superu

SEPTEMBER 2017

Evaluation terms for the social sector

A glossary

Using Evidence for Impact

SEPTEMBER 2017

Evaluation terms for the social sector

A glossary

Using Evidence for Impact

Introduction

This glossary was designed to accompany the handbook *Making Sense of Evaluation*, but it is a useful resource on its own.

Who is this for?

This glossary is for those who need to understand evaluation terminology, and also for those who are looking for certainty about what an evaluation term might mean in a particular context.

Where there are different possible uses or interpretations, we have indicated this. If in doubt, check how a term is being used in your specific case.

What we cover in this glossary

The glossary has been developed to provide people new to evaluation with an overview of evaluation concepts and tools so that they are sufficiently equipped to understand the methods and ideas that are being put forward.

It is divided into two parts – the first part looks at terms relating to concepts, while the second explains some key evaluation and analysis methods and when they are most useful.

Evaluation Concepts

Technical term	What it is	Terms that might also be used
Additionality	This is what happened because of the initiative, which would not have happened anyway – it describes the difference made by what you did.	Added value
Attribution	Assigning a causal link between observed (or expected) changes and a specific intervention.	
Audit	A verification of whether the resources have been used in line with the funding rules and legislation. This is about compliance. It's about checking if you have done it correctly.	Inspection
Baseline	The situation before the initiative. This might be about individuals or the population. You need to have this information to be able to see how far you have progressed towards achieving your outcomes, or to carry out pre/post analysis.	Starting point Point of departure
Benchmark	Benchmarking is about comparing your results with others, or to some external standard. To do this you need comparable information, which could be quantitative or qualitative.	Standard Reference point
Beneficiary	Technically, anyone who benefits from the initiative, but mainly people who do so directly. The term has other meanings, so different words are often preferred.	Client Customer Participant
Bias	The extent to which a measurement, sampling, or analytic method systematically underestimates or overestimates the true value of a variable or attribute. In evaluation, bias does not refer to the concept of fairness or prejudice as it may in common usage.	
Causal effect	The change or consequence that can be attributed to a policy or intervention. This is about asking "did this cause the effect?" not about unravelling all the possible causes of the effect.	
Collective Impact	Collective impact is an approach that brings together organisations from different sectors to agree to solve a specific social problem using a common agenda, aligning their efforts, and using common measures of success.	

Technical term	What it is	Terms that might also be used
Comparability	If you want to compare several interventions, you need to use the same measurement and the same population. Standard indicators can help, but you may still need to be careful to take account of underlying differences between the groups being compared.	
Completion rate	The proportion of the planned outputs that have been completed. For example, what proportion of clients completed a training course or an addiction programme.	
Contextual evidence	Information about whether a strategy fits with the context in which it is to be implemented – is it likely to be feasible to implement, useful and likely to be accepted by a particular community.	
Counterfactual situation	What would have happened if the intervention had not taken place. This might be the "do nothing" scenario or "no change to current practice". You need this to judge the scale of the change that is a result of your initiative. However, it is often difficult to establish, so in the real world it tends to be estimated on the basis of comparing with a similar group which has not had the intervention (where that is possible), by extrapolating past trends, or if all else fails, by using "before and after" comparisons within the same group.	Counterfactual
Deadweight	Some of the change observed might have happened even without your intervention. This needs to be acknowledged. Generally, it's very difficult to measure – use your best reasonable estimate here.	
Displacement effect	Effect that happened in your area at the expense of another area. For example, crime reduction in one area is great, but not if you've just moved the problem into another neighbourhood.	
Distance travelled	"Distance travelled" is used, particularly in the UK non-profit sector, to refer to the relative progress made by people towards health, well-being, and developmental goals, such as from addiction to recovery, educational achievement, or living independently of others. It is an alternative to measuring success by the achievement of absolute goals that tends to be insensitive to progress made that may fall short of the absolute goal.	
Economic evaluation	Looks at what resources are being used in an intervention and their costs (direct and indirect) compared to outcomes. This is the evaluation that looks at "how much bang for your buck?"	Impact evaluation (when it's being used as shorthand for "economic impact")

