Centre for Development Impact
PRACTICE PAPER ANNEX
Innovation and learning in impact evaluation

Applying Process Tracing in Five Steps

Abstract This Practice Paper Annex describes the main steps in applying process tracing, as adapted from *Process-Tracing Methods: Foundations and Guidelines* (Beach and Pedersen 2013). It also provides some examples of how these steps might be applied in practice, drawing on a case study discussed in CDI Practice Paper 10.

Introduction

This Practice Paper Annex is the companion to CDI Practice Paper 10, both of which focus on process tracing. CDI Practice Paper 10 discusses what process tracing can offer to impact evaluation and draws on two early applications of process tracing for assessing impact in international development interventions as case studies. This annex describes the main steps in applying process tracing, as adapted from a recent book by Derek Beach and Rasmus Brun Pedersen: *Process-Tracing Methods: Foundations and Guidelines* (2013). It also provides some examples of how these steps might be applied in practice, drawing on the case study of the Universal Health Care Campaign evaluation (Stedman-Bryce 2013), as discussed in CDI Practice Paper 10.

A few caveats are required. First, this annex is based on the authors' interpretation of Beach and Pedersen's method, and we encourage readers to refer to the book if they decide to apply it. Second, several of the steps described below may be difficult to apply in the context of impact evaluation and/or international development – as discussed further in CDI Practice Paper 10. To our knowledge, no completed impact evaluations have used this method in full; and the illustrative examples are drawn from an evaluation that used a different variant of process tracing (described in Oxfam GB 2011 and discussed in CDI Practice Paper 10). This annex does not therefore represent a tried-and-tested guide. However, it may provide a useful starting-point for evaluators who are considering using process tracing, and facilitate further debate on its applicability within impact evaluation and international development.

What is a 'case' in process tracing?

Process tracing involves an in-depth analysis of a single case. A case could be (for example) an intervention, a project, a country, an individual, or an organisation. A case in process tracing *must* include:

- The effect under investigation (in impact evaluation, likely to be an outcome or impact – these terms are used interchangeably in this annex);
- The hypothesised cause (likely to be an intervention, or part of an intervention); and
- The processes or events that link the hypothesised cause and the effect (for example, the activities and intermediate outcomes of an intervention).

1 Three types of process tracing

Beach and Pedersen (2013) outline three distinct types of process tracing: theory-testing, theory-building and explaining-outcome. Each uses a different approach to analysing how a specific cause (A) led to a given outcome (B) within a case.

Theory-testing (TT) process tracing is used when:

- We know what A and B are (for example, we know that we ran an intervention, A, and we know that outcome B has occurred).
- We think there is a causal link between A and B (for example, we think the outcome occurred at least in part because of our intervention).
- We think we know why A led to B (for example, we have a theory of change that explains why our intervention should lead to outcome B).

Theory-building (TB) process tracing is used when:

- **Either (version 1)** we know what both A and B are (for example, we know that we ran an intervention, A, and we know that outcome B has occurred), and we think there is a causal link between A and B;
  or (version 2) we know B but do not know A (for example, we know that outcome B occurred, but we are not sure what caused it).
- We are in the dark about why A led to B (for example, we do not have a theory of change that explains why our intervention might lead to outcome B).

Explaining-outcome process tracing is used when:

- We know what B is (there is an interesting outcome that we want to investigate). But we do not know what A is (we do not know what caused B).
- We are interested in *fully explaining* why B happened – working out all the various factors that contributed to it in order to craft a 'minimally sufficient' explanation for B.

In all three versions, the researcher already knows B (the outcome). This has some implications for impact evaluation, discussed further in CDI Practice Paper 10.

Both theory-testing and theory-building process tracing methods aim to develop *theory* that is generalisable to other interventions or situations. In contrast, explaining-outcome process tracing is *case-centric* rather than *theory-centric*. The main implication of this is that the explanation it generates is *only relevant* to the specific case being investigated, and so cannot be generalised to other situations. This approach is the most common form of process tracing that has been used to date, and is often used to provide explanations for historical events (for example, the causes of the First World War).

This annex largely focuses on articulating the steps in theory-building and theory-testing process tracing. These seem most applicable to impact evaluation because:

- When evaluating impact, evaluators almost always know A (i.e. that an intervention was conducted) and are usually interested in the relationship between A and an outcome (B).
- Impact evaluations are often interested in generating findings that can be applied in other programmes, rather than generating an explanation of an outcome that is deeply case-specific and only relevant to one individual case.

