Starting shortly

Please wait!

ActivityInfo

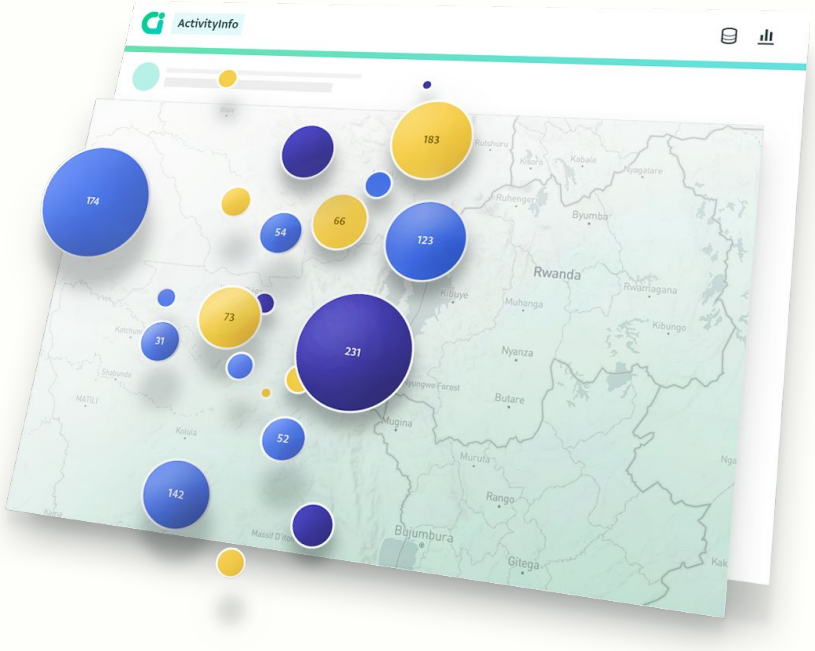
Implementation of Evaluation in Humanitarian Action

ActivityInfo

Presented by the ActivityInfo Team

Software for Monitoring & Evaluation

- ★ Track activities, outcomes
- ★ Beneficiary management
- ★ Surveys
- ★ Work offline/online



ActivityInfo

Webinar series outline

11/05

First Webinar

Getting started with Evaluation in Humanitarian Action

8/06

Second Webinar

Evaluation planning and design in Humanitarian Action

6/07

Third webinar

Implementation of evaluation in Humanitarian Action

ActivityInfo

Today's session outline

- ***Introduction***
  - *Reminder! Key information needed for today's session*
- ***Evaluation designs***
  - *Which are the different design categories?*
  - *How can we select the most appropriate design?*
  - *Cross-cutting considerations: engagement with the affected population*
  - *Cross-cutting considerations: the importance of bias*
- ***Sampling***
  - *What are common sampling techniques?*
  - *What are the challenges in sampling in Humanitarian action?*
- ***Field methods***
  - *Which are the most commonly used field methods? qualitative and quantitative field methods*
- ***Analysis***
  - *Which is the recommended analysis based on the type of evaluation question?*
- ***Case study: evaluation of education programme in Western Tanzania - how to use***
- QandAs

Key information to remember for this section

The evaluation questions determine the evaluation design, the data-collection and analysis methods, and the sampling approaches. Evaluation has two main purposes; accountability and learning.

Types of questions:

- Descriptive**:** How did affected people use the shelter kits provided?
- Normative: To what extent did the shelter provided meet the Sphere Standards?
- Causal: To what extent did the provision of assistance at the village level discourage migration to the regional capital?
- Evaluative: Was our policy to only provide shelter kits to those with proof of plot ownership appropriate?
- Action- oriented: How could we better support vulnerable persons to rebuild their shelters?