Technical term	What it is	Terms that might also be used
Effects	A generic term for the changes (intended or unintended, positive or negative) that happened as a result of what you did.	Outcomes Impacts
Effectiveness	The extent to which you have achieved the changes you set out to make for the people you work with. These might be changes for individuals or changes to groups or systems, depending on your objectives.	Success Efficacy
Efficiency	How far you are delivering your programme with the least possible use of resources to the area of greatest need. OR How far you are getting the best possible results for the resources that you have.	Actually people are pretty much agreed on this – some people go further and talk about whether this is improving over time or finding new ways to deliver.
Evaluability assessment	A process before evaluation that looks at available knowledge and information to see whether an evaluation will be possible. In practical terms this means checking whether an evaluator, using appropriate methods and techniques will be able, in the time and budget available, to answer the evaluation questions with a strong likelihood of reaching useful and robust conclusions.	
Evaluation	The systematic and objective assessment of a project, programme or policy, its design, implementation and results. The aim is to determine the relevance and fulfilment of objectives, efficiency, effectiveness, impact and sustainability. It should provide information that is credible and useful, enabling the incorporation of lessons learned into the decision-making process. Evaluation in some instances involves the definition of appropriate standards, the examination of performance against those standards, an assessment of actual and expected results and the identification of relevant lessons.	Impact assessment Review Value
Evaluation capacity	The extent to which an organisation's staff and leadership have the skills, knowledge, attitude and resources to do useful evaluations.	
Evaluation design	The technical part of an evaluation plan that sets out the evaluation questions, how the data will be collected and analysed, and on what basis judgements will be made.	

Technical term	What it is	Terms that might also be used
Ex-ante evaluation	Evaluation that is carried out before intervention implementation. It is intended to help ensure that an intervention is as relevant and coherent as possible, so it looks at issues such as whether the problem has been diagnosed properly, whether the objectives are clear and whether the expected impacts are realistic. It might also set your baseline.	Prior appraisal Prior assessment
Experiential evidence	The collective experience and expertise of those who have practised or lived in a particular setting. It also includes the knowledge of subject matter experts.	Intuitive knowledge Tacit knowledge
Ex-post evaluation	Evaluation carried out when the intervention is over. It would normally start to look at emerging impact, and to understand the factors leading to success or failure.	
External validity	The degree to which you can generalise from the thing that has been evaluated to something that has not. The extent to which findings, conclusions and recommendations are applicable to other settings.	
Formative evaluation	An evaluation that is primarily about learning. Most often carried out during the lifetime of a programme, with feedback to managers and stakeholders. Focuses largely on implementation procedures, their effectiveness and relevance.	See also process evaluation
Impact	The change that can credibly be attributed to an intervention. In New Zealand the definition of this very much depends on who is using it – there are many and quite different interpretations. Always check what is meant if this word is being used.	Effect Contribution to change Long-term outcome
Impact assessment	The process of identifying the future consequences of a current or proposed action.	
Impact evaluation	Assesses intervention effectiveness in achieving its ultimate goals. This is going to be done a number of years after it is implemented (how many will depend on the nature of the change concerned).	
Indicator	Something you can use to measure or to demonstrate a change that that you have probably influenced. Some people use indicator when it's not possible to directly measure an outcome; others use the term interchangeably with measure. You should check with users what they mean by this term in any given context.	Measure Key performance indicator (KPI) Metric