2 Causal mechanisms in process tracing

CDI Practice Paper 10 introduces causal mechanisms in process tracing. As mentioned, the concept of mechanisms is also used elsewhere, for example in realist evaluation (Pawson and Tilley 2004; Westhorp 2014). The concept can be confusing – one recent paper argues that there is no clear consensus in the literature on what exactly constitutes a mechanism (Shaffer 2014). This section expands on the Practice Paper, providing further detail on Beach and Pedersen's interpretation of causal mechanisms in process tracing.

One intuitive way to understand mechanisms for the purposes of process tracing is using the 'black box' analogy commonly used in international development. Many impact evaluation designs focus on verifying whether project activities took place (A), and attempting to verify (and sometimes measure the size of) the expected impact (B). But this does not necessarily tell us anything about **what it is about A that leads to B** – how and why the intervention led to a specific outcome. In many impact evaluation methods, there is therefore a black box between the intervention and the outcome that remains closed (see Figure 1).

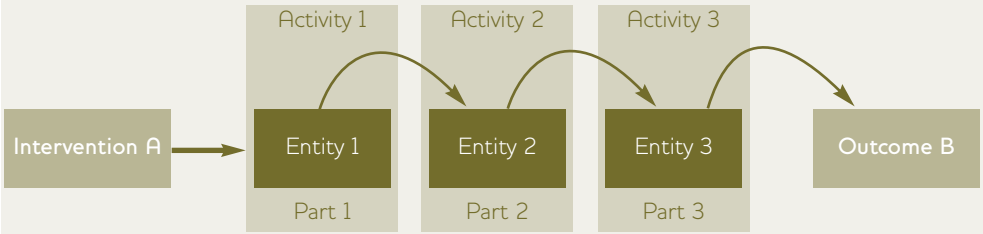
Figure 1 **The black box in impact evaluation**

Source: Authors' own.

Process tracing methods open up this black box, and the causal mechanism is what is inside. This mechanism can be understood as a *force* or a *power* – the thing that causes event A to give rise to outcome B. Using gravity as an analogy: if I drop a tennis ball and it falls to the ground, gravity is the 'mechanism' that explains why A (opening my hand) leads to B (the tennis ball falling). Mechanisms are underpinned by a generative model of causal inference, discussed in CDI Practice Paper 10, Box 1.

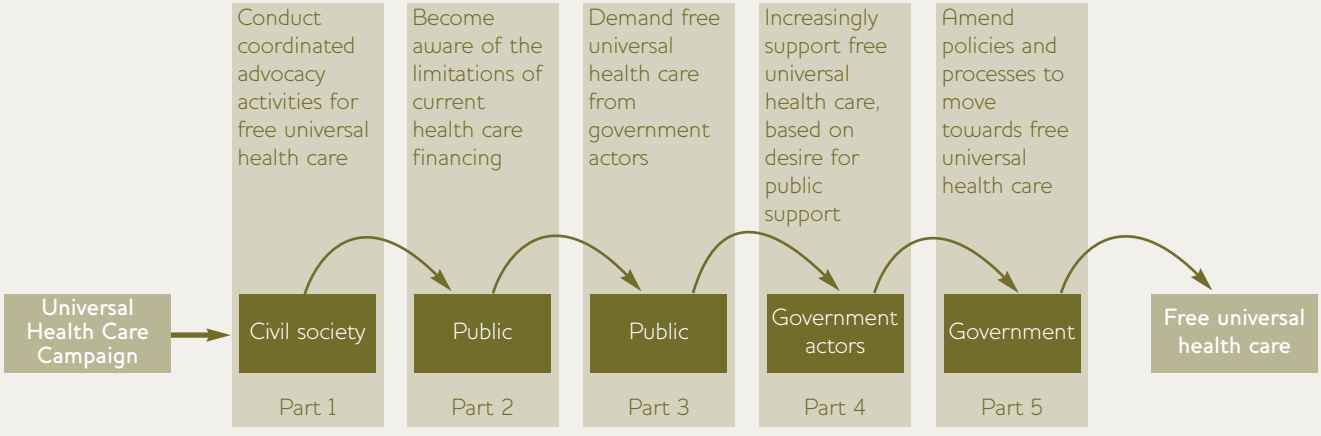
In Beach and Pedersen's description of process tracing, mechanisms are conceptualised as being made up of a number of 'parts', composed of *entities* (for example, people, organisations, systems – nouns) that engage in *activities* (for example, protesting, researching, campaigning – verbs); see Figure 2. This is demonstrated in Figure 3 using a simplified example drawn from an evaluation of an advocacy campaign for free universal health care in Ghana (Stedman-Bryce 2013), as discussed in CDI Practice Paper 10.

