead to greater impact.

Collecting clean water during the response in Democratic Republic of Congo, evaluated in 2013/14. Photo: Kate Holt/Oxfam

EVALUATION DESIGN

Large-scale humanitarian responses are evaluated using the Humanitarian Indicator Tool (HIT), a methodology designed to estimate the degree to which the programme meets the dimensions of the sector-wide and internationally recognised Core Humanitarian Standard. The HIT has 15 quality standards, each with defined benchmarks, which allow external evaluators to assess and score whether the standard was 'met', "almost met", 'partially met' or 'not met'. Standards include issues such as timeliness; coverage of affected populations; technical aspects measured against Sphere standards; partner relationships; and the needs of women, girls, men, boys and vulnerable groups, and are weighted depending on their relative importance. The HIT is carried out as a desk study by an external evaluator using documented evidence to generate a score against each standard and a cumulative total expressed as a percentage. If there are gaps in the documentation, or clarity is needed, supplementary interviews are carried out by phone or Skype.

ACCOUNTABILITY

WHAT WE ARE MEASURING

Oxfam defines accountability as the process through which an organisation balances the needs of stakeholders in its decision-making and activities, and delivers against this commitment. This reflects Oxfam's belief that *how* work is done is as important as *what* is done.

Accountability Reviews consider the extent to which projects are meeting their accountability commitments to partners and communities. They assess sampled projects that are approximately a quarter to half way through implementation – enough time for accountable relationships to have been developed and with enough time remaining to put learning into practice during the lifetime of the project.

The assignments consider three key dimensions of accountability:

- Transparency – being transparent in our relationships.
- Feedback/listening – that feedback mechanisms exist and lead to change.
- Participation – key stakeholders participate in decision-making and implementation throughout the lifetime of a project.

They also examine the leadership, systems and practices of Oxfam and partner staff, as well as the perceptions of staff, partners and communities. Questions about partnership practices, staff attitudes, and satisfaction (how useful the project is to the people and how wisely the money on the project has been spent) are asked where appropriate.

EVALUATION DESIGN

A series of set exercises are carried out by external evaluators in order to collect qualitative and quantitative data from country and project teams, and partners and communities. Accountability Reviews usually consist of a document review, workshops, on-site focus group discussions with beneficiary communities and structured interviews with Oxfam staff, partners and other stakeholders. The evaluator triangulates this information to reach conclusions on the degree to which Oxfam's work meets its own standards for accountability.

Youth agents of change in Georgia – My Rights, My Voice programme. Accountability Review 2013/14. Photo: Khatuna Khubitia/Oxfam

COMMUNITY DEVELOPMENT PROJECTS

WHAT WE ARE MEASURING

LIVELIHOODS SUPPORT

Projects are evaluated for their impact on overall household income. Asking directly for all the details of a household's income (which includes income from many different sources) and agricultural production (which may be consumed in the household rather than sold for cash; bartering, etc.) would be complicated and prone to error. Instead, measurement is made on the basis of household consumption and expenditure, which are normally found to be closely linked to income in low- and medium-income countries. Survey respondents are asked for details of the food consumed in the household during the seven days prior to the survey, as well as for expenditure on non-food items made by the household in the last month and 12 months. These data are aggregated to provide an indication of the household's overall consumption per person, per day.

Respondents are also asked about their ownership of various types of household goods and assets, and the condition of their housing. These data are then used to create a wealth index.

Effectiveness Reviews that focus on livelihoods support also include measures of the projects' impacts on intermediate outcomes. Depending on the nature of the project this may, for example, involve looking at the extent to which it has resulted in changes in the adoption of improved agricultural practices, the level of agricultural production or sales, the profitability of household businesses, or households' saving behaviour or use of credit. Examining these intermediate outcomes enables the evaluation to identify which aspects of the project's theory of change have been successful, and where further intervention may be needed.

Increasing smallholder producers' access to markets in Colombia, evaluated under livelihoods in 2013/14. Photo: Rob Fuller/Oxfam

Drought Management Initiative: Livestock, Kenya. Evaluated under resilience in 2012/13. Photo: Alejandro Chaskielberg/Oxfam

RESILIENCE

These Effectiveness Reviews seek to understand the extent to which Oxfam projects have enabled people to minimise risk from shocks (whether natural disasters, conflict, or more local-level or household-level crises) and adapt to emerging trends and uncertainty (such as changing climate conditions). A full evaluation of the outcomes of increased resilience can be conducted only after shocks, stresses and uncertainty have been experienced, where people's and communities' actual responses can be assessed. In practice, Effectiveness Reviews are often carried out in the absence of shocks, meaning that such an assessment cannot be made. Instead, the evaluation team seeks to measure context-specific indicators for characteristics that are thought to be associated with resilience in the context in question. A framework of five dimensions is used to categorise these characteristics:

- **Livelihood viability** – the extent to which a household's livelihood can continue and thrive in spite of shocks, stresses, and uncertainty.
- **Innovation potential** – the ability to take appropriate risks and positively adjust to change.
- **Access to contingency resources and support** – access to back-up resources and appropriate assistance in times of crisis.
- **Integrity of the natural and built environment** – the health of local ecosystems, natural resource management practices, and presence of appropriate physical infrastructure.
- **Social and institutional capability** – the extent to which formal and informal institutions are able to ensure equitable access to essential information, services, and resources to reduce risk and support positive adaptation.

