

Helpdesk Research Report: Theory-based evaluation approach

19.12.2012 (revised 21.01.2012)

Query: Does the literature on theory-based evaluation for evaluation suggest a common and clearly defined approach to evaluation and impact evaluation? Does theory-based evaluation provide an analytical tool for carrying out the social and political analysis that should underpin impact evaluation of social development (SD) and governance programmes?

Enquirer: DFID

Author: Becky Carter (becky@gsdrc.org)

Contents

1. Overview
2. Rapid review of literature on theory-based evaluation (TBE)
 - 2.1 Background
 - 2.2 Lack of consensus
 - 2.3 General principles
 - 2.4 Some variations in TBE approaches
 - 2.5 Methods
 - 2.6 Practice
3. TBE tools for impact evaluation of SD and governance programmes
4. References
5. Additional Information

1. Overview

There is a wealth of literature available on theory-based evaluation and impact evaluation. For at least 40 years or more (some reviewers trace its origins back to the 1930s; most mention Suchard's work in the 1960s; and with rising prominence in the 1990s), academics and evaluators have debated and developed the concept. Given the time limitations, this report has attempted to identify and review a selection of the seminal studies and reviews from this large body of literature, guided by contributing experts' recommendations.

The proliferation of terms used on this topic can be confusing, in particular as they do not always have consistently distinct definitions (Rogers 2007). Terms used include: programme theory evaluation, theory-based evaluation, theory-guided evaluation, theory of action, theory of change, programme logic, logical frameworks, outcomes hierarchies, realist or realistic, and, more recently, programme theory-driven evaluation science (Coryn et al 2011). There are also various ways the terms evaluation 'design' and 'approach' are used. In this report, approach is used to refer to an evaluation strategy that specifies how the fundamental logic of the evaluation is implemented through the appropriate research methodology, and methods and tools (drawing on Stern et al 2012, 15).

For simplicity this report uses the label 'TBE approach' as shorthand to cover any evaluation approach that examines 'the assumptions underlying the evaluated intervention's causal chain from inputs to outcomes and impact' (White 2009b, 3).

Despite the large body of literature on TBE **experts disagree on whether TBE is a common and clearly defined approach**. Some think a common conceptual and operational understanding has been elusive (Coryn et al 2011), while others point to a completely consistent basic concept regardless of slight differences in the use of terminology (Booth expert comment; White 2009b).

There are some **core features of the TBE approach that appear consistent** across the main accounts of the approach:

- Opening up the black box to answer not simply the question of *what* works, but also *why* and *how* it worked. This is key to producing policy relevant evaluation.
- Understanding the transformational relations between treatment and outcomes, as well as contextual factors.
- Defining theory as the causal model or theory of change that underlies a programme.
- Having two key parts: *conceptual* (developing the causal model and using this model to guide the evaluation); and *empirical* (testing the causal model to investigate how programme cause intended or observed outcomes).
- Being issues led, and therefore, methods neutral.

Some of the **variations in TBE strategies** are:

- Approach to types of theory: whether the black box is empty, full of theories or inhabited by people, and the implications for how to accumulate knowledge and establish the theory of change.
- Approach to causal inference: the realist evaluation approach adopts a generative approach to attribution seen by some as distinct from other (i.e. experimental) designs; other approaches promote the use of a range of techniques and tools to make counterfactual comparisons under the TBE approach.

This review highlights the following key points from the literature.

- Some proponents of TBE promote the **benefits of applying a TBE approach to experimental designs**.
- Much of the guidance proposes the use of **mixed methods**, using both quantitative and qualitative data, but leaves open exactly how to go about choosing the appropriate design of mixed methods.
- Despite interest and rich literature on the TBE approach, **few studies apply the approach in practice** and additional exemplars of TBEs are seriously needed, including reports of

successes and failures, methods and analytic techniques, and evaluation outcomes and consequences (Coryn et al 2011).

- SD and governance programmes tend to be complicated and complex, and difficult to evaluate and there is little agreement on what a well designed evaluation looks like in these cases. A number of **recent works develop guidance on the TBE approach and tools specifically for the evaluation of complex and complicated programmes.**

This report summarises findings from the literature on whether TBE is a common and clearly defined approach, concluding with a focus on TBE tools for impact evaluation of SD and governance programmes.

2. Rapid review of literature on theory based evaluation

2.1 Background

Over the decades different approaches to evaluation have been prominent. Atheoretical experimentation was popular from the late 1950s (Chen and Rossi 1989), but by the 1980s method-oriented evaluation approaches were under attack for their inability to ‘open the black box’ (the space between the actual input and expected output of a programme) (Stame 2004, 58). Academics identified a paradigmatic shift by the late 1980s to looking at not just what works but why (Chen and Rossi 1989; Mcloughlin and Walton 2011).¹

There is interest **‘to establish and promote a credible and robust expanded set of designs and methods that are suitable for assessing the impact of complex development programmes’** (Stern et al 2012, 1).

