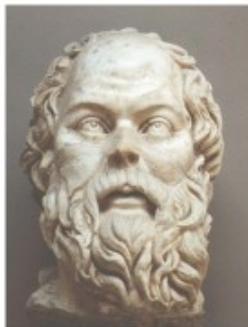




The  
**IGUANA Project**  
*Improving Education Together*

## IGUANA LEARNING PROGRAMME CORE TEXT

<b>MODULE</b>	<b>Organisational Intelligence</b>
<b>SUB-MODULE</b>	<b>Evaluation</b>
<b>VERSION</b>	<b>1</b>
<b>DATE</b>	<b>January 2014</b>



How do we know we make a  
difference?

### A BRIEF INTRODUCTION TO EVALUATION

18/02/2014

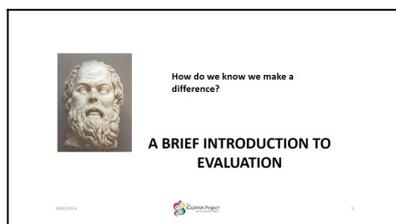


3

## **About this Document**

This document accompanies the powerpoint presentation 'Evaluation Overview (A brief introduction to evaluation). It provides a more detailed explanation of the slides that make up the presentation. The document structure and content follows the numbering and sequence of the slides in the presentation.

## A Brief Introduction to Evaluation – what this core text covers



The evaluation component of the learning programme is critical to supporting innovation and change. Embedding evaluation in a school innovation programme will generate iterative feedback loops to support ‘double loop learning’ for individuals, participating schools and IGUANA as a whole (Argyris & Schön, 1996).

Yet many organisations fear evaluation. They think it’s all about measuring success and about punishing them for failing to meet their targets. This is particularly the case for schools which, in many countries, face the constant anxiety of being evaluated by government inspectors. The worry that inspectors may arrive with little prior warning and downgrade their school’s status from ‘outstanding’ to ‘good’ – or worse – is a constant nightmare for staff and students. In this context, evaluation is often seen in a negative way, as one of the factors that drive organisational performance anxiety.

But when it is used to help organisations to learn how to do things better, evaluation is a very powerful tool to support change and innovation. To help organisations learn, evaluation needs to be used not just as a retrospective tool to assess performance, for example at the end of a school year (‘ex-post’ evaluation). Rather, it needs to be embedded within the organisational process to support a cycle of continuous improvement. This means that evaluation can and should be used to help schools design and develop their change planning (‘ex-ante’ evaluation). It also means that evaluation can and should be applied to monitor and assess progress as the school moves forward to implement its plan for change (‘formative’ evaluation). In this sense, evaluation has an important role to play in helping schools calculate how far they have travelled on their ‘change journey’.

Essentially, the role of evaluation in organisations is not to drive perfection but to understand what is relevant, what can be controlled and what can’t, what is good enough and above all what can be applied from learned experience to help the organisation change for the better. This core text helps schools to do this by introducing an evaluation approach based on ‘scientific realist evaluation’. This combines robust data collection with participatory methods. The core text introduces a method known as ‘theory of change’ to implement this approach. The text provides instructions on how to do this in practice.

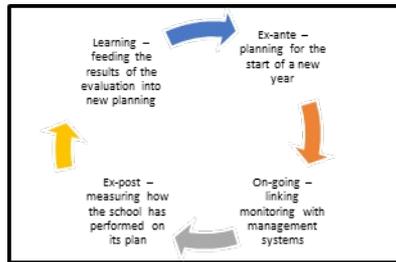
### Why do evaluation?

Evaluation is important for a number of reasons. It helps schools to plan for change. It helps them choose benchmarks against which they can assess how far they are progressing with their plan. It supports accountability within the organisation. It helps schools to learn from what they achieve and from what doesn’t go according to plan. There are four main reasons – or purposes – for doing evaluation:

- a *developmental* purpose - to help the school develop a plan for innovation and change (ex-ante evaluation)
- an *operational* purpose - to help the school keep track of how it is progressing in relation to its current strategies and plans (on-going or ‘formative’ evaluation)
- a *summative* purpose - to help the school measure what it has achieved (ex-post evaluation)

- a *sustainability* purpose - to help the school 'learn' and become a 'learning organisation'

## When to do evaluation



The four main evaluation purposes described above broadly correspond to what might be termed 'evaluation moments', i.e. stages in the life cycle of the 'object of evaluation' (the 'thing' that needs to be evaluated). This life cycle can take a number of forms, for example a school year; a strategic plan; a specific change project.

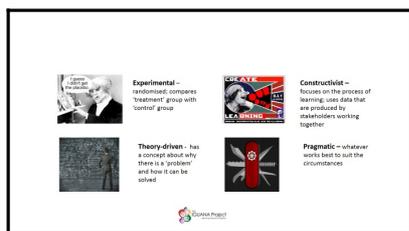
Ex-ante evaluation is used to help design a strategy, programme or project. It corresponds to the *developmental* evaluation purpose. The evaluation focus here is on identifying what needs to be changed and on forecasting what might happen if changes were put into place through, for example, a new policy strategy or a concrete programme. In the school setting, ex-ante evaluation might combine a needs analysis of school staff with a literature review on 'what works'.

