Brief Introduction into Monitoring and Evaluation

Goals – Requirements – Designs – Preparation & Implementation – Assessment
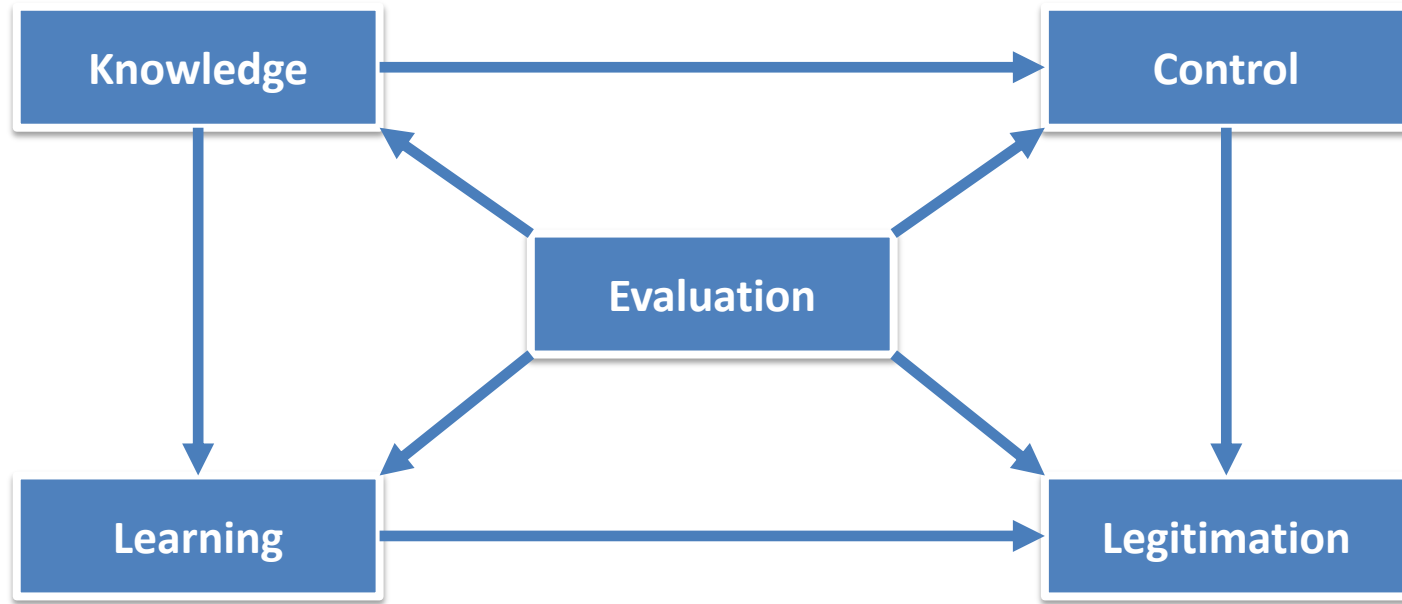
Dr. Stefan Silvestrini

bengo Thementag

Online Workshop, November 21, 2023

Centrum für Evaluation

CEval

Center for Evaluation

Goals of Evaluation

Knowledge	→	**Control**
	Evaluation	
Learning	→	**Legitimation**

© Stockmann, 2006

Questions:
- ✓ Primary purpose
- ✓ Selection of evaluation criteria
- ✓ Tasks of the evaluation
- ✓ Resposibilities
- ✓ Design and methodology

Seven Questions when Planning an Evaluation

✓ **Who** needs the information (for decision-making)?

✓ **When** is the information required?

✓ **Which** information/data is required, which indicators are available?

✓ **Who** has to collect the data, who is responsible for providing it?

✓ **How** should the data be collected and analyzed?

✓ **Why**, i.e. **for which purpose** is the information required, which decisions will be drawn upon their basis?

✓ **How much** will it cost to provide the information, **which resources** are required to collect the data?