2.2 Lack of consensus

TBE came to prominence two decades ago with Chen’s book ‘Theory-Driven Evaluations’ (Coryn et al 2011) followed by Weiss’ work (1997). Since then a number of articles, guidelines and textbooks have been published, **going some way to develop TBE into a detailed methodological framework.** These include: Chen (2005); Donaldson (2001, 2007); Funnell and Rogers (2011); Leeuw and Vaessen (2009); Rogers (2007, 2008, 2009); Pawson and Tilley (1997); Weiss (2001); White (2009b); and White and Phillips (2012). Some provide step-by-step guidance for implementing a TBE approach; others outline the logic of TBE rather than presenting clear methodological steps, leaving them open to a variety of interpretations in practice (White and Phillips 2012 on the realist evaluation approach).

Experts disagree on whether TBE is a common and clearly defined approach. Coryn et al (2011, 200) find that there is ‘little consensus on its nomenclature and central features’ and ‘a common vocabulary, definition, and shared conceptual and operational understanding has largely been elusive’. On the other hand, other experts think TBE is a well established approach (White 2009b) and that while ‘the language differs slightly among practitioners... the basic concept is completely consistent’ (Booth, expert comment).

¹ While beyond the scope of this report it is worth noting many other evaluation approaches – e.g. pragmatic evaluation, naturalistic evaluation, pluralist evaluation (Pawson and Tilley 1997) – have been popular over the years.

There is also inconsistency in understanding the scope of TBE – further complicated by varying interpretations of what is meant by the terms evaluation ‘approach’ and ‘design’. TBE has been described variously as being an underlying logic, broad concept or framework. For example, Rogers (2007) sees TBE as a design principle under which sit a variety of ways of developing a causal model and using this model to guide the evaluation. Some see these different ways of developing and using the causal model as revealing a lack of consistency in TBE, others consider that these more specific approaches fall under the broad TBE logic (Stern et al 2012).

2.3 General principles

This rapid review identifies some core features of the TBE approach that appear consistent across the main accounts of the approach.

- **A theory-based approach TBE opens up the black box to answer not simply the question of *what works*, but also *why* an intervention achieved its intended impact – or why it did not – and *how it worked* – or did not work.** Without greater understanding of the how and why an impact has occurred, evaluators are ‘helpless to improve on it in any way’ and in this vacuum, ‘efforts to change and improve the intervention may actually have adverse consequence’ (Scott and Seacrest 1989, 329). Adopting a TBE approach is the best way of ensuring evaluation policy relevance, ‘since it will yield information on how the program is working not just if it is working’ (White 2006, 20).
- TBE is concerned with impact (in terms of investigating final results and attribution to the intervention being evaluated) (Booth, expert comment), and therefore, ‘committed to internal validity’ (Stame 2004, 63). However, ‘it does not allow this concern to interfere with **the main task of establishing how any impacts are achieved, and why exactly they are not achieved when they are not** – which are the keys to doing better next time’ (Booth, expert comment).
- TBE emphasises **an understanding of the transformational relations between treatment and outcomes** (Chen and Rossi 1989). TBE aims to identify these ‘mechanisms’ that make things happen. This goes from asking whether a programme works to understanding what it is about the programme that makes it work (Stern et al 2012).
- **TBE considers programmes in their context, which includes actors’ environments (embeddedness) and the culture and behaviour of the wider programme context** (Stame 2004). It bases the evaluation on ‘an account of what may happen, as understood by actors and/or interpreted by evaluators: **values are accounted for in the way they help frame the actors’ views, and are not ignored**’ (Ibid. 2004, 63).
- TBE has two vital components (Rogers et al 2000 cited in Coryn et al 2011).
 - 1) Conceptual: TBE is about developing the causal model that links programme inputs and activities to a chain of intended or observed outcomes, and then using this model to guide the evaluation.** Any evaluation that explicates the theory behind a programme but does not use it to guide the evaluation is not a TBE (Ibid.). Other terms for the causal model are professional logic (Rogers 2007, 70) and theory of change (popularised by Weiss), or also programme logic and intervention logic (Funnel and Rogers 2011). Leeuw (2003, 6) explains there is an important difference between the concepts of programme logic and programme theory: the former rarely outlines the

responsible mechanisms for the linkage between inputs and outcomes while the latter explicitly identifies how a programme causes the outcomes. It is worth noting that there is no single definition or methodology for this programme theory (Vogel 2012). Vogel (2012, 4) sets out that as a minimum, theory of change is considered to encompass a discussion of the following elements: context, long-term change, process/sequence of change, assumptions, a diagram and narrative summary.

