Katie Naeve
Julia Fischer-Mackey
Jyotsna Puri
Raag Bhatia
Rosaine N Yegbemey

Evaluating advocacy: an exploration of evidence and tools to understand what works and why

December 2017

Working Paper 29

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programs. We are the global leader in funding, producing and synthesising high-quality evidence of what works, for whom, how, why and at what cost. We believe that using better and policy-relevant evidence helps to make development more effective and improve people's lives.

3ie working papers

These papers cover a range of content. They may focus on current issues, debates and enduring challenges facing development policymakers and practitioners and the impact evaluation and systematic review communities. Policy-relevant papers draw on relevant findings from impact evaluations and systematic reviews funded by 3ie, as well as other rigorous evidence to offer insights, new analyses, findings and recommendations. Papers focusing on methods and technical guides draw on similar sources to help advance understanding, design and use of rigorous and appropriate evaluations and reviews. 3ie also uses this series to publish lessons learned from 3ie grant-making.

About this working paper

Evaluating advocacy: an exploration of evidence and tools to understand what works and why aims to examine and discuss challenges related to evaluating advocacy initiatives, identify factors associated with successful advocacy interventions, and provide additional methods and tools to use in evaluating them. This paper is a result of a collaborative partnership between 3ie and the Water Supply and Sanitation Collaborative Council (WSSCC), called the Evidence Programme on Sanitation and Hygiene. This collaboration was inspired by an effort to address the prevalent problems surrounding water, sanitation and hygiene.

WSSCC provided funding to 3ie for this working paper.

The content of this paper is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors and omissions are also the sole responsibility of the authors. Please direct any comments or queries to the corresponding author, Raag Bhatia, rbhatia@3ieimpact.org.

Suggested citation: Naeve, K, Fischer-Mackey, J, Puri, J, Bhatia, R and Yegbemey, R, 2017. *Evaluating advocacy: an exploration of evidence and tools to understand what works and why*. 3ie Working Paper 29. New Delhi: International Initiative for Impact Evaluation (3ie).

Executive editors: Emmanuel Jimenez
Production manager: Angel Kharya
Assistant production manager: Akarsh Gupta
Copy editor: Jaime L Jarvis
Proofreader: Bulletproof
Cover design: John F McGill and Akarsh Gupta

© International Initiative for Impact Evaluation (3ie), 2017

Evaluating advocacy: an exploration of evidence and tools to understand what works and why

Katie Naeve
Root Capital

Julia Fischer-Mackey
American University

Jyotsna Puri
Green Climate Fund

Raag Bhatia
International Initiative for Impact Evaluation (3ie)

Rosaine N Yegbemey
3ie

3ie Working Paper 29

December 2017

Acknowledgements

We thank 3ie and the Water Supply and Sanitation Collaborative Council (WSSCC) for research and grant support for this study and prompting and contributing significantly to this important discussion. We are grateful to Nilesh Chatterjee for his extensive comments and review work. We also thank Shagun Sabarwal, Fiona Underwood, Archana Patekar, Chaitali Chattopadhyay, Anna Hakobyan and Steven Chapman for their valuable insights for this work. We thank Juan Obach, Lorena Caro-Zambrano and Diana Goldemberg for their excellent research assistance throughout this project. We also appreciate comments from participants in seminars organised by BBC Media Action in London in 2015, participants in the joint seminar hosted by Children's Investment Fund Foundation, WSSCC and 3ie in London in November 2015, and participants at a learning workshop hosted by WSSCC and the Geneva Evaluation Network at the United Nations in April 2016. We also thank the advocacy and evaluation experts we interviewed for this study, who provided direction and crucial insights.

Summary

Advocacy commonly refers to individuals or groups undertaking initiatives, often speaking on behalf of vulnerable populations, to bring about a change in society. Advocacy takes numerous forms, ranging from actions that target individuals or institutions in positions of power (i.e. policy advocacy) to those that target individuals (i.e. advocacy for behaviour change).

In recent decades, the importance of advocacy in bringing about positive social change has been widely recognised. With the advent of social media, technology and the Internet, investment and reliance on advocacy efforts have increased exponentially. In this study, we examine methods and techniques to evaluate advocacy programs for their effectiveness and identify factors associated with successful programs. The audience for this study is advocacy professionals, funders, and institutions that rely on and programme for advocacy.

Since advocacy efforts often unfold in dynamic sociopolitical environments and have nonlinear impact trajectories with a diversity of variables that influence outcomes, it is often difficult to evaluate advocacy programs and initiatives.

The objectives of this paper are:

- To examine and discuss challenges related to evaluating advocacy initiatives;
- To identify factors associated with successful advocacy interventions; and
- To identify possible evaluation methodologies; compare and contrast various tools available for advocacy evaluation; and present a toolkit of methods to be used by evaluators and practitioners of advocacy.

Methods

We employ a mixed-method approach in this paper. We review advocacy evaluation studies that measure the causal effects of advocacy interventions, using a positive deviance approach to analyse factors associated with programme success. Further, we use in-depth qualitative interviews with 14 advocacy and advocacy evaluation experts and review relevant literature to inspect challenges and identify prospective methodologies associated with comprehensively evaluating advocacy initiatives.

Results

Challenges to advocacy evaluations relate primarily to the multiple objectives of advocacy-related actions, difficulties in reliable data collection, and the aggregative approach that most advocacy actions take, as they build on previous successes while discarding unsuccessful aspects. This underscores the need for flexible evaluation tools. Determining the timing of evaluations (prospective and concurrent versus subsequent to intervention) is also a concern, as is determining the role of the evaluator. It is difficult to establish causality (Is the advocacy intervention what *caused* the change?) and measure attributable change or measure contribution, because varied actors are involved in successful advocacy initiatives.

We discuss standards that can be used to evaluate different types of advocacy intervention. It is important to have evidence that can examine the efficacy of links in

theories of change and have high specificity and sensitivity. We conclude that certainty and uniqueness of evidence are important attributes for evaluations of advocacy programs.

A positive deviance analysis of 56 impact evaluations that evaluate advocacy interventions shows that important correlates of successful advocacy programs are *who* advocates, if *incentives* are offered, whether the target group is compared to another population in advocacy messages, *who delivers* messages, and the *channel* for information dissemination. We find no theory-based impact evaluations that examine policy advocacy initiatives.

Methods such as the case study, process tracing, outcome mapping and qualitative comparative analysis provide some solutions to challenges encountered in evaluating advocacy initiatives, but each has its limitations.

Conclusion

We found no evaluations of policy advocacy that used theory-based impact evaluations. All the evaluations that used such methods examined advocacy efforts to change behaviour at the individual or group level. We highlight that evaluating advocacy actions requires a combination of methods, and that any single method is too limited to evaluate advocacy initiatives. Before selecting an evaluation method, it is essential to perform a careful study of the nature and purpose of the initiative and factors influencing it. Building retrospective theories of change is one way to begin. Mapping stakeholders and contributors and identifying critical nodes in the overall theory of change can then indicate which methods could be most usefully employed to obtain a wholesome perspective of effectiveness and efficiency.

Identifying key links in the theory of change that are critical to overall success allows researchers to undertake important and timely efficacy evaluations, while using rigorous identification methods. For assessing overall effectiveness, process tracing – combined with techniques such as Bayesian updating and with special attention to specificity and sensitivity of evidence – may be useful. A clear idea of what data to collect is a potentially important approach that may help to combine the virtues of qualitative and quantitative methods. We also briefly underscore the importance of having high certainty in evidence and uniqueness, borrowing from the forensic science. Lastly, we present a toolkit that we hope will be useful to those interested in evaluating advocacy programs and initiatives.

Contents

Acknowledgements ..... **i**
Summary ..... **ii**
List of figures and tables ..... **iv**
Abbreviations and acronyms ..... **v**
1. Introduction **1**
1.1 Importance of advocacy in the present context 2
1.2 Rationale for this study 5
1.3 Challenges relating to advocacy evaluation 6
1.4 Objectives 7
2. Challenges and opportunities in advocacy evaluation **7**
2.1 Methods 7
2.2 Results 8
2.3 Main findings 13
3. A review of impact evaluations of advocacy programmes **14**
3.1 Data and methods 15
3.2 Data analysis 16
3.3 Results 17
3.4 Discussion 27
4. Additional methods for advocacy evaluation **31**
4.1 Methods 31
4.2 Results 32
4.3 Tools 40
4.4 Discussion 46
5. Conclusion **46**
Appendix A: Advocacy evaluation resources **48**
Online appendixes **52**
References **53**

List of figures and tables

Figure 1: An example theory of change for advocacy aimed at influencing policy 3
Figure 2: An example theory of change for advocacy aimed at influencing attitudes or behaviours 4

Table 1: Search for the evidence review 16
Table 2: Impact evaluations included in thematic evidence review 18
Table 3: Geographical regions of interventions evaluated 23
Table 4: Income levels of countries included in interventions evaluated 23
Table 5: Methods used in the studies reviewed to evaluate advocacy initiatives 24
Table 6: Goal of advocacy initiatives evaluated 24
Table 7: Type of advocacy interventions evaluated 25
Table 8: Success rates of interventions based on how information was provided 27
Table 9: Examples of interim outcomes 38
Table 10: Additional methods and tools for advocacy evaluation 43

Abbreviations and acronyms

GEM	General elimination method
RCT	Randomised controlled trial
WSSCC	Water Supply and Sanitation Collaborative Council

1. Introduction

Initiatives undertaken by individuals or groups to influence critical decisions in sociopolitical and economic spheres are commonly referred to as advocacy. Essentially seen as the process of speaking on behalf of individuals, particularly vulnerable groups, with the aim of bringing about systemic changes, advocacy often entails mobilising potentially affected individuals and organisations to participate along with advocates in campaigning. In many cases, advocacy initiatives also aim to empower individuals to fight for themselves by speaking up (Mayoux 2003).

The word *advocacy* is derived from the Latin word *advocare*, which means 'to call out for aid'. Advocacy involves an advocate, usually an individual or group involved in the advocacy activity who communicates with a target group and is responsible for bringing about the change. The target group has the potential (power) to bring about the change and is advocated to. We distinguish the target group from beneficiaries, the individuals who will likely benefit from advocacy actions. For example, a not-for-profit organisation may be an advocate that campaigns for stricter laws and punitive measures related to child labor. The target group for this advocacy initiative is policymakers who are expected to enact stringent laws to safeguard children's rights, and children are the beneficiaries of advocacy.

However, the term *advocacy* is wide ranging. It has come to include voluntary initiatives by individuals (or organisations) to bring about social change, and includes support for and adoption of a particular cause by beneficiaries. In this study, we do not examine advocacy initiatives that may be used for negative outcomes.

Nearly every organisation or researcher studying advocacy uses a different definition of advocacy. We have developed the following definition:

> Advocacy is the deliberate process of managing information and knowledge[1] with the clear goal to influence and/or change the policies, practices, power dynamics, attitudes and/or actions that directly and positively affect lives. Advocacy can be employed by a directly or indirectly affected population, or by a third party on behalf of an affected population, including those inside and outside of government or other positions of power.

This definition includes advocacy initiatives in which the targeted outcome is change in policies (policy advocacy) or changes in behaviour, attitude, or knowledge an individual or group possesses. The definition does not include, however, initiatives in which the targeted outcome is advocacy.[2]

Advocacy initiatives may take numerous forms. We distinguish between the following types of advocacy (Coffman 2009):

- Policy advocacy: This group of actions includes initiatives or programs that target changes in policies or legislation that, in turn, may affect entire sociopolitical systems. Policy advocacy programs typically target members of the

[1] In this definition, information and knowledge management includes information dissemination, evidence-building and education.

[2] We define advocacy as a means to an end, not the end itself.

administration, such as legislators and elected officials. Policy advocacy aims to create new policies or change or refine existing ones, frequently benefiting disadvantaged populations. Policy advocacy actions usually engage individuals in positions of power (e.g. legislators).

- Systems advocacy: These are advocacy initiatives aimed at bringing about positive change in programs and practices at the organisational or community level to benefit the intended population. This form of advocacy targets established local bodies for bringing about a change.
- Advocacy for attitude and behaviour change: These are advocacy programs that engage and target individuals or communities. The desired outcome is behavioural or attitudinal change among individuals or communities.