CEval GmbH

Acceptance

Appropriate commu-nication

Formulation of distinctive objectives

Embedment in intervention structure

Assurance of results quality

Evaluation Standards

Selection of Evaluation Designs

- ✓ Experimental and quasi-experimental designs
- ✓ **Pipeline designs**
- ✓ **Regression discontinuity designs**
- ✓ Panel and time-series designs
- ✓ Qualitative designs (e.g. Grounded Theory Method, Qualitative Comparative Analysis)
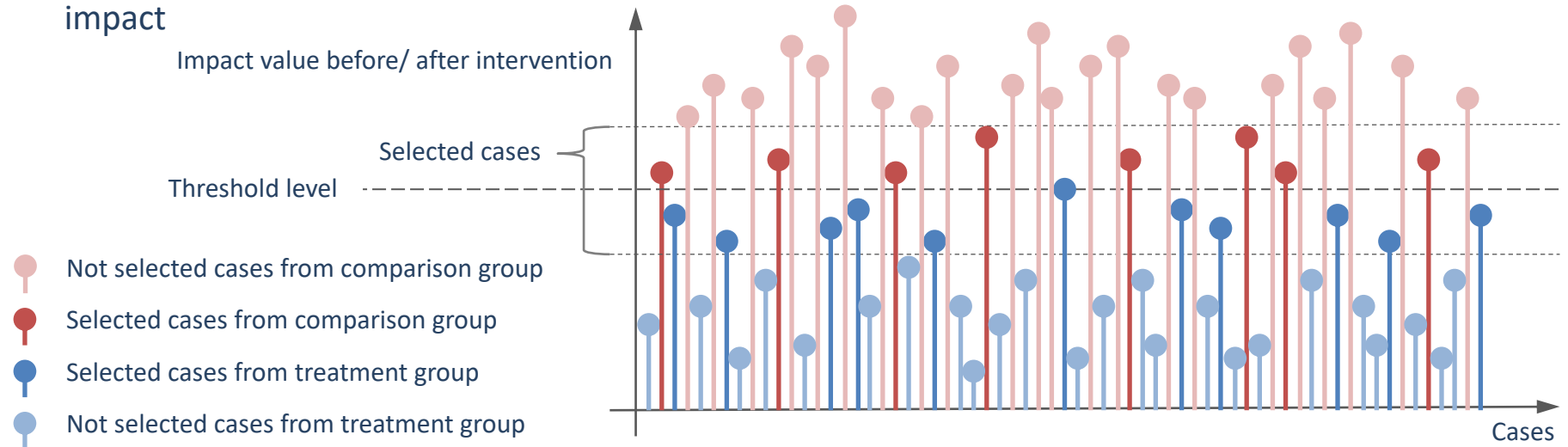
To be distinguished:

- ✓ Terms occasionally referred to as 'evaluation designs' such as e.g. Rapid Rural Appraisal, MAPP, Most Significant Change
- → **Not evaluation designs in a methodological sense!**

Pipeline Design

✓ Applicable for evaluating interventions with measures being implemented in several phases with similar target groups in different regions/ communities etc.

✓ Requires random selection of target groups/ regions to implementation phases and relative stability of environmental conditions during the intervention period