2) Empirical: TBE involves testing the causal model to investigate how programme cause intended or observed outcomes, by collecting evidence to validate, invalidate or revise the hypothesised explanations (also called assumptions) with the goal of rigorously evidencing the links in the actual causal chain (White and Philips 2012).

- **TBE is ‘issues-led not methods-led’**, as all evaluations should be according to White (2010, 162). The TBE approach uses all suitable methods without privileging or depending on any of them (Stame 2004).
- **TBE can be a comprehensive evaluation covering the whole programme theory or tailored to focus on only one aspect, element or chain of the programme theory** (Weiss and Chen cited in Coryn et al 2011). The latter options may be ‘process’ rather than ‘impact’ evaluations.

Coryn et al (2011) have produced a framework of TBE principles, derived from a systematic analysis of the theory of the approach. They found this exercise formidable given that TBE: 1) has no obvious ideological basis; and 2) a wide variety of practitioners claim to be theory-driven in some capacity. Their derived principles are a mix of general rules of conduct and qualities, and methodological action. They are situational; no evaluation will necessarily cover the entire framework as each is contingent on a variety of factors including the evaluation’s nature, purpose and intended use/users. See Table 1 below.

Table 1: Core principles and sub principles of theory-driven evaluation (Coryn et al 2011)

<ol style="list-style-type: none"> 1. Theory-driven evaluations/evaluators should formulate a plausible programme theory from: <ol style="list-style-type: none"> a. existing theory and research (e.g. social science theory) b. implicit theory (e.g. stakeholder theory) c. observation of the programme in operation/exploratory research (e.g. emergent theory) d. a combination of any of the above (i.e. mixed/integrated theory) 2. Theory-driven evaluations/evaluators should formulate and prioritise evaluation questions around a programme theory: <ol style="list-style-type: none"> a. formulate around programme theory b. prioritise evaluation questions 3. Programme theory should be used to guide planning, design and execution of the evaluation under consideration of relevant contingencies: <ol style="list-style-type: none"> a. design, plan and conduct evaluation around a plausible programme theory b. design, plan and conduct evaluation considering relevant contingencies (e.g. time, budget and use) c. determine whether evaluation is to be tailored (i.e. only part of the program theory) or comprehensive 4. Theory-driven evaluations/evaluators should measure constructs postulated in programme theory: <ol style="list-style-type: none"> a. measure process constructs postulated in programme theory b. measure outcome constructs postulated in programme theory c. measure contextual constructs postulated in programme theory 5. Theory-driven evaluations/evaluators should identify breakdowns, side effects, determine programme effectiveness

(or efficacy), and explain cause-and-effect associations between theoretical constructs:

- a. identify breakdowns, if they exist (e.g. poor implementation, unsuitable context and theory failure)
- b. identify anticipated (and unanticipated), unintended outcomes (both positive and negative) not postulated by programme theory)
- c. describe cause-and-effect associations between theoretical constructs (i.e. causal description)
- d. explain cause-and-effect associations between theoretical constructs (i.e. causal explanation)
 - i. explain differences in direction and/or strength of relationship between programme and outcomes attributable to moderating factors/variables
 - ii. explain the extent to which one construct (e.g. intermediate outcome) accounts for/mediates the relationship between other constructs

2.4 Some variations in TBE approaches

Reviews of TBE have pointed to some variations in the different developments of the approach. These differences include how they deal with theory and how they tackle causal inference.

Approaches to theory

Blamey and Mackenzie (2007, 442) explain that '**a key problem in getting to grips with the literature on TBE is the fundamental lack of consistency on how different types of theory are described**', and **the implications for approaches to knowledge accumulation**. They identify that different terms are used to describe the same type of theory and similar labels are given to epistemologically separate kinds of theory.

Stame (2004) gives a summary of **the differences in how TBE theories deal with the 'black box problem'**.

- Chen and Rossi (1989) see the **black box as an empty box** – such programmes have no theory and TBE should provide it by studying treatment; discussing stakeholders' and evaluators' views on outcomes; explaining why and how a programme fares as it does (following both normative and causal theories).
- Weiss (1987) sees the **black box as full of theories** – or 'theories of change' – all of which have to be brought to light to reach a consensus on which deserve to be tested. TBE should make the mechanisms clear and use data of different kinds to test them.
- Pawson and Tilly (1997) see the **black box as inhabited by people** – it is people, embedded in their context who, when exposed to programmes, do something to activate mechanisms and change. Stame (2004) says this makes for a completely different design of evaluation as different views are not obtained through consensus (Weiss) but through 'adjudication', ie. establishing what may be more worthy.