Advocacy initiatives may also be classified according to the channels of communication they employ. For example, media advocacy makes use of mass media and communication strategically for bringing about social change. We discuss this in Section 2.

1.1 Importance of advocacy in the present context

The act of advocating to influence those in power has occurred for centuries. Citizens have been advocating on their own behalf, as illustrated by the anti-taxation movements in colonised countries (Westin 2012). Similarly, individuals and groups have been advocating on behalf of others, such as during the universal suffrage or civil rights movements (Smiltneek n.d.; Paden 2011). In the past century, advocacy for behaviour change related to health, political participation and pro-environmental causes has also become very common (Cummins 2015).

In recent decades, advocacy efforts have grown in number, sophistication and scope. In some cases, advocacy has been found to be a useful tool in bringing about desirable social changes. In order to promote desirable behaviours among people in a community, advocacy strategies need to account for laws and policies in the community and gauge the extent to which these enable desired behaviour – i.e. if existing policies allow and aid the adoption of desirable behaviours or if there are punishments for noncompliance. Successful behaviour change programs often target changes in social norms. Advocacy actions therefore frequently target social mobilisation at the community level and policy change simultaneously (UNICEF 2010).

Policy advocacy aims to establish or change legislation. This influences social or group behaviour, while communications associated with the advocacy campaigns usually aim to inform, change knowledge, and therefore influence the attitudes and practices of individuals. Advocacy aims to change the social and political environment to facilitate and maintain this change at the level of the individual (UNICEF 2010).

Modern advocacy initiatives have witnessed phenomenal growth in size, frequency and substance in recent years. Profiling, use of technology and social media, and use of insights from psychology, economics and anthropology have all enriched this space, and advocacy initiatives ranging from grassroots protests to transnational coalitions have all used these methods and insights. Perhaps as a consequence, advocacy has come to be

recognised as a specific discipline, and organisations have responded by increasing monetary outlays for advocacy initiatives (Mayoux 2003).

A key to understanding the impact of advocacy initiatives is identifying the process through which changes occur in attitudes, behaviour or policy, while recognising that this is not necessarily a linear process. Below, we present two theories of change to illustrate how advocacy actions may create change at the policy and individual levels.

Figure 1: An example theory of change for advocacy aimed at influencing policy

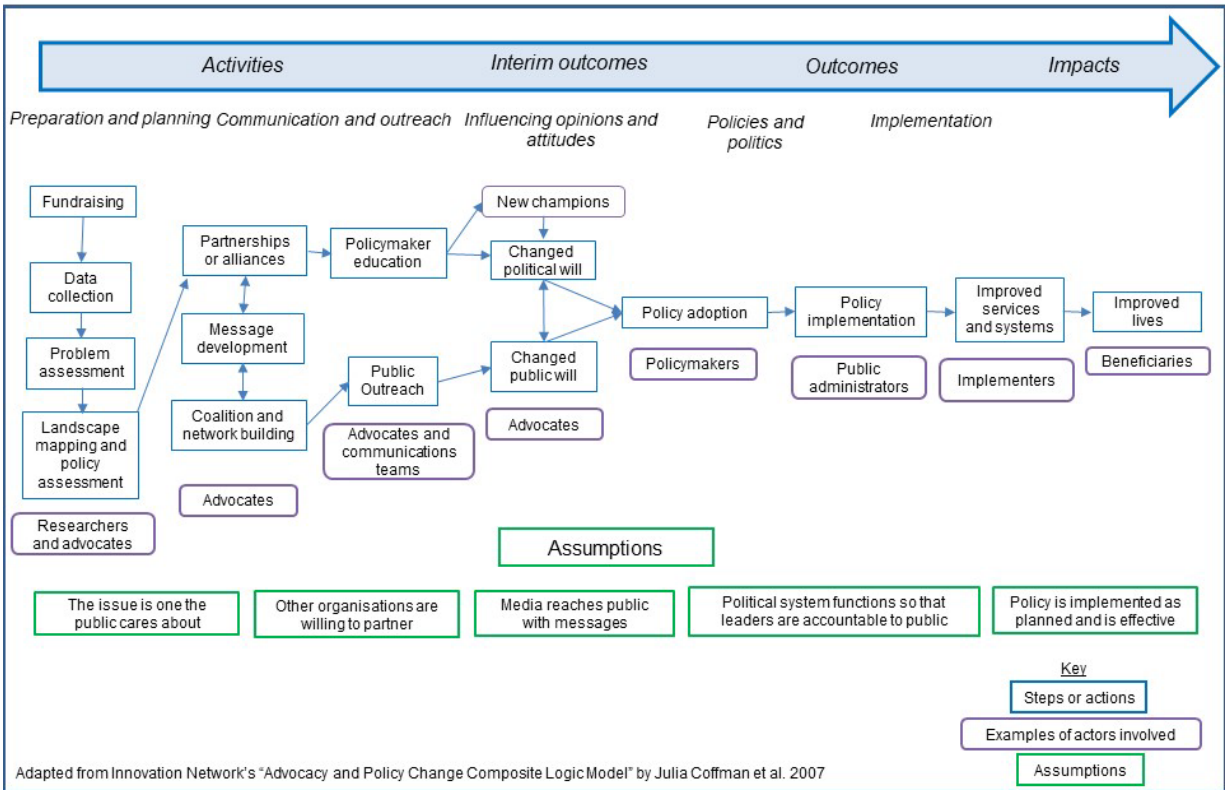

Figure 1 shows key elements of an advocacy effort designed to influence policy to provide improved public services. During the preparation and planning stage, the advocate or researcher typically conducts a policy assessment, maps the landscape of potential supporters and opponents, and develops appropriate strategies and messages to bring about the desired policy change. The messages vary depending on the target audience and purpose. For example, certain messages are designed to convince other organisations to join the effort. With the help of new allies, messaging is aimed towards two groups – the public and policymakers. These efforts are designed to build public will and political will to support the intended policy change. Mediated by the media and influenced by peers and constituency members, the theory is that this will influence policymakers to make a change (Coffman et al. 2007). Ideally, policymakers or other prominent figures become champions for the new policy and assist in bringing attention to the problem while negotiating the content of new policy. Once the intended policy is adopted through legislative vote or executive action, it is implemented and maintained with the help of public administrators and implementers.

This theory of change for policy advocacy relies on some critical assumptions:

- Targeting: Communication or information reaches its target audiences – allies, the public and policymakers (i.e. the strategy is well-targeted);
- Engagement and resources: The problem advocacy initiatives are targeting is one that people care about and one that other organisations are willing to devote time and resources to;
- Feedback loops and accountability: Political leaders listen to the public and civil society and are in some way accountable to them (i.e. there is a feedback loop); and
- Efficacy: The policy is well-designed and, if implemented well, will lead to improved and more effective services.

Figure 2: An example theory of change for advocacy aimed at influencing attitudes or behaviours

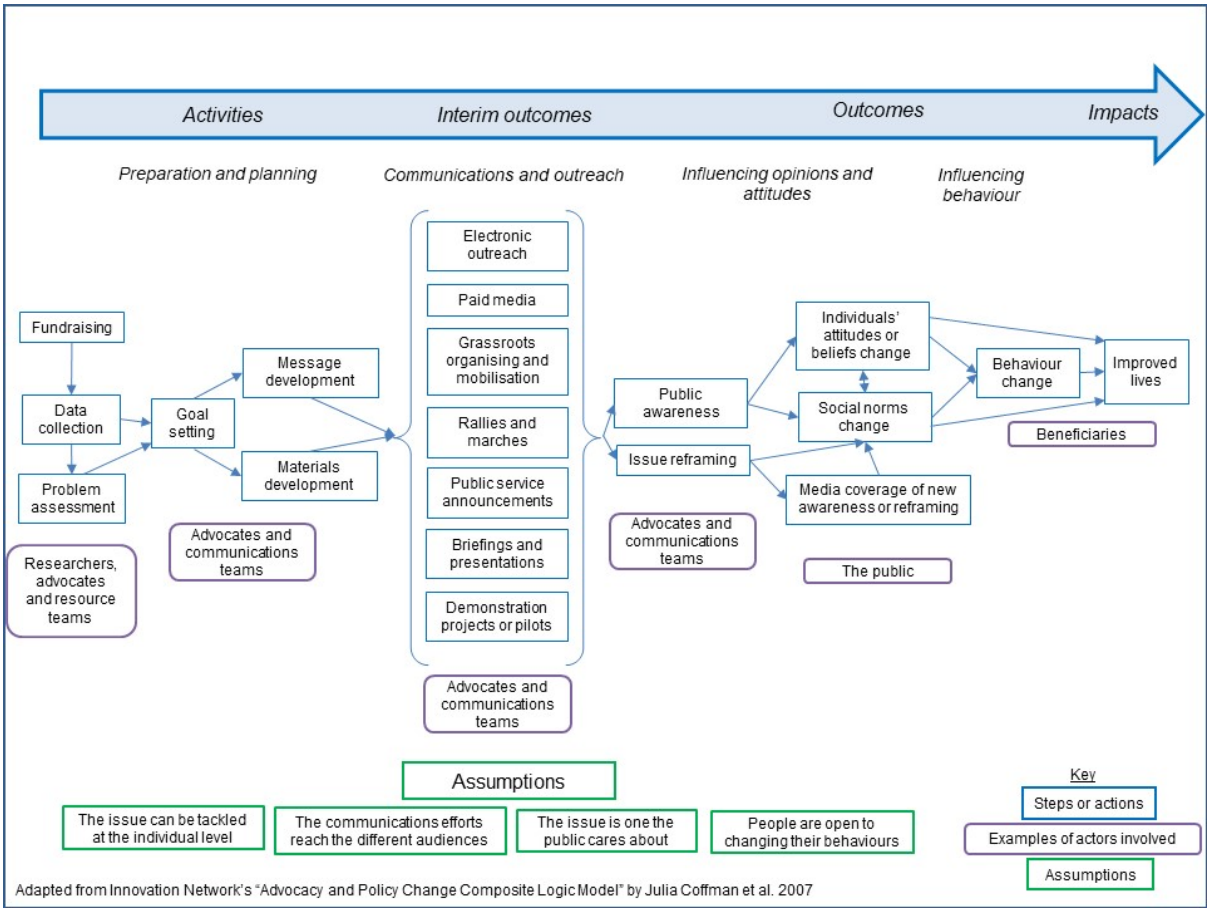

Figure 2 illustrates a potential theory of change for advocacy initiatives that target behaviour change. Resource mobilisation is typically a key initial step. Then, advocates and researchers use evidence and other information to assess the depth and reach of the problem. The goal for the campaign is consequently narrowed and refined. Based on the goal, the advocacy team develops messages and materials customised for different audiences. Target audiences include not only people whose behaviour the campaign intends to change (in this case, the beneficiaries and the target group merge), but also another target group that includes public role models, educators or healthcare professionals. The theory is that support for the campaign from these influencers will

likely influence the target population. Awareness-raising activities undertaken as part of this overall action may include Internet outreach, paid media spots, grassroots mobilisation, rallies or marches, briefings or presentations to experts, public service announcements, and demonstration projects or pilots.

The intended outcome of such efforts is to make the beneficiary (i.e. the public) aware of the concern. For example, an advocacy campaign that promotes handwashing may either provide evidence of the link between handwashing and individual health or reframe it as a community or collective good (because it stops the spread of disease). Issue reframing is often part of raising public awareness and motivating people to change their behaviours.

The final stage of this theory of change is the impact, in which the behaviour change translates into real improvement in people's lives. In the handwashing example, the desired impact is a decrease in the incidence of sanitation-related diseases (e.g. diarrhoea) in the population.

This theory of change rests on a few key assumptions:

- Efficacy: It is possible to address the problem at the individual level (i.e. if implemented well, handwashing will lead to a reduction in disease incidence);
- Targeting: Communication efforts reach their intended audiences;
- Trust and engagement: ultimately, beneficiaries will believe that the solution works and engage with it; and
- Self-realisation: People (beneficiaries) are open to changing behaviours and able to change them.