Technical term	What it is	Terms that might also be used
Inputs	The resources available for the implementation of an intervention. These might be financial, but also include human resources, organisational capacity, infrastructure and knowledge among other potential inputs.	Resources
Internal validity	The extent to which the evidence collected supports the conclusions about causal linkages. Do the conclusions stack up?	
Intervention	In this context, the "thing" being evaluated. It may be a policy, a programme or group of programmes or a project.	Policy Programme Project
Key performance indicator (KPI)	The most important of a list of potential indicators, that can give you the key elements of what you need to know	Key performance measure
Logic model	A diagram or picture of your intervention that shows in simple terms how what you are doing is expected to lead to the outcomes that you intend. It usually identifies the resources or inputs, the activities, outputs and outcomes for the intervention.	Intervention logic Theory of change Logical framework Logframe Benefits map (software and business types use this one) Results chain Impact map
Logical framework	A specific tool based on a matrix, which highlights the needs, objectives, inputs outputs and results of a given intervention. The simplest form of a logical framework is a 4x4 table with 16 cells, although this isn't a strict format. It uses words, not diagrams, and is mainly used in the field of international development.	Logframe
Measure	Something you can measure or count that is the direct result of your activity.	Indicator Performance indicator Key performance indicator (KPI) Metric Performance measure
Meta-evaluation	Evaluation of another evaluation or a series of evaluations. The concept is that you can learn more from the synthesised results of a series of evaluations than from relying on the findings of a single evaluation. There are several ways of judging the reliability of evidence in this way.	Synthesis Systematic review Meta-analysis

Technical term	What it is	Terms that might also be used
Monitoring	The process of consistently measuring over time and reviewing the delivery and results of an initiative so that you can adapt it in the light of findings, or correct it where problems are identified.	
Net effect	The extent of the changes observed that can realistically be attributed to the intervention – ie without the things that would probably have happened anyway (the counterfactual).	
Objective	A clear statement of the anticipated results to be achieved by the intervention. These might be quantitative (with targets) or qualitative (with descriptors).	Desired outcome
Outcome	The changes (eg in attitudes or behaviours) that are likely or achieved as the result of your intervention in the short-, medium- and long-term. An outcome statement is sometimes called an objective.	Results Impacts Note: both these terms are used in varying ways to mean slightly different things. Use with care and be clear how you are doing so.
Outcome evaluation	Focuses on the effectiveness of the intervention in achieving its outcomes. Bear in mind that outcomes can be produced in the short term, a longer period or in the long term. You can't do an outcome evaluation too early in the life of an intervention (because there won't be many!).	
Outcomes framework	A tool that helps you to link what you want to achieve and how you will measure success.	Most people use this term. However what's included in the framework might be broader or narrower according to the situation or how your sector or organisation traditionally does this.
Output	What the intervention actually produced. Could refer to, for example, hours of counselling, extra teaching hours or number of clients supported.	
Performance management	An assessment to gauge progress toward predetermined goals. It is aimed at improving performance within finite resources.	

Technical term	What it is	Terms that might also be used
Primary data	Data which is directly collected in the field, for example by carrying out a survey as part of an evaluation. (See also secondary data.)	
Process evaluation	A method of assessing how a programme is being implemented. Process evaluation focuses on the programme's operations, implementation, and service delivery.	See also formative evaluation
Programme	Any set of activities managed together over a period of time that aim to achieve some sort of positive change for a person or group of people you work with.	Project Initiative Intervention (People might also refer to policy, which might also be a mechanism for achieving change.)
Qualitative data	Qualitative information or data is about qualities or characteristics rather than quantities or measured values. It is broad ranging and essentially non-numerical in nature.	
Quantitative data	Information that can be expressed in numerical terms, counted, or compared on a scale.	
Relevance	How the explicit objectives of an intervention relate to the problem, issue, need or opportunity that the intervention is meant to address. This is particularly important in the design phase of an intervention or an ex-ante evaluation where the focus is on the choice of strategy.	
Reliable	Consistency or dependability of data and evaluation judgments, particularly looking at the quality of the instruments, procedures and analyses used to collect and interpret them. Evaluation information is reliable when repeated observations using similar instruments under similar conditions produce similar results.	
Result	What happens as a consequence of the policy or programme. This word is used in many ways in different contexts and should be treated with care – always check how it is being used in each specific case.	Outcome
Result Measurement Framework	A measurement system which provides a clear line of sight between the investments we make and what happens as a result	In New Zealand this may mean the link between investment and priority results of the Social Investment Strategy, and the Better Public Service targets.