On-going or 'formative' evaluation is carried out regularly throughout the life cycle of the intervention that is being implemented. The evaluation focus combines two things: firstly, assessing outcomes over a period of time at different steps in a process and, secondly, in evaluating the process itself. So in the context of the school, formative evaluation has been used extensively over many years to monitor and assess teaching performance in order to improve student attainment, as well as to evaluate teaching and learning methods themselves.

Ex-post evaluation is normally applied at the end of an intervention in order to identify outcomes and the factors that have contributed to these outcomes. An obvious example of ex-post evaluation in the school environment is the 'examination' which has played a central role in the traditional education system for centuries. Students are assessed at the end of the school year to determine the effect their learning has had on their knowledge and skills. By comparing the results of these tests with the results of the previous year, this kind of ex-post evaluation enables a quantitative measure of improvement (or lack of improvement) to be calculated. A more complex ex-post evaluation design might take the form of an educational initiative in which a new teaching programme based on the use of interactive games is introduced into a school. The impact of the intervention on teaching and learning outcomes is then evaluated at the end of the pilot, comparing the evaluation results with an assessment of teaching and learning outcomes in a similar school that didn't take part in the pilot. This kind of ex-post evaluation is called 'counter-factual evaluation' - a comparison between what actually happened and what would have happened in the absence of the intervention.

The end point of the evaluation process focuses on learning from experience. For evaluation to add maximum value, its results need to be used to improve practice. At a basic level, this entails dissemination – making sure the outcomes of an evaluation are made available to all the stakeholders who have an interest in them. However, learning from evaluation implies a broader *sustainability* purpose. This means applying evaluation results to help support innovation, and the transfer and replication of results to similar interventions elsewhere, and to help design future interventions. In the school context, learning from evaluation helps the school become a 'learning organisation' by embedding evaluation results in changed systems and practice.

## What kind of evaluation?



There are many kinds of evaluation approaches, methods, tools and techniques. Each draws on a particular philosophical tradition and each has its own position on what constitutes 'evidence'. Evaluation has long been dominated by the 'experimental' (sometimes called 'positivist') perspective.

Randomised controlled trials (RCT's) are a key tool in experimental evaluation and are seen as the 'gold standard' in evaluation.<sup>1</sup> They establish causal relationships between an intervention and its effects. They predict outcomes and impacts. However, they are very difficult to use in practice. RCTs work best in situations where it is possible to easily control for 'intervening factors'. In the laboratory or in pharmaceutical trials it is relatively straightforward to separate participants in an intervention into a 'treatment group' (for example participants who receive the drug being trialled) and a 'control group' (participants who receive a placebo) and then compare the effects of the experiment on the two groups. In the complex and messy real world, however, it is virtually impossible to maintain the assumption that a suspected cause precedes an event (for example, in clinical trials that the application of a particular drug will 'cause' the relief of particular symptoms). In the real world, social and cultural factors intervene. Things like history effects (the effects of 'external' variables that may have an influence on the outcome of the intervention); selection effects (statistical bias in the treatment and control groups); instrumentation effects (for example using measurement tools in different settings); attrition (uneven loss of participating subjects in treatment and control groups). In short, the range and complexity of 'intervening variables' that may influence the effects of a social intervention are potentially unmanageable.<sup>2</sup>

Because of these challenges, and also because it was recognised that traditional evaluation often failed to represent the voices of stakeholders who were not powerful, evaluation began to develop alternative approaches to experimental evaluation. Constructivist evaluation emphasised the 'social construction' of reality, where behaviour is subject to a continuous process of redefinition and 'social shaping'. It considers all relevant actors - or 'stakeholders' - as having a bone fide interest and engagement in the process of evaluation. Guba and Lincoln, for instance, distinguish between three broad groups of stakeholders: agents, beneficiaries and victims, all of whose perspectives and needs have to be identified and, if necessary, reconciled in order that a proper evaluation can be carried out.<sup>3</sup> Pragmatic evaluation recognised that any evaluation is carried out within constraints of time, resources and skills, and so evaluation design should be based on what is possible and do-able in a given situation.<sup>4</sup> Other approaches focused on identifying the 'theory' behind an intervention. Evaluators need to have in their heads a conceptual model of the programme or project they are evaluating in order to explain subsequent outcomes that can be tested by observation.<sup>5</sup>

<sup>1</sup> CAMPBELL, D.T. AND J.C. STANLEY, *EXPERIMENTAL AND QUASI-EXPERIMENTAL DESIGNS FOR RESEARCH*. CHICAGO, RAND-McNALLY, 1973.