Qualitative versus quantitative: quantitative methods collect numerical data and qualitative methods collect non-numerical data:

- Quantitative: Produce numerical values as they measure amount or quantity.
- Qualitative: intend to explore and describe judgments, opinions, perceptions and attitudes toward a given situation or subject

Triangulation: findings are based on multiple sources

Steps from evaluation design to analysis

Basic Steps

1. Choose of appropriate evaluation design
2. Choose appropriate sample strategy and sample size
3. Identify the appropriate field method
4. Identify how best to analyze the results

ActivityInfo

Evaluation design

Evaluation design

Types of evaluation designs

01 Experimental designs

- Assignment to the assisted or control groups is done before the start of the assistance with **random assignment**
- Ethical considerations due to random assignment
- Considered as the most rigorous design
- Randomised control trial (RCT) the most common design

random assignment: each unit (e.g person) has equal chance of being assigned to the assisted or control group

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Types of evaluation designs

02

Non experimental designs

- With neither comparison nor control groups
- The most common type of design in EHA - need to move beyond using only this type of designs
- Example: case studies

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Types of evaluation designs

03 Quasi-experimental design

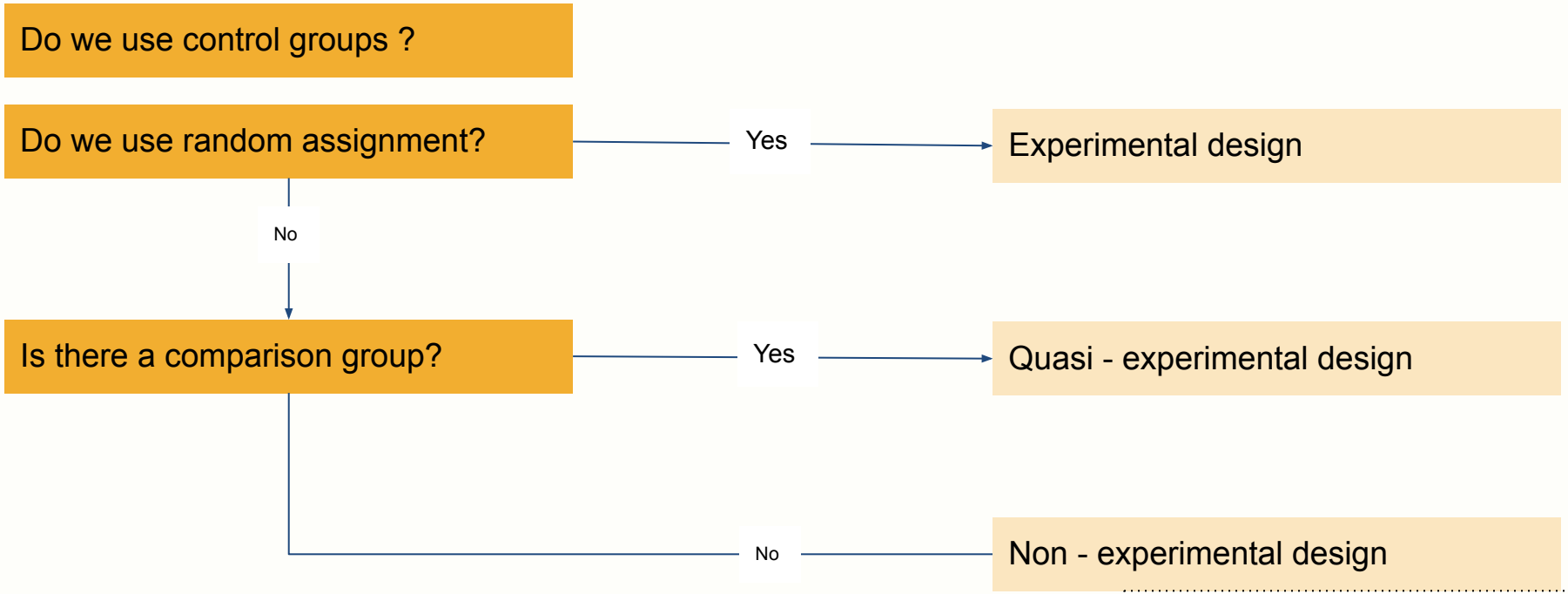
- Comparisons are made either of the assisted group over time, or between the assisted and a **comparison group** selected after the start of the assistance.
- A single group is compared over time
- Main strength: provide rigorous evidence, while avoiding the ethical problems of experimental designs

