v.2.0

HUMAN RIGHTS & SCIENCE

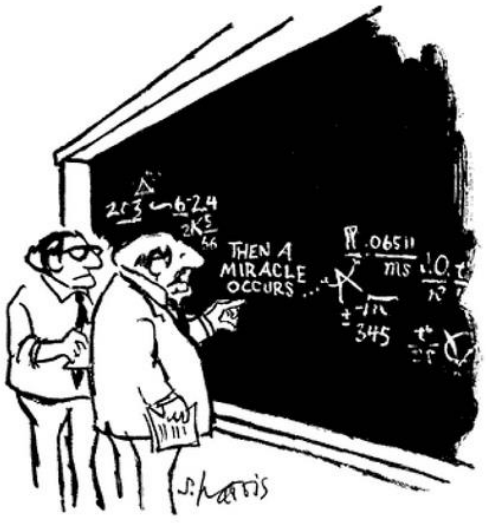
ACTION10

Programme Partner logo

Name of programme

Real-time Outcome Planning and Evaluation (ROPE)

Programme journal MONITORING, Template with guidelines



"I think you should be more explicit here in step two."

Image credit: Sidney Harris

Introduction

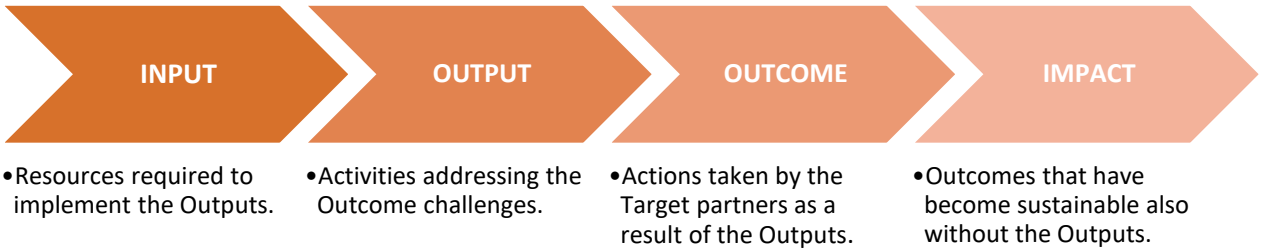
ROPE-Monitoring benefits from the HR&S Action Principles (TAct), Cross-cultural partnership tool (CUP) Transparency and accountability tool (TrAcc) and ROPE-Design. Those tools and principles were combined with the principles of Process tracing, Contribution tracing and Bayesian updating.

Acknowledgement

We are grateful to Beach and Pedersen (2013) for outlining Process tracing, to Befani and Gavin Stedman-Bryce (2016) for developing the Contribution Tracing approach and to CARE International in Ghana and their evaluation team leadar, Samuel Addai-Boateng for sharing about “Contribution Tracing” experiences on different social media platforms.

About ROPE-Monitoring

"It is better to be wise than to be smart."



Definitions by HR&S.

Efficiency

The ROPE is ambitious. The consequence is obviously that it consumes resources (time, money, attention….). The way HR&S takes to handle this challenges is not to try to simplify beyond quality but to develop tools that make the tool doable for the outer circle of stakeholders combined with a core team within HR&S that handles the tool and compiles the data collected from the outer circle.

The process is also planned to be supported with an IT platform. The platform shall also hold information about for example partners, previous communication and meetings.

Institutional memory

The ROPE programme journals (PJ) contributes to the Institutional memory of HR&S and Action10.

The PJs have been prepared in the form of templates, supported by written policies and institutional guidelines. All documents are filed in by the HR&S / Action10 EP team, whereas the Programme and Target partners provide monitoring data in the most convenient way possible.

TAct, CUP, TrAcc and ROPE-Design

Unlike Process tracing, Contribution tracing and Bayesian updating, ROPE address needs driven approach, truth, trust, equity, harmony, resilience, financial accountability, cross-cultural understanding, equal partnership, knowledge sharing, sustainable economy, and institutional capacity. This is all ensured through ROPE- Design. The ROPE Theory of change capture the Outcome challenge / Outputs required for all section in ROPE –Design

Process tracing, Contribution tracing and Bayesian updating

Process tracing

Process tracing involves unpacking the causal mechanism that explains what it is that linked cause A to outcome B. The investigator establishes a causal chain linking A to B and tests the strength of the evidence at each step in the chain by applying a number of probability tests ('straw in the wind', 'hoop, 'smoking gun' and 'doubly decisive' tests) underpinned by Bayesian logic.

Thus, process tracing is a qualitative method that uses probability tests to assess the strength of evidence for specified causal relationships. The method is used within a single-case design and without a control group. It offers the potential to evaluate impact through establishing confidence in how and why an effect occurred.

Process tracing involves articulating the steps between a hypothesised cause (for example, a development intervention) and an outcome. A causal chain linking cause A and outcome B is developed, and Bayesian probability logic is followed in order to assess the strength of the evidence of each part of the chain.

Contrary to statistical methods, the quality of the evidence is not judged by sample size (the number of observations) but rather the probability of observing certain pieces of evidence. Assessments of probability in process tracing are not necessarily quantitative. Rather, **evidence consists of empirical observations combined with knowledge of contextual factors (such as prior knowledge, timing, and the ways in which facts emerge).** The investigator looks for evidence to increase confidence that an outcome was caused in a particular way. Using probability logic, the investigator then systematically assesses the evidence in order to test hypotheses at each stage of the theory, including hypotheses representing alternative causal explanations.