Coryn et al (2011) present **further variations within TBE approaches on how to establish the programme theory**. Patton (2008) favours deductive, inductive, or user-oriented approaches to developing programme theory; Donaldson (2001, 2007) describes four potential sources (prior theory and research, implicit theories of those close to the program, observations of the programme in operation and exploratory research to test critical assumptions); and Chen (2005) advocates a stakeholder-oriented approach, with the evaluator playing the role of facilitator.

Leeuw and Vaessen (2009) describe **two steps (identify intervention theory and reconstruct underlying assumptions) in establishing a programme theory** and provide guidance on the type of evidence and methods for data collection. For reconstruction of the underlying assumption, they identify three strategies: policy-scientific method (focuses on interviews, documents, argumentation analysis); strategic assessment method (focuses on group dynamics and dialogue); and elicitation method (focuses on cognitive and organisational psychology). Leeuw (2003) also stresses the need to make the underlying programme theories more transparent so that they are open to scrutiny and to facilitate reconstruction.

Approaches to causal inference

Blamey and Mackenzie (2007) undertake **a detailed comparison of two TBE approaches – realist or realistic evaluation** (Pawson and Tilly 1997) **and theories of change** (Weiss 1995) **and conclude that although they ‘may both be from the same stable, they are in practice very different horses’** (Ibid., 452). They reach this conclusion based on the differences in conceptualising theory (as discussed in previous section) and differences in causal attribution. According to Blamey and Mackenzie (2007, 449-450), differences in causal attribution are:

- The theories of change approach argues that the attribution problem can be partly addressed through the process of building consensus amongst a wide group of stakeholders about a programme’s theory and then testing the extent to which anticipated thresholds, timelines and outcomes are achieved.
- Realists, on the other hand, adopt a generative approach to attribution. This is explicitly focused on a cumulative and iterative process of theory building, testing and refinement in relation to specific programme subcomponents. It seeks patterns between interventions and their outcomes, and focuses on the generative mechanism by which the relationship is established.

They suggest that **many policy programmes lend themselves to the explicit testing of a dual theories of change/ realistic evaluation model**. Theories of change could be the means of explicating implementation theory for the purpose of programme planning, improvement and the development of robust monitoring systems at a macro programme level, while realistic evaluation approaches might then be brought to bear on more micro-level aspects of the most promising programme theories (Blamey and Mackenzie 2007).

Another set of detailed guidance on understanding causal inference using a TBE approach is presented by Funnell and Rogers (2011). Their framework has three steps: 1) congruence with the programme theory (do the results match the programme theory?); 2) counterfactual comparisons (what would have happened without the intervention?); and 3) critical review (are there other plausible explanations of the results?). They describe a range of techniques and tools to implement the approach. Critical points in their guidance are:

- Counterfactual comparisons can use different designs and techniques, ‘including informant assessment, experimental design, quasi-experimental designs, and qualitative comparative analysis’ (Ibid, 474).
- Understanding associated concepts such as sufficiency and ‘necessariness’ to assist with refining counterfactual comparisons

- A good programme theory identifies both programme and non-programme factors that may influence outcomes.

However, a recent report by Stern et al defines theory-based evaluation as distinct from experimental and other evaluation designs, according to a taxonomy based on the approaches' different bases for causation inference. They define TBE as having an inference basis of 'generative causation', which depends on identifying the 'mechanisms' that explain effects, as compared with the different causation inference basis of other evaluation approaches (statistical, experimental and quasi-experimental, configurational) – see Table 2.

Table 2: Taxonomy of design approaches, variants and causal inference (Stern et al 2012)

Design approaches	Specific Variants	Basis for Causal Inference
Experimental	RCTs, Quasi experiments natural Experiments	Counterfactuals; the copresence of cause and effects
Statistical	Statistical Modelling, Longitudinal studies, Econometrics	Correlation between cause and effect or between variables, influence of (usually) isolatable multiple causes on a single effect Control for 'confounders'
Theory-based	<i>Causal process designs</i> : Theory of change, Process tracing, Contribution Analysis, Impact pathways <i>Causal mechanism designs</i> : Realist evaluation, Congruence analysis	Identification/confirmation of causal processes or 'chains' Supporting factors and mechanisms at work in context
'Case-based' approaches	<i>Interpretative</i> : Naturalistic, Grounded theory, Ethnography <i>Structured</i> : Configurations, QCA, Within-case-analysis, Simulations and Network analysis	Comparison across and within cases of combinations of causal factors Analytic generalisation based on theory
Participatory	<i>Normative designs</i> : Participatory or democratic evaluation, Empowerment evaluation <i>Agency designs</i> : Learning by doing, Policy dialogue, Collaborative action research	Validation by participants that their actions and experienced effects are 'caused' by programme Adoption, customisation and commitment to a goal
Synthesis studies	Meta analysis, Narrative synthesis, Realist based synthesis	Accumulation and aggregation within a number of perspectives (statistical, theory based, ethnographic etc.)