Comparison groups: are very susceptible to experimental contamination due to the large number of actors and other support networks. This applies both to experimental and quasi-experimental designs

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Types of evaluation designs: Key questions driving the decision



https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Example of an evaluation design: comparison group

Definition: Comparison group designs compare the assisted group with a selected comparison group. This is a quasi-experimental design

Use in EHA: Weak comparison groups are sometimes found in EHA. Stronger comparison groups can be established.

Strong points:

- Comparison groups reduce the risk of mistaking background changes for the impact of the intervention

Weak points:

- Very difficult to avoid contamination in humanitarian settings
- Strong statistical skills and good data on the assisted group needed for methods such as propensity matching

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

The bias: Key considerations

What: Threat to the accuracy of the evaluation findings
Source of bias: The choice of design, methods and sampling approaches influences the potential types of bias

Selection bias: occurs when the sample elements chosen are biased in some way. One example would be if beneficiary distribution lists were used as the sampling frame from which to select a sample

Bias in data collection: gender bias - for example only men field staff interviewing women respondents

Evaluator bias: Evaluator biases may include a dislike of a particular agency or programme approach, or the temptation to repress highly critical findings for fear of losing future contracts with the client

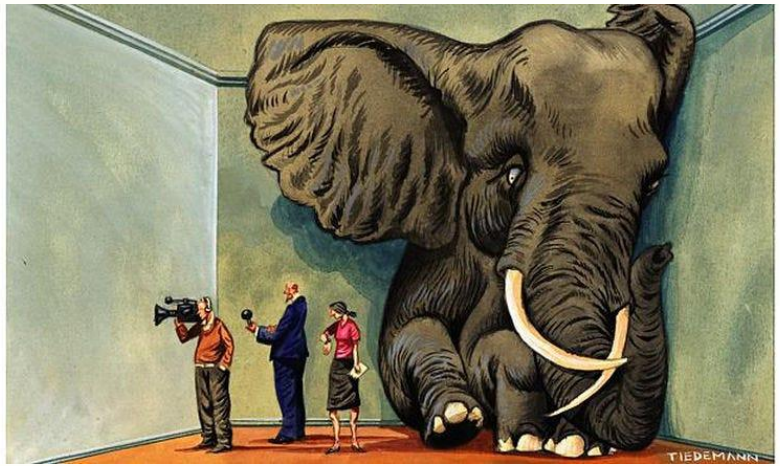
https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

The bias: Key considerations

Example:
Organisations are more likely to evaluate what are perceived to be successful programmes than failed ones

The 2015 State of the Humanitarian System report states that
"*Over 50% of evaluations rate the performance of the evaluation subject as good*"