1.2 Rationale for this study

Despite growing interest in and use of advocacy action, the evidence base of what makes advocacy initiatives effective is limited. With the advent of myriad advocacy techniques and with advocacy becoming a component of almost every political or social-behavioural change program, there is an important need to evaluate the effectiveness of advocacy: do advocacy initiatives work, and if so, how much? What are the characteristics of successful advocacy efforts? Evaluation of advocacy initiatives or advocacy evaluation has the potential to inform those engaging in, benefiting from or funding advocacy actions. It can therefore help future advocates learn advocacy, validate investments in advocacy work for donors and promote accountability (Coe and Schlangen 2011).

Advocacy evaluation can also be instrumental in informing learning and strategy, both internally (within organisations) and across the advocacy field. Prospective or formative evaluation can help organisations modify their strategies or change course as a result of measured outcomes (Coffman 2009). Advocacy evaluation can guide allocation of resources by identifying the effectiveness and value for dollar of different initiatives and assist donors in assessing the impact of their investments to justify funding. Finally, advocacy evaluation may help in understanding the impact of an initiative in the context of specific programme goals or outcomes: did advocacy ultimately help the programme reach its intended goals? Did advocacy bring about change in a shorter amount of time? Did advocacy create conditions that made change sustainable?

1.3 Challenges relating to advocacy evaluation

Advocacy initiatives are of numerous types and require a diverse toolkit for evaluation. When the effectiveness of an intervention is being investigated and the target of the intervention is a large enough group of identifiable units, it may be possible to use experimental methods. This large-*n* scenario (e.g. people, villages, classes) presumes that a comparison group is identifiable (Boaz et al. 2008). Advocacy initiatives that aim to change behaviour or attitudes of beneficiaries usually fall into this category. For situations in which random assignment to a control group is not feasible, there is a range of quasi-experimental methods available that attempt to construct credible counterfactual groups through alternative means. This group of interventions requires a large number of units of assignment (or units at which an initiative is implemented). If there are few units of assignment of the advocacy initiative, such as a nationwide policy change, these conditions may not hold, requiring other methods to understand the initiative's effectiveness.

In the case of advocacy interventions, the intent of the intervention is not always to change attitudes or behaviours of a large number of individuals. For advocacy approaches such as those that aim to change policy, small-*n* evaluation approaches, as defined in White and Phillips (2012), are typically called for. It is common for this family of evaluations to focus on the specification of a theory of change, together with a number of alternative causal hypotheses. Causation is then established by collecting evidence to validate, invalidate or revise the hypothesised explanations, with the goal of evidencing the links in the actual causal chain. If, rather than focusing on effectiveness questions, evaluations aim to understand the relevance of advocacy interventions, investigate channels through which an effect is achieved, or examine the extent to which programs have been implemented well (implementation fidelity), then a range of qualitative and process evaluation techniques are available.

A challenge in advocacy evaluation is that it may not always be possible to quantify the effects of advocacy initiatives. This adds to the difficulty in designing indicators to measure the success of an advocacy program. In addition, the trajectory connecting advocacy initiatives and the end results or markers of progress is complex and nonlinear. This calls for a nuanced approach to advocacy evaluation that takes into account understanding the finer aspects in the process of policy change (Teles and Schmitt 2011).

Advocacy initiatives operate in a fluid sociopolitical landscape, necessitating the use of evaluation techniques that are flexible and responsive to variations in the sociopolitical milieu (Lempert 2009). This requires that advocacy evaluation be viewed as a form of trained judgements, calling for a comprehensive understanding of the politics of the issue, mapping the players involved and knowing the time frame against which achievements may be gauged (Teles and Schmitt 2011; IIED 2017). Finally, while evaluating advocacy programmes undertaken by coalitions of advocacy groups, evaluation frameworks must be equipped to assess each group member's investment in the programme (Lempert 2009). We summarise that evaluations must be equipped to assess each group member's investment to the programme in order to understand the starting point (baseline) and intended outcomes to draw conclusions.

1.4 Objectives

This paper discusses challenges posed by advocacy evaluation; showcases some attributes of successful advocacy initiatives, as evinced by rigorous impact evaluations; and discusses available evaluation tools for advocacy evaluation. In Section 2, we discuss issues related to evaluating advocacy efforts. Section 3 presents findings from evaluations that have used theory-based experimental and/or quasi-experimental methods to identify change and measure it. We then use a positive deviance approach to discuss attributes of successful advocacy actions. Section 4 discusses additional methods and tools for undertaking advocacy evaluations (where impact evaluations cannot be executed for various reasons, or to supplement impact evaluations), addressing some challenges posed by advocacy initiatives that target policy or systems change.

This paper has the following goals:

- Examine and discuss challenges related to evaluating advocacy initiatives;
- Identify factors associated with successful advocacy interventions, using a positive deviance approach;
- Identify possible evaluation methodologies;
- Compare and contrast tools available for advocacy evaluation; and
- Present a toolkit of methods to be used by evaluators and advocates.

2. Challenges and opportunities in advocacy evaluation

Evaluations are crucial for assessing the effectiveness of advocacy investments and initiatives. They can inform organisations and individuals undertaking, investing in and benefiting from advocacy initiatives of the process and results of advocacy initiatives, and inform future initiatives.

2.1 Methods

In this section, we present findings from a desk review of the literature and interviews with advocacy experts to explore and examine the challenges associated with evaluating advocacy programs (see online Appendix B for the interview protocol). We first reviewed the published scientific and grey literature on studies evaluating advocacy initiatives. For this, we searched the 3ie repository and Google Scholar databases using the search terms ‘advocacy AND evaluation’, ‘evaluation theory’ and ‘advocacy AND evaluation AND methods’. We shortlisted and saved relevant articles and grey literature, which we then reviewed and analysed, organising key findings by theme.

We also interviewed 14 advocacy and advocacy evaluation experts from international advocacy and evaluation agencies, non-governmental organizations (NGOs), consulting firms and educational institutions (see online Appendix C). Respondents were purposively identified and informed about study objectives. After gaining consent, we set an appointment for an in-depth interview. We developed two interview guides (see online Appendix B), one for use with independent advocacy evaluation experts and the second for the use with experts from advocacy organisations. Both guides consisted of open-ended questions about advocacy evaluation approaches, strengths and drawbacks of these approaches and specific evaluation techniques, factors associated with successful

and unsuccessful advocacy programs, and future implications for the evaluation of advocacy. Questions moved from a general probe to specific prompts.

Researchers trained in interview methods, advocacy and evaluation conducted the interviews. Interviewers read out a consent form to each respondent, describing the purpose of the study and how findings would help policymakers and practitioners. Respondents were informed about confidentiality and the voluntary nature of participation. The interviewer proceeded with the questions only if the respondent gave verbal consent. Interviews lasted 30 to 45 minutes on average. Data collection was conducted from August 2015 to October 2015.

Interviews were not recorded, but interviewers took comprehensive notes during the interviews. Researchers organised responses into thematic categories. Further examination into the commonalities and differences in respondents' answers to similar questions or topics led to the identification of prevalent trends. The researchers checked the categories and themes for agreement and to ensure trustworthiness of interpretation. These categories and themes were then classified under different areas of interest to advocacy evaluation and are reported in the next section. Prior literature review of the topic and familiarity with the broad areas and challenges of advocacy evaluation helped in analysing and reporting interview findings.

2.2 Results

This section teases out the significant themes identified through interviews with evaluation experts and review of relevant literature.[3]

Advocacy evaluations can have different aims: For many respondents, the primary function of advocacy evaluation is to aid and enhance the advocacy initiative. Although demonstrating causality and quantifying impact are sometimes objectives, these are frequently not the main objectives of advocacy evaluations. The reasons for undertaking advocacy evaluation may include learning, improving effectiveness, developing or assessing strategy, validating an approach, assessing impact, or understanding the relative contributions of different stakeholders. This calls for more nuanced and diverse approaches to advocacy evaluation.

There is a need for a diverse set of advocacy evaluation methods: All of the interviewed experts underscored the need for a diverse set of methods that allow for flexibility and comprehensiveness. According to advocacy evaluation experts Coe and Schlangen, 'Tool-driven approaches that condense assessment to the relationship between inputs and outcomes, or shoehorn analysis into frameworks that do not quite fit, can give a false sense of hope that advocacy can be precisely measured' (2011, pp. 1-2). Though theory-based impact evaluations that employ quasi-experimental and/or experimental methods lend themselves well to assessing certain advocacy interventions, other initiatives (e.g. policy advocacy evaluations) are less amenable to these methods.

Evaluation expert Gabrielle Watson emphasised that advocacy is experimental by nature, where there is no stable or static initiative or initiative environment. This requires evaluators to be thoughtful, flexible and comprehensive in the approach they select to

[3] The names of the respondents have been kept anonymous.

evaluate such complex and adaptive systems.To understand how a coalition of diverse stakeholders come to agree on a policy objective at a particular point in time and use their positions in society to influence decision makers, a variety of qualitative methods will likely provide more insights and useful knowledge than experimental ones. Moreover, theory-based impact evaluation methodologies have certain requirements, such as control over implementation, environment, the ability to randomise assignment of treatment (e.g. a service or program) and the ability to create valid counterfactuals, which may not always be feasible in evaluations of advocacy initiatives.

Causality is difficult to establish in advocacy evaluations, but it is possible nonetheless: One key challenge for advocacy evaluation is the lack of consensus on valid methods. In theory-based impact evaluations, a number of tools are used for identification and measurement. These are well-accepted (White and Sabarwal 2014a; Rogers 2014), although they too have their detractors (Pritchett 2013; Vivalt 2017). These methods include randomised assignment and other quasi-experimental methods, such as instrumental variables, regression discontinuity and difference-in-difference using matching techniques. Key to undertaking these evaluations is ensuring that there are a number of prerequisites to be fulfilled (Jimenez and Puri 2017). These include a theory of change; pre-analysis plans; data that is as objectively measured as possible, including survey data at baseline and endline (and is predicated on good pilots and formative work); a good understanding of outcome(s) and possible indicators; good monitoring data and information on implementation fidelity; a good identification strategy, sufficient data size for statistical confidence; and high-quality analyses that mitigate a multitude of possible biases that may creep in over and above the bias of programme placement and selection. Arguably, these methods are best used when the suite of interventions is well-known, when there is a credible theory of change and when these theories of change are non-complex.

We argue that theory-based impact evaluations are best used to examine subparts of theories of change in three cases, as related to advocacy evaluations: first, if there is a critical constraint or a bottleneck in the theory of change that is ill-informed by past evidence (the critical bottleneck condition); second, if the intention is to test the efficacy of the intervention (e.g. does increased handwashing *really* lead to reduced disease incidence?); and third, if there is a need to show measurable change (e.g. *how much* was the percentage change in disease incidence as a consequence of handwashing?). In all other cases, we argue, theory-based impact evaluations will probably take too long or are not required, and their use should be examined on a case-by-case basis.

All evaluations need to deal with different sorts of biases (Jimenez and Puri 2017). As we discuss in Section 3, advocacy evaluation has traditionally used qualitative approaches, with some exceptions. Qualitative approaches emphasise interviews with key informants and stakeholders. This places significant importance on the data source and reporting being independent, unbiased and truthful. Social desirability bias, confirmation bias, availability bias, and anchoring bias are all examples of biases that may creep into perception data or self-reported data (Barooah et al. 2017; Kahneman et al. 1991). Triangulation of data can help (BetterEvaluation 2014a). Another method is to collect data from key independent sources using the Bellwether Methodology, developed by the

Harvard Family Research Project (Coffman and Reed 2009).[4] Information collected from bellwethers can illustrate the effectiveness of advocates in communicating their messages and moving their issues onto the agenda of policymakers. Biases also creep in during analyses (Coates and David 2002). This is mainly because of incentives and a strategic need for confidentiality (Jones 2011). Clearly, advocacy evaluations need to take these biases into account and mitigate them to the extent possible.