Technical term	What it is	Terms that might also be used
Review	An assessment of the performance of an intervention, periodically or on an ad hoc basis. Frequently "evaluation" is used for a more comprehensive and/or more in-depth assessment than "review". The terms "review" and "evaluation" are often used to mean the same thing.	"Review" could also be confused with literature reviews or systematic reviews that are specific research and evaluation methods.
Secondary data	Existing data, including previous research or evaluations and administrative data, and information from monitoring. (See also primary data.)	
Social change	A change in the customs, institutions, or culture of a society, especially due to ideological or technological factors.	
Social impact	Social impact is the intended and unintended social consequences, both positive and negative, of planned interventions and any social change resulting from those interventions. In some respects it is the counterpart of environmental impact.	
Social investment	In New Zealand, social investment relates to: • Putting the needs of people who rely on public services at the centre of decisions on planning, programmes and resourcing. • Purchasing outcomes rather than specific inputs, and moving funding to the most effective services irrespective of whether they are provided by government or non-government organisations (NGOs).	
Stakeholder	People, groups or organisations that have a direct or indirect interest in the intervention being evaluated. This would include funders, managers and people representing the participants in a programme. There may also be indirect stakeholders if the intervention might have an effect on a wider public.	
Steering group	A group that steers and guides an evaluation. It provides feedback to the evaluators, and is ideally involved in the evaluation design. A good steering group brings both expertise and buy-in, helping ensure take-up of the findings of an evaluation. It may include the commissioner of the evaluation and some or all of the key stakeholders.	

Technical term	What it is	Terms that might also be used
Summative evaluation	An evaluation completed (often) after the intervention ends which looks back at what has been achieved. It is generally for an external audience (eg a funder). It looks at the extent to which the planned objectives were achieved, and is often about assessing the value of a programme. (See also formative evaluation.)	Impact evaluation
Target	A word with many meanings. It can be used to mean the people that you aim to assist (your target audience) or the level of change you aim to achieve (target outcome). In the latter case it will be a quantitative measure (a number or proportion).	
Theory of change	A comprehensive description of how and why a desired change is expected to happen in a particular context. It tries to understand the causal links between what you do and what change occurs – why you believe it will happen. This is probably bigger than an intervention logic which is more focused on how your programme is expected to lead to the desired outcomes.	Programme story Programme theory Intervention logic
Utility	How far the impacts achieved by the intervention correspond to the problem to be solved and society's needs. It looks at what actually happened, rather than the stated objectives, so is helpful when objectives are poorly defined or where results are unexpected.	
Validity	The extent to which the data collection strategies and instruments measure what they claim to measure. Validity can also refer to the extent to which findings are backed by evidence.	
Value for money	A judgement on the extent to which sufficient impact is achieved for the money invested. Various methods are used, which enable comparison of programmes. Value for money looks at both qualitative and quantitative measures, so it is important that your criteria are explicit.	

Tools, methods and associated terms

Tool or method	What it is	When it is most useful
Action research	Any research into practice undertaken by those involved in that practice, with an aim to change and improve it. What distinguishes action research from other forms of enquiry is the combination of the two aspects and the links between them. Action research is generally cyclical, with four inter-related stages: *plan, act, observe, reflect*. It tends to be collaborative in two senses: • many action research activities are best carried out with colleagues • action research always involves the participants, at least in knowing what is being explored and why It also tends to be qualitative rather than quantitative with the emphasis on language rather than numbers, and reflective, involving critical reflection on both the process and the outcomes.	Used in many field settings where it is not possible to use more traditional quasi-experimental research methods as they can't be readily adjusted to the demands of the situation. If you do alter them in midstream you may have to abandon the data collected up to that point.