Figure 2 **A causal mechanism in process tracing**



Source: Adapted from Beach and Pedersen (2013).

Figure 3 **A simplified causal mechanism based on the Universal Health Care Campaign in Ghana**



Source: Adapted from Beach and Pedersen (2013) using the example of the Universal Health Care Campaign case study.

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

- **More than empirical events or 'intervening variables'.** Describing a sequence of events between A and B or articulating intervening variables is not enough. Rather, mechanisms are **theories** about *how* and *why* one event leads to another. They represent the causal force or power that leads event A to give rise to outcome B.
- **Systematic.** Mechanisms in theory-building and theory-testing process tracing are held to reflect independent facts about how change happens in the world. As such, mechanisms exist independently of any particular event – whether or not they are operating at a particular moment in time. This has many similarities with how mechanisms are conceptualised in realist evaluation (see e.g. Westhorp 2014). This means that a mechanism identified and tested in one case will (if it is framed correctly, at an appropriate level of abstraction from the particular case) apply to other cases (within a specific bounded context – for example a country, a state, a sector or a time period).
- **Singular.** Within a specific case in process tracing, *a single causal mechanism* is examined (and the parts that make up that mechanism). This differs to how mechanisms are conceptualised elsewhere. For example, realist evaluation approaches also understand mechanisms as causal powers, but do not necessarily view them as systems made up of parts; instead the researcher may examine multiple different mechanisms within a single case (Westhorp 2014; Wong *et al.* 2013). These differences once again reflect the fact that there is no clear consensus in the literature on what exactly a mechanism is (Shaffer 2014).
- **Macro-level, micro-level or somewhere in between**, but always studied at a level that most makes sense within a specific case. For example, one process tracing case may examine a micro-level mechanism, relating to how the decisions and behaviour of individuals lead to change in wellbeing. Another may examine a macro-level mechanism, relating to how structural shifts in a social system lead to change within a society or institutional environment.

3 Five steps in process tracing

This section outlines the steps in *theory-building* and *theory-testing* process tracing as described by Beach and Pedersen (2013).

Step 1. Developing a hypothesised causal mechanism

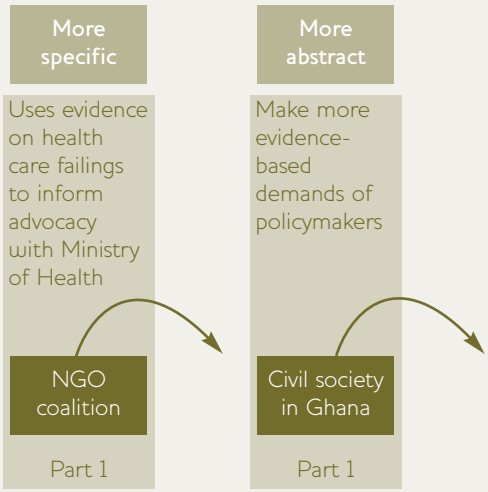
The first step in theory-testing process tracing is elaborating the mechanism to be tested (in contrast, theory-building process tracing starts at Step 3 – see below). This may involve revising or adding detail to an existing theory, such as a project theory of change. Developing a hypothesised mechanism involves clearly elaborating all the steps between A (the hypothesised cause) and B (the outcome of interest). As demonstrated in Figure 2 and Figure 3, Beach and Pedersen recommend that each *part* in the mechanism should specify which *entities* (for example, individuals, organisations, groups – nouns) are expected to conduct which *activities* (for example, protesting, researching, advocating – verbs). Each part of the mechanism can therefore be framed as a hypothesis (for example, 'civil society organisations will consult with citizens') and can be tested.