While a particular variant of TBE – realist evaluation – is 'epistemologically antagonistic to the use of controlled trials' (Blamey and Mackenzie 2007, 450), others see this dichotomy between TBE and experimental designs as '**false**' (Cook 2000 cited in Rogers 2007, 66) **and discuss the benefits of applying a TBE approach to experimental designs** (White 2009; Leeuw and Vaessen 2009). White (2009b, 3) says 'elaborations of program theory have long been used by some practitioners of experimental and quasi-experimental approaches as a way of explaining their findings'.

White's (2009b) **guidance for undertaking a theory-based impact evaluation** sets out six general principles: 1) map out the causal chain (programme theory); 2) understand context; 3) anticipate heterogeneity; 4) rigorous evaluation of impact using a credible counterfactual; 5) rigorous factual analysis; and 6) use mixed methods. He also suggests appropriate research designs, techniques and methods, including how to define an appropriate counterfactual usually using a control group through either experimental or quasi-experimental approaches

2.5 Methods

TBE proponents emphasise that the approach is 'method neutral'. TBE guidance by Leeuw and Vaessen (2009) explains that the theory acts as template for method choice, variable selection, and other data collection and analysis issues.

Much of the guidance highlights the advantages of using mixed methods, but leaves open exactly how to go about choosing the appropriate design of mixed methods. While it is common to discuss the importance of 'mixed methods' at a general level it is 'harder to define what and how and in which context these could be mixed and combined in impact evaluation' (Passen, expert comment). Chen (2006) underlines the 'great need for systematically developing mixed method 'use' strategies, as well as establishing its own standards and criteria for assessing the method use' (quoted in Riche 2012, 12). White (2012, 4) concludes that 'while the logic underlying the methodologies is usually well developed, less has been done to set out how evaluation methods could be systematically applied to promote the validity of conclusions'.

It has not been possible in the time limits and scope of this research question to review the literature on mixed methods. How to select and design the optimum mix of methods is another area of work with its own body of literature.

2.6 Practice

TBE has attracted many supporters as well as detractors (Coryn et al 2011). Organisations that have increasingly promoted a TBE approach in international development settings include the Overseas Development Institute (ODI), the International Initiative for Impact Evaluation (3ie), the United Nations Evaluation Group (UNEG) and the Independent Evaluation Group (IEG) of the World Bank for evaluating humanitarian efforts (Ibid.).

However, despite interest and rich literature on the TBE approach, few studies apply the approach in practice (McLoughlin and Walton 2010). Scriven (1998, 59) finds that much of what passes as theory-based evaluation today is simply a form of 'analytic evaluation [which] involves no theory in anything like a proper use of that term' (cited in Leeuw and Vaessen 2009). Stern et al (2012) also reviewed existing evaluation examples and found that theories of change were not routinely articulated even when this would have helped draw causal inferences.

Coryn et al (2011), who undertook a systematic review of TBE practice from 1990 to 2009, concluded that **additional exemplars of TBES are seriously needed**, including reports of successes and failures, methods and analytic techniques, and evaluation outcomes and consequences.

3. TBE tools for impact evaluation of SD and governance programmes

SD and governance programmes tend to be complicated and complex, and ‘difficult to evaluate’ (Stern et al 2012, 11). It is beyond the scope of this report to go into the detail of these challenges; a useful summary is provided by Vogel (2012, 49-50).

There appears to be less agreement on what a well designed study looks like in the case of small n qualitative assessment, which is likely to be required for evaluation of SD and governance programmes. According to White and Phillips (2012), there is broad agreement on what a well designed study looks like for large n impact evaluations (involving tests of statistical difference in outcomes between treatment and comparison groups) and small n modelling-based approaches. There is not the same agreement for qualitative assessments.

In addition, while over time linear models to describe programme theories have developed into more contextualised, comprehensive, ecological program theory models, **some still question whether these models can adequately represent ‘complex realities and unpredictable, continuously changing, open and adaptive systems’** (Patton 2010 cited in Coryn et al 2011).