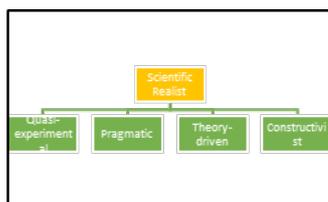
<sup>2</sup> STROMSDORFER, E.W. EVALUATING CETA: ADVANCES IN ASSESSING NET PROGRAM IMPACT. *EVALUATION REVIEW*, 11:4, 1987

<sup>3</sup> GUBA AND LINCOLN (1989) *FOURTH GENERATION EVALUATION*, SAGE PUBLICATIONS

<sup>4</sup> PATTON M Q (1986) *UTILISATION-FOCUSED EVALUATION*, SAGE, BEVERLY HILLS

<sup>5</sup> PAWSON R AND N TILLEY (1997) *REALISTIC EVALUATION*, SAGE, LONDON

## The IGUANA Evaluation Approach



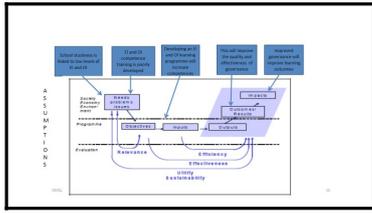
The evaluation approach recommended for schools participating in the IGUANA project is based on 'scientific realist evaluation'. This combines some of the 'rigour' of experimentalism with approaches that can reflect the context of the intervention, and the perspectives of the stakeholders involved.

Realist evaluation allows for context to be taken into consideration when evaluating interventions.<sup>6</sup> The process looks at how something is supposed to work, with the goal of finding out what strategies work for which people, in what circumstances, and how. The key features of the approach are as follows:

- Interventions are viewed as an attempt to address an existing social problem – that is, to create some level of social change. The focus of evaluation should therefore be on assessing whether and how this change has occurred.
- Interventions work by enabling participants to make different choices, so a key objective of evaluation is to capture how and why these choices are made.
- Making and sustaining different choices requires a change in a participant's reasoning and/or in the resources they have available to them. This combination of 'reasoning and resources' is what enables the programme to 'work' and is defined as a program 'mechanism'.
- Interventions work in different ways for different people - a key task of evaluation is therefore to capture 'what works, for whom under what conditions'.
- The contexts in which programs and interventions operate make a difference to the outcomes they achieve. Mapping context and how it affects outcomes is crucial to the evaluation.
- The evaluation design needs to reflect a number of 'pragmatic' considerations: the 'object' of the evaluation; the purposes of the evaluation; the resources available to carry it out.
- A realist approach is essentially about testing a theory about what 'might cause change', even though that theory may not be explicit.
- Data collection and analysis needs to reflect the different positions of stakeholders and the information these stakeholders will have.
- Learning is key to collecting and measuring data on evaluation outcomes and impacts, but more importantly it is key to understanding whether the 'theory of change' underlying the intervention is the 'right' one. In this sense, evaluation is similar to 'action research', where a 'change hypothesis' is tested by observing how the theory works in practice.

<sup>6</sup> Sullivan H and M Stewart (2006), Who Owns the Theory of Change? Evaluation, Vol. 12, No. 2, 179-199

## Theory of change

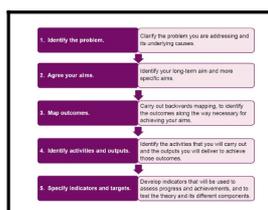


'Theory of change' is a tool to put realist evaluation into practice. It can be defined as a systematic and cumulative study of the links between the vision and context of a project and its objectives, activities and results. It involves the specification of an explicit theory of how and why an intervention might cause or have caused an effect.

Theory of change can be used to collect evidence in situations where experimental methods are difficult to apply.<sup>7</sup> The theory of change for an evaluation of intervention will specify:

- The 'problem' that needs to be changed
- The underlying factors that 'cause' the problem
- The 'model of action' that is expected to 'solve' the problem
- The objectives that are expected to be carried out to solve the problem
- The activities to implement the objectives
- The short-term outcomes that are expected to happen as a result of implementing the objectives
- The longer-term impacts that are expected to happen through the continued achievement of these outcomes
- The indicators used to measure the outcomes and impacts
- The 'means of verification' (the evaluation methods and tools used to collect data for the indicators)
- The 'assumptions' that need to be in place if the theory is to work

## How to develop a Theory of Change



Developing a theory of change to evaluate an intervention in practice means ensuring the specification shown above is put into place. This involves five steps.

**Step 1: Identify the problem.** Clarify the problem that is being addressed by the intervention you are evaluating and identify what is causing the problem. For example, the problem might be 'A high rate of non-attendance in the school. The suspected cause might be defined as 'Poor motivation of students'.

**Step 2: Identify the aims.** These can be defined as the long term expected changes (impacts) that the intervention that is going to be implemented will deliver. In the above example, this might be 'An increase in levels of school attendance'.