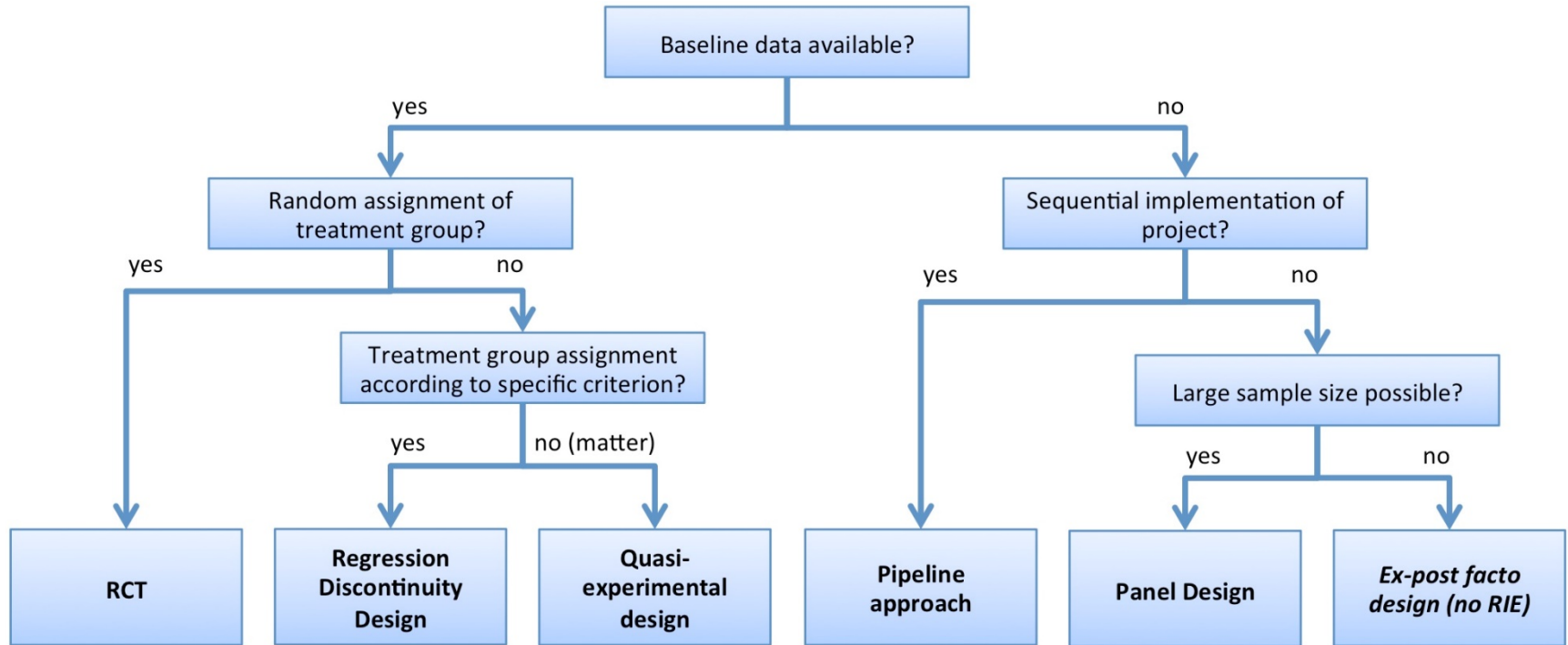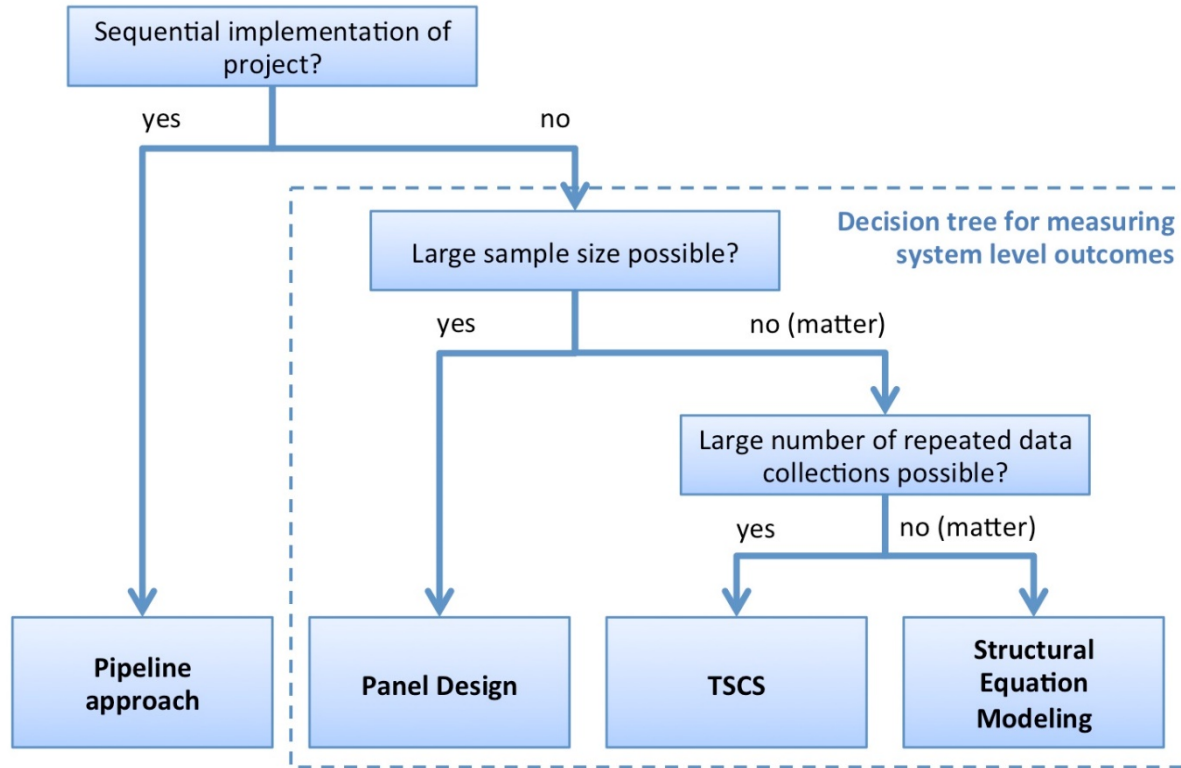✓ Comparison group of the first phase becomes target group of the second phase etc.