Some proponents of TBE see it as essential for evaluating all interventions that involve changing institutions: ‘because we understand so little about how useful institutional changes occur, so we need to be forced to make explicit what we are assuming about how results are achieved and use evaluations to see whether we are right or not’ (Booth, expert comment).

A number of recent works develop guidance on the TBE approach and tools specifically for the evaluation of complex and complicated programmes. Some of these take examples of evaluations of governance and social development interventions and explain how a TBE approach worked:

- Vogel (2012) provides **suggestions on how theory of change thinking can help the analyses of complicated and complex aspects of programmes**, through looking at theory of change thinking as a ‘learning lens’ that invites dialogue and triangulation from a number of viewpoints and sources of evidence.
- Rogers’ (2008) guidance on using a theory-based evaluation approach to evaluate complicated and complex programme interventions **delineates the characteristics of these types of programmes and gives successful (and unsuccessful) examples of implementing TBE approach and tools.**
- Carvalho and White (2004) lay out the **application of a theory-based approach to the evaluation of social fund projects**, with a focus on the issues of subproject sustainability and institutional development impact. This approach rests on making explicit the assumptions which underlie the way in which a programme is meant to work.
- White and Masset (2004) **present a theory-based impact evaluation of the Community-Based Nutrition Component of the Bangladesh Integrated Nutrition Project** which illustrates the benefits of a theory-based impact evaluation, i.e. one which traces the links from inputs through to impacts, rather than only look for evidence of impact. Rigorous quantitative methods were used.

- White and Phillips (2012) have set out **general steps for examining the causal relationship in cases (commonly small *n*) where a credible counterfactual cannot be measured using experimental or quasi-experimental approaches**. They also identify a set of approaches that can systematically consider how an outcome might have occurred and what evidence and targeted data collection is needed, through drawing on the theory of change and being alert to variations in implementation and external factors. The approaches include: realist evaluation; general elimination methodology (aka the Modus Operandi method); process tracing; contribution analysis. They also identified another set of complementary approaches that place stakeholder participation at the heart of data collection: most significant change; success case method; outcome mapping; method for impact assessment of programs and projects (MAPP).

4. References

Blamey, A. and Mackenzie, M., 2007, 'Theories of Change and Realistic Evaluation: Peas in a Pod or Apples and Oranges', *Evaluation* 2007 13: 439 <http://evi.sagepub.com/content/13/4/439.full.pdf+html>

Carvalho S. & White H., 2004, 'Theory-based Evaluation: The Case of Social Funds' *American Journal of Evaluation*, 25(2), pp.141–60 <http://aje.sagepub.com/content/25/2/141.full.pdf>

Chen, H.T., 1990, 'Theory-driven evaluations', Thousand Oaks, CA, Sage
http://books.google.co.uk/books/about/Theory_Driven_Evaluations.html?id=-hRsvHu85jkC

Chen, H.T., 1994, 'Theory-driven Evaluations: Need, Difficulties, and Options', *American Journal of Evaluation* February 1994 15: 79-82 <http://aje.sagepub.com/content/15/1/79.refs>

Chen, H. T., 2005, 'Practical program evaluation: Assessing and improving planning, implementation, and effectiveness', Thousand Oaks, CA, Sage
http://books.google.co.uk/books/about/Practical_Program_Evaluation.html?id=2rcl2C0FBPAC

Chen, H. T., 2006, 'A theory-driven evaluation perspective on mixed methods research', *Research in the Schools*, Vol. 13(1), 2006 <http://tinyurl.com/bqtuce9>

Chen, H. & Rossi, P., 1989, 'Issues in the Theory-driven Perspective', *Evaluation and Program Planning* (12)4, pp.299–306
http://econpapers.repec.org/article/eeeepplan/v_3a12_3ay_3a1989_3ai_3a4_3ap_3a299-306.htm

Coryn, C., Noakes, L., Westine, C. and Schroter, D., 2011, 'A Systematic Review of Theory-Driven Evaluation Practice From 1990 to 2009', <http://tinyurl.com/cbvbc5h>

Donaldson, S., 2001, 'Mediator and moderator analysis in program development', In S. Sussman (Ed.), 'Handbook of program development for health behavior research' (pp. 470-496), Newbury Park, CA, Sage <http://www.uk.sagepub.com/books/Book10856>

Donaldson, S., 2007, 'Program Theory-Driven Evaluation Science: Strategies and Applications', Thousand Oaks, Sage <http://tinyurl.com/ckv6vxv>

Funnell, S., & Rogers, P., 2011, 'Purposeful Program Theory: Effective Use of Theories of Change and Logic Models', San Francisco, Wiley & Sons <http://tinyurl.com/btcksyu>