Tool or method	What it is	When it is most useful
Administrative data	Administrative data is collected by government agencies or private organisations in the course of conducting their business or services, for operations such as delivering a service, or legal requirements to register events or as a record of transactions or events.	It is relatively cheap and easy to analyse and use your own administrative data to make better decisions, eg in allocating resources to programmes or services. This data may also help you to keep track of what is happening, to build historical knowledge and assist in planning. Comparing your own data to external (eg government) datasets can be helpful if the datasets contain comparable information on the same population of interest. The statistical analyses performed using your own data and an external set of data may give you a better, objective assessment of how well your programme/service works for your clients.
Case study	Case studies are used to arrive at an in-depth understanding of real world situations. It is usual for multiple sources of evidence to be used to examine the issue from a range of perspectives.	This approach may help you to understand, or may illustrate what has worked well, what has been achieved and what have been the issues.
Comparison group	A non-randomly selected group that does not receive the services, products or activities of the programme being evaluated.	
Control group	A randomly selected group that does not receive the services, products or activities of the programme being evaluated.	Used in randomised control trials, to reduce the possibilities that differences are caused by factors other than the initiative being examined.
Cost-benefit analysis (CBA)	A method of reaching economic decisions by comparing the costs of doing something with its benefits. A monetary value is assigned to both costs and benefits (positive and negative consequences). Various methods can be used to assign values when no market value is available. In a CBA the unit used will be dollars.	CBA judges the advantages of an intervention from the points of view of all the groups involved.

Tool or method	What it is	When it is most useful
Cost-effectiveness analysis	Cost-effectiveness considers the outputs produced by a project, which are not measured in monetary terms. In a cost-effectiveness analysis, units other than money will be used, such as reduction in sick days, or quality adjusted life years (QALYs)	Health care research, for example, often uses cost-effectiveness analyses to analyse the cost of health care interventions in terms of lives saved, illnesses prevented or years of life gained.
Decay/Drop-off	The effect of interventions, especially when aimed at behaviour change, may not last forever as people revert back to old behaviours. This needs to be taken into account when assessing future effects.	When looking at benefits over a long period, you should factor in some level of attrition.
Delphi Panel	Data is collected from a selected panel through a series of questionnaires. These are analysed and refined, so that the group starts to converge towards an agreed decision. This technique is often applied in a public policy setting.	A way for groups to build consensus, for example on what is appropriate evidence.
Developmental evaluation	This is an evaluation approach that can help develop social change initiatives in complex or uncertain environments. It enables real-time, or close to real-time, feedback to programme staff thus facilitating a continuous development loop.	Particularly suited to initiatives involving innovation, radical programme re-design, replication, complex issues or crises. It is not suitable for every situation.
Discounting	A technique used in economics to compare costs and benefits that occur in different time periods. It is a separate concept from inflation, and is based on the principle that, generally, people prefer to receive goods and services now rather than later. This is known as 'time preference'. The rate used in these calculations is referred to as the discount rate.	Used particularly when comparing initiatives producing benefits over different timescales or over long periods. It can also be used to calculate how long an investment will take to pay for itself or what return it will generate.
Ethnography	Detailed observations and interviews to explore social interactions, behaviours and perceptions that occur within groups, teams, organisations and communities. Ethnography provides rich, holistic insights into people's views.	Helpful in dealing with the complexity in society by better understanding people in the context of their lives. However, small sample sizes may not be seen as valid for decisions affecting large populations.
Experimental study or method	A methodology in which research subjects are randomly assigned to either a treatment or control group, data is collected both before and after the intervention, and results for the treatment group are benchmarked against a counterfactual established by results from the control group.	

Tool or method	What it is	When it is most useful
Focus group	A small but diverse group of people whose reactions are studied in open or guided discussion. Often used in market research or political analysis. Focus groups are similar to, but should not be confused with in-depth interviews. The moderator uses a discussion guide that has been prepared in advance of the focus group to guide the discussion.	The discussion in a focus group can produce insights or data that would be difficult to access outside a group setting.
Influence mapping	A process that involves identifying the individuals and groups with the power to effect a key decision or change, including the position and motives of each person and the best channels to communicate with them. Also known as stakeholder influence mapping, power mapping or the arena of influence.	
Longitudinal study	An observational study which follows the same group of people over a period of time to gather repeated information about their lives.	Longitudinal data can be used to: • look at the relationship between past behaviour and future outcomes • describe patterns over time (eg can track employment history) • examine before and after comparisons of an intervention
Net present value (NPV)	The discounted value of a stream of either future costs or benefits. The term Net Present Value (NPV) is used to describe the difference between the present value of a stream of costs and a stream of benefits.	
Participatory evaluation	Evaluation method in which representatives of agencies and stakeholders (including beneficiaries) work together in designing, carrying out and interpreting an evaluation.	When an evaluation is designed to provide insights to steer an initiative, then involvement of stakeholders in the evaluation process can provide useful buy-in.
Participatory evaluation	An evaluation in which the key stakeholders and participants are actively involved. They might be part of the agenda setting, design of evaluation or interpretation of results. A well designed participatory evaluation will still satisfy requirements for objectivity.	