<sup>7</sup> SULLIVAN H AND M STEWART (2006), WHO OWNS THE THEORY OF CHANGE? EVALUATION, VOL. 12, NO. 2, 179-199

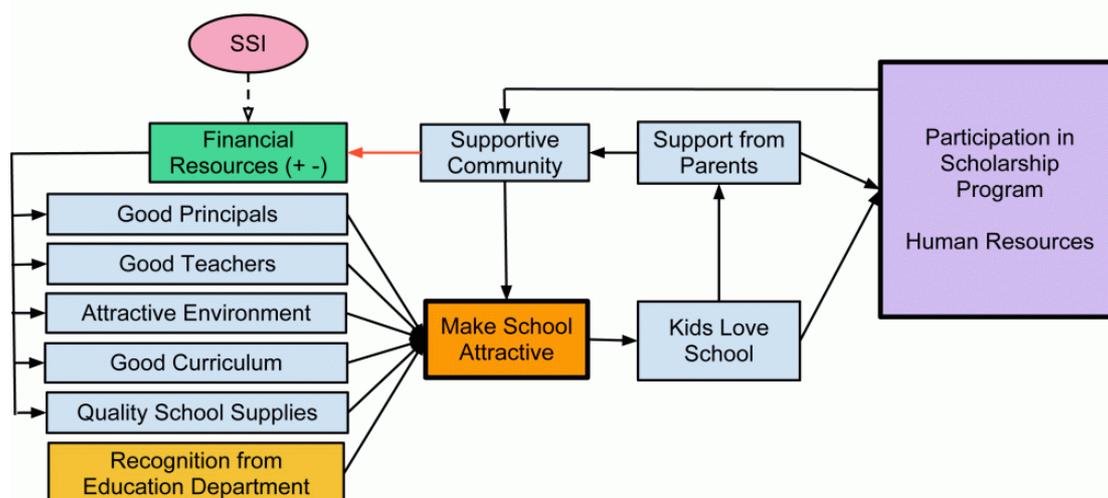
Step 3: Map the expected 'change journey'. Work backwards from the expected impacts to identify the expected outcomes along the way that the intervention is expected to deliver. In this example, these might be – in sequence – 'An increased understanding of the issues that lead to poor student motivation'; 'Improved teaching and learning skills'; 'Increased student motivation'.

Step 4: Identify the activities and outputs necessary to deliver the expected outcomes. In this example, these might be – in sequence – 'Carrying out a needs assessment with students and teachers to identify what are the gaps in current teaching and learning provision'; 'On the basis of the needs assessment, developing a new curriculum for the school'; 'Putting the new curriculum into practice'.

Step 5: Specify indicators and targets. Develop indicators that will be used to assess progress and achievements and test the theory of change and its component parts. In this example, the indicators could be '% of teachers and students reporting poor motivation'; 'Increase in self-reported motivation after one year of adopting the new curriculum'; 'Increase in student examination scores after one year of adopting the new curriculum'.

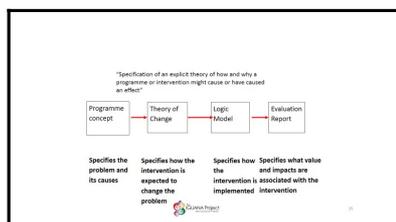
An illustration of a real-world theory of change application is shown below in the case of the Cambodian Rural Schools Programme.

### SSI Theory of School Success Logic Model



As the illustration shows, the School Success Initiative aims to address the problem of poor learning results in the country – particularly in rural areas. The programme Theory of Change identifies a number of key factors that are considered to be critical in creating this situation. The most important factor is a lack of financial resources. This leads to: a lack of skilled school leaders and teachers; poor teaching environments; a curriculum that is not fit for purpose; poor quality schools supplies and lack of engagement at ministry level. The changes required to address these problems are the mirror images of the problems, i.e. developing skilled school leaders and teachers; supporting an attractive teaching and learning environment and so on. The key aims of the intervention are to increase the attractiveness of the school; to then increase the motivation of students to attend, and the key activities or inputs are a scholarship programme; effective human resources; support from parents and the community.

## Using theory of change to identify 'cause and effect'



As noted above, one of the attractions of using experimental methods like RCTs in evaluation is that they provide evidence of whether or not an intervention has caused an effect.

If we are evaluating the effect of a new drug on malaria, for example, we can compare the proportion of those within the 'treatment group' (people who have been given the new drug) who have contracted malaria during the trial with those in the 'control group' (people who have not been given the new drug) who have contracted malaria during the trial. If there is a statistically significant difference in the numbers of people with malaria between the two groups, then it is reasonable to assume that the new drug has caused a reduction in malaria. Outside the clinical trial or the laboratory it is very difficult to prove these causal relationships. Theory of change enables us to make a reasonable judgement about what actions are likely to have caused particular effects because it allows us 'tell the story' of an intervention. For example, the theory of change might specify that activity A affects short-term outcomes A and C, which in turn affect medium-term outcomes E and F, and long-term outcomes A and D. The theory of change might also include feedback loops. For instance, changing the results of a short-term outcome could trigger a change in another short-term outcome that then reverberates in or feeds back to the first outcome

To help make the story of the intervention more clear, theory of change is usually combined with a 'logical framework' that specifies how the different components of the theory of change are linked together in practice. **Logic models** provide a way of linking the theory of change to the 'intervention logic' of an intervention, i.e. how the theory of change is translated into objectives, activities, and then outputs and outcomes<sup>8</sup>. But whereas 'logic models' on their own are essentially a descriptive device for mapping project components and the relationships between them, theory of change adds an explanatory framework for assessing how and why a particular component of the intervention is expected to lead to a particular outcome (change) for a particular target group. It enables us to:

- more clearly specify the project's target users, what are their needs and how these can be addressed
- more clearly specify the activities that need to be done in order to address these needs
- more clearly specify the expected outcomes the project activities will deliver and how we can accurately assess how far the project is meeting these expected outcomes
- identify the assumptions that underlie these expectations
- provide a systematic way of monitoring progress on outcomes

A simple combined theory of change and logic model is shown in the illustration below.