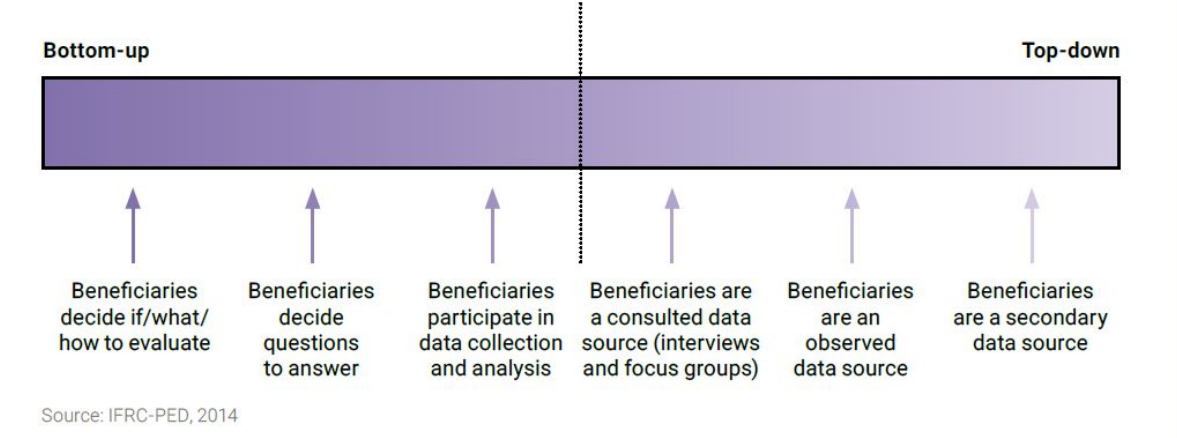


https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Engagement with the affected population

The 2015 State of the Humanitarian System reported: Greater awareness at the field level of the importance of engaging with affected people…so that conflict- and disaster-affected populations are not seen “purely as recipients”, and that interventions are designed to centre more on their needs and preferences.



Source: IFRC-PED, 2014

ActivityInfo

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Engagement with the affected population

Evaluation Planning:

Planning for consultation with them during the evaluation process

Evaluation Design

- identify which evaluation questions are relevant for consultation with the affected population
- review what already exists in terms of feedback from the affected population
- consider which groups among the affected population are to be consulted, and thus how the population is to be disaggregated

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation design

Engagement with the affected population

Good practice example: Listening to affected people about their experience of agency efforts to be accountable to them

Through research carried out between November 2014 and February 2015 as part of the Pamati Kita Common Services project, communities affected by Typhoon Haiyan were consulted about their perspectives and experiences of humanitarian agencies' efforts to be accountable to them. The consultations revealed that affected people had a strong preference for face-to-face communication over more technological means of communication, which had been favoured by many agencies.

Sources: Ong et al. (2015); Buchanan-Smith et al. (2015)

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Evaluation Design

Key messages

- The design is determined solely by the evaluation questions, and no evaluation design is perfect. The constraints imposed by timing, budget, data availability, and so on limit the options. The options chosen, and the reasons for doing so should be noted in both the inception and final reports
- The choice of design, methods and sampling approaches influences the potential types of bias. We need to acknowledge the bias and the potential limitations associated with this.
- Consider how best to engage with affected population from the planning stages of evaluation

Sampling

Sampling

What?

What?

The selection of a subset of a population for inclusion in a study instead of the entire population

Decide the sampling strategy

Decide the sample size

ActivityInfo

Sampling

Strategy: From simple to complex

Simple case:
sample is drawn by convenience from the affected population

Complex case:
samples can be randomly drawn from different social strata

Good enough approach

Non-random sampling for qualitative data collection
Random sampling for quantitative data collection

ActivityInfo

Sampling

Random Vs non random sampling

Non-random sampling selects the sample based on some property of the sample.

Important: non-random samples are not representative of the population as a whole

Random sampling draws a sample from a population where each member of the population has an equal chance of being selected

Important: samples are said to be representative

Sampling

Non random (non probability) sampling in detail

Appropriate for:

- Qualitative methods
- Explanatory research
- Quick decisions and/or targeting specific cases
- For cases that time and money are limited
- For cases that access is limited
- Providing illustrative examples
- Target size is small

Not appropriate for:

- For cases that results generalization is needed

Sample size:

Creswell and Clark (2011) suggest that four to six cases are adequate for a case study and 20-30 interviews for an interview-based study

Sampling

Non random sampling: examples

Purposive sampling
Selects the sample based purposively so that the sampled elements can provide the most information for the study.

Prons: inexpensive, simple, irrelevant responses are filtered out to give you a focused perspective

Cons: Because of the specificity, the validity of the results is questionable, not effective at large scale.