Regression Discontinuity Design
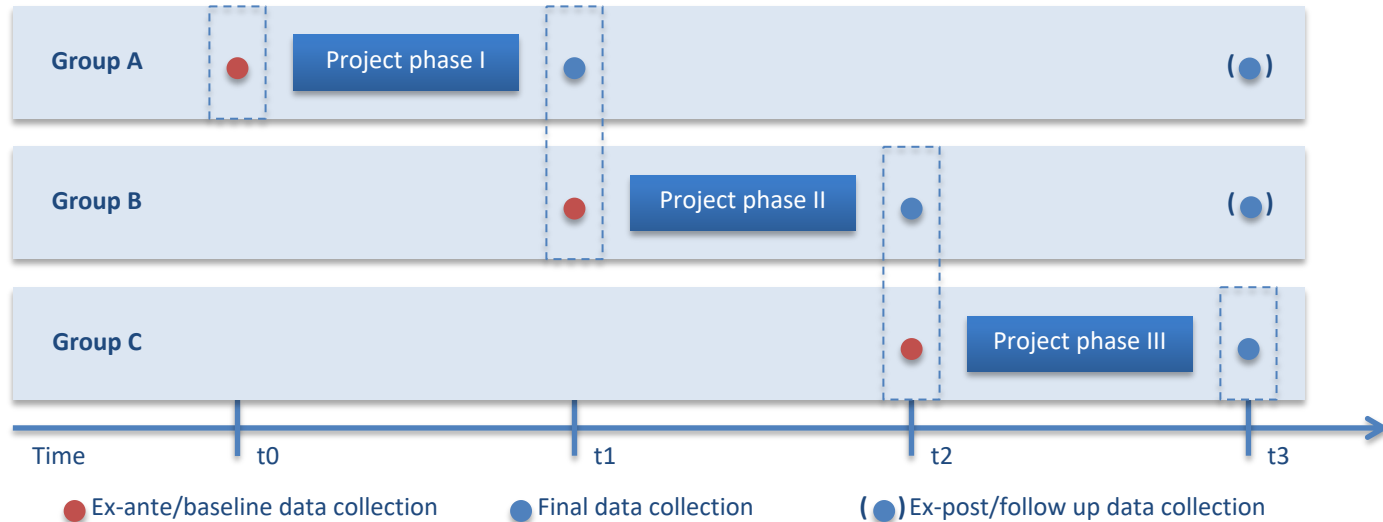
✓ Applicable for evaluating interventions whose target groups are selected based on a certain threshold level