Tool or method	What it is	When it is most useful
Quasi-experimental study or method	A methodology in which research subjects are assigned to treatment and comparison groups typically through some sort of matching strategy or natural experiment that attempts to minimise the differences between the two groups in order to approximate random assignment.	Can be applied when research subjects have not been randomly assigned to treatment and control by the researchers.
Randomised control trial (RCT)	A process in which participants are assigned to control or treatment groups at random, meaning that all members of the sample must have an equal chance of being selected for the control or treatment groups. This enables the assumption that the two groups are on average equivalent and there are no systematic differences between them. Through this the likelihood that any differences observed are due to the programme, rather than some other factor, are increased. An RCT won't tell you why the intervention worked or didn't work, or who benefits the most in your sample.	Used to test cause and effect. RCTs are typically much harder to design and carry out in the case of complex social problems. RCTs can present ethical issues in some contexts, eg withholding an intervention from some groups based on random allocation.
Realist evaluation	Realist evaluation asks the question: "What works, for whom, in what respects, to what extent, in what contexts, and how?" In order to answer that question, realist evaluators aim to identify the underlying generative mechanisms that explain 'how' the outcomes were caused and the influence of context.	
Sensitivity analysis	Analysis of the effects of varying the projected values of important variables in an analysis.	When you want to know how robust your results are, you can test them by changing one of your variables to see if the answer still holds. For example, in a cost benefit analysis, you might vary the discount rate used, to see what the effect is.
Social Return on Investment	A method for measuring values that are not traditionally reflected in financial statements, including social, economic and environmental factors, which can identify how effectively an organisation uses its capital and other resources to create value for the community.	SROI exists within and has strong similarities to traditional cost-benefit analysis but was first documented as a distinct methodology for the social enterprise/venture philanthropy context.

Tool or method	What it is	When it is most useful
Survey	A technique for gathering the opinions or experience of a group of people. They are often used to summarise results at a larger population level. Survey data can be used to: • Describe the characteristics of a population, at a given point-in-time • Describe associations between factors (eg that a relationship between x and y exists)	A well designed survey can tell the what, when and how and can cover a range of topics and issues. Surveys can also be set up quite quickly. However, they are not very good at determining why something happened. Very large surveys tend to be updated at quite wide time intervals. It is very easy to do a bad survey (poor sampling, ill-constructed questions).
Systematic review	The assembly, critical appraisal and synthesis of all relevant studies of a specific programme, practice or policy to assess its overall effectiveness, feasibility and "best practices" in its implementation.	This is a rigorous methodology which can provide a fast, high quality summary of a complete body of evidence. An advantage is that it reduces personal bias in identifying the evidence gaps. However, it requires a substantial number of robust studies or publications on a given topic, and if specially commissioned, can take quite a long time to produce.
Utilisation-Focused Evaluation	Utilisation-Focused Evaluation is an approach based on the principle that an evaluation should be judged on its usefulness to its intended users. Therefore evaluations should be planned and conducted in ways that enhance the likely utilisation of both the findings and of the process itself to inform decisions and improve performance.	

Sources

This glossary draws on a number of internationally recognised sources, such as:

DAC/OECD Glossary of Key Terms in Evaluation and Results Based Management

USAID Glossary of Evaluation Terms

BetterEvaluation.org

HM Treasury Green Book: Appraisal and Evaluation in Central Government

Guide to the Evaluation of Socioeconomic Development (European Commission)