Building a causal mechanism in process tracing is in some ways similar to the process of developing a theory of change. Like a theory of change, a mechanism can be seen as a theory about how change happens in a particular context. As such, Beach and Pedersen recommend conducting a thorough review of existing literature and evidence to inform the development of the mechanism, to ensure it incorporates previous thinking and learning about change in similar situations.

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

- The mechanism to be tested should be broken down into the smallest feasible number of *parts*, which each directly cause the subsequent part. There should not be any leaps in logic (such as 'training will lead to new policies being developed').
- At the same time, every part should be *necessary* – there should be no superfluous parts which are not absolutely required for the mechanism to work.
- It should be possible to empirically measure each part of the mechanism.
- The mechanism should be framed at a suitable level of abstraction from the specific case, depending on how important (and realistic) it is for the findings to be generalisable to other cases (see Figure 4). If a mechanism is framed in more abstract language, this allows the theory to be applied to other situations and contexts (although it will still need to be tested using context-specific evidence in other situations, to see if it does in fact apply).

Figure 4 **Two different ways of framing the same part of a mechanism**



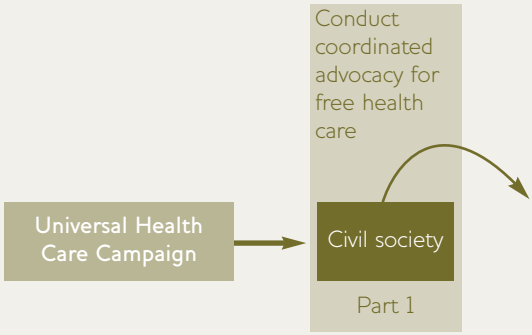
Source: Authors' own.

Step 2. Operationalising the causal mechanism

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

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

Box 1 **Hypothesis 1: Civil society organisations (CSOs) began conducting coordinated advocacy for free health care as a result of the Universal Health Care Campaign**



Alternative plausible explanations:

- Before the campaign, CSOs had already planned and worked together as advocates for free universal health care.
- Other campaigns or movements were also promoting coordinated advocacy for health-care reform.

Observable manifestations that would support Hypothesis 1

- Examples of Universal Health Care Campaign advocacy activities involving multiple CSOs working together (events, platforms, meetings) [trace evidence].
- Interviews with civil society actors suggest good relationships and communication between different CSOs working on the Universal Health Care Campaign [account evidence].
- A number of civil society workers attribute coordinated advocacy activities/good communication to Universal Health Care Campaign activities [account evidence].

Observable manifestations that would eliminate alternative plausible explanations

- No or few examples of coordinated advocacy prior to the launch of the Campaign [trace and sequence evidence].
- No or limited other campaigns or activities promoting health-care reform at the time of the Campaign [trace evidence].

Source: Adapted from Beach and Pedersen (2013) using the example of the Universal Health Care Campaign case study.

Operationalising the mechanism should involve identifying evidence for *causal links* between one part of a mechanism and another. What is the evidence that a part of the mechanism happened *because of* the previous part, rather than for some other reason? This requires thinking through the plausible alternative explanations that might explain each part of the mechanism, and looking for observable manifestations of these. The example in Box 1 demonstrates this using a part of the mechanism from the Universal Health Care Campaign evaluation.

Step 3. Collecting evidence

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

In theory-building (rather than theory-testing) process tracing, Step 3 is applied *first*. If a theory does not exist about how and why A leads to B, the researcher begins by collecting observable evidence about the steps linking the two. The evidence is then used to infer a causal mechanism that explains the facts (Steps 2 and then 1).

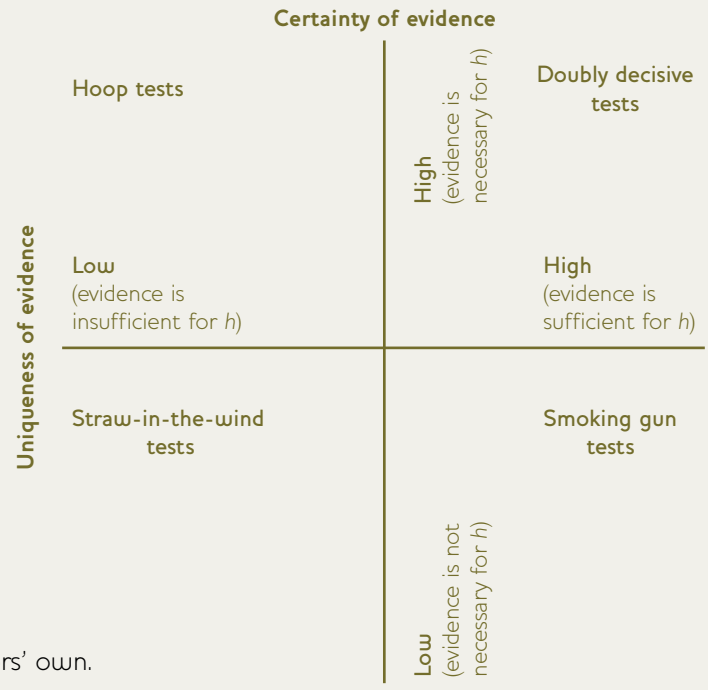
Step 4. Assessing the inferential weight of evidence