<sup>8</sup> McLAUGHLIN, J.A. AND G.B. JORDAN. 1999. LOGIC MODELS: A TOOL FOR TELLING YOUR PROGRAM'S PERFORMANCE STORY. EVALUATION AND PLANNING 22:65-72

A	B	C	D	E	F	G	H
<b>If....</b> <b>Initial issue/ context</b>	<b>Then...</b> <b>Rationale for intervention</b>	<b>Then....</b> <b>Actions Taken</b>	<b>Then...</b> <b>Short-term outputs</b>	<b>Then...</b> <b>Interim outcomes</b>	<b>Then...</b> <b>Long-term Objectives (impacts)</b>	<b>Outputs &amp; outcomes Indicators</b>	<b>Means of Verification</b>
There is a high level of student non-attendance	Need for action to increase student attendance	Implement student and teacher needs assessment	New teaching and learning curriculum	Increased student and teacher motivation	Increase in quality of teaching and learning at school	% students and teachers reporting increase in motivation  Difference in average examination results	Student and teacher survey  Statistical analysis of examination results before and after intervention

### Measuring the ‘distance travelled’ in a change journey

As noted above, theory of change presents interventions as a ‘change journey’ that consists of a sequential progression of ‘step-changes’, each of which has an effect on subsequent steps. In our school example above, the journey begins with an initial ‘theory of change’ - the theory that poor school attendance is caused by low student motivation.

Following on from this initial theory of change, each step in the change journey then has its own ‘theory of change’, i.e.:

If it is true poor school attendance is caused by low student motivation then it follows that a systematic mapping of the factors that influence poor motivation, and of the gaps in teaching and learning skills and knowledge that are associated with increasing motivation, would highlight what changes need to be supported (Step 2).

Having identified the changes that need to be delivered, the theory of change for Step 3 is ‘Teaching and learning motivation can be improved by introducing a new curriculum which includes the use of interactive games to deliver educational content’.

The theory of change for Step 4 is then ‘If student and teacher motivation is increased through the adoption of a new curriculum and the use of interactive games, then this will lead to an improvement in teaching and learning outcomes’.

Finally, to complete the change journey, the theory of change for Step 5 is ‘An improvement in teaching and learning outcomes leads to an improvement in the school attendance rates’.

Each of these theories of change is associated with a set of actions. For example, the theory of change for Step 2 of the change journey - ‘a systematic mapping of the factors that influence poor motivation, and of the gaps in teaching and learning skills and knowledge that are associated with increasing motivation, would highlight what changes need to be supported’ – requires an activity that involves a review of the literature on school motivation together with a needs analysis of teachers and students. Each of these activities has:

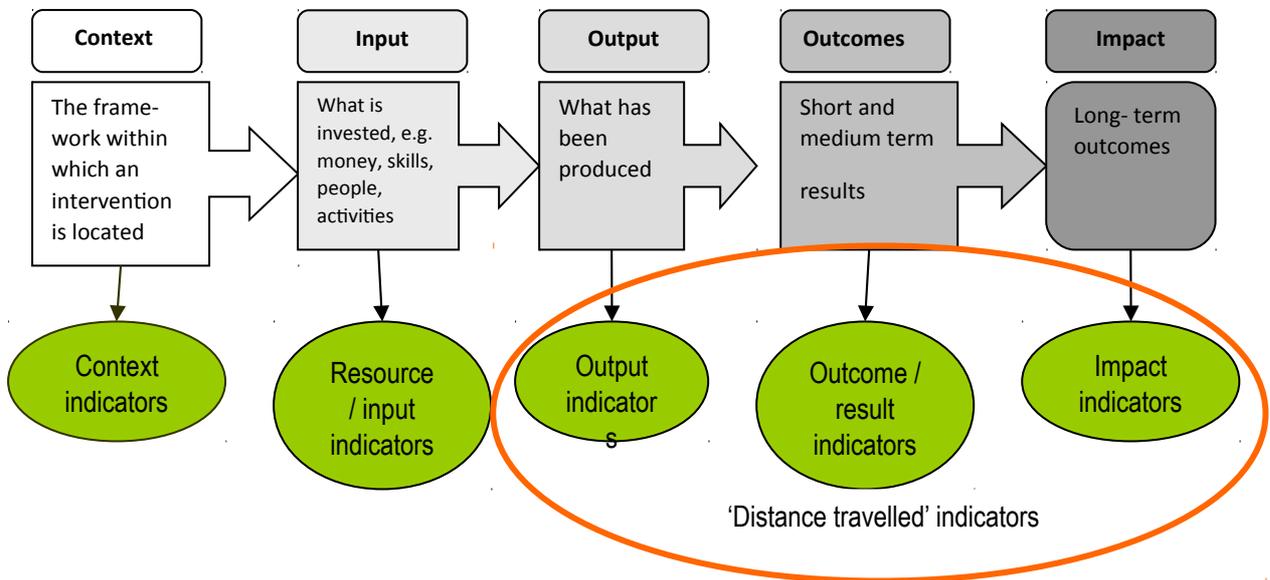
- a set of objectives the activity is designed to achieve,
- a set of outputs the activity is expected to produce,
- a set of outcomes that are expected to occur as a result of the utilisation of the outputs (and potential longer-term impacts that may result from these outcomes),
- a set of indicators that can be used to verify whether and in what ways these expected outcomes have been achieved,
- a 'means of verification' (the method used to collect and analyse data to verify the outcomes).