More than one causal chain may contribute to the effect under investigation and factors external to the intervention may also contributed to the outcome. The use of process tracing in impact evaluation allows judgements on contribution.

Probability tests

Straw-in-the-wind test (low uniqueness, low certainty).
This is the weakest of the four tests, neither necessary nor sufficient to confirm a hypothesis.

- What happens if the hypothesis passes the test (i.e. reliable evidence of this type exists)?
  The investigator can be slightly more confident in the hypothesis, but this is not enough to conclusively prove it or to disprove alternative hypotheses. However, straw-in-the-wind tests can provide a valuable benchmark, and if a hypothesis passes multiple tests this can add up to important evidence.
- What happens if the hypothesis fails the test (i.e. reliable evidence of this type does not exist)?
  This slightly raises doubts about the truth of the hypothesis, but is not enough to rule it out.

Hoop test (high certainty: necessary to confirm hypothesis).

- What happens if the hypothesis passes the test?
  It does not significantly raise the investigator's confidence that the hypothesis is true.
- What happens if the hypothesis fails the test?
  It disconfirms the hypothesis. Because of this, hoop tests are often used to exclude alternative hypotheses.

Smoking gun test (high uniqueness: sufficient to confirm hypothesis).

- What happens if the hypothesis passes the test?
  The investigator can be confident that the hypothesis is true.
- What happens if the hypothesis fails the test?
  It does not significantly decrease confidence in the hypothesis.

Doubly decisive test (high certainty, high uniqueness).
This is the most demanding test, both necessary and sufficient to confirm a hypothesis.

- What happens if the hypothesis passes the test?
  We can be confident that the hypothesis is true, and that all alternative hypotheses are false.
- What happens if the hypothesis fails the test?
  It depends on the nature of the test.

Contribution tracing

According to Contribution tracing (CT), a reasonable contribution causal claim can be made if :

- There is a reasoned theory of change for the intervention, in the sense that the key assumptions behind why the intervention is expected to work, make sense, are plausible, may be supported by evidence and/or existing research, and are agreed upon by at least some of the key stakeholders.
- The activities of the intervention were implemented as set out in the theory of change.
- The theory of change (or key elements thereof) is supported by and confirmed by evidence of observed results and underlying assumptions, thus, **the chain of expected results occurred.** The theory of change has not been disproved.
- Other influencing factors have been assessed and either shown not to have made a significant contribution or their relative role in contributing to the desired result has been recognized.

Bayesian probability

Bayesian[1] probability is an interpretation of the concept of probability, in which, instead of frequency or propensity of some phenomenon, probability is interpreted as reasonable expectation representing a state of knowledge or as quantification of a personal belief.

Bayesian probability belongs to the category of evidential probabilities; to evaluate the probability of a hypothesis, the Bayesian probabilist specifies some prior probability, which is then updated to a posterior probability in the light of new, relevant data (evidence). The Bayesian interpretation provides a standard set of procedures and formulae to perform this calculation.

Bayesian methods are characterized by concepts and procedures as follows:

- While for the frequentist a hypothesis is a proposition (which must be either true or false), so that the frequentist probability of a hypothesis is either 0 or 1, in Bayesian statistics the probability that can be assigned to a hypothesis can also be in a range from 0 to 1 if the truth value is uncertain.
- The use of random variables, or more generally unknown quantities, to model all sources of uncertainty in statistical models including uncertainty resulting from lack of information.
- The need to determine the prior probability distribution taking into account the available (prior) information.

[1] The term Bayesian derives from the 18th century mathematician and theologian Thomas Bayes, who provided the first mathematical treatment of a non-trivial problem of Bayesian inference.

- The sequential use of Bayes' formula: when more data become available, calculate the posterior distribution using Bayes' formula; subsequently, the posterior distribution becomes the next prior.

Calculations

Bayesian Updating is based on joint probability - the probability of two things happening together. Consider two events, A and B, then perform arithmetic on that relationship to provide a updated (posterior) estimate of a prior probability statement.

Bayesian updating begins with the conditional probabilities p(B|A) as given,
when what is desired is the other conditional probability, p(A|B).

Bayes rule: updating probabilities as new information is acquired.

About probability

Probability is the measure of the likelihood that an event will occur. Probability is quantified as a number between 0 and 1. The higher the probability of an event, the more likely it is that the event will occur; 0 indicates impossibility and 1 indicates certainty.

(A simple example is the tossing of a fair (unbiased) coin. Since the coin is fair, the two outcomes ("heads" and "tails") are both equally probable; the probability of "heads" equals the probability of "tails"; and since no other outcomes are possible, the probability of either "heads" or "tails" is 1/2 (which could also be written as 0.5 or 50%).)

The probability of an event A is here expressed as p(A).

If two events, **A and B are independent** then the joint probability p (A and B) = p(A) x p(B).
(For example, if two coins are flipped the chance of both being heads is 0.5 × 0.5 = 0.25)

Conditional probability is the probability of some event A, given the occurrence of some other event B. Conditional probability is written p (A|B) and is read "the probability of A, given B".
It is defined by p (A|B) = p(A and B) x p(B) .