As discussed further in CDI Practice Paper 10, the way that evidence is assessed in process tracing is analogous to a criminal trial. For each part of the mechanism the evidence from various sources is weighed in the attempt to put together a case that gives a reasonable degree of confidence that each part of the mechanism exists or does not exist in the particular case. Bayesian probability logic is followed in order to assess the strength of the evidence of each part of the chain.

Four 'tests' have been developed to assist with this process: 'straw-in-the-wind' tests, 'hoop' tests, 'smoking gun' tests and 'doubly decisive' tests (Bennett 2010; Collier 2011; Van Evera 1997). These tests are based on the principles of certainty and uniqueness; in other words, whether the tests are *necessary* and/or *sufficient* for inferring the evidence. Uniqueness and certainty exist along two dimensions – which can be presented on a matrix (see Figure 5).

Each of the tests is discussed in more detail in CDI Practice Paper 10 using the intuitive example of a criminal trial. The tests are further illustrated on page 7 with examples derived from the Universal Health Care Campaign evaluation. Further detail on applying this logic of inference (using a practical example) is provided in Befani and Mayne (2014).

Figure 5 **Matrix for assessing the certainty and uniqueness of evidence**



Source: Authors' own.

Example of tests in practice, based on the Universal Health Care Campaign case study

Straw-in-the-wind test

- **Hypothesis** The campaign increased the capacity of member CSOs to plan and work together as advocates for free universal health care.
- **Evidence constituting straw-in-the-wind tests** *The key findings of a report on the complex National Health Insurance Scheme (NHIS) were presented to campaign members, giving them the opportunity to ask questions of the authors and increase their knowledge on these issues.*
  The range and volume of advocacy activities for free universal health care increased significantly following publication of the campaign's report.
  Both of these examples increase the plausibility of the hypothesis, but do not firmly prove it. Nor do they firmly disprove alternative hypotheses (for example, that other campaigns or activities were improving civil society capacity). However, together they provide stronger evidence than they do alone.

Hoop test

- **Hypothesis** The Universal Health Care Campaign significantly increased the capacity of member CSOs to plan and work together as advocates for free universal health care.
- **Evidence constituting a hoop test** *Until the campaign's report was published, there was little notable coordinated advocacy for free universal health care.* This piece of evidence constitutes a hoop test because it is necessary to keep the hypothesis under consideration. If the evidence suggested that considerable advocacy *had* taken place before the campaign, this would seriously weaken the hypothesis. The evidence also disproves an alternative hypothesis: that before the campaign, CSOs had already planned and worked together on advocacy for free universal health care. Therefore, together with the straw-in-the-wind tests discussed above, the evidence provides reasonable confidence in the hypothesis.

Smoking gun test

- **Hypothesis** The National Health Insurance Authority (NHIA) in Ghana revised its methodology for calculating active membership of the NHIS because of pressure created by the campaign.
- **Evidence constituting a smoking gun test** *The Ghana delegation at an international meeting on Universal Health Care in Geneva stated that the campaign's report 'was very helpful and prompted us to revise our figures'.* It is highly unlikely that the delegation would make this statement if the report had not influenced them, particularly since the NHIA had initially dismissed the report during the national health sector debate that ensued in Ghana after its publication. This piece of evidence constitutes a smoking gun because it is compelling enough to give high confidence that the hypothesis is true and alternative hypotheses (such as that the methodology was being revised anyway) are false. However, it is unusual to find evidence so clearly linking a policy change to a particular advocacy campaign – suggesting that although evaluators should certainly look for smoking guns, they should not necessarily hold out for this type of evidence.