Evaluations need to be rigorous, and a variety of tools are needed: Advocacy campaigns usually take place in complex and dynamic contexts, and different factors may contribute to or inhibit desired change. Evaluations that aim to establish and measure causal relationships (say, between x and y) need to build on first principles. In the nineteenth century, John Stuart Mill (1893) posited that the following relationships need to be established to understand causality (Rogers 2014):

- The efficacy condition: Is there a credible causal mechanism between x and y? Could x cause y?
- The endogeneity condition: Could y cause x?
- The sensitivity: Can we *measure* the caused condition – i.e. the percentage of the population that was affected or the incidence of the effect? (Kellstedt and Whitten 2013)
- Specificity: Can we also measure what proportion of the population was *not* affected or the absence of the effect, truly?

The efficacy condition may be answered if there is pre-existing evidence. It usually becomes important to understand if a reverse causal relationship is also possible (the endogeneity condition). Furthermore, knowing the sensitivity and specificity of these relationships is also key. In responding to the specificity question, it also becomes important to identify confounding variables. We argue that theory-based impact evaluations take care of a subpart of this overall menu of relationships.

Should random or random-like assignment be possible, theory-based impact evaluations help to examine the efficacy condition, for instance, but their ability to provide high-sensitivity and high-specificity evidence depends, to a large extent, on the data available. The audience for theory-based impact evaluations generally includes the research community, donors and implementing organisations, and their utility may be to prove the impact of programs, learn on a micro-level about what works or does not work to achieve given outcomes, improve or replicate programs, or provide justification for ending ineffective programs.

Process tracing methods can aid advocacy evaluations effectively. Befani and Mayne (2014) discuss two elements of process tracing that can help impact evaluations and are relevant to advocacy evaluations. These include establishing uniqueness and certainty: in any hypothesised causal chain, an evaluator must compare alternative causal sequences through (a) reviewing the evidence under the assumption that the

[4] The Bellwether Methodology was created to determine, through structured interviews with 'bellwethers', where a policy issue or proposal is positioned on the policy agenda by investigating how decision makers and other influential individuals are thinking and talking about it and their likelihood to act on it. It can also be used as a tool to inform advocates about specific knowledge gaps related to their issue among bellwethers.

hypothesised causal relationship holds – i.e. x led to y – in the theorised way, and (b) reviewing the evidence with the assumption that the causal relationship does *not* hold – i.e. an alternative causal relationship explains the sequence.

The test of strength of evidence in a theory of change is often built on probabilities. With knowledge of contextual factors (prior knowledge, timing, the ways in which facts emerge), evaluations often depend on the investigator's logic to deduce causality. Subsequently, the evaluator examines the available evidence to compare the inferential weight of evidence of each of these alternative potential causal relationships. As Collier (2011) and Befani (2012) explain, this has the advantages of 'regularity' (i.e. knowledge of frequency of co-occurrence of two events) and 'configuration' (i.e. with the evidence the evaluator is able to judge necessity and sufficiency of the causal relationships) and it can be 'generative' (i.e. can explain how an event occurred), but it may not have the advantages of counterfactuals (i.e. may not be able to inform how *much* was the effect and how much of it was caused by a specific factor)

Contribution, rather than attribution, may be the goal of advocacy evaluations: Organisations frequently focus on establishing contribution rather than attribution of actions.[5] This is especially true for advocacy work, since advocacy initiatives typically require a multitude of actors to work together. Most advocacy organisations recognise that they may not be solely responsible for an outcome. This is especially true for policy advocacy (Gardner and Brindis 2017). While the overall effect of advocacy efforts in putting the issue on the agenda may be easily verifiable, it is harder to pinpoint the contribution of any one organisation or initiative to the change in policy. In some cases, attribution may also be undesirable (Zandniapour and Brennan 2010; Coates and David 2002; Cluster Munition Coalition n.d.).

We argue that it is still useful to measure the overall effectiveness of a coalition of organisations, even though attributable change cannot be mapped to each member of the coalition. This can also be done using theory-based impact evaluation methods. Indeed, advocacy evaluations can also contribute significantly to learning whether certain coalitions and coalition types work better than others (UN OCHA n.d.).

Timing of the evaluation is critical: Given the dual focus on learning and improving the outcomes of the advocacy effort, an increasing number of advocacy evaluations are either prospective or conducted in real time. Prospective evaluations are assessments undertaken before the action to help researchers understand the likelihood of success, and frequently also to enable strategic and timely learning about advocacy and policy change efforts, so that immediate course shifts can be made if needed. So, whereas an evaluation might primarily look back to assess impact at the end of an intervention, the prospective approach assesses progress along the way to long-term impact goals

[5] It is important to note that contribution does not mean 'likely attribution', for which less rigorous methods can be used. Contribution refers to proving overall attribution to a set of interventions, and a given agency A may have contributed one of the interventions to that set. In other words, we can demonstrate that the overall program worked (or not) through rigorous counterfactual methods and that agency A contributed to that program with inputs. However, we do not know for sure whether agency A's contribution made the overall program more effective, but other methods may help us theorize about this.

(Gienapp and Cohen 2011). According to Julia Coffman, advocacy evaluation expert and founder of the Center for Evaluation Innovation,

> The main benefit of a prospective approach is that it positions the evaluation to be useful for both learning and accountability purposes. It delivers feedback to refine advocacy strategy and implementation, and encourages advocate engagement in the evaluation process. Coffman (2009 p.9).

Additionally, the long-term nature of advocacy initiatives and length of time between the start of an initiative and an outcome is often so long that the causal chain becomes difficult to piece together. For this reason, incremental changes, or interim outcomes, may be more fruitful to focus on in such cases (Chapman and Wameyo 2001).

Frameworks for data analysis: The role of analysis in advocacy evaluation presents another challenge. Different evaluation methods do not provide prescriptions for how to analyse the data collected. A key goal in the analysis of data is to make a plausible case that a specific initiative contributed to a specific policy outcome in a specific way.

Collier (2011) and Beach and Pederson (2013) illustrate the use of 'uniqueness' and 'certainty' as conditions of evidence for how data may be analysed and interpreted. They compare finding evidence to doing forensic analyses after a crime has been committed (a murder). They divide this into four categories of tests, according to whether and the extent to which the evidence presented is unique and certain:

- The straw-in-the-wind test (low uniqueness and low certainty);
- The hoop test (high certainty, necessary to confirm hypothesis);
- The smoking gun test (high uniqueness, sufficient to confirm hypothesis); and
- The doubly decisive test (high certainty and high uniqueness).

Similarly, advocacy evaluation experts Jim Coe and Jeremy Smith use the concept of 'strategic plausibility'. They suggest that results-based frameworks are not useful when applied to complex social and political environments, and that 'a demand for predictable, quantifiable results may lead to a misrepresentation of reality'. Instead, to demonstrate accountability to partners, beneficiaries and funders, evaluators need to take an 'expansive but realistic view of how changes come about and the role an individual organization can have within the wider processes' (Coe and Smith 2015). This allows an evaluator to address the complexity of an advocacy initiative rather than assume that it follows a linear trajectory.

Strategic plausibility is the idea that an evaluator should interrogate and assess the explicit and implicit logic of how change occurs and through which actions. Like Collier and Beach and Pederson, Smith likens this to a criminal trial in the absence of strong scientific evidence, in which the prosecutor's case rests on the quality of his logic and his ability to disprove alternative explanations beyond a reasonable doubt (Smith 2015).

Role of the evaluator: Advocacy evaluators often play a similar role to programme evaluators, in that they are situated outside of the initiative or programme being evaluated, and they are expected to be objective in collecting and judging the evidence before them.

Many social science disciplines hold that the researcher or evaluator should be objective and distanced from the subject being studied or evaluated. With developmental evaluation evaluators embedded in the programme (USAID 2016), their role is not to make independent assessments, but rather to facilitate evidence collection and use for improving ongoing work. Developmental evaluation, pioneered by Michael Quinn Patton and others to evaluate complex, evolving efforts, can be very useful for advocacy initiatives:

> Developmental evaluation refers to long-term, partnering relationships between evaluators and those engaged in innovative initiatives and development. … Evaluators become part of a team whose members collaborate to conceptualize, design and test new approaches in a long-term, ongoing process of continuous improvement, adaptation, and intentional change. The evaluator's primary function in the team is to elucidate team discussions with evaluative questions, data and logic, and to facilitate data-based assessments and decision-making in the unfolding and developmental processes of innovation. (Coffman 2009 p.7; Patton 2011).

Coffman explains:

> Developmental evaluation is different from traditional evaluation in that evaluators do not make definitive judgments about success or failure. Rather, like with prospective evaluation, they provide feedback, generate learning, and either support strategy decisions or affirm changes to them (Coffman 2009 p.11; Patton 2011).

2.3 Main findings

Interviews with experts and review of literature help us identify some unique challenges that must be taken into account while evaluating advocacy programmes. Advocacy initiatives often operate in dynamic, rapidly changing and highly uncontrolled environments. There is usually a diversity of actors implementing different and perhaps publicly unobservable treatments (or campaigns), which may be difficult to untangle. The phenomena are unique and hence comparisons cannot be made easily.

Advocates' efforts, or treatments, do not operate in a vacuum, nor can they anticipate, plan for or even identify the confluence of factors at play for a specific issue. Furthermore, advocacy evaluations are sometimes begun only after an initiative has begun, rendering the use of some evaluation designs impossible (Befani and Mayne 2014).

One challenge in advocacy evaluation is in establishing and measuring causality. Often, the process of ruling out all other plausible explanations for an outcome is difficult, if not impossible. Because of this, the degree to which one can prove that other factors outside the intervention are not responsible for the changes observed is often partial and somewhat subjective and is not always prioritised. This means that, in advocacy evaluation, a lower level of certainty is often accepted when it comes to demonstrating impact. After speaking with experts and reviewing the literature, we reached the following conclusions:

- Advocacy evaluation should determine whether a *plausible and defensible case* can be made that an advocacy effort has had an impact on the policy process or *contributed to* a policy change, using means for establishing certainty and uniqueness of evidence. Advocacy evaluation should focus on and assess big outcomes that precede policy change – e.g. public will, political will or shifts in social norms. Evaluation should ideally document the long-term impact of advocacy and policy change on people's lives – e.g. on the environment or the economy (Coffman 2009).
- Theories of change can be useful starting points for most advocacy evaluations. They can help in understanding impact and establishing uniqueness and certainty of evidence. Theory-based impact evaluations have several uses in advocacy evaluation. We argue that they are useful if the goal is to establish efficacy, measure overall change, or learn about how a critical bottleneck in a (large and complex) theory of change may be removed or reduced.
- There are many ways to generate evaluative evidence that has high specificity and high sensitivity that are applicable to advocacy evaluations. Ensuring that bias is dealt with appropriately is clearly important in evaluations. Additionally, ensuring that evaluative evidence, whether generated through process tracing, strategic plausibility or theory-based impact evaluations, should help to assess competing claims of causal relationships. Advocacy evaluations should generate evidence that has high uniqueness and certainty.

3. A review of impact evaluations of advocacy programmes

This section aims to identify factors contributing to successful advocacy by reviewing behaviour and attitude change advocacy initiatives that have been evaluated using experimental and/or quasi-experimental methods. Typical identification methods include randomised assignment and other quasi-experimental methods, such as instrumental variables, regression discontinuity and difference-in-difference using matching techniques. In many cases, modeling techniques are also used for theory-based impact evaluations (White and Sabarwal 2014a). Theory-based impact evaluations use a theory of change and an identification strategy to identify causal linkages and measure the effects of interventions and programmes. We also argue that theory-based impact evaluations are best used to examine subparts of theories of change and in three cases, as related to advocacy evaluation: first, if there is a critical constraint or a bottleneck in the theory of change that is ill-informed by past evidence (the critical bottleneck condition); second, if the intention is to test the efficacy of the intervention; and third, if there is a need to show measurable change. In this sense, theory-based impact evaluations may also assist in identifying factors *contributing* to the effects of an intervention while explaining why an impact occurred. Identification strategies help to deal with a variety of biases, including selection bias (or programme placement bias) and bias arising due to endogeneity, as well as confounding factors (Barooah et al. 2017). Randomised controlled (or quasi-experimental) trial methods use random assignment to understand and measure the effect of the intervention on the treatment group. Similarly, quasi-experimental methods assess these causal impact by mimicking random assignment through various methods, such as matching or using regression discontinuity designs or instrumental variables (White and Sabarwal 2014b). Since we look at positive or successful experiences, this is a positive deviance analysis (BetterEvaluation n.d.b).