(For example; in a bag of 2 red balls and 2 blue balls (4 balls in total), the probability of taking a red ball is 0.5 ; however, when taking a second ball, the probability of it being either a red ball or a blue ball depends on the ball previously taken, such as, if a red ball was taken, the probability of picking a red ball again would be 1 / 3 since only 1 red and 2 blue balls would have been remaining.)

Inverse probability: In probability theory and applications, Bayes' rule relates the odds of event A1 to event A2 before (prior to) and after (posterior to) conditioning on another event B. The odds on A1 to event A2 is simply the ratio of the probabilities of the two events.

Given a probability distribution $p(x|\theta)$
for an observable quantity x conditional on an unobserved variable θ
the "inverse probability" is the posterior distribution $p(\theta|x)$
which depends both on the likelihood function (the inversion of the probability distribution) and a prior distribution.
The distribution $p(x|\theta)$ itself is called the direct probability.

$p(\theta|x) = p(x|\theta) \times p(\theta) \times p(x)$

$p(\theta)$ is the prior,
$p(x|\theta)$ is the likelihood (i.e. the evidence that you use to update the prior).
$p(\theta|x)$ is the posterior probability.

Notice that the posterior probability is a probability given the evidence.

Numerical Example
$p(A) = 0.5$
$p(B) = 0.000001$
$p(B|A) = 0.00000198$

Wanted: an updated (a posteriori) probability $p(A|B)$

$p(A|B) = p(A \text{ and } B) / p(B) = p(B|A) \times p(A) / p(B) =$
$0.00000198 \times 0.5 / 0.000001 = 0.99$

Estimate subjective probabilities

In most forms of probability, quantitative information is gathered and interpreted to help determine this likelihood through a mathematical mechanism, normally relating to the mathematical field of statistics. The percentage chance of a flipped coin landing on heads or tails can be interpreted as a probability, expressed as a 50% chance that it will land heads up, and a 50% chance it will land tails up.

A set of three possibilities are:
1) use empirical data if available;
2) use computer-based simulations if possible;
3) estimate subjective probabilities using either consensus based or mathematical pooling approaches.

Deciding probability value
Subjective probability is a probability derived from an **individual's personal judgment about whether a specific outcome is likely to occur**. It contains no formal calculations and only reflects the subject's opinions and past experience. Subjective probabilities differ from person to person, and they contains a high degree of personal bias. Allow the facilitator to use various tools to help the expert decide on a probability value, for example comparing the current situation to lotteries with known probabilities such as coin tosses.

Multiple view points

Where possible and appropriate, process tracing methodology should incorporate multiple viewpoints and data sources from a range of stakeholders. This can include the perspectives of the powerful and not so powerful. Understanding and probing around power could reveal significant, and often overlooked, insights into how change occurred in a specific case. To ensure that we systematically incorporate issues of power, we could use stakeholder analysis or power mapping to guide our process tracing inquiries.

Table. Bayesian updating of contribution claim x:

Contribution claim and components	**Bayesian updating**		**Comment**
	From (prior)	**To (posterior)**	
Contribution claim 1: xxx	y	z	
Components of contribution claim 1			
Component 1,1: xx			
Component 1,2: xx			

The power mapping process entails use of a visual tool to conceptualize the sphere of a person or group's influence. The stakeholder analysis explores in more detail the nature of the power and its position and the interests that give it that position. The stakeholder analysis of an issue consists of weighing and balancing all of the competing demands on a intervention by each of those who have a claim on it, in order to arrive at the intervention's obligation in a particular case.