Taken together, the theories of change that express the steps of the project 'change journey', the activities carried out at these steps, the activity objectives, outputs, outcomes, potential impacts, indicators and means of verification comprise the 'intervention logic' of the intervention.

The task of the evaluation is to explore how this intervention logic has evolved in practice, and how far along the 'change journey' the intervention has travelled.

The evaluation approach and methodology should thus be focused on identifying and evaluating the changes associated with the key outputs produced by the activities at each step. The data collection tools (i.e. the 'means of verification') therefore need to be designed to assess the outcomes of these outputs using evaluation measures (i.e. indicators) that are contextualised to the particular characteristics of each activity and its outputs. These can then be used to measure how far the intervention has travelled along its planned 'change journey', as the illustration below shows.

#### 'Distance travelled' stages in a theory of change



The end point of a theory of change will specify the long-term impacts that are anticipated as a result of the implementation of the intervention – and the indicators that should be used to assess whether these anticipated impacts have been realised. This change 'journey' can be demarcated at a range of points along its trajectory - from 'context', through 'inputs' through to 'outputs', then 'outcomes' and finally 'impacts'. At the overall project level, the 'distance travelled' refers to the stage the intervention has reached along the 'theory of change' pathway – in our example along the road to reducing the level of non-attendance in the school.

The key indicators in measuring distance travelled are outputs, outcomes and impacts. These are often (mis)used inter-changeably in evaluation, and it is important at the outset to clearly specify what is meant by are outputs, outcomes and impacts. These are shown in the Box below:

**Definitions of outputs, outcomes and impacts indicators**

**Output** indicators relate to activity. They are measured in physical or monetary units (e.g., number of training programmes produced.).

**Result (outcome)** indicators relate to the direct and immediate effect on direct beneficiaries brought about by a programme. They provide information on changes to, for example, the behaviour, capacity or performance of beneficiaries (e.g. increase in emotional intelligence competences of participants in IGUANA).

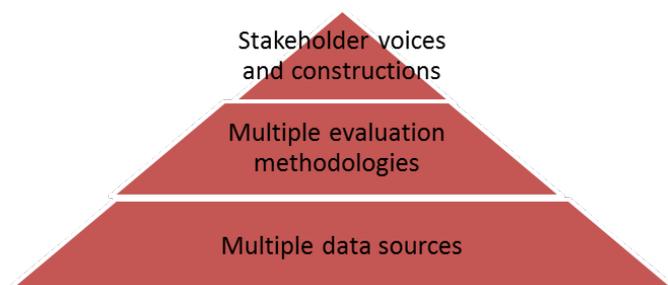
**Impact** indicators refer to the consequences and broader and longer-term social and economic changes of the programme beyond the immediate effects. Two concepts of impact can be defined: **Specific** impacts are those effects occurring after a certain lapse of time but which are, nonetheless, directly linked to the action taken and the direct beneficiaries (e.g. increase in learning performance of a school participating in IGUANA). **Global** impacts are longer-term effects affecting a wider population (e.g. reduction in school drop-out rates in EU)

There are different methods and techniques that can be used to collect and analyse data to measure distance travelled. Each have different advantages and disadvantages. Typical methods and techniques used are:

- User surveys – these provide lots of numbers, but provide low quality data
- Ethnographic methods (e.g. video diaries) – these provide lots of rich data; but lots of analysis and skills are required to use them. They are also seen by some as ‘unscientific’ and ‘subjective’
- Statistical modelling (e.g. regression analysis) – this needs high quality data and good analytical skills
- Action research – evaluators actively participate in the evaluation and work to support change, then measure the change