3.1 Data and methods

We conducted a systematic search of published impact evaluation studies to examine the effectiveness of advocacy programmes. We then identified factors associated with successful advocacy interventions using the positive deviance method. Criteria for inclusion and exclusion of impact evaluation studies in this review were developed. To be included in the review, studies had to meet the following inclusion criteria:

- First, studies had to evaluate a single advocacy intervention or campaign, or have separate evaluations for each intervention described. Advocacy interventions in each selected study had to meet the definitions of advocacy used in this paper. Selected studies largely included advocacy interventions in which an individual, group or organisation undertook a behaviour change cause voluntarily (without receiving any monetary or non-monetary incentives). Examples of such interventions included encouraging community members to use bed nets to prevent mosquito-borne diseases (e.g. Bowen 2013) and promoting handwashing behaviour (e.g. Chase and Do 2012).
- Second, the study's evaluation design had to meet the (3ie) definition of impact evaluation, which requires the research design address the attribution challenge and establish cause and effect between programmatic activities and specified outcomes. Studies that were included used identification methods that used either random assignment or quasi-experimental methods.
- Third, we included only studies over a 20-year period, excluding studies published before 1995. This was necessitated by the time frame in which we needed to accomplish this work. However, we believe that this did not leave out many theory-based impact evaluations. As Miranda and colleagues (2016) show, there were very few impact evaluations before 1995.

Studies were excluded from this review if they had no clear or strong identification strategy. Evaluations of advocacy as a field or a sector were also excluded in the review for this part, because these are neither specific nor time-bound initiatives.

We conducted the systematic search using five databases – Google Scholar, JStor, the 3ie Impact Evaluation Repository, the Abdul Lateef Jameel Poverty Action Lab (J-PAL) repository of evaluations, and the Innovations for Poverty Action (IPA) Lab repository of evaluations.[6] We selected these five databases because they were known to include studies that met the level of rigor demanded by the inclusion criteria or had sufficiently large archives containing studies that may have been missed by more targeted databases.[7]

[6] We reviewed and coded 56 studies, but one study (Gerber and Green 2005) contained two separate evaluations, which we treated as independent of each other.

[7] To inform other parts of this study, we also included databases that included grey or unpublished literature (online databases of WSSCC, Overseas Development Institute [ODI], Center for Evaluation Innovation and the Innovation Network); we also searched databases of leading advocacy organizations, including Amnesty International, PATH, Oxfam America, Oxfam Great Britain and CARE International. However, although many of the studies and papers found through searches in sources other than the five databases were relevant for other segments of this paper, they were not included in the review for this part of the study.

Table 1 lists the search terms used in each database, along with the outcomes of each search. The results were carefully assessed using the inclusion and exclusion criteria, and the most relevant studies were included in the final review. This search process was exhaustive and continued until we arrived at a saturation point with no new results. A total of 56 evaluation studies were included in the final review.

Table 1: Search for the evidence review

Database	Search terms used	Studies turned up in search	New studies downloaded for final review	Studies screened in*
3ie repository	Advocacy	3	0	33
	Campaign	41	11	
	Information	420	5	
	Behav*	1,376	198	
	Media	55	1	
	Educat*	1,142	1	
	Knowledge	138	4	
	Communicat*	108	6	
J-PAL repository	Advocacy	0	0	22
	Campaign	90	20	
	Information	317	3	
IPA repository	Advocacy	0	0	1
	Campaign	1	1	
	Information	16	0	
Google Scholar	Impact evaluation of advocacy	622,000	1	0
JStor	Advocacy (in full text) and evaluation (in title)	90	2	0
Total		**625,797**	**254**	**56**

Note: Some studies turned up in multiple searches. The final screened in studies represent unique studies and are not duplicated. We attributed each screened in study to the database in which it was first located.

We found no studies evaluating initiatives to change policy that met our inclusion criteria. Studies included in the review were all evaluations of advocacy interventions to address behaviour change in beneficiaries. This underscores the finding in the first part of this study – evaluations that examine policy advocacy have not traditionally used experimental and/or quasi-experimental methods.

3.2 Data analysis

We used a coding protocol to examine and extract information from all 56 evaluation studies included for the final analysis. Each selected study was treated as an independent case, and the coding sheet for each study (see online Appendix D) included variables that captured publication details, target audience, contextual information, description of the advocacy intervention, methodology, outcomes, effectiveness and limitations of the intervention. These variables allowed us to create a comprehensive analysis of trends and characteristics of advocacy interventions associated with effective

outcomes. Using the coding protocol, all the selected evaluation studies were coded in a Microsoft Excel spreadsheet. When possible, information was coded using a numeric system to facilitate analysis.

Data analysis followed the positive deviance approach, coding advocacy interventions into a binary variable – 1 if the advocacy evaluation found that the advocacy initiative had been successful, and 0 otherwise. We then analysed characteristics of successful advocacy initiatives to draw out trends or correlations. The positive deviance approach is useful in explaining phenomena observed in a small number of studies that prohibit a more quantitative approach to the analysis. Findings from the analysis were categorised in meaningful themes or observations, as we describe in the Section 3.3.

3.3 Results

3.3.1 Description of the sample of reviewed advocacy evaluation studies

Table 2 provides a brief summary of studies reviewed for this study. This table provides the complete list of studies, along with authors' names and date of publication, country where advocacy intervention was implemented, the policy area (e.g. health, political participation, water and sanitation), the type of advocacy initiative (information or media campaign), the evaluation methodology (randomised controlled trial [RCT] or quasi-experimental design), and the sample type and size. The last column states whether the advocacy intervention was successful.

Table 2: Impact evaluations included in thematic evidence review

	Author(s) and date	Country	Policy area	Type of advocacy initiative	Evaluation methodology	Sample size	Advocacy initiative successful
1	Ahmed, Kazi et al. 2013	Bangladesh	Water, sanitation and hygiene	Information campaign	RCT	533 households	No
2	Allcott, Hunt 2011	USA	Environment	Information campaign	RCT	588 households	Yes
3	Ashraf, Nava et al. 2013	Zambia	Health, water, sanitation and Hygiene	Information campaign, Other	RCT	487 households	Yes
4	Banerjee, Abhijit et al. 2010	India	Education, politics or political participation (including voting and corruption)	Information campaign, Organising	RCT	280 individuals	Yes
5	Banerjee, Abhijit et al. 2011	India	Politics or political participation (including voting and corruption)	Information campaign, Media campaign	RCT	775 polling station areas	Yes
6	Banerjee, Abhijit et al. 2014	India	Politics or political participation (including voting and corruption)	Information campaign, Media campaign	RCT	Varied by intervention	Yes
7	Bidwell, Kelly et al. 2015	Sierra Leon	Politics or political participation (including voting and corruption)	Information campaign, Media campaign	RCT	224 individuals	Yes
8	Bowen, Hannah L 2013	Cameroon	Health	Media campaign	Quasi-experimental	1,717 individuals	Yes
9	Céspedes, Jaime et al. 2013	Colombia	Health	Information campaign	RCT	1,216 children, 928 parents, 135 teachers	Yes
10	Chase, Claire and Do, Quy-Toan 2012	Viet Nam	Health, water and sanitation	Information campaign, Demonstration/Protest/ Pressure, Media campaign	RCT	3,104 households	No
11	Chong, Alberto et al. 2011	Mexico	Environment	Information campaign	RCT	199 supermarkets	No
12	Chong, Alberto et al. 2013	Peru	Environment	Information campaign, Other	RCT	6,718 households	No

	Author(s) and date	Country	Policy area	Type of advocacy initiative	Evaluation methodology	Sample size	Advocacy initiative successful
13	Chong, Alberto et al. 2015	Mexico	Politics or political participation (including voting and corruption)	Information campaign	RCT	2,360 voting precincts	Yes
14	Citrin, Jack et al. 2014	USA	Politics or political participation (including voting and corruption)	Information campaign	RCT	45,408 individuals	Yes
15	Creel, Alisha H et al. 2011	Malawi	Health	Information campaign, Media campaign	RCT	300 individuals	Mixed
16	Dinkelman, Taryn and Martinez, Claudia 2014	Chile	Education	Information campaign	RCT	226 individuals	Yes
17	Duflo, Esther et al. 2012	Cameroon	Health, Education	Information campaign	RCT	318 individuals	Mixed
18	Dupas, Pascaline 2011	Kenya	Health, Education	Information campaign	RCT	328 schools	Yes
19	Espinoza-Gómez, F et al. 2002	Mexico	Health	Information campaign	RCT	187 households	Yes
20	Essien, E et al. 2011	Nigeria	Health	Information campaign	RCT	346 Individuals	Mixed
21	Fathelrahman, A et al. 2010	Malaysia	Health	Information campaign	RCT	140 individuals	Yes
22	Ferraz, Claudio and Finan, Frederico 2008	Brazil	Politics or political participation (including voting and corruption)	Information campaign, Media campaign	RCT	373 municipalities	Yes
23	Fryer Jr., Roland G 2013	USA	Education	Information campaign	RCT	1,907 students	Yes
24	Galiani, S et al. 2012	Peru	Health, water and sanitation	Information campaign, Demonstration/Protest/ Pressure, Media campaign	RCT	2,847 households	No
25	Gerber, Alan S and Green, Donald P 2005 (Intervention in West Haven)	USA	Politics or political participation (including voting and corruption)	Media campaign	RCT	17,866 households	No

	Author(s) and date	Country	Policy area	Type of advocacy initiative	Evaluation methodology	Sample size	Advocacy initiative successful
26	Gerber, Alan S and Green, Donald P 2005 (Intervention in Iowa and Michigan)	USA	Politics or political participation (including voting and corruption)	Media campaign	RCT	1,905,320 voters	No
27	Gerber, Alan S et al. 2008	USA	Politics or political participation (including voting and corruption)	Other	RCT	180,002 individuals	Yes
28	Godlonton, Susan et al. 2015	Malawi	Health	Information campaign	RCT	937 individuals	Yes
29	Green, Donald and Vasudevan, Srinivasan 2015	India	Politics or political participation (including voting and corruption)	Information campaign, Media campaign	RCT	60 radio station areas	Yes
30	Gutieras, Raymond et al. 2015	Bangladesh	Health	Information campaign	RCT	650 compounds (areas) of slums	No
31	Humphreys, Macartan and Weinstein, Jeremy 2010	Uganda	Politics or political participation (including voting and corruption)	Information campaign	Quasi-experimental	215 individuals	No
32	Hutchinson, Paul and Meekers, Dominique 2012	Egypt	Health	Information campaign, Media campaign	Quasi-experimental	2,086 families	Yes
33	Jamison, Julian et al. 2013	Uganda	Health	Information campaign	RCT	1,791 individuals	No
34	Jensen, Robert 2010	Dominican Republic	Education	Information campaign	RCT	2,250 students	Yes
35	Kalichman, Seth et al. 2008	South Africa	Health	Information campaign	RCT	353 individuals	Yes
36	Karlan, Dean and Wood, Daniel H 2015	USA	Health, Other (global aid)	Information campaign	RCT	16,889 individuals	Mixed
37	Kazemi, Ashraf et al. 2012	Iran	Health	Information campaign	RCT	91 individuals	Mixed
38	Luo, Renfu et al. 2012	China	Health	Information campaign	RCT	3,661 children	Mixed