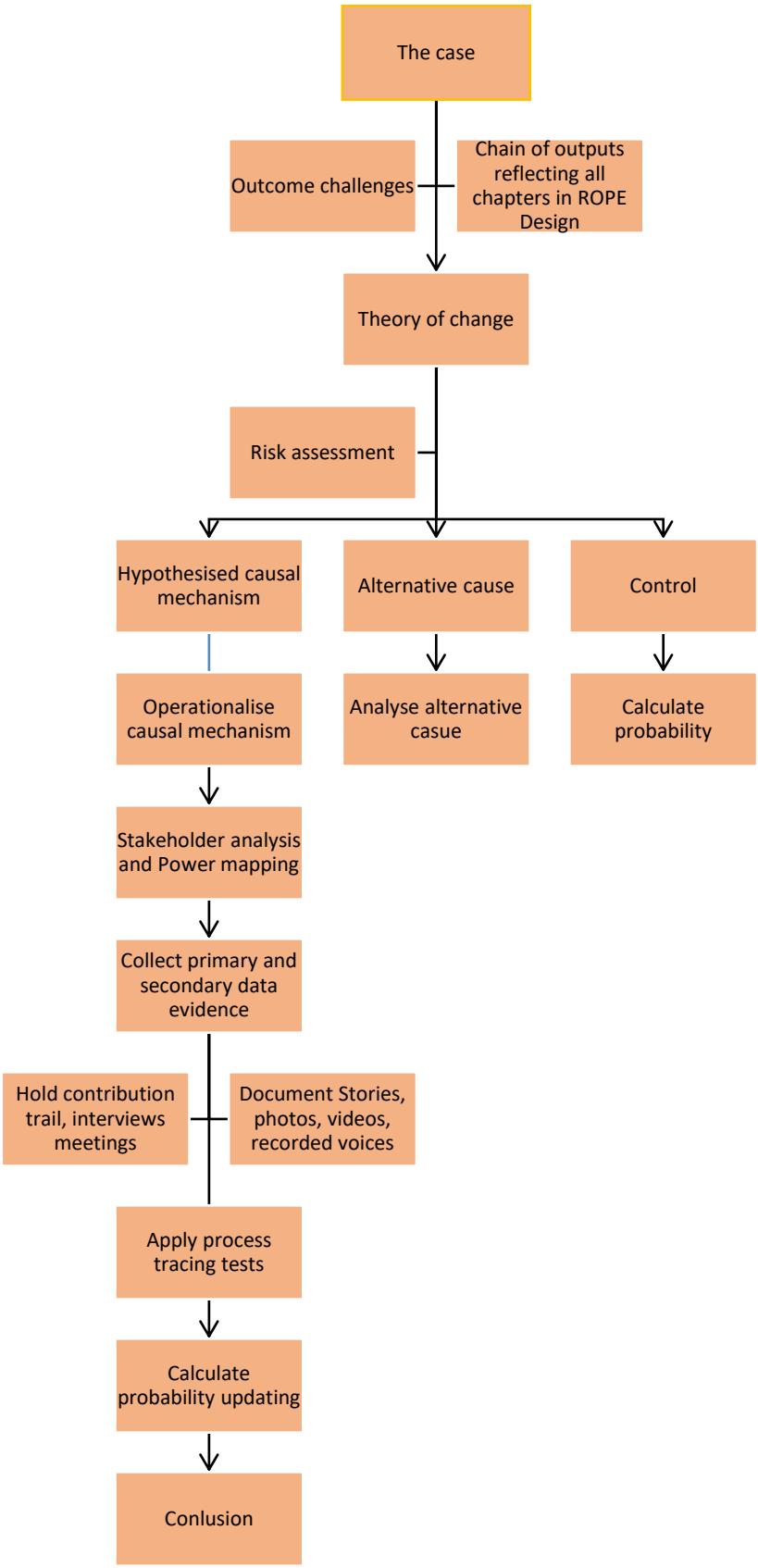
The ROPE – Monitoring Method

"Honest Reporting between partners is TRUST dependent."

The ROPE-Monitoring is a frame-work and supporting tool and shall never become a sheet with numbers, but rather a structured set of stories complemented with a compilation of outputs.

ROPE-Monitoring is participatory and involves all stakeholders at every step. The full team is involved with formulating the case (contribution claim), determining what evidence to look for, testing the evidence and calculating the confidence levels of a contribution claim (all with the support of someone who understands the approach).

The goal in "not to assess impact in terms of net change, but more specifically to assess changes in our 'confidence' that an intervention had impact" (Befani and Stedman-Bryce, 2016). The approach includes unpacking the causal mechanism that explains what it is that linked cause A to outcome B. The investigator establishes a causal chain linking A to B and tests the strength of the evidence at each step in the chain by applying a number of probability tests ('straw in the wind', 'hoop, 'smoking gun' and 'doubly decisive' tests) underpinned by Bayesian logic.


The case
Outcome challenges
Chain of outputs reflecting all chapters in ROPE Design
Theory of change
Risk assessment
Hypothesised causal mechanism
Alternative cause
Control
Operationalise causal mechanism
Analyse alternative casue
Calculate probability
Stakeholder analysis and Power mapping
Collect primary and secondary data evidence
Hold contribution trail, interviews meetings
Document Stories, photos, videos, recorded voices
Apply process tracing tests
Calculate probability updating
Conlusion

Analysing the Case

ROPE-Monitoring involves an in-depth analysis of a single case. A case could be an intervention, a project or a programme. Defining the Case in ROPE-Monitoring means compiling data from ROPE-Design, in order to ensure that a needs driven Case is identified. With ROPE it ALWAYS the TPs that identifies the Case and NEVER the donors.

Analysing a case means analysing how a specific cause, led to a given outcome within a case. The ROPE-Monitoring explains the causal mechanism; the drive that causes event A to give rise to outcome B. Mechanisms are underpinned by a generative model of causal inference, **mechanisms are** conceptualised as being made up of a number of 'parts'**, composed of entities** (for example, people, organisations, systems – nouns) **that engage in activities** (for example, protesting, researching, campaigning – verbs). Mechanisms must be framed correctly, that is at an appropriate level of abstraction from the particular case.

A **case** includes:

- An **outcome** or an impact under investigation, the effect.
- The **hypothesised cause**.
- The processes or **events** that link the hypothesised cause and the effect, the ROPE-Design outputs.

In theory-testing process tracing, Beach and Pedersen describe mechanisms as:

- **More than empirical events**. Describing a sequence of events between A and B or articulating intervening variables is not enough. Rather, mechanisms are theories about how and why one event leads to another. They represent the causal force or power that leads event A to give rise to outcome B.
- **Systematic**. Mechanisms reflect independent facts about how change happens. As such, mechanisms exist independently of any particular event – whether or not they are operating at a particular moment in time. This means that a mechanism identified and tested in one case will (if it is framed correctly, at an appropriate level of abstraction from the particular case) apply to other cases (within a specific bounded context – for example a country, a state, a sector or a time period).
- **Singular.** Within a specific case in process tracing, a single causal mechanism is examined (and the parts that make up that mechanism)[2].
- **Macro-level, micro-level or somewhere in between**, but always studied at a level that most makes sense within a specific case. For example, one process tracing case may examine a micro-level mechanism, relating to how the decisions and behaviour of individuals lead to change in

[2] This differs to how mechanisms are conceptualised elsewhere. For example, realist evaluation approaches also understand mechanisms as causal powers, but do not necessarily view them as systems made up of parts; instead the researcher may examine multiple different mechanisms within a single case (Westhorp 2014; Wong et al . 2013). These differences reflect the fact that there is no clear consensus in the literature on what exactly a mechanism is (Shaffer 2014).

wellbeing. Another may examine a macro-level mechanism, relating to how structural shifts in a social system lead to change within a society or institutional environment.