Safe Age of Marriage and Women's Economic Empowerment in Yemen, evaluated in 2012/13. Photo: Wolfgang Greesman/Oxfam

WHAT WE ARE MEASURING

WOMEN'S EMPOWERMENT

These Effectiveness Reviews seek to understand the extent to which Oxfam projects have contributed to women's empowerment, broadly defined as a process whereby a woman is transformed from a situation with limited power to a situation where power is enhanced. Oxfam recognises women's empowerment to be a complex, multi-dimensional concept. While not arguing for a standard set of women's empowerment indicators that are applicable to all contexts, Oxfam has developed a multi-dimensional index to support measurement of this hard-to-measure area that sets out five dimensions of women's empowerment that the organisation considers to be important in all contexts.

Women's economic leadership through horticulture planting-material business, Rwanda. Evaluated in 2013/14. Photo: Simon Rawles/Oxfam

The framework is structured as follows:

- ***Personal*** – this level of change includes characteristics of empowerment taking place within the person. It is divided into two dimensions:
  - **Power from within**:
    This dimension includes measures of change in personal self-confidence and self-esteem, personal opinions, personal attitudes and believes.
  - **Power to**:
    This dimension includes measures of changes in personal autonomy and individual capability to decide actions and carry them out.
- ***Relational*** – this level of change includes characteristics of change taking place in the power relations of the woman's network. It is divided into two dimensions:
  - **Power with**:
    This dimension includes measures that recognise empowerment as a collective *process*, which requires the support and interaction of other peers and organisations. It therefore includes measures of social capital, and participation in community groups.
  - **Power over**:
    This dimension measures changes taking place in the power relationship between individuals, including changes both within the household and within the community.
- ***Environmental*** – this level of change includes measure of change taking place in the broader environment.

While the framework is fixed, context specific indicators are developed under these dimensions in a participatory manner with local Oxfam staff, partners and communities in order to capture what women's empowerment means in the particular socio-economic contexts under analysis. These are then combined into a composite index, which provides an overall measure for empowerment. The framework recognises three possible levels where change can take place: at individual level; in the relationship with others; and changes in the broader environmental. These levels of change are then combined within a structure of power (within, to, with, and over).

In cases where they are not directly included in the overall measure of empowerment, Effectiveness Reviews selected under Women's Empowerment also seek to investigate unexpected project effects such as incidences of violence against women or unpaid care work. A specific module has been developed to measure the latter, which includes how time is used, the intensity of care work and the redistribution of activities or responsibilities within the household.

COMMUNITY DEVELOPMENT PROJECTS

EVALUATION DESIGN

The impact of development projects that are implemented at an individual, household or community level is assessed through the use of ex post quasi-experimental impact evaluation techniques, to assess the level of impact a project has had. Quasi-experimental impact evaluation methods involve using individual or household surveys to collect data from a sample of project participants (referred as the intervention group), as well as from a sample of individuals or households with similar characteristics that did not receive direct or indirect support from the project (referred as the comparison group). The intervention and the comparison groups are then compared using statistical tools of propensity-score matching and multivariate regression in order to estimate differences in outcomes that can be causally attributed to the project.

Much thought goes into selecting appropriate comparison groups, and methods for selecting a comparison group necessarily vary for each evaluation. In many cases this involves visiting nearby communities with similar characteristics to those where the project was implemented, considering the projects' originally targeting criteria and/or comparing between communities or households that have been selected for participation in a different phase of the project. Particular attention is paid to reducing self-selection bias. This is done by considering observable characteristics (such as geographic and socio-economic status) as well as unobservable characteristics (such as willingness, availability and desire to participate into development projects). The statistical tool of propensity-score matching provides additional confidence that the two groups (intervention and comparison) do not differ based on a number of observable recalled baseline characteristics.

Wherever possible, quantitative household surveys are accompanied by qualitative interviews or focus group discussions, in order to provide more depth of understanding about the changes resulting from the project in question. Where appropriate, individual and household surveys are conducted using mobile phones or tablets. This improves the efficiency and precision of the data collected, as well as allowing the prompt sharing of initial summary statistics with the communities involved in the data collection. This approach has proved to be a valid tool for data validation and community feedback.

This method is used to estimate what impact the project in question has had in terms of Oxfam's global outcome indicators, as well in other outcomes relevant to the specific project.

Women's beekeeping and access to financial services, Ethiopia. Evaluated under Women's Empowerment in 2013/14. Photo: Tom Pietrasik/Oxfam

Women's citizenship for change in Bolivia, evaluated in 2013/14. Photo: IFFI

SMALL N PROJECTS (often under the Good Governance thematic area)

WHAT WE ARE MEASURING

Interventions undertaken under the good governance thematic area seek to empower citizens and support them and their organisations in influencing decisions that affect their lives. Evaluations consider intermediate and final outcomes that are central to the selected project's theory of change. These might, for example, include increased ability to participate meaningfully in governance processes, changes in the practices of target state institutions, or policy change.

EVALUATION DESIGN

Such interventions often work to achieve specific intermediary and final outcomes where there are too few units of assignment to permit tests of statistical significance between intervention and comparison groups. In order to assess the effectiveness of this work, Oxfam draws from a theory-based non-counterfactual evaluation approach designed to investigate causal inference, known as process tracing, and has developed an evaluation protocol that is implemented by external researchers/evaluators. Process tracing supports the assessment of a project's contribution to achieving defined outcomes by: investigating whether and to what extent the outcome has materialised; considering the evidence supporting different "causal stories" or "explanatory hypotheses" to build an understanding of "how change happened", with the intervention just one possible contributing factor; and, where the evidence supports a causal relationship between the intervention and the outcome, determine the significance of the intervention's contribution in light of other contributing factors.

The approach should accomplish three things:

1. Shortlist one or more evidenced explanations for the outcome in question.
2. Rule out alternative competing explanations incompatible with the evidence.
3. If more than one explanation is supported by the evidence, estimate the level of influence each has had on bringing about the change in question.

Health for All campaign in Ghana, evaluated in 2012/13. Photo: Alliance for Reproductive Health Rights (ARHR)