	Author(s) and date	Country	Policy area	Type of advocacy initiative	Evaluation methodology	Sample size	Advocacy initiative successful
39	McConnell, Margaret 2012	Ghana	Other	Information campaign, Organising	RCT	1,219 individuals	No
40	Miller, Grant et al. 2012	China	Health	Information campaign	RCT	3,553 students	Yes
41	NIMH Collaborative 2010	China, India, Peru, Russia and Zimbabwe	Health	Information campaign	RCT	18,147 individuals	Yes
42	Oreopoulos, Philip and Dunn, Ryan 2013	Canada	Education	Information campaign	RCT	975 students	Yes
43	Panagopoulos, Costas and Green, Donald P 2011	USA	Politics or political participation (including voting and corruption)	Information campaign	RCT	206 congressional districts	Yes
44	Paluck, E 2009	Rwanda	Other (conflict)	Information campaign, Media campaign	RCT	480 individuals	Mixed
45	Pandey, Priyanka et al. 2009.	India	Education, politics or political Participation (including voting and corruption)	Information campaign	RCT	610 individuals	Yes
46	Pattanayak, S et al. 2007	India	Water and sanitation	Information campaign	Quasi-experimental	529 households	Mixed
47	Reinikka, R and Svensson, J 2004	Uganda	Politics or political participation (including voting and corruption)	Information campaign	Quasi-experimental	202 schools	Yes
48	Shamagonam, J et al. 2005	South Africa	Health	Information campaign	RCT	1,168 students	Yes
49	Smith, E et al. 2008	South Africa	Health	Information campaign	RCT	2,383 students	Mixed
50	Thornton, R et al. 2014	Malawi	Health	Information campaign	RCT	1,634 individuals	Mixed
51	Tiwari, A et al. 2010	China	Health	Information campaign	RCT	200 individuals	Mixed

	Author(s) and date	Country	Policy area	Type of advocacy initiative	Evaluation methodology	Sample size	Advocacy initiative successful
52	Wantchekon, Leonard and Fujiwara, Thomas 2013	Benin	Politics or political participation (including voting and corruption)	Information campaign, Other	RCT	100 villages	Yes
53	Wen, X et al. 2010	China	Health	Information campaign	RCT	2,343 students	Mixed
54	Wu, Z et al. 2007	China	Health	Information campaign, Demonstration/Protest/ Pressure	RCT	825 individuals	Yes
55	Yazdani, R et al. 2009	Iran	Health	Information campaign	RCT	417 students	Yes
56	Young, S et al. 2011	Peru	Health	Information campaign	RCT	3,039 individuals	Yes

Geographical regions covered in the reviewed studies
Selected studies evaluated interventions across 4 continents, with 13 in North America, 8 in South America, 18 in Sub-Saharan Africa and 16 in Asia. The geographic distribution is clearly a function of the databases used for this review.

Table 3: Geographical regions of interventions evaluated

Region*	Number (and percentage) of initiatives
North America	13 (23%)
South America	8 (14%)
Europe	1 (2%)
Sub-Saharan Africa	18 (32%)
North Africa/Middle East	3 (5%)
Asia	16 (29%)

Note: *A single study can cover multiple regions.

Income levels of countries
Nearly half of the interventions (46%) were conducted in middle-income countries.

Table 4: Income levels of countries included in interventions evaluated

Income level*	Number (and percentage) of initiatives
High	14 (25%)
Middle	26 (46%)
Low	18 (32%)

Note: *A single study can cover income levels in multiple regions.

Sample size
Selected studies included individuals, households and physical locations (districts, municipalities, slums, polling stations, radio stations and supermarkets) in their sampling strategy. Sample size of individual participants ranged from 91 respondents in a study in Iran (Kazemi et al. 2012) to 1,905,320 voters in a study in Iowa and Michigan, United States (Gerber and Green 2005). Household sample size ranged from 187 in a study in Mexico to 6,718 in an evaluation in West Haven, Connecticut, United States (Gerber and Green 2005).

Evaluation design
A majority of the evaluation studies included in this review were RCTs (see Table 5). Some of the evaluations reviewed (n=4) followed a quasi-experimental design, including an instrumental variables approach, propensity score matching and difference-in-difference. Three studies self-identified as using mixed-method approaches that combined experimental and/or quasi-experimental approaches with non-experimental methods.

Table 5: Methods used in the studies reviewed to evaluate advocacy initiatives

Study design	Number (and percentage) of studies (%)
RCT	49 (88%)
Quasi-experimental	4 (7%)
Mixed methods	3 (5%)

Goal of the interventions studied

Table 6 shows that advocacy interventions in almost all selected evaluation studies (55 out of 56) primarily targeted behaviour change – e.g. reducing the usage of plastic bags by individuals (Chong et al. 2011). Interventions in 24 studies also aimed to influence knowledge, opinions and attitudes of the beneficiaries – e.g. reducing the stigma related to HIV (Creel et al. 2011). In 10 of the selected studies, advocacy interventions also aimed at promoting democratic change or encouraging citizen involvement in the political process.

Table 6: Goal of advocacy initiatives evaluated

Goal of the advocacy initiative	Number (and percentage) of initiatives
Behaviour change of beneficiaries	55 (98%)
Opinion, attitude or knowledge change of beneficiaries	24 (43%)
Improved material situation of individuals	12 (21%)
Democratic change or channels for citizen involvement in political process	10 (18%)
Behaviour change of policymakers or public servants	5 (9%)
Civil society capacity building, organising or building of alliances	2 (4%)
Budget allocation or reallocation	1 (2%)
Change in public policy/implementation of policy	-
Private sector change	-

Note: *A single study can cover multiple initiatives, which is why the percentage sum exceeds 100.

Types of advocacy evaluations reviewed

Advocacy interventions included information and media campaigns; public campaigns or protests; efforts to organise citizens; and other types of initiatives, including citizen debate and influential mailings. Table 7 provides the numbers and percentages of the types of advocacy evaluations reviewed (see also Table 2).

Table 7: Type of advocacy interventions evaluated

Types of advocacy interventions	Number (and percentage) of initiatives
Information campaign	55 (98)
Media campaign	11 (20)
Public campaign or protest	2 (4)
Citizen organising	2(4)
Direct contact or lobbying with policymakers	-
Litigation or use of legal system	-
Community budget review	-
Other	7 (13)

Note: *A single study can cover multiple interventions, which is why the percentage sum exceeds 100.

3.3.2 Effectiveness of reviewed interventions and analysis of success factors

Initiatives were coded as successful, unsuccessful or having mixed results, according to the authors' reported findings against stated goals and the research design used. Thirty-four of the evaluation studies concluded that their advocacy interventions had been successful, 10 were unsuccessful and 12 had mixed results. Many studies evaluated multiple interventions; however, we only reviewed the effectiveness of the advocacy intervention in each study.[8]

Almost all 34 studies with successful advocacy interventions had conducted information campaigns or combined information campaigns with another component. The high number could also be because the vast majority (55 of the 56 studies) of studies in the review had some component of information dissemination. Media campaigns in 9 out of the 11 evaluation studies involving such campaigns were successful or had mixed results, and most of these campaigns were narrowly targeted. This indicates that there may be certain factors or characteristics that were different between interventions that were successful and those that were not. Since all the studies had a strong evaluation design, it is possible to tease out significant factors associated with successful advocacy interventions.

In information campaigns, the message is highly important to the success of the initiative. We reviewed studies that evaluated information campaigns and we closely examined the type of information, source of information and channels used to provide information, although we did not assess the quality or appropriateness of the message. Based on this analysis, five observations emerged with respect to factors associated with successful advocacy interventions for behaviour or attitude change of beneficiaries. We discuss these here.

[8] We do not provide judgements on the effectiveness of the messages in the advocacy campaigns, but rather use an overall success rating for the intervention as it was evaluated, using its specific research design.

Observation 1: The source of advocacy matters for the success of advocacy intervention.

Different programme outcomes were observed for different programme advocates (different from information providers). Interventions with a government official as a programme advocate were found to be 83 per cent more likely to be effective than the comparison group (e.g. Ferraz and Finan 2008). When research teams were advocates, interventions were 72% more likely to be effective, and when the advocate was a local NGO or civil society organisation, the intervention was 67% likely to be successful or have mixed results. The least effective interventions were ones in which the advocates were service providers (46%).[9]

Observation 2: Incentives in some form (even small) are associated with successful advocacy initiatives.

Fourteen advocacy interventions offered some kind of monetary or in-kind support to beneficiaries or targets; of these, 13 were reported as successful or having mixed results (e.g. Miller et al. 2012). The success rate of interventions that provided incentives is 93% (higher than the overall success rate of all interventions studies, which is 61%). The monetary or in-kind support offered to participants was small, including such things as free cellular phones for receiving information, or sweets, stickers or calendars during informational visits (e.g. Fryer 2013; Espinoza-Gómez et al. 2002).

Observation 3: Comparing helps. Comparison information about related individuals, groups, or communities matters for advocacy.

Initiatives aiming to benefit a physical community (e.g. a village or neighbourhood) were slightly more effective (89%) than those targeting individuals (84%). But the difference is minor and needs further examination in future studies. What is interesting is that three out of four (75%) advocacy interventions that provided comparison information about another community with respect to the beneficiary's group (i.e. comparison of an intervention in one community with that of another community) were successful, as compared with 61% of advocacy efforts that provided only absolute information about a beneficiary's group or community. Similarly, only 54 per cent of efforts that provided absolute information about beneficiaries were successful.

Observation 4: The information provider matters for successful advocacy interventions for behaviour change.

An information provider is simply the messenger in an intervention. These information providers may or may not be the advocate or the individual or group that designed the initiative or message. Information provided directly by government representatives or NGOs yielded more positive results (73%) than any other source of information for nearly every method of information delivery. Nearly two thirds of the interventions (61%) in which information was provided by a research team were judged successful. However, only 43 per cent of the initiatives in which information was provided by a surveyor were successful. Only 33% of initiatives where information was provided by a service provider were successful.

[9] Service providers are part of the delivery mechanism of the program or project.

Observation 5: The media channel through which messages or information is provided matters.
Advocacy interventions in which the information campaign used newspaper and other mass media channels were the most successful (73%), followed by campaigns that used leaflets or pamphlets (64%) and those that provided information in a group setting (57%). (See Table 8.) Half of the campaigns that provided information to individuals interpersonally (53%) and via text message (50%) were reported as successful. Only one third (33%) of information campaigns that provided information via phone calls were successful.

Table 8: Success rates of interventions based on how information was provided

Channels	Successful	Unsuccessful	Mixed results
Written in a newspaper or in other mass media	73%	9%	18%
Written in a leaflet/pamphlet	64%	12%	24%
Provided in a group setting or public event	57%	24%	19%
Provided orally in person	53%	26%	21%
Provided by text message	50%	50%	0%
Provided orally by phone	33%	42%	25%

3.4 Discussion

We reviewed 56 studies that evaluated advocacy interventions using experimental and/or quasi-experimental research methods to understand the effectiveness of these interventions and the factors associated with positive programme outcomes. Almost all advocacy interventions evaluated by the selected studies centred on information provision for behaviour or attitude change in beneficiaries. This underscores the ease with which experimental and/or quasi-experimental methods may be used for understanding the effectiveness of behaviour change programmes. It also underscores the fact that it is harder to employ these techniques in other sorts of advocacy interventions and indeed, that the sub-discipline of theory-based impact evaluation will need to innovate methodologically if it is to stay relevant to larger and different advocacy initiatives.

Our review found that about two thirds of the information campaigns in the evaluated studies were successful. The most successful information campaigns demonstrated one or more of the following characteristics: they had government officials as programme advocates; they provided target audience groups with some incentives; they offered audience comparison information about another (comparison) community; they provided information via government officials and NGOs; and they often used mass media channels.

The source of advocacy emerged as a salient correlate of successful initiatives. When government officials took up behaviour change issues for advocacy, the programme was more likely to be successful as compared to when local service providers were the advocates. The same held true for research teams. We argue that this is probably related to the fact that when a government official advocates for a specific behaviour

change, people are more likely to take the matter and the required change seriously. We do not cross-correlate this with trust levels of government officials in different countries, but suggest that that may be a useful analysis in a bigger sample of studies. Research teams also seem to be quite persuasive – arguably because they are usually composed of scholars who are generally well-respected, educated and well-connected.

Similarly, initiatives in which the information was provided directly by an NGO or government representatives were most successful. For example, in Brazil, Ferraz and Finan (2008) studied a successful advocacy intervention in which the government provided information via federal audit reports and assessed expenditures of federally transferred funds by municipalities. This finding could have implications for evaluations aiming to test advocacy campaigns that provide information through a survey team as opposed to through a local organisation or government partner. Future research should also examine why behaviour change advocacy initiatives by local service providers were relatively unsuccessful. Arguably, people may perceive a vested interest of the local provider.