Step 2 – Developing the Theory of Change

Key stakeholders in the programme explain what makes them believe that what they did (their activities, etc.), contributed to the changes that they claim. By doing the Theory of change of the programme is presented. Theories of Change must be precise and testable.

Theory of Change is essentially a comprehensive description and illustration of how and why a desired change is expected to happen in a particular context. It is focused in particular on mapping out or "filling in" what has been described as the "missing middle" between what a programme or change initiative does (its activities or interventions) and how these lead to desired goals being achieved. It does this by first identifying the desired long-term goals and then works back from these to identify all the conditions (outcomes) that must be in place (and how these related to one another causally) for the goals to occur.

These are all mapped out in an Outcomes Framework. The Outcomes Framework then provides the basis for identifying what type of activity or intervention will lead to the outcomes identified as preconditions for achieving the long-term goal. Through this approach the precise link between activities and the achievement of the long-term goals are more fully understood. This leads to better planning, in that activities are linked to a detailed understanding of how change actually happens. It also leads to better evaluation, as it is possible to measure progress towards the achievement of longer-term goals that goes beyond the identification of program outputs.

Risk assessment

Make a risk assessment related to the Theory of Change.

Step 3 Developing a hypothesised causal mechanism

The next step is elaborating the mechanism to be tested. This may involve revising or adding detail to the Theory of change. Developing a hypothesised causual mechanism involves clearly elaborating all the steps between A (the hypothesised cause) and B (the outcome of interest).

Each part in the mechanism should specify:

- **which entities** (for example, individuals, organisations, groups – nouns) are expected to conduct
- **which activities** (for example, protesting, researching, advocating – verbs).

Each part of the mechanism can therefore be framed as a hypothesis (for example, 'civil society organisations will consult with citizens') and can be tested.

Building a causal mechanism in process tracing is in some ways similar to the process of developing a theory of change. Like a theory of change, a mechanism can be seen as a theory about how change happens in a particular context. As such, Beach and Pedersen recommend conducting a thorough review

of existing literature and evidence to inform the development of the mechanism, to ensure it incorporates previous thinking and learning about change in similar situations.

However, there are a few important differences between a mechanism and a theory of change:

- The mechanism to be tested should be **broken down into the smallest feasible number of parts**, which each directly cause the subsequent part. There should not be any leaps in logic (such as 'training will lead to new policies being developed').
- At the same time, **every part should be necessary** – there should be no superfluous parts which are not absolutely required for the mechanism to work.
- It should be **possible to empirically measure** each part of the mechanism.
- The mechanism should be framed at a **suitable level of abstraction** from the specific case, depending on how important (and realistic) it is for the findings to be generalisable to other cases. If a mechanism is framed in more abstract language, this allows the theory to be applied to other situations and contexts (although it will still need to be tested using context-specific evidence in other situations, to see if it does in fact apply).

Step 4 Operationalising the causal mechanism

Operationalising the mechanism involves working out what each part of the mechanism will look like in practice. Operationalising the
mechanism involves specifying observable manifestations – empirical evidence that it will be possible to collect in order to determine whether each part of the mechanism happened or did not happen.

Evidence might include:

- **Account evidence**: the content of empirical material (interviews, focus groups, observational evidence, meeting minutes, oral accounts).
- **Trace evidence**: evidence whose mere existence provides proof that a part of a hypothesised mechanism exists (for example, official meeting minutes demonstrate that a meeting did in fact take place).
- **Pattern evidence:** statistical patterns. Classic statistical probabilities can be relevant when evaluating this evidence (for example, employment statistics in a mechanism relating to racial discrimination).
- **Sequence evidence**: the chronology of temporal and spatial events (for example, we may expect to see events happening in a particular order if a specific part of a mechanism exists).

Alternative causes

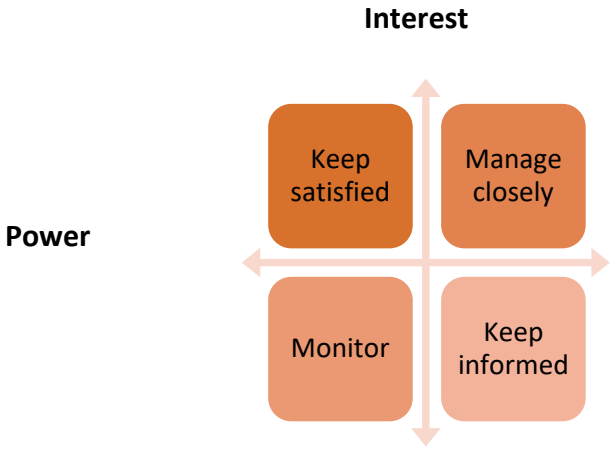
Operationalising the mechanism should involve **identifying evidence for causal links** between one part of a mechanism and another. What is the evidence that a part of the mechanism happened because of the previous part, rather than for some other reason?