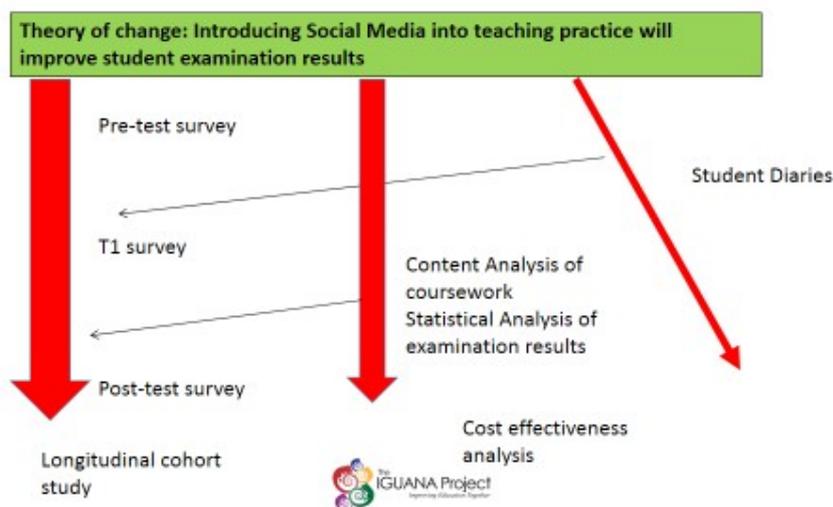
More information on evaluation methods and techniques is provided below.

In practice, evaluators tend to use a combination of different methods and techniques. This is known as ‘triangulation’. Triangulation makes it easier to identify ‘causal pathways’ between activities in an intervention and their effects. Triangulation entails the collection, analysis and synthesis of evidence of different types and from different sources, drawn from different kinds of evaluation activities, in order to arrive at conclusions in situations where attributing causality is difficult. In particular, a key aim of triangulation is to capture and reflect the ‘voice’ of different stakeholders in order to identify and understand their different positions and perspectives, as shown in the illustration below.



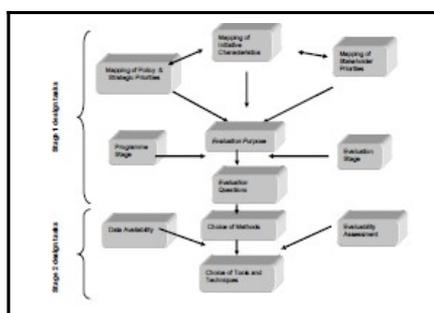
For example, say a school wants to improve its examination results. It has a theory that introducing Social Media into teaching practice will improve student examination results. It develops a pilot project to run a teaching module including social media. An evaluation plan is developed and implemented to assess the pilot project, how it works and what results it delivers. The evaluation plan combines a number of methods and techniques, as shown in the illustration below:

- A longitudinal cohort study. This takes a representative sample of students and follows their progress over the duration of the pilot. It compares the cohort's examination results before the pilot with results after the pilot has finished.
- A student and teacher survey. This gets students and teachers to self-report on things like their skills and knowledge before the pilot (pre-test) and compares these with the answers provided after the pilot has finished (post-test).
- Content analysis of coursework. This carries out a qualitative analysis of coursework produced by the students to identify whether the use of social media has improved its quality.
- Student diaries. This records the experiences of a small sample of students participating in the pilot.
- Cost effectiveness analysis. This compares the costs and the effects of the social media supported teaching programme with conventional programmes.



At the end of the pilot, the data collected and analysed using the different methods and techniques are compared with each other and then combined to provide an overview of the evaluation results.

## Designing the evaluation



Just as the intervention being evaluated has a life cycle and progresses through different stages, so does its evaluation, and the methods and tools appropriate for each stage of the evaluation differ. The key stages of the life cycle of an evaluation are:

Stage 1: Mapping and planning

Stage 2: Implementation

Stage 3: Reporting and dissemination

### Stage 1: Mapping and planning

At the outset, the evaluation needs to identify: what are the purposes of the evaluation, who are the audience, and what kinds of things need to be focused on. It also needs to consider the logistics of carrying out the evaluation: what are the settings in which evaluation will be carried out; what people are available to implement it and what skills are available; what communications channels need to be put into place. Following this initial assessment, an evaluation plan should be drawn up which will outline the evaluator's decisions on the choices available. An initial checklist might therefore include:

- what the purposes of the evaluation should be (e.g. ex-ante; formative; summative etc.)
- which main audiences or 'stakeholders' will be interested in the results
- how the evaluation fits into the project 'life cycle'
- what types of questions will be asked
- what kinds of methods and techniques will be suitable
- how evaluation will be integrated into the overall project
- what role the evaluator ought to play
- what involvement users should have in the evaluation process

An evaluation plan is the outcome of the exploration and planning process. It establishes priorities, specifies the phasing of evaluation activities, identifies how the different stakeholders and other participants will be involved, and determines appropriate methods for data capture.

An evaluation plan is not a once-and-for-all time activity. Projects will need to prepare an initial plan in the start-up phase and to keep it under periodic review. At key intervals or milestones, it may be appropriate to develop a new plan for the coming phase of the work and to show how the activities will feed into and inform later phases.

The results of these activities will be an evaluation plan that:

- ◆ Defines the priority areas.
- ◆ Specifies the key evaluation questions that the evaluation will endeavour to answer.
- ◆ Broadly defines the methods and tools to be used to answer these questions.
- ◆ Shows how the timing of evaluation outputs will inform the key decisions of the project.
- ◆ Indicates the mechanisms and procedures to ensure regular feedback to all stakeholders and especially project managers.