DG Regional Policy, no date, 'Theory based evaluation', Based on material produced for DG Regional Policy by Frans L. Leeuw

http://ec.europa.eu/regional_policy/information/evaluations/pdf/impact/theory_impact_guidance.pdf

Garbarino, S. and Holland, J., 2009, 'Quantitative and Qualitative Methods in Impact Evaluation and Measuring Results', March 2009, GSDRC Issues Paper,

<http://www.gsdrc.org/docs/open/EIRS4.pdf>

Leeuw, F., 2003, 'Reconstructing Program Theories: Methods Available and Problems to be Solved', American Journal of Evaluation 2003; 24; 5 <http://ics.uda.ub.rug.nl/FILES/root/Articles/2003/LeeuwFL-Recon/LeeuwFL-Reconstructing-UU--2003.pdf>

Leeuw, F., 2012, 'Linking theory-based evaluation and contribution analysis: Three problems and a few solutions', Evaluation 2012 18: 348

<http://evi.sagepub.com/content/18/3/348.abstract>

Leeuw, F. and Vaessen, J., 2009, 'Impact Evaluations and Development: NoNIE Guidance on Impact Evaluation', The Network of Networks on Impact Evaluation, Washington DC

<http://www.gsdrc.org/go/display&type=Document&id=4099>

Mansuri, G., and Vijayendra R., 2013, 'Localizing Development: Does Participation Work?', Washington, DC, World Bank <http://tinyurl.com/cu8xp9f>

Mcloughlin, C. and Walton, O., 2012, 'Topic Guide: Measuring Results', GSDRC, University of Birmingham

<http://www.gsdrc.org/go/topic-guides/measuring-results>

OECD DAC, 2002, 'Glossary of Key Terms in Evaluation and Results Based Management', OECD-DAC, Paris. <http://www.oecd.org/development/peerreviewsofdacmembers/2754804.pdf>

Patton, M. Q., 2008, 'Utilization-focused evaluation (4th ed.)', Thousand Oaks, CA, Sage

http://books.google.co.uk/books?id=UKGHN5UuDjC&source=gbs_similarbooks

Patton, M. Q., 2010, 'Developmental evaluation: Applying complexity concepts to enhance innovation and use', New York, NY, Guilford

http://books.google.co.uk/books/about/Developmental_Evaluation.html?id=s5okv_bZ8EQC

Pawson, R., & Tilley, N., 1997, 'Realistic Evaluation', London, Sage

http://books.google.co.uk/books/about/Realistic_Evaluation.html?id=GXagvZwC2bcC

Ramalingham, 2011, 'Learning how to learn: eight lessons for impact evaluations that make a difference', Background Note, April 2011, ODI <http://www.odi.org.uk/sites/odi.org.uk/files/odi-assets/publications-opinion-files/7096.pdf>

Rodrik, D., 2008, 'The new development economics: we shall experiment, but how shall we learn?', Revised version of paper presented at Brookings Development Conference, 29-30 May

<http://tinyurl.com/bpzdlwb>

- Rogers, P., 2007, 'Theory-Based Evaluation: Reflections Ten Years On', *New Directions for Evaluation* Volume 2007, Issue 114, pages 63–81 <http://tinyurl.com/cvwr9z2>
- Rogers, P., 2008, 'Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions', *Evaluation* 2008 14: 29 <http://evi.sagepub.com/content/14/1/29.refs>
- Rogers, P., 2009, 'Matching impact evaluation design to the nature of the intervention and the purpose of the evaluation', *Journal of Development Effectiveness*, 1:3, 217-226 <http://ideas.repec.org/a/taf/jdevef/v1y2009i3p217-226.html>
- Riché, M., 2012, 'Theory Based Evaluation: A wealth of approaches and an untapped potential', European Commission http://ec.europa.eu/regional_policy/impact/evaluation/conf_doc/helsinki_mri_2012.pdf
- Rogers, P. J., Petrosino, A., Huebner, T. A., & Hacsí, T. A., 2000, 'Program theory evaluation: Practice, promise, and problems', In P. J. Rogers, T. A. Hacsí, A. Petrosino, & T. A. Huebner (Eds.), 'Program theory in evaluation: Challenges and opportunities' (pp. 5–14), *New Directions for Evaluation*, No. 87. San Francisco, CA: Jossey-Bass. <http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0787954322.html>
- Savedoff, W.D., Levine, R., and Birdsall, N., 2006, 'When Will We Ever Learn? Improving Lives Through Impact Evaluation', *The Evaluation Gap Working Group*, Centre for Global Development, May 2006 <http://www.cgdev.org/content/publications/detail/7973>
- Scott, A. and Sechrest, L., 1989, 'Strength of Theory and Theory of Strength', *Evaluation and Program Planning* (12)4, pp.329-336 <http://www.egadconnection.org/Strength%20of%20theory%20and%20theory%20of%20strength.pdf>
- Stame, N., 2004, 'Theory-based Evaluation and Types of Complexity', *Evaluation*, Vol 10(1): 58–76 http://www.stes-apes.med.ulg.ac.be/Documents_electroniques/EVA/EVA-GEN/ELE%20EVA-GEN%207360.pdf
- Stern, E., Stame, N., Mayne, J., Forss, K., Davies, R. and Befani, B., 2012, 'Broadening The Range Of Designs And Methods For Impact Evaluations. Report of a study commissioned by the Department for International Development', April 2012, Working Paper 38, DFID <http://www.dfid.gov.uk/Documents/publications1/design-method-impact-eval.pdf>
- Ton, G., 'The mixing of methods: A three-step process for improving rigour in impact', *Evaluation* 2012 18: 5 <http://evi.sagepub.com/content/18/1/5.full.pdf+html>
- Vogel, V., 2012, 'Review of the use of 'Theory of Change' in international development', *Review Report for DFID*, April 2012 http://www.dfid.gov.uk/r4d/pdf/outputs/mis_spc/DFID_ToC_Review_VogelV7.pdf
- Weiss, C., 1987, 'Where Politics and Evaluation Research Meet', in D. Palumbo (ed.) 'The Politics of Program Evaluation', Newbury Park, CA, Sage <http://tinyurl.com/d8x9ucr>
- Weiss, C., 1995, 'Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community-Based Initiatives for Children and Families', in J. P. Connell, A. C.