Sampling

Random (probability) sampling in detail

Appropriate for:

- large-n methods where there is a need to generalise from the sample to the whole population

Not appropriate for:

- For cases that we have access in small sample size - highly misleading estimate of the situation

True random sampling requires a sampling frame and a list of the whole population from which the random sample can be drawn. In the absence of such list we use pseudo random sampling

Estimating the sample size need depends on:

- Whether a sample is being used to generalise about a population or to compare two populations

Sampling

Random sampling: examples

Cluster sampling
is sampling where a number of locations are sampled, each with a cluster of a particular number of cases.

Prons: it can be used without a sampling frame for individuals in the population.

Cons: larger overall sample

Sampling

Mixed method case

Mixed methods require both random and non-random sampling approaches, but whether they use large-n or small-n methods, humanitarian evaluators should explain their sampling strategy and choices. They should also explain any potential biases or other limitations inherent in their choices.

Why are those used in EHA?
The different strengths of large-n and small-n methods to indicate both what happened and why, explain why the best EHAs use mixed methods.

Sampling

Common sampling challenges in humanitarian action

- Over-reliance on availability or convenience sampling in small n-studies.
- Insufficient random sample size to enable statistically valid generalisations from the sample.
- Inappropriate use of random sampling in small-n studies.
- Failure to make the sampling approach clear in the evaluation reports, especially when small-n methods are used.

Sampling

Key messages

- Almost every evaluation uses sampling. even the smallest evaluation can make good sampling choices to improve the quality of the evaluation.
  - Purposive sampling is probably the strongest sampling type for small-n studies because the members of the sample are deliberately chosen for the knowledge they can contribute to the research.
  - True random sampling requires a sampling frame and a list of the whole population from which the random sample can be drawn. In the absence of such list we use pseudo random sampling.
- Consider the use of mixed methods whenever this is possible.

Field Methods

Field methods

Initial considerations

- The evaluation questions determine the methods used
- Practical considerations of budget, time, ethics, and logistics may prevent the most appropriate methods being used for a particular question
- Evaluations can use a variety of methods to measure what they want

Important: It is important not to confuse designs and methods. Design refers to the structuring of the data gathering and analysis, and method refers to how the data is gathered.

Field methods

Broad field methods categories

Interviewing can be divided into:
- Structured interviews
- Semi-structured interviews
- Unstructured interviews

Observation
- Structured
- Unstructured

Learning oriented methods
- Story-telling
- Most significant change

Unobtrusive measures
- Monitoring of page visits
- Social media

Field methods

Most commonly used field methods categories

Interviewing can be divided into:

- Structured interviews
- Semi-structured interviews
- Unstructured interviews

Qualitative interview methods

- Key informant Interviews
- Group Interviews/Focus Groups Discussions

Survey methods

- Face-to-face surveys
- Online surveys

Household interviews

- Both quantitative and qualitative

Field methods

Methods in the Spotlight

Focus group discussions

- Based around a topic guide with 3-5 discussion topics.
- Most effective with partner agency staff, but can also be used with affected population.
- Generates qualitative data, good for exploring the views of particular groups
- Need a comfortable, controlled space.
- Need a facilitator and a note-taker.
- Samples may be randomly drawn from within groups to form the focus group. The group must be similar to each other without status or other differences.
- Medium cost.

Field methods

Methods in the Spotlight: Focus groups discussions

Good practice example: World Vision's use of focus group discussions

"For Focus Group Discussions (FGD), we typically have a team leader on every team, though on occasion we are unable to do that so we use the regional structure….. But our FGD teams are comprised of four people at a bare minimum, and up to six: team leader, translator, two note takers, observer and facilitator. The observer and facilitator are opposite genders and switch roles as facilitator depending on the gender of the FGD; the observer, team leader and translator can also jump in to become note takers if the discussion is very lively."