Thoughts on constructing tests in impact evaluation

Beach and Pedersen suggest that tests should be specified in advance of data collection, and researchers should then collect specific evidence that will either pass or fail the test. This could involve thinking about the type of evidence that might have high uniqueness (a smoking gun) or high certainty (a hoop test) in the particular case, and then deciding data sources and designing data collection tools in order to look for this evidence. However, it seems likely that some tests can only be constructed after data collection as it is not always possible to know exactly what to look for in advance. Once evidence has been collected, it seems plausible to categorise and sort it retrospectively into different types of tests.

The construction of tests is a highly analytical and contextual process. It is only possible to work out whether evidence is surprising based on a good knowledge of the context. The process is also inevitably to some extent subjective, as one person's straw-in-the-wind test might look more like a smoking gun to another researcher. This suggests the importance of retaining transparency in the process, and being clear about the sources and nature of the evidence and tests used.

Step 5. The conclusions of a process tracing exercise

Once the process is complete, the researcher should be able to assert a degree of confidence in each part of the hypothesised mechanism, based on the evidence collected and the tests applied. Because all the parts in a mechanism are *necessary* for the mechanism to operate (see Step 1), this means that the evidence for the mechanism as a whole is only as strong as the evidence for its weakest links. It is therefore important to clearly flag where evidence is

weakest (for example, where a part of the mechanism is supported only by a straw-in-the-wind test). The evidence on each part of the mechanism will then enable the researcher to either *accept* or *reject* the mechanism as a whole.

If the mechanism is accepted, this means there is sufficient evidence for the researcher to have confidence that cause A led to the effect B through the process described in the mechanism. It therefore provides evidence about *how* and *why* change happened in a particular case. However, if the mechanism is *not* accepted, this does not mean that cause A did not lead to effect B – simply that there is insufficient evidence to show that it did so through the specified mechanism. A may have led to B in a completely different (or even slightly different) way. This emphasises the crucial importance of Step 1, and the risks associated if the mechanism is not carefully developed.

Mechanisms that have been accepted in one particular case may also allow generalisations to other cases – depending on the level of abstraction used when framing the parts of the mechanism (see Step 1). For example, a mechanism linking an advocacy programme with policy change in Ghana may also be relevant to advocacy interventions in Malawi. Therefore, a mechanism that has been accepted in one case may help when designing new programmes or developing similar mechanisms elsewhere. However, a mechanism that has been tested and accepted in one case cannot simply be held to predict success elsewhere without further testing.

References

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

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

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

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

Oxfam GB (2011) 'Process Tracing: Draft Protocol', Oxfam GB

Pawson, R. and Tilley, N. (2004) 'Realist Evaluation', paper funded by the British Cabinet Office

Shaffer, P. (2014) 'Two Concepts of Causation: Implications for Poverty', *Development and Change* 46.1: 148–66

Stedman-Bryce, G. (2013) *Health for All: Towards Free Universal Health Care in Ghana, End of Campaign Evaluation Report*, Oxford: Oxfam GB

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

Westhorp, G. (2014) *Realist Impact Evaluation: An Introduction*, London: Methods Lab/Overseas Development Institute

Wong, G.; Westhorp, G.; Pawson, R. and Greenhalgh, T. (2013) *Realist Synthesis: Rameses Training Materials*, London: Health Research Health Services and Delivery Research Program: 1–55

Centre for Development Impact (CDI)

The Centre is a collaboration between IDS (www.ids.ac.uk) and ITAD (www.itad.com).

The Centre aims to contribute to innovation and excellence in the areas of impact assessment, evaluation and learning in development. The Centre's work is presently focused on:

(1) Exploring a broader range of evaluation designs and methods, and approaches to causal inference.
(2) Designing appropriate ways to assess the impact of complex interventions in challenging contexts.
(3) Better understanding the political dynamics and other factors in the evaluation process, including the use of evaluation evidence.

This CDI Practice Paper Annex was written by **Melanie Punton** and **Katharina Welle**.

ISSN: 2053-0536
AG Level 2 Output ID: 313

Institute of Development Studies, Brighton BN1 9RE, UK
T +44 (0) 1273 915637 **F** +44 (0) 1273 621202 **E** ids@ids.ac.uk **W** www.ids.ac.uk