✓ Comparison of observable changes over time of target group and comparison group members with impact values close to this threshold level

✓ Allows only the measurement of 'local effects', close to the threshold level

✓ Danger: If non-linear regression functions are applied, curve progression can be misinterpreted as impact

Impact value before/ after intervention

Selected cases

Threshold level

Not selected cases from comparison group

Selected cases from comparison group

Selected cases from treatment group

Not selected cases from treatment group

Cases

Decision Tree for Selecting Evaluation Designs

Decision Tree for Selecting Evaluation Designs

Preparing and Implementing an Evaluation

CEval GmbH

Planning phase	Implementation phase	Reporting phase

Planning phase
- ✓ Definition of evaluation subject and goals
- ✓ Identification of addressees and further stakeholders

  _____

- ✓ Development of evaluation questions
- ✓ Selection of an appropriate evaluation design

Implementation phase
- ✓ Development of an analysis grid and data collection plan
- ✓ Selection and development of data collection instruments and analysis methods
- ✓ Document/ secondary analysis
- ✓ Empirical data collection

Reporting phase
- ✓ Data cleaning and aggregation
- ✓ Data analysis
- ✓ Results assessment, drawing of conclusions and recommendations
- ✓ Documentation and presentation of evaluation results

  _____

- ✓ Follow-up and management response

Scoping an Evaluation

✓ Defining evaluation subject and objectives

✓ Necessity of a temporal, spatial and organizational scoping of the evaluation subject and of the evaluation results

✓ Identification of stakeholders and addressees

✓ Developing 'Terms of Reference'

    ✓ Description of the scope of the evaluation, its addressees and further stakeholders

    ✓ Outline of the rationale of the evaluation as well as its tasks and goals and goals

    ✓ Definition of the analysis dimensions, main evaluation questions and sometimes also evaluation criteria

    ✓ Provision of the evaluation schedule/ deadlines, tasks and responsibilities

    ✓ Expected qualifications/ competences and experiences of the evaluator(s)

    ✓ Sometimes including suggestions for evaluation design and its practical implementation

    ✓ Usually also containing an estimated or maximum budget

Practical Implementation

✓ Preparation phase

  ✓ Kick-off workshop

  ✓ Development and agreement on analysis grid, data collection plan and data collection instruments

  ✓ Drafting of an inception report

✓ Data collection phase

  ✓ Document analysis

  ✓ Explorative data collection (e.g. expert interviews)

  ✓ Standardized data collection (e.g. surveys)

  ✓ If necessary: interim workshop/ report

✓ Reporting phase

  ✓ Drafting of evaluation report

  ✓ Results presentation/ discussion workshop

  ✓ Finalization of evaluation report

Data Collection

- ✓ Identification of data sources
- ✓ Informing all stakeholders about the evaluation, its purpose and proceedings
- ✓ Organization of logistics
- ✓ Pre-testing of data collection instruments and adaptation if necessary
- ✓ Important: realistic time planning!
- ✓ Preparation of interview
  - ✓ Early information and contacting of interviewees
  - ✓ Consideration of travel time, preparation and transcription of interview
- ✓ Preparation of surveys
  - ✓ Early information of respondents
  - ✓ If face-to-face survey: consider interviewer training
- ✓ Regular data backup and assessment
- ✓ Intermediate data analysis

Data management und -analysis

- ✓ Assessment of data quality by plausibility/cross-checks
- ✓ Analysis of contradicting findings
- ✓ Data disaggregation
  - ✓ Differentiation of findings from different data sources, instruments and stakeholder groups according to evaluation questions and stakeholder groups
- ✓ Data aggregation
  - ✓ Synthesis of information from different data sources, instruments and stakeholder groups
- ✓ Peer review of findings and assessments (→ researcher triangulation)
- ✓ Ordering of results according to analysis grid/ report structure
- ✓ Deduction of recommendations
  - ✓ Evidence based, no speculations, no new findings!