This also requires thinking through the plausible alternative explanations that might explain each part of the mechanism, and looking for observable manifestations of these.
Identify alternative causes. Compile related activities by other actors such as the Strategic partners, Government and other Stakeholders.

Step 5: Stakeholder analysis

Stakeholders are organisations, groups, departments, structures, networks or individuals, which have something to gain or lose through the outcomes of a planning process or project. Stakeholder analysis is the process of the assessing a decision's impact on relevant parties. Stakeholder analysis in ROPE-Monitoring is also a Chain of Actors.

The first step is to identify all the stakeholders associated with the objective, project, problem or issue. Then the stakeholders are organised into a grid with different matrices according to their interest and power. 'Interest' measures to what degree they are likely to be affected by the research project or policy change, and what degree of interest or concern they have in or about it. 'Power' measures the influence they have over the project or policy, and to what degree they can help achieve, or block, the desired change.



Stakeholders with high power, and interests aligned with the project, are the people or organisations it is important to fully engage and bring on board. At the very top of the 'power' list will be the 'decision-makers', for example members of the government. Beneath these are people whose opinion matters - the 'opinion leaders'. This creates a pyramid sometimes known as an Influence Map. Stakeholders with high interest but low power need to be kept informed but, if organised, they may form the basis of an interest group or coalition which can lobby for change. Those with high power but low interest should be kept satisfied and ideally brought around as patrons or supporters for the proposed policy change. The final step is to develop a strategy for how best to engage different stakeholders in a project, how to

'frame' or present the message or information so it is useful to them, and how to maintain a relationship with them. Identify who will make each contact and how, what message they will communicate and how they will follow-up.

Multiple view points

Where possible and appropriate, process tracing methodology should incorporate multiple viewpoints and data sources from a range of stakeholders. This can include the perspectives of the powerful and not so powerful. Understanding and probing around power could reveal significant, and often overlooked, insights into how change occurred in a specific case. To ensure that we systematically incorporate issues of power, we could use stakeholder analysis or power mapping to guide our process tracing inquiries.

The power mapping process entails use of a visual tool to conceptualize the sphere of a person or group's influence. The stakeholder analysis explores in more detail the nature of the power and its position and the interests that give it that position. The stakeholder analysis of an issue consists of weighing and balancing all of the competing demands on a intervention by each of those who have a claim on it, in order to arrive at the intervention's obligation in a particular case.

Step 6 – Identify the Control

The control identifies the situation when the programme was initiated, against which progress is measured.

Collect evidence

ROPE – Monitoring follows the procedures of CT and seeks the most powerful evidence. The 'right questions' are asked of the 'right people' and the 'right evidence' is sought in the right places. The developers of CT call this "evidence with the highest probative value". Probative value is the power of specific items of evidence to increase or decrease our confidence in a specific claim (Befani and Stedman-Bryce, 2016, p.2).

Step 8 – Collecting secondary and primary data evidence

This involves gathering evidence (secondary and primary) for each observable manifestation of each part of the mechanism. As in any evaluation, the evaluator should consider the reliability of each source and its potential limitations and biases, and take appropriate steps to maximise the reliability and validity of the evidence used.

Secondary data collection and analysis

Common sources of secondary data include censuses (information about the members of a given population), information compiled by the government such as national development plans, information compiled by the government such as national development plans, information collected by government departments, organisational records and data that was originally collected for other research purposes.

- Secondary data analysis is expected to save time that would otherwise be spent on collecting data and, particularly in the case of quantitative data, and can provide larger and higher-quality databases. In addition, secondary data is essential, since it in most cases can capture past change and developments.
- Secondary data can be compiled through a desk review that will include an analysis of relevant documents, information, data and statistics. Triangulation shall be used when applicable to facilitates validation of data through cross verification from two or more sources.
- Data collected will be disaggregated (by for example gender, age and location) where possible.

Primary data collection

Primary data is collected through the ROPE-Design and Monitoring and includes the below as is relevant to the specific programme. The interventions includes all stakeholders, partners and selected beneficiaries who participated in the programmes. The Programme, Target and Strategic partners as well as other stakeholders gather with the monitoring team to discuss what evidence would prove, and disprove, the existence of the mechanism they jointly developed before; and hence prove or disprove their claim. This is the evidence they expect to see. All the stakeholders also work together to think about all the possible things that could also have caused the programme result.

<u>Type of events:</u>

- Participatory observations.
- Field visits to communicate generally with partners

- Individual interviews with key partners and stakeholders.
- Focus group discussions.
- Contribution trails.
- Photo-voice.
- Story telling.
- Individual interviews with key informants.
- Validation workshops. Maybe 2-4 day ROPE workshops.

Individual interviews

HR&S defines ROPE individual interviews as a conversation between two people that has a structure and a purpose, and feeds into the ROPE-Monitoring. It is designed to elicit the interviewee's knowledge or perspective on a topic relevant for the programme.

Individual interviews, which can include key informant interviews, are useful for exploring an individual's beliefs, values, understandings, feelings, experiences and perspectives of an issue. Individual interviews also allow the researcher to ask around a complex issue, learning more about the contextual factors that govern individual experiences.

The interviews are recorded.

Focus group discussions

HR&S defines a ROPE focus group discussion as an organised discussion within a small group of maybe six to eight people. The discussions are organised to feed into the ROPE – Monitoring scheme.