- ◆ Breaks down evaluation activities into its component activities, allocating responsibilities and making clear how the evaluation will be organised.
- ◆ Identifies additional design and planning tasks which will need to be undertaken for evaluation in later stages of the project.

## Sage 2: Implementation

Having developed an evaluation plan, the next stage of the evaluation will inevitably focus on carrying that plan out. The main stages involved in implementation are:

- Establishing the evaluation criteria that need to be assessed
- Deciding on what methods and techniques are to be used for data capture
- Managing and co-ordinating data collection, including analysing the results

The initial exploratory and planning stage will have identified the sorts of questions different stakeholders want to ask. These questions need to be converted or translated into evaluation criteria to enable indicators to be defined and measurable data to be collected. There are no hard and fast rules on how to formulate evaluation criteria. The main influences on determining which criteria to specify are the 'object' and 'purposes' of the evaluation. Examples of evaluation criteria are things like 'usability', 'effectiveness', value for money'.

As with evaluation criteria, the selection of methods and techniques will be highly dependent on the **object and purposes** of evaluation. In this respect, some broad 'rules of thumb' to consider are:

- Operational evaluations that concentrate on providing real-time monitoring and support for project management typically entail the on-going collection of data from a limited number of key actors in the project (for example project managers; materials designers and representative groups of target users). Commonly-used techniques are structured interviews; focus groups; diaries and logs and task analysis (getting key actors to 'think aloud' about why they make a particular decision).
- Summative evaluations, that consider, for example, the ways in which target users responded to a learning intervention, and in what ways their behaviours changed, typically utilise: questionnaire surveys, interviews and focus groups. These methods are normally used retrospectively. In contrast, observation, diaries and logs and content analysis (of, for example, effectiveness and user satisfaction with regard to curriculum materials) are normally used in real-time as the intervention develops.

It should also be remembered that the selection of particular methods and techniques has **skill and data resource** implications. Some techniques (for example ethnographic methods) involve lengthy, in-depth field work producing copious amounts of data that require highly skilled analysis and interpretation.

The Table below summarises the main sets of methods (and associated techniques) commonly used in evaluation.

<b>METHOD</b>	<b>Typical techniques</b>	<b>Typical context of use</b>	<b>Pros and Cons</b>
SURVEYS	Interviews Mapping Questionnaires	All-purpose. Operational: mapping interactions between actors; Summative: user satisfaction; user impacts. Learning: surveys of participants experiences.	Easy to carry out. Can produce large numbers of responses. Limited depth in questionnaire surveys (less in interviews and focus groups). Good in outcome-linked evaluations.
FIELD STUDIES	Observation Task Analysis Critical incidents Ethnography Case studies Diaries	All-purpose. Summative: how users respond to intervention. Operational: how institutional structures operate. Learning: retrospective analysis of what happened. Comparison of different settings.	In-depth data, giving insights on social construction of intervention. Time-consuming and skill-intensive. Difficult to utilise in outcome-linked evaluations.
INTERPRETATIVE	Content analysis	All purpose. Used in operational (analysis of meetings etc.); summative (analysis of materials or reports) and learning (deconstruction of programme reports)	Deconstruction of 'hidden' meanings and agendas. Rich interpretation of phenomena. Inherent risk of ideological bias.
PARTICIPATORY	Action research	Typically in developmental evaluation mode.	Encourages real engagement of 'subjects' of intervention. Good in highly uncertain contexts. Evaluators sometimes get 'too involved' in intervention itself.

It should be borne in mind that the selection of particular methods and techniques implies also using the appropriate type of data analysis (which has its own resource and skills implications). In general, large data sets (such as derived from surveys) normally need statistical software systems such as SPSS. Interpretative data (derived for example from content analysis) can be analysed with proprietary qualitative software packages such as N-VIVO. In any case, a clear coding frame to analyse such data is necessary.

### **Stage 3: Reporting and Dissemination**

Dissemination should not be restricted to the circulation of a final report - especially in the case of 'developmental' evaluations. Different stakeholders may require different communication approaches. These might include:

- Short summaries of the evaluation, tailored to different audiences
- Journal articles for other researchers
- Topical articles in the trade press (e.g teaching magazines)
- Workshops for specific audiences
- Feedback seminars for key decision-makers.

The use that is made of dissemination outputs should be consistent with the 'purposes' of the evaluation as defined in the initial Exploratory and Planning phase. In other words, evaluations should be designed in terms of the decisions and actions they will inform. It is not always easy to reflect this in recommendations, especially when the relevance of such recommendations may not be easily recognised by some stakeholders. The art of making useful recommendations lies in:

- understanding the context in which the audience operates
- addressing future realities rather than dwelling on the past
- clarifying choices based on realistic options
- showing how in practice recommendations can be implemented.

In the final analysis, dissemination is only the first part in the process of using evaluation results to support the organisation in learning from its experiences, and in applying learning to promoting change and innovation. In order to maximise learning, the school needs to adopt the systems and processes of a learning organisation. This is covered in another sub-module of the IGUANA Learning Programme.