Management Response

✓ Objective: Initiation of steering decisions

✓ Lies in the responsibility of the commissioner of the evaluation and represents the results of the internal discussion of the evaluation results; acceptance or refusal of recommendations should be reasoned and explained

✓ Provides feedback to the program staff, partners and other organizational units for which the evaluation provides relevant information

✓ Should include lessons learned and a follow-up plan

✓ Guiding questions for management response

   ✓ Which results are relevant for your work? Which are the most important fields of action?

   ✓ Which conclusions do you draw from the evaluation results? To what extent do they coincide with the conclusions in the monitoring?

   ✓ Which issues are most important to you? Which are not? For what reasons?

   ✓ What should be done by whom until when in order to change what?

How can the quality of an evaluation be assessed?

Caveats
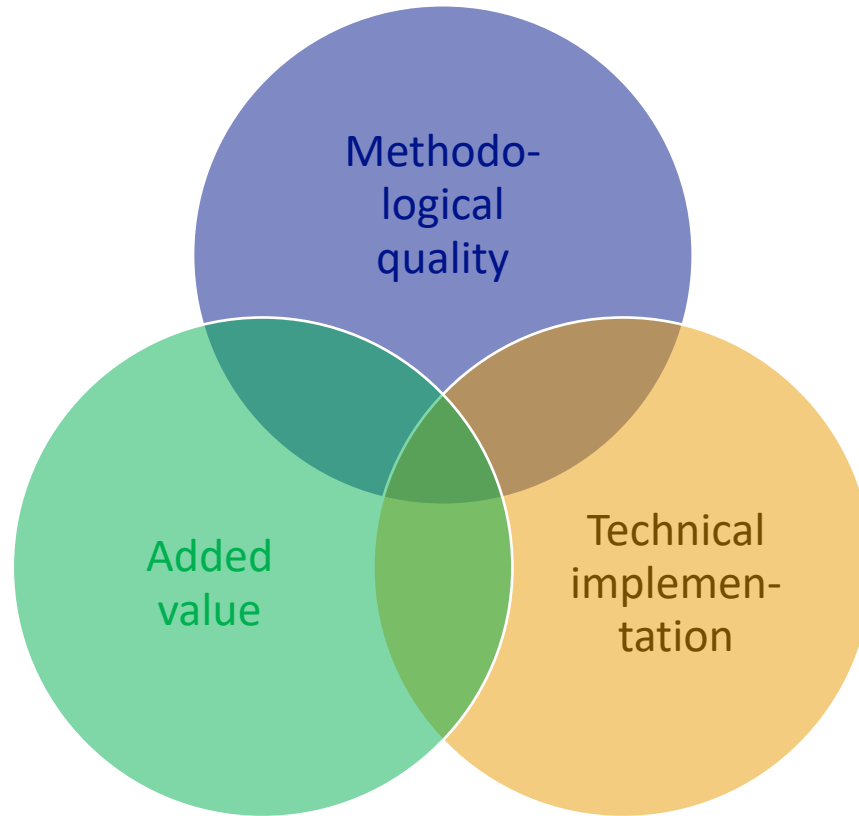
Aspects to consider when comparing the quality of evaluations:

- ✓ Different scopes

- ✓ Comparability of methodologies of evaluations may be limited

- ✓ When working in a team inter-rater reliability can be an issue

- ✓ Partly 'strong' aggregation of results required

- ✓ Albeit orientation at international recognized standards, indicators and assessments may be subject to discussion

Thank you very much for your attention!

Dr. Stefan Silvestrini

T +49 (0)681 387539 66

E s.silvestrini@ceval.de

W www.ceval.de

Centrum für Evaluation

CEval

Center for Evaluation

Digital Theme Days 2023
Minutes – Evaluations in working with non-government organizations

The event took place as part of the first **Digital Theme Days 2023**. The Digital Theme Days provide an opportunity for all active stakeholders in the funding system of projects of importance to development under the responsibility of private German executing agencies to exchange ideas and network on current and relevant topics.