Source: Kathy Duryee, WVI, personal communication, 2014

Field methods

Methods in the Spotlight

Surveys

- Use a survey instrument that the interviewers follow rigidly.
- The duration of each interview depends on the length and complexity of the survey instrument.
- Gather quantitative information – good for answering ‘How many?’ questions. Can also ask closed qualitative questions as open questions are very demanding in terms of coding and analysis.
- Data recording on a survey form or digital device. Electronic recording is far superior.
- Interviewees are randomly selected so that the survey results can be generalised to the whole population. Every departure from true random sampling (e.g. pseudo-random or cluster sampling) increases the sample size needed to give statistically valid results.
- Expensive because of the need to test and validate the survey and large number of interviews needed

Field methods

Methods in the Spotlight

Good practice example: World Vision International practice on data-collection tools
Once the English version of data-collection tools is ready, the subsequent process is followed:
1. A translator from the region where the survey will be administered is enlisted to translate them into the local language or dialect.
2. The tools are back-translated. This is not always possible, but is an ideal.
3. A training session is held with enumerators, during which they verify the accuracy of the translation, and ensure they all understand the questions in the same way. This is a very iterative process.
4. Also during enumerators' training, a field test is done in communities using the same language as the survey tools.
5. Following the field test, all survey responses are reviewed and any final changes to translations are made.
Source: Kathy Duryee, WVI, personal communication, 2014

Field methods

Key messages

- The evaluation questions determine the methods used
- Practical considerations of budget, time, ethics, and logistics may prevent the most appropriate methods being used for a particular question
- Evaluations can use a variety of methods
- It is crucial to account for enumerators training and tool field testing prior to data collection

Analysis

Analysis

What?

Analysis should be a rigorous and logical process that turns the evaluation data into findings

Primary Data: Primary data is data collected for the purpose of the evaluation
Secondary Data: Secondary data is data collected for other purposes but is used by the evaluation.

Analysis

Broad categories of analysis

Big-n or quantitative data → statistical analysis

Small-n or qualitative data → coding

Analysis

Broad categories of analysis

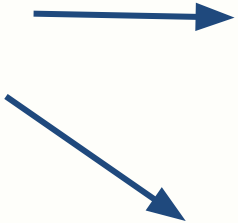
Coding

- Formal qualitative data analysis usually depends on attaching categories to particular pieces of evidence and then drawing those pieces of evidence together
- In EHA the codes usually take the form of keywords and complex codes are seldom used

Analysis

Broad categories of analysis

Statistical Analysis



Descriptive statistics

Used to summarise key aspects of a population.

Inferential Statistics

used either to make inferences about a population from a sample, or to make inferences about hypotheses

Analysis

Type of analysis

Normative Questions

- Use of triangulated evidence of breaches of or compliance with standards.
- Break down the relevant norm or standard applicable into separate elements.
- Examining the extent to which an intervention met each of the standards makes it possible to answer the overall question.

Analysis

Type of analysis

Descriptive Questions

- Potential use of coding
  - to identify overarching themes and issues in the intervention
- Potential use statistical methods
  - describe the intervention by identifying averages and so on.

Analysis

Type of analysis

Evaluative questions

- break them down to separate descriptive, normative, and other elements from the purely evaluative element
- frequent use of evaluative rubric

Performance	Excellent	Very good	Good	Adequate	Poor
Descriptors of performance	Clear example of exemplary performance or best practice for food distribution; no weaknesses.	Very good or excellent performance in virtually all aspects. Strong overall but not exemplary. No weaknesses of any real consequence.	Reasonably good performance overall, might be a few slight weaknesses but nothing serious.	Fair performance, some serious but not fatal weaknesses.	Clear evidence of unsatisfactory functioning; serious weaknesses across the board on crucial aspects.