- Focus group discussions provide participants with a space to discuss a particular topic, in a context where people are allowed to agree or disagree with each other. Focus group discussions allow exploration of how a group thinks about an issue, the range of opinions and ideas, and the inconsistencies and variations that exist in a particular community in terms of beliefs and their experiences and practices.
- The participants of a ROPE Focus group discussion are carefully recruited to be those for whom the issue is relevant. Reflections are made over the benefits and limitations of the recruited participants. For example if they represent either one population (e.g. school going girls) or a mix (e.g. school going boys and girls), and whether or not they know each other.
- Group discussions are recorded.

Contribution Trial

The partners and other stakeholders gather with the monitoring team for a symbolic 'contribution trial' to discuss what evidence would prove, and disprove, the existence of the mechanism and to think about all the possible things that could also have caused the programme result.

Photo-voice

HR&S defines a ROPE Photo-voice as a participatory method that enables people to identify, represent and enhance their community, life circumstances or engagement with a programme through photography and / or videos and accompanying written captions. Photo-voice involves giving a group of

participant's access to cameras, enabling them to capture, discuss and share the stories they find significant.

Story telling

HR&S defines a ROPE picture story method enables the participants to communicate their perspectives on particular issues through a series of drawings or short texts they have made. The story telling can either be done in writing, or verbally with a HR&S representative. The picture story method provides a non-threatening way to explore views on a particular issue (e.g. barriers to girl's education). Photographers, video camera staff and voice recording staff, record the interviews.

Interview with key informants

Local government officials are often identified as key informants as well as staff of civil society organisations.

Validation workshops

The validation workshop is often a 2-4 day ROPE workshops.

Collect data also for alternative causes

ROPE-Monitoring minimizes the confirmation bias of only looking for evidence that supports one's claims. This is done by collecting data also for alternative causes such as Interventions by other organisations, international and national as well as the university. The inclusion of 'critical friends', during for example the 'contribution trial' informs about who may represent other plausible sources and / or explanations for the observed change.

Assessing the strength of evidence

- How credible is the measurement?
- Do reasonable people agree with the measurement?
- Does the pattern of results observed validate the results chain?
- What has been the role of external influences and other contributing factors?
- Where are the main weaknesses in the measurement?

Step 9 - Assessing the strength of evidence by applying probability tests

For each part of the mechanism, the inferential evidence from various sources is weighed in the attempt to put together a case that gives a **reasonable degree of confidence that each part of the mechanism exists or does not exist in the particular case.**

The investigator examines the available evidence to test the inferential weight of evidence for each of the causal sequences. Four 'tests' have been developed to assist with this process: 'straw-in-the-wind' tests, 'hoop' tests, 'smoking gun' tests and 'doubly decisive' tests (Bennett 2010; Collier 2011; Van Evera 1997). These tests are based on the principles of certainty and uniqueness; in other words, whether the

tests are necessary and/or sufficient for inferring the evidence (Beach and Pedersen 2013; Befani and Mayne 2014).

Comparing alternative causal sequences

In assessing the probability that the hypothesised causal chain led to an isolated effect, the investigator compares alternative causal sequences, through:

A. Reviewing the evidence under the assumption that the hypothesised causal sequence holds: cause A led to outcome B in the theorised way.
B. Reviewing the evidence under the assumption that the hypothesised causal sequence does not hold: an alternative causal sequence explains the outcome.

Assessing evidence if necessary and/or sufficient

- Tests with high uniqueness help to strengthen the confirmatory evidence for a particular hypothesis, by showing that a given piece of evidence was sufficient to confirm it.
- Tests with high certainty help to rule out alternative explanations by demonstrating that a piece of evidence is necessary for the hypothesis to hold.

Straw-in-the-wind test (low uniqueness, low certainty).

This is the weakest of the four tests**, neither necessary nor sufficient** to confirm a hypothesis.

- What happens if the hypothesis passes the test (i.e. reliable evidence of this type exists)?
  The investigator can be slightly more confident in the hypothesis, but this is not enough to conclusively prove it or to disprove alternative hypotheses.
  However, straw-in-the-wind tests can provide a valuable benchmark, and if a hypothesis passes multiple tests this can add up to important evidence.
- What happens if the hypothesis fails the test (i.e. reliable evidence of this type does not exist)?
  This slightly raises doubts about the truth of the hypothesis, but is not enough to rule it out.

Hoop test (high certainty: necessary to confirm hypothesis).

Hoop tests are often used to **exclude alternative hypotheses**.

- What happens if the hypothesis passes the test?
  It does not significantly raise the investigator's confidence that the hypothesis is true.
- What happens if the hypothesis fails the test?
  It disconfirms the hypothesis.

Smoking gun test (high uniqueness: sufficient to confirm hypothesis).

- What happens if the hypothesis passes the test?
  The investigator can be **confident that the hypothesis is true**.
- What happens if the hypothesis fails the test?
  It does not significantly decrease confidence in the hypothesis.

Doubly decisive test (high certainty, high uniqueness).

This is the most demanding test, both necessary and sufficient to confirm a hypothesis.

- What happens if the hypothesis passes the test?
  We can be **confident that the hypothesis is true, and that all alternative hypotheses are false**.
- What happens if the hypothesis fails the test?
  It depends on the nature of the test.

Step 10 – Put a number on the confidence of the claim

It is now time to put a number on the confidence of the claim through applying Bayesian Confidence Updating, that is, to decide a probability value.