Interventions that provided in-kind or monetary incentives to beneficiaries or advocates to engage in an issue leveraged greater participation and tended to be successful. A systematic review and meta-analysis of studies that examined the role of personal financial incentives in changing habitual health-related behaviours found that although financial incentives changed habitual health behaviours and may help reduce health inequalities, the impact of financial incentives was not sustained long after the incentive disappeared (Mantzaria et al. 2015). We do not suggest that incentives are either necessary or sufficient for advocacy programmes to be successful. However, we discuss possible interpretations of this result while underlining that incentives vary in value and meaning:

- First, advocacy interventions in which gifts were used to incentivise participation in a meeting, visit or event where one received information may have been effective because they lowered the opportunity cost of attending meetings. But at the margin, the incentive is likely to have attracted individuals who were otherwise less interested or uninterested in the subject matter. Organisers thus stood to see a greater change among this population than those already familiar with or acting on the issue.
- A second possible interpretation is that the incentive was small and therefore only attracted those already interested. Indeed, it is possible that something else eased the translation into action. In a study in Peru (Chong et al. 2013), beneficiaries who were randomly assigned to received recycling bins were much more likely to recycle, compared with others who were randomly selected to only receive information about recycling.
- A third interpretation is that the provision of informational materials in a form other than the traditional pamphlet or leaflet – e.g. on a calendar or sticker – may make the medium itself interesting for the beneficiary (e.g. Banerjee et al. 2014). Similarly, Chong and colleagues (2013) report that providing informational stickers in addition to recycling bins increased recycling behaviour.

Providing comparison information about one's group or community (in relation with another group or community) could be another facilitating factor in successful advocacy

interventions. Allcott (2011), for example, found that providing comparative information that showed households' energy usage compared with their neighbours was effective at reducing the energy consumption of poorer-performing households without negatively influencing the better-performing households. Telling people that their neighbours are conserving energy is three times more effective in curtailing energy expenditure in a household than simply telling the household they can cut their bills by using certain methods. In social psychology, this phenomenon is called conformity – the act of matching attitudes, beliefs and behaviours to group norms. This tendency to conform occurs in small groups and/or society as a whole, and may result from subtle unconscious influences or direct and overt social pressure, and often arises from a desire for security within the group.[10]

We also found that successful advocacy interventions frequently use mass media campaigns. Information campaigns met with greater success when information was provided in a newspaper (Reinikka and Svensson 2004), by leaflet or pamphlet (Yazdani et al. 2009), in person, or in a group setting. Banerjee and colleagues (2011) published report cards for politicians in leading local newspapers that were delivered door to door. In this case, although the information was disseminated in a relatively impersonal manner, it was highly relevant to the population. We hypothesise that mass media channels address the entire target group and therefore aim to transform social norms. Mass media channels also reduce the isolation or persecution that an otherwise lone or small subgroup may feel if they are targeted by themselves, while highlighting the importance of the topic.

Advocacy programmes can be implemented at three levels: individual, community-wide and national. In our sample, we find that individual and community-level interventions are not as successful as national interventions, but determining why is difficult. We present four theories:

1. Intensity: People prefer not to change their behaviour. Inertia is strong, and changing behaviours requires significant changes in thinking and action. Messages are easier to ignore when the intervention is small and there is no pressing need to respond to each such message, but when information permeates widely and is spread through mass media, it becomes difficult to ignore. This is especially true for health-related advocacy interventions. Examples are the cases of recycling and smoking cessation.
2. Negative externalities: Many national advocacy campaigns underscore the negative externality of inaction – i.e. the harm that not changing will do to *others.* Examples include the movement against drunk driving (drunk driving kills children) and the argument against smoking (passive smoking has adverse health consequences).
3. Peer effects: Changes in the share of people who engage in a certain behaviour, e.g. smoking, may affect other people's decision to quit (Cutler 2004).

[10] Robert Cialdini, professor emeritus of psychology at Arizona State University and best known for the 1984 book *Influence: The Psychology of Persuasion*, calls this concept 'social proof' and includes it in his six principles of persuasion. For Cialdini, social proof is the phenomenon that people do things that they see other people doing. In one of his experiments, one or more participants would look up into the sky. They found that bystanders would quickly gather and also look up to the sky to look for what the others were seeing. At one point, this experiment had to be stopped because too many people started looking up and it stopped traffic.

4. Scale-up: It is possible that the national campaigns were scaled-up versions of locally successful advocacy initiatives that targeted behaviour change, and that the theory-based impact evaluations we examined presented evaluations of national initiatives rather than more modest ones.

Advocacy initiatives in which information was provided orally by phone or by text message were not as successful in bringing about the desired change (e.g. Jamison et al. 2013). This may be because text messaging is an impersonal form of communication. Furthermore, the recipient of a targeted text message may feel singled out and may deny the behaviour; this method also creates ambiguity about who is providing the information and raise doubts about the truthfulness of the information, as compared with information that comes from mass media channels. However, this finding about text and by-phone oral messages needs examination in future studies. It is also important to note that several evaluations of mobile phone initiatives used for various economic and social development programmes have shown promising results (Ishola and Chipps 2015), suggesting that the type of target, recipient and supplier of the information, as well as the context in which the information is given, is correlated with the outcome of the initiative (Borkum et al. 2015; Cole and Fernando 2012; Kremer et al. 2009).

Other advocacy initiatives use a combination of approaches. Banerjee and colleagues (2014) present an evaluation of a successful advocacy campaign in India, in which local organisations undertook a pre-election awareness campaign in the lead-up to Gram Panchayat elections to help slum dwellers hold local politicians electorally accountable. The campaign included activities such as calendar distribution, door-to-door canvassing, street theatre and using report cards to convey information about elected officials' responsibilities in providing public goods and employment to the villages. The printed calendar highlighted district statistics, and a multi-year report card was released in a newspaper, providing information on councilors' or village leaders' performance. This advocacy campaign was found to have influenced councilors to redirect spending towards slum-relevant infrastructure.

3.4.1 Limitations of this study

This paper is a starting point, and more research is needed in this area of advocacy evaluation. Future research should also focus on factors such as context, geographic region and economic status of the country, and examine which advocacy or information campaigns work best by context.

This study identifies countries or regions, or even the national income grouping, but does not do a detailed analysis of which advocacy interventions worked best in which context, due to the small sample of studies in each context. Studies need to examine if context (including cultural norms, political participation, type of governance, severity of targeted issues, and socio-economic factors of targeted population) is associated with successful advocacy interventions for behaviour change, as well as the characteristics or factors associated with success.

This is a review of a limited number of studies, and we do not undertake a meta-analysis. The positive deviance approach we used means we are able to present only some attributes of successful programmes and are not able to indicate what might be determining factors for programmes to be successful.

We only examine published studies in five databases. We do not look at unpublished studies or the grey literature. Thus, this part of the study is likely to suffer from some degree of publication bias: it is unlikely that many programmes or advocacy campaigns with no effect or negative effects were written up for publication in journals. We do not undertake a meta-analysis of these studies.

Research designs used in these evaluation studies relied on experimental and/or quasi-experimental designs. This also presents limitations. Experimental methods are unable to deal with changing and dynamic contexts or rapidly changing interventions or target groups. This contextual element is especially important in advocacy interventions that aim to influence policy, systems, or organisations, and communities, and even for behaviour change.

4. Additional methods for advocacy evaluation

In this section, we describe additional methods of evaluations of advocacy interventions that may especially be employed for policy advocacy evaluations. We also provide a toolkit for evaluators of advocacy interventions.

One limitation of the review that we discuss in section 2 is that it excludes evaluations of advocacy programmes that aim to influence policy, because there are few theory-based impact evaluations that examine such advocacy programmes. This is likely because it is difficult to use impact evaluation methods for evaluating policy advocacy interventions as it is difficult to construct a counterfactual in policy settings.

So, what methods can we use to evaluate policy advocacy initiatives? Advocacy evaluation is a young discipline without agreed-upon methodologies. Some consensus on best practices is emerging, and these have influenced the methods we highlight and discuss in this part of the study. In section 4, we focus on an array of evaluation methods and approaches for advocacy initiatives that aim to change policy or policy advocacy evaluation. Interviews with experts and review of literature have helped us compile a set of additional approaches to policy advocacy evaluation. These methods not only contribute to efforts in evaluating policy advocacy; they also aid in overcoming the challenges discussed to evaluation approaches, which we discuss in sections 2 and 3.

4.1 Methods

We employed a mixed approach that used qualitative interviews and a literature review to identify other methods for evaluating advocacy initiatives. We conducted qualitative interviews with 14 advocacy and advocacy evaluation experts from international advocacy and evaluation-based agencies, NGOs, consulting firms and educational institutions. (See online Appendix C for a list of experts consulted.) This interview and literature review process revealed a list of methods to evaluate advocacy interventions.

Interview data was complemented by a comprehensive literature review of advocacy evaluation methods, including academic literature and publications and tools from evaluators and advocacy institutions. Brief summaries of evaluation methods are presented here, along with concrete examples of application. Please refer to Appendix A for a list of published references gathered for each of the presented evaluation methods. This is intended to be a suggestive list, not an exhaustive one. We aim to show how the

listed methods can address some of the specific issues faced in evaluating advocacy. Understanding the key questions addressed by these approaches will also be useful in guiding further discussion on their areas of complementarity with impact evaluation approaches.

4.2 Results

This section describes advocacy evaluation methods, additional to experimental and/or quasi-experimental methods, which may be used for policy advocacy evaluations.

4.2.1 Case study

Advocacy never operates in a policy vacuum. Political, cultural and social contexts are continuously changing, making evaluation in this uncontrolled environment difficult. One way to understand and document an advocacy initiative operating in a complex context is the case study method. The term ‘case study’ is often misunderstood and misused to mean anecdote, qualitative method, brief synopsis of experience or program, or an unscientific process. There is a lack of consensus among different disciplines about how it should be used and what makes it a credible form of evidence or evaluation tool. As a method, it has improved significantly in recent years, as has its credibility.

Although the term ‘case study’ is used broadly to describe single-observation events, largely due to the long history of the case study as an in-depth research tool, case study as a methodology is quite specific. Swanborn offers a comprehensive definition in *Case Study Research: What, Why, How?* (2010), highlighting aspects that are often overlooked when the term is used incorrectly.

A case study refers to the study of a social phenomenon:

- It is carried out within the boundaries of one social system (the case) or within the boundaries of a few social systems (the cases), such as people, organisations, groups, individuals, local communities or nation states that are the case’s natural context;
- It monitors the phenomenon during a certain period or, alternatively, by collecting information afterward in the context of a certain period;
- The researcher focuses on describing and explaining social processes that unfold between people participating in the process, including their values, expectations, opinions, perceptions, resources, controversies, decisions, mutual relations and behaviour;
- The researcher is guided initially by a broad research question; explores the data; and, only after some time, formulates more precise research questions, keeping an eye open for unexpected aspects of the process;
- Uses several data sources, such as documents, interviews with informants and participant observation; and
- Optionally in the final stage, the investigator invites the studied persons and stakeholders to present their perspectives, or to share with them the preliminary research conclusions, not only to attain a more solid base for the final report, but also sometimes to clear up misunderstandings, ameliorate internal social relations and ‘point everyone in the same direction’ (Swanborn 2010 p.13).

With case studies, the training of the researcher and the quality of implementation of the method are extremely important to the validity and usefulness of the evaluation. Unacknowledged bias and poor analysis skills can render the case study less useful.

4.2.2 Process tracing

Process tracing was first developed in 1979 for political science research to analyse historical events, and it is often an approach used in a case study. Process tracing focuses on discerning causation through identifying possible causal mechanisms between observed variables by examining the fit of a theory to the intervention or initiative's causal steps. Process tracing is used to explain 'how "X" produces a series of conditions that come together in some way (or not) to produce "Y". By emphasising the causal process that leads to certain outcomes, process tracing lends itself to validating theoretical predictions and hypotheses' (Wesleyan University n.d. p.32). As such, process tracing does not aim to measure impact, but rather to increase our confidence in the identified causal mechanisms that led an intervention to the observed outcome (Befani and Mayne 2014). By identifying the causal mechanisms at play, process tracing allows for the observation of the avenues through which change occurred and how the various mechanisms led to change.