Topic: **Evaluations in working with non-government organizations**
Date and Time: **November 21, 2023, from 2:00 PM to 4:00 PM**
Workshop Location: Online
Speaker: Dr. Stefan Silvestrini, CEO of the CEval GmbH
Participants: 28 individuals (representatives from German and local non-governmental organizations, staff of the German Federal Ministry for Economic Cooperation and Development and Engagement Global/ bengo)

1. Objective

The workshop is designed to advance the knowledge and implementation of effective monitoring and evaluation practices within the Private Träger (PT) context of German non-governmental organizations. Drawing on expertise from Dr. Stefan Silvestrini, CEO at CEval GmbH, participants will delve into the goals and methodologies of evaluations, exploring topics such as the importance of defined criteria, collection of primary empirical data, and the creation of knowledge to assess project success. The workshop will also cover key aspects of evaluation success, such as stakeholder involvement, clear task division, and the assessment of technical and methodological quality.

2. Agenda

1. Welcome, Introduction, and Program Overview
2. Content Input "Brief Introduction into Monitoring and Evaluation" by CEval
3. Open Discussion and Questions
4. Breakout Groups

3. Presentations and Discussion

3.1 Presentation by CEval (Dr. Stefan Silvestrini)

The PowerPoint presentation is attached.

Additional Notes
Dr. Stefan Silvestrini, CEO at CEval GmbH, presented insights into monitoring and evaluation, focusing on a results-based monitoring system at Engagement Global. He emphasized speaking from a commissioning perspective rather than an implementer viewpoint. The goals of evaluation were discussed, highlighting the application of empirical methods to assess projects and create knowledge. Criteria definition, collection of primary empirical data, and the role of evaluation as a learning tool were emphasized.

Key Points

Purpose of Evaluation:
- Conducting an evaluation involves assessing the extent to which a project has met its goals, essentially performing an assessment.
- Distinguishing criterion: Define criteria beforehand, such as DAC-criteria in the NGO context of international cooperation.

Primary Goals:
- Create knowledge to understand project success and reasons for outcomes.
- Enable organizations to align attributes, make informed decisions and improve approaches based on the insights gained from the assessments.

Learning Tool:
- Emphasis on the role of evaluation as a learning tool for organizational and long-term learning.
- The Legitimization of the projects in front of the funders is a potential advantage, but the learning should remain the focus.

Seven Questions before Drafting:
- Identify information recipients, timing of evaluation (formative, midterm, or ex-post), baseline data availability, responsibilities, logistics, data collection methods, and cost considerations.

Success Factors:
- Stakeholder involvement and acceptance are crucial.
- Establishing a culture of learning is fundamental; evaluations should be usable.
- Clear division of tasks, effective information flow, and well-defined goals contribute to success.
- Collection of comparison data should be integral to daily work.

Matrix of four Evaluation Standards:
- Clarity of purpose, feasibility within time and resource constraints, fairness in considering different perspectives, and accuracy in collecting and interpreting data.

Evaluation Designs:
- It is necessary to distinguish between design, methodology, and approach.
- Brief overview of Pipeline Design and Regression Discontinuity Design, emphasizing their suitability in NGO contexts.

Implementation Phases:
- The planning phase involves decision-making, defining evaluative questions, and contracting.
- The implementation phase includes breaking down criteria into questions, developing indicators, and commencing data collection.
- The Evaluation design, methodology and approach should be adaptable and able to be changed based on unexpected observations

Reporting Phase:
- Evaluation reports should be adaptable, subject to change based on unexpected observations.
- Before finalizing the report, the drafted results should be presented to the stakeholders in discussion workshops to give them the opportunity to express their perception of these results.
- The opinion expressed in these workshops should be integrated in the final report.

Assessment of Quality:
- Methodological quality assessment involves evaluating the bias, objective alignment, the understanding of the evaluation subject, and the suitability of the framework.
- All data collection methods and analysis grids used in evaluations should be included in the report annex.