Subjective probability estimation

The main critical issue with Bayesian Updating rests in the estimation of probabilities. The probability of an event is based on the likelihood of that event occurring. In most forms of probability, quantitative information is gathered and interpreted to help determine this likelihood through a mathematical mechanism, normally relating to the mathematical field of statistics.

Subjective probability is a probability derived from an **individual's personal judgment about whether a specific outcome is likely to occur**. It contains no formal calculations and only reflects the subject's opinions and past experience. Subjective probabilities differ from person to person, and they contain a high degree of personal bias.

Most approaches to subjective probability assessment try to generate a single number that represents a synthesis of the group's opinions.

Tools

Allow the facilitator to use various tools to help the expert decide on a probability value. One useful technique is to use a probability scale, for example, the probability scale developed by IPCC for assessing climate change evidence[3]:

Term	Likelihood of the outcome (%)
Virtually certain	99-100
Very likely	90-100
Likely	66-100
About as likely as not	33-66
Unlikely	0-33
Very unlikely	0-10
Exceptionally unlikely	0-1

[3] Source: Technical Summary In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change.

Step 11 – Bayesian confidence updating

Bayesian confidence updating is used to quantify our post-observation confidence (of the specific pieces of evidence identified in relation to a contribution claim). Bayesian Updating is based on joint probability - the probability of two things happening together.

Compilation

Table. Bayesian updating of contribution claim x, compiling multiple view points.

Contribution claim and components	**Bayesian updating**		**Comment**
	From (prior)	**To (posterior)**	
Contribution claim 1: xxx	y	z	
Components of contribution claim 1			
Component 1,1: xx			
Component 1,2: xx			

Assess confidence

Step 12: Assess confidence / What does the scores mean?

Calculating the posterior probability using Bayes' formula gives us a numerical value (see Table below) to assess the confidence we have in each component of a causal mechanism contributing to the process.

Table. Bayesian updating: numeric confidence levels. Source: Befani and Stedman-Bryce (2016)

Described as	**Numeric confidence levels**
Practical certainty	0.99+
Reasonable certainty	0.95–0.99
High confidence	0.85–0.95
Cautious confidence	0.70–0.85
More confident than not	0.50–0.70
No information	0.50

Once the process is complete, the researcher should be able to assert a degree of confidence in each part of the hypothesised mechanism, based on the evidence collected.

Monitoring Team

"Watertight Partner communication"

When outputs have been offered to the Target partners, their perception of the same shall be compiled in a structured manner. The purpose is obviously to get support with lessons to be learned and the related programme redesign and improvement. By doing so we are also provided with data for evaluation sessions and we can report on the status of the programme to external stake-holders.

The monitoring centres on collecting information for testing a claim by investigating whether evidence exists to support or reject it. A claim is simply a statement about the role a project may have played in bringing about change. Guided by the various types of evidence used in the four process tracing tests – the hoop test, smoking gun, doubly-decisive and straw-in-the-wind – the team is on a mission to build a watertight dossier for the case on hand.

Preparations

- Train the Team in communicating well with the Target partners.
- Formulate and agree on a testable contribution claim.
- Identify the data to collect.
- Develop an evaluation plan.
- Create the interview questionnaires.
- Prepare the ROPE-Monitoring 'workbook'.

Implementation

With the support of a semi-structured interview guide, recording apps on their phones, and other resources the team interact with the Target partners. Interview after interview, community after community, the exercise shall progress smoothly and a lot of interesting evidence – videos, photos, documents, recordings, etc – shall be gathered. At the end of a data collection exercise, it will be obvious that a data collection team has learned several lessons.

Sometimes, questionnaires are distributed at the end of a larger (about five days) output event such as a workshop, using IT support (f.ex. Survey Monkey). The results are analysed during the evaluation sessions and compiled in the evaluation journals.

Classification of evidence

The evidence is then classified under the various types of probability test.

Database

Databases are constructed.

References

Befani, B and Stedman-Bryce, G (2016) Process tracing and Bayesian updating for impact evaluation. Evaluation, 23, 1, 42–60. http://evi.sagepub.com/content/early/2016/06/24/1356389016654584.abstract

Beach, D. and Pedersen, R.B. (2013) Process-Tracing Methods: Foundations and Guidelines, Ann Arbor MI: University of Michigan Press.

Befani, B. and Mayne, J. (2014) 'Process Tracing and Contribution Analysis: A Combined Approach to Generative Causal Inference for Impact Evaluation', IDS Bulletin 45.6: 17–36

Bennett, A. (2010) 'Process Tracing and Causal Inference', in H.E. Brady and D. Collier (eds), Rethinking Social Inquiry: Diverse Tools, Shared Standards , 2nd ed., Lanham MD: Rowman and Littlefield: 207–19

Brundin, C., (2014). Ownership and Equal Partnership, A study of donor-receiver relationships in two development programs in rural Togo. Independent Research Project in Political Science, International Master's Programme in Political Science, Department of Political Science, Stockholm University.

Collier, D. (2011) 'Understanding Process Tracing', Political Science & Politics 44.4: 823–30

Earl, S., et al. (2001). Outcome Mapping; Building Learning and Reflection into Development Programs. I. D. R. C. (IDRC). Ottawa, Canada.

Van Evera, S. (1997) Guide to Methods for Students of Political Science , Ithaca NY: Cornell University Press.