ActivityInfo

Analysis

Type of analysis

Causal questions

- Causal inference: establishing of a relationship between a cause and an effect.
- Use of inferential statistics usually

Strategies:

(1) Ask those who have observed or experienced the causation first hand.
(2) Check if the content of the intervention matches the nature of the outcome
(3) Look for distinctive effect patterns
(4) Check whether the timing of outcomes makes sense
(5) Look at the relationship between ‘dose’ and ‘response
(6) Use a comparison or control
(7) Use statistical models to eliminate other potential causative factors
(8) Identify and check the causal mechanisms - Theory of change

Analysis

Key messages

- Analysis should be a rigorous and logical process that turns the evaluation data into findings
- Statistical analysis is recommended for big - samples and coding for small samples (qualitative information)
- Statistical analysis can take various forms - the most frequent categorization is descriptive and inferential statistics

Case Study

Case study

Motive

- Research suggests that using PDAs (Personal Digital Assistants-small size computing devices) has the potential to reduce the logistics burden, cost and error rate of data collection.
- Mobile phones offer similar advantages to PDAs in reducing recording and entry errors.
- This also allows for real-time quality control and supervision of the enumerator, to reduce the risk of data fabrication.

https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide

Case study

An evaluation of the education programme in refugee camps in Western Tanzania

Purpose: to evaluate the Education Programme in the refugee camps in Western Tanzania. Specifically the consultancy assignment was concerned with assessment of the implementation process, quality of education, impact and efficiency of educational inputs in each camp.

Context: UNICEF is one among the key actors in ensuring that quality education is provided as a basic human right for every child. UNICEF has been supporting the provision of basic and non-formal education in refugee camps in Western Tanzania since 1994.

In each camp there is an established education system. Both formal and non formal education are provided.The formal education comprises primary education and post primary education while pre-schools, vocational training and adult education fall under non formal education. As of February 2004, there were 125,188 children enrolled in primary schools in refugee camps. The teacher pupil ratios ranged from 1:48 in Mkugwa to 1:111 in Lugufu II.

https://www.alnap.org/system/files/content/resource/files/main/tanzania-2004-004-education-refugee-camps.pdf

Case study

An evaluation of the education programme in refugee camps in Western Tanzania

Methodology

Evaluation design: Non experimental approach

The evaluation of the education programme adopted both quantitative and qualitative approaches.

The quantitative approach helped establish the status of education based on indicators such as enrolment, teacher pupils ratio and performance in both school-based and regional examinations.

The qualitative approach was used to gain insights into the approaches and processes used in carrying out education activities in refugee camps.

https://www.alnap.org/system/files/content/resource/files/main/tanzania-2004-004-education-refugee-camps.pdf

Case study

An evaluation of the education programme in refugee camps in Western Tanzania

Field methods

Documentary analysis: Several documents were reviewed to gain understanding and perspectives of the environment in which education is provided in the camps.

Questionnaire: A questionnaire was developed to solicit statistical information from schools. It was meant to be completed by head teachers and pre-school coordinators

Interviews: Interviews were held with two NGOs implementing education activities

https://www.alnap.org/system/files/content/resource/files/main/tanzania-2004-004-education-refugee-camps.pdf

Case study

An evaluation of the education programme in refugee camps in Western Tanzania

Analysis used

- Descriptive statistics to present to population profile
- Descriptive statistics for core indicators (enrollment rate, attendance etc.)
- Qualitative analysis (coding) was used to identify underlying factors for children dropout and identify core aspects of school environment.
- Qualitative analysis and descriptive analysis was used to illustrate teachers' capabilities.

https://www.alnap.org/system/files/content/resource/files/main/tanzania-2004-004-education-refugee-camps.pdf

Case study

How can we use ActivityInfo?