Caveats:
- Team inter-rater reliability may be a challenge due to different expert backgrounds.
- The fact that Evaluation outcomes are known and disseminated may not always lead to changes.

The presentation also touched on the crucial importance of involving stakeholders in all stages of the evaluation, of clearly attributing the different tasks and of a well-structured and transparent flow of information. Furthermore, it shared some considerations for technical implementation and methodological quality. In the discussions that the participants engaged in, key questions related to evaluation quality, stakeholder involvement, and data provision to clients were addressed. Overall, the workshop aimed to provide valuable insights and practical strategies for effective monitoring and evaluation within the NGO sector.

3.2 Discussions

- **If you check the quality after handing in the evaluation report, it is too late. What can you do beforehand?**
  The inception phase is important. There, you can create a kind of summary; this document should be very concise. On the other hand, you should not intervene in ongoing data collection—this leads to bias. Let yourself be shown the design and inception report; there may be possibilities for improvement.

- **Should you provide raw data to the client?**
  Yes, if the database is large enough (remove personal information), you can provide raw data. If only 10 persons take part in your evaluation, it can be difficult to anonymize data; in this case, I would hesitate to provide information.

- **Do you have examples of how participants can be included in the evaluation? In planning or drafting?**
  Some organizations extensively discuss the evaluation with their partners before; an example is the German Agency for International Cooperation (GIZ, Gesellschaft für Internationale Zusammenarbeit), which conducts administrative frame talks (Auftragsklärungsgespräche). Therefore, an inception meeting is usually held with partners in the field and all partners to address their questions.

- **What is your advice when you have planned an evaluation, but a war breakout or an ongoing crisis prevents you from going to the field?**
  We do not send anyone to acute lethal dangers. If there is no access to the data, then do not conduct an evaluation. We still have remote evaluations (especially in the setting of the Covid-19 pandemic). In recent decades, capacities have increased in remote evaluations; meanwhile, Uganda and Kenya have enormously professional experts in evaluations, and they do not need international experts.

- **When it turns out that the quality is not as we hoped, or it turns out that expectations are not realistic, what does bengo do?**
  We would rather see if the evaluation fits your report. If you find out that the evaluation does not have good quality, we would ask you to have your evaluator rework it. If you have done everything you could, we would probably accept it. There is plenty of evidence that the more precise the instructions, the more accurate the evaluation. Therefore, write the Terms of Reference (ToR) as accurately as you can. Secondly, delve into your descriptive data, start some kind of exchange (via e-mail, workshops, talks, etc.), and only then write the report. Before writing the report, you need to have a sounding board for your evaluation.

- **At what point in time should you start a participation process?**
  Participation should be first, but not during the steps of collecting and analyzing the data. When it comes to assessing qualitative data, stakeholders should not be involved. On the other hand, in discussing the findings, they surely have to be involved.

- We have been collecting a lot of data since 2008, evaluating how our different activities work out, but we don't do evaluations that build up our database. We collect kind of fragmented data. How can we put the pieces together?
  If you have had something consistent since 2008, you could timeline evaluations. Do you have random samples? What if not, with what criteria has the sample been drawn? A link to evaluation standards literature from West Michigan University may be helpful; this document is a good starting point on the topic: https://wmich.edu/sites/default/files/attachments/u350/2021/program-eval-standards-jc.pdf

4. Conclusion from the breakout group

- For funding in the NGO sector in the field of international cooperation, it would be helpful to delve deeper into designs of specific evaluations important for funding Private Träger (PT), German non-governmental organizations, such as ex-post evaluations; there is no handout or information paper on this, and this issue has been somewhat neglected.
- Funding guidelines for evaluations should be presented and discussed further in the system of federal funding for Private Träger (PT) - German non-governmental organizations,.
- NGOs have access to background information on Evaluations in general, but need more information on the specific requirements for evaluations done in the case of federal funding for German non-governmental organizations. The Input of this session led in this direction but in future sessions more emphasis should be put on this specific need.