Kubisch, L. B. Schorr and C. H. Weiss (eds), 1995, *New Approaches to Evaluating Community Initiatives*, vol. 1, Concepts, Methods and Contexts, Washington, DC, Aspen Institute

Weiss, C., 1997, 'Theory-Based Evaluation: Past, Present, and Future', *Evaluation*, Volume 1997, Issue 76, pages 41–55 <http://tinyurl.com/cwtk7zd>

Weiss, C., 2001, 'Theory-based Evaluation: Theories of Change for Poverty Based Programs', In O. Feinstein & R. Picciotto, eds. *Evaluation and Poverty Reduction*. Brunswick, Transaction
<http://tinyurl.com/aww234w>

White, H., 2006, 'Impact Evaluation: the Experience of the Independent Evaluation Group of the World Bank', Independent Evaluation Group, Washington D.C..
<http://ideas.repec.org/p/pramprapa/1111.html>

White, H., 2009a, 'Some Reflection on Current Debates in Impact Evaluation', April 2009, Working Paper 3, International Initiative for Impact Evaluation 3ie
http://www.3ieimpact.org/media/filer/2012/05/07/Working_Paper_1.pdf

White, H., 2009b, 'Theory Based Impact Evaluation: Principles and Practice', June 2009, Working Paper 3, International Initiative for Impact Evaluation 3ie
http://www.3ieimpact.org/media/filer/2012/05/07/Working_Paper_3.pdf

White, H., 2010, 'A Contribution to Current Debates in Impact Evaluation', *Evaluation* 2010 16: 153
<http://tinyurl.com/cvhhr8j>

White, H. and Phillips, D., 2012, 'Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework', Working Paper 15, International Initiative for Impact Evaluation 3ie http://www.3ieimpact.org/media/filer/2012/06/29/working_paper_15.pdf

5. Additional Information

Key Websites:

3ie International Initiative for Impact Evaluation <http://www.3ieimpact.org/>

Development Impact Evaluation Initiative (World Bank) <http://tinyurl.com/36h9mg3>

Methods Lab <http://www.odi.org.uk/projects/2599-methods-lab-building-relevant-timely-tools-impact-evaluation-assessment>

Experts consulted:

David Booth, ODI

Catherine Dom, Mokoro Ltd

Maren Duvendack ODI

Frans Leeuw, Maastricht University

Arianna Legovini, World Bank

Stephen Lister, Mokoro Ltd

Erika McAslan Fraser, Social Development Direct

Tiina Pasanen, ODI

Louise Shaxson, ODI

Howard White, 3ie International Initiative for Impact Evaluation

Suggested citation:

Carter, R. (2012), 'Theory-based evaluation approach (GSDRC Helpdesk Research Report)', Birmingham, UK: Governance and Social Development Resource Centre, University of Birmingham.

About Helpdesk research reports: This helpdesk report is based on 3.5 days of desk-based research. Helpdesk research reports are designed to provide a brief overview of the key issues, and a summary of some of the best literature available. Experts are contacted during the course of the research, and those able to provide input within the short time-frame are acknowledged.