Data collection for Monitoring purposes

→

Real - time monitoring reports

Facilitate targeting and sampling for evaluation

→

Data collection for evaluation purposes

↓

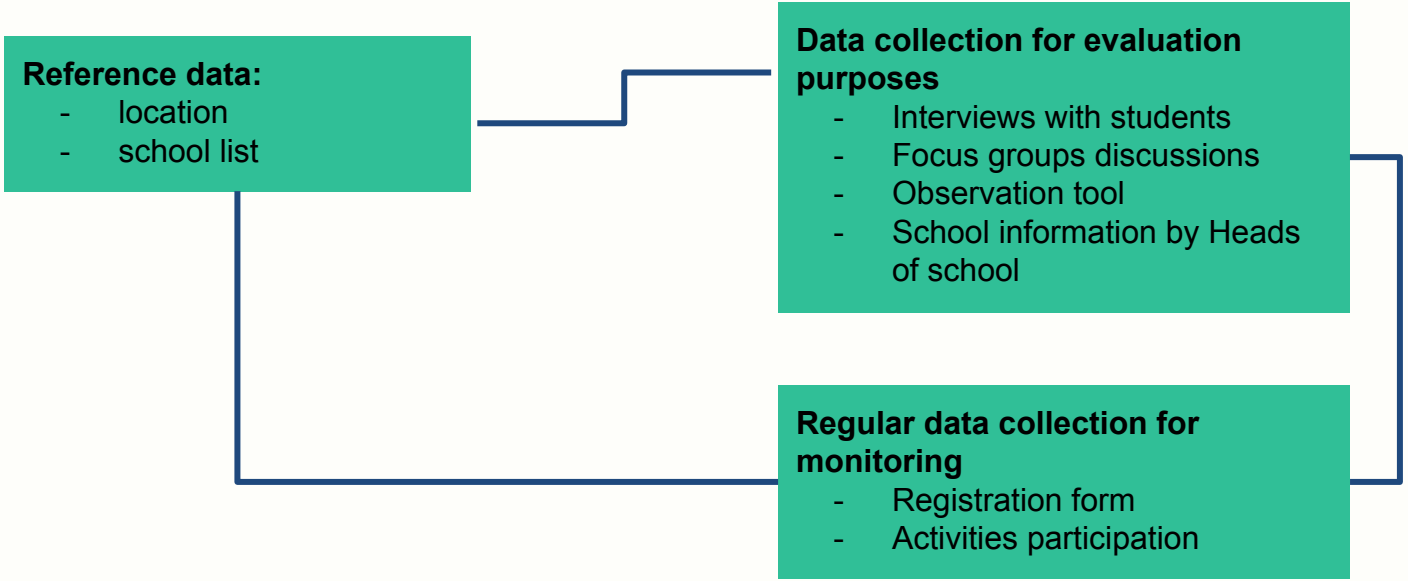
Real time report evaluation reports

Facilitate real time monitoring of the implementation of the evaluation and descriptive analysis

ActivityInfo

Case study

How this is operationalized in ActivityInfo?



ActivityInfo

Case study

What are the advantages?

- Streamline of data collection from monitoring to evaluation
- Use of monitoring information to better specify:
  - evaluation questions
  - as a secondary data source
  - facilitate sample choice and targeting
- Facilitates evaluation implementation monitoring at real time
- Enables collection of responses via collection links
- Facilitates the descriptive analysis of evaluation data collected
- Facilitate the inferential analysis via the integration with R

Implementation of evaluation in the Humanitarian sector

Resources

- *Evolving evaluation practice: past, present and future challenges*
  https://www.alnap.org/help-library/evolving-evaluation-practice-past-present-and-future-challenges
- *State of humanitarian system*
  https://sohs.alnap.org/help-library/2022-the-state-of-the-humanitarian-system-sohs-%E2%80%93-summary
- *Evaluation of Humanitarian Action Guide*
  https://www.alnap.org/help-library/evaluation-of-humanitarian-action-eha-guide
- *Using Evaluation for a Change: Insights from humanitarian practitioners*
  https://www.alnap.org/help-library/using-evaluation-for-a-change-insights-from-humanitarian-practitioners
- *Strategies on causal attribution*
  https://www.unicef-irc.org/publications/pdf/brief_6_overview_strategies_causal_attribution_eng.pdf
- *Evaluation of education programme in Western Tanzania*
  https://www.alnap.org/system/files/content/resource/files/main/tanzania-2004-004-education-refugee-camps.pdf

Time for Q&A!

Thank you!