Process tracing uses an intervention's theory of change and the assumptions it holds about the actions, events and conditions that will occur as the initiative progresses from inputs to outcomes and impacts (usually in a nonlinear fashion). The accuracy of these assumptions are what process tracing tests. It includes four key steps:

1. Build confidence in the causal mechanisms using knowledge of the context and intervention alone (pre-data collection);
2. Using the data and information collected, test whether the intervention indeed contributed to the outcome;
3. Test whether other factors outside the intervention may have contributed in some regard to the outcome; and
4. Test the entire theory of change to examine whether the anticipated chain of effects unfolded as expected in reality (Befani and Mayne 2014).

Box 1: An example of process tracing

Health for all: towards free universal health care in Ghana 2012/13
Evaluator: Gavin Stedman-Bryce

Oxfam Great Britain uses an adapted version of process tracing to evaluate its advocacy initiatives. An example is Stedman-Bryce's project effectiveness review, *Health for all: towards free universal health care in Ghana 2012/13*. This campaign aimed at improving the National Health Insurance Scheme in Ghana by increasing awareness of the weaknesses of the scheme and proposing specific solutions. To assess the project's contribution to key policy outcomes, Stedman-Bryce examined (1) the extent to which interim or longer-term outcomes were successfully achieved and (2) the extent to which the programme contributed to those outcomes.

He reconstructed the campaign's theory of change with stakeholders; identified interim and final outcomes; and obtained and assessed evidence on whether the outcomes materialised – and, if they did, the plausible causal explanations for those outcomes.

Stedman-Bryce conducted 21 carefully selected key informant interviews, which were systematically coded to parse key evidence, recurring regularities and emerging themes. He used triangulation to validate the data and findings. He also investigated 50 relevant documents from a variety of sources and worked with the campaign team to identify three key interim outcomes, which the campaign was working to achieve. He used participatory methods to facilitate identification of key outcomes and information.

The process tracing study concluded that there was strong causal evidence that the campaign's cornerstone report, despite the significant backlash it received from Ghanaian politicians, had led to a revision in the way national authorities calculated and reported the impact of the program, which had been previously grossly overstated. The campaign also was found to have mobilised civil society organisations working on health access issues. The campaign was not, however, found to have had an impact on the support for free universal healthcare in the lead-up to the following elections.

For more information about this evaluation, see Stedman-Bryce (2013).

4.2.3 Contribution analysis

Contribution analysis (Beer and Coffman 2015, citing Mayne 2008) is often used in conjunction with or as a starting point for process tracing (Befani and Mayne 2014). Different from process tracing, contribution analysis is a tool created by evaluators for evaluation. However, similar to process tracing, it is used to assess the accuracy of a theory of change using collected data, results and logic to test whether the assumptions articulated in the theory of change hold in reality. Therefore, the goal of contribution analysis is to establish a level of confidence that an initiative contributed to observed results. Importantly, contribution analysis also has the power to identify why results occurred (or did not) and how the actions of the initiative evaluated played a role in the outcomes, which makes it a powerful tool not only in the assessment of an initiative's contribution, but also in the dynamics of that contribution, which can be useful for learning and future implementation.

Contribution analysis is typically retrospective, conducted after policy outcomes are observed. The process has six iterative steps:

1. Mapping advocacy results using a logic model, outcome chain or similar approach;
2. Gathering existing evidence on those results;
3. Exploring alternative explanations for the results to determine if they might provide a better explanation of the observed results than the advocacy effort being examined;
4. Developing a performance story that lays out the context, planned and actual accomplishments, lessons learned, and main alternative explanations for the results, along with why those alternative explanations should not be accepted;
5. Seeking additional evidence where alternative evidence cannot be discounted or where the contribution argument is questionable; and
6. Revising and strengthening the performance story where possible.

If this cannot be done, either more data and analysis is required or the conclusion is that a plausible and defensible case cannot be made that the advocacy effort contributed to the observed results.

4.2.4 General elimination method

One of the processes for ruling out other plausible explanations for an outcome is the general elimination method (GEM). GEM is a case study method that is used to determine whether a plausible and defensible case can be made that the advocacy effort had an impact by eliminating alternative or rival explanations of cause and effect. When using GEM, one gathers evidence to eliminate alternative ideas about what caused an outcome until the most convincing and evidence-based explanation remains (Coffman 2010).

GEM involves three steps:

1. Identifying the list of possible causes for the outcomes and impacts of interest;
2. Identifying the conditions necessary for each possible cause in the list to have an effect on outcomes or impacts; and
3. Working out whether the conditions for each possible cause are present or not. If these elimination steps are followed, the final set of causes should only include those whose requisite conditions are completely present (Scriven 2008).

One of the key conditions for GEM to be effective is that the list of possible causes must be exhaustive in order to rule out all alternative hypotheses. This can be difficult, if not impossible, in some settings (Alexander and Bonino 2015 p.10).

Box 2: An application of GEM

Evaluation of the Supreme Court advocacy campaign
Evaluator: Michael Quinn Patton

Michael Quinn Patton used GEM to evaluate whether an advocacy campaign implemented over nine months, with more than $2 million in support from a number of foundations, influenced a Supreme Court decision.

In his application of GEM, Patton searched for connections between an effect (the Supreme Court decision) and the intervention (the advocacy campaign) using a retrospective case study. He gathered evidence through 45 interviews with individuals directly involved in the campaign or the case, document analysis, a review of court arguments and decisions in more than 30 court documents and 20 scholarly publications and books, and news analysis and documentation from the campaign. He then applied the forensic method, in which he eliminated other plausible explanations until the most compelling, evidence-based explanation remained.

By ruling out other possible theories, Patton was able to conclude that the campaign contributed significantly to the Supreme Court's decision.

For more information about this evaluation, see Patton (2008).

4.2.5 Episode study

One case study approach for understanding a policy change is the episode study. An episode study is especially useful for identifying the impact of research on policy change (ODI 2009). In such a study, an episode of policy change or influence is the starting point for the case study (Carden 2009). It is then used to go back and track what impact research had among the variety of issues or events that led to the policy change. An episode study could focus on a single episode or several comparable episodes. It differs from other case study approaches, which usually take an initiative as the starting point and look forward (BetterEvaluation 2014b).

Steps to undertake an episode study:

1. Identify the policy change that is to be evaluated as the outcome of interest;
2. Identify the research questions related to the policy issue;
3. Construct a historical narrative that follows the events leading up to the policy change, including a timeline of key decisions, practices, and events and the individuals involved. This requires using multiple sources of information; and
4. Use the historical narrative to explore how and why the important policy decisions and practices occurred, and the (potential) role played by advocacy or research efforts (ODI 2009).

Box 3: An application of episode study

The PRSP initiative: multilateral policy change and the role of research
Evaluators: Karin Christiansen and Ingie Hovland

In this report, the authors use episode study to track the evolution and influence of the World Bank's Poverty Reduction Strategy Papers (PRSP) initiative, along with related factors such as the role of research on policy reform. They test three hypotheses about the influence of research on evidence-based pro-poor policy:

1. It is more effective if it fits within or has the capacity to challenge the political and institutional limits of policymakers;
2. It is more likely to contribute to policy change if outputs are based on inclusive processes and communication of the evidence is context-appropriate; and
3. It will be more effective if researchers and policymakers share common networks in particular policy areas by constructing a historical narrative in the lead-up to the policy changes that were observed in each case study they analysed.

They created a timeline of key policy decisions and practices; identified relevant events, documents and actors; and investigated why those policy decisions and practices occurred and how research might have influenced them. Data came from interviews with important stakeholders, literature and document reviews, and triangulation of information.

The authors concluded that policy change within the World Bank and International Monetary Fund to adopt the new PRSP initiative was due, indirectly, to relevant academic and applied policy research. They also found that research is more likely to be influential if it is undertaken with input from policymakers, is rigorous, and evidence is disseminated through research and policy networks.

For more information about this evaluation, see Christiansen and Hovland (2003).

4.2.6 System mapping and network diagrams

Increasingly sophisticated coalitions and alliances are collaborating around advocacy efforts, some of which are global in nature (Chapman and Wameyo 2001). Coalitions may be built around a key issue that is agreed upon, but the groups involved have distinct missions. In other words, rather than just having one goal, the effort may have several goals that overlap in some way. Often, a programme evaluator's first task is to clarify the objectives of the programme being evaluated.

Network diagram and system mapping are tools used to facilitate this process. These tools may help the evaluator to map out different potential roles in advocacy efforts (e.g. lobbying, protesting, advising or advocating). A network diagram is made up of a set of nodes and lines that connect those nodes, and is often created using participatory social network analysis methods. Such maps or diagrams can be part of an articulated theory of change or can be done independent of theory of change (BetterEvaluation 2014c; Durland and Fredericks 2005). Tools exist that can facilitate this analysis, including lists of policy stakeholders that may be considered (Start and Hovland 2004).

4.2.7 Outcome mapping

Outcome mapping can be used to visualise a theory of change and incorporate relationships and intended outcomes during planning.[11] It can also be used as an evaluation approach to analyse the outcomes of an initiative (Earl et al. 2001). It can be combined with other evaluation tools to help assess contribution claims, such as after-action reviews (Smutylo 2005). Since it facilitates an in-depth examination into an initiative's theory of change, it provides a framework to collect data on interim outcomes that can lead to longer-term, more transformative changes and allows for the plausible assessment of the initiative's contribution to identified results.

Identifying interim outcomes is essential in advocacy evaluation, because interim outcomes themselves can be game-changing: in many cases, longer-term goals are not realised and are distant. Classifying interim outcomes can help map progress towards longer-term outcomes, especially when an initiative's path or strategy is modified mid-course. It is important, however, that the interim outcomes measured be the most meaningful, as opposed to the most easily identified, so that the most influential steps are followed and can be traced to later outcomes (Beer and Coffman 2015). Table 9 provides examples of interim outcomes with corresponding descriptions.

Outcome mapping consists of three main steps:

1. Intentional design, which helps a project team identify and agree upon the high-level changes they aim to achieve and plan their strategy accordingly;
2. Outcome and performance monitoring, which provides a framework for monitoring activities undertaken and follows actors' progress towards stated goals; and
3. Evaluation planning, which helps actors effectively allocate effort and resources towards priorities (Smutylo 2005).

Table 9: Examples of interim outcomes

Interim outcome	Description
Attitudes or beliefs changes	Target audiences' feelings or affect about an issue or policy proposal
Awareness raising	Audience recognition that a problem exists or familiarity with a policy proposal
Constituency or support base growth	Increase in the number of individuals who can be counted on for sustained advocacy or action on an issue
Issue reframing	Changes in how an issue is presented, discussed or perceived
Media coverage	Quantity and/or quality of coverage generated in print, broadcast or electronic media
New advocates	Previously unengaged individuals who take action in support of an issue or position
New champions	High-profile individuals who adopt an issue and publicly advocate for it
New donors	New public or private funders or individuals who contribute funds or other resources for a cause

[11] A related method that may be of interest is RAPID outcomes assessment. See <http://betterevaluation.org/evaluation-options/rapid_outcome_assessment>.

Interim outcome	Description
Organisational capacity	The ability of an organisation or coalition to lead, adapt, manage and technically implement an advocacy strategy
Partnerships or alliances	The formation or growth of mutually beneficial relationships with other organisations or individuals who support or participate in an advocacy strategy
Political will	The willingness of policymakers to act in support of an issue or policy proposal
Public will	The willingness of a (non-policymaker) target audience to act in support of an issue or policy proposal
Salience	An increase in importance a target audience assigns an issue or policy proposal

4.2.8 Outcome harvesting

Similar to outcome mapping, outcome harvesting reviews an array of sources to document how an initiative contributed to observed outcomes. Uniquely, outcome harvesting uses evidence to identify what has been achieved, and then assesses how a specific initiative did or did not contribute to that achievement, making a verifiable connection (Wilson-Grau and Britt 2012). Outcome harvesting does not rely on predetermined or suggestive outcomes and track an initiative's progress against them, but rather builds the case for which outcomes occurred and how an initiative is linked to them.

Outcome harvesting follows the following process:

1. Once key questions are developed, data is gathered on actions and outcomes through a variety of primary and secondary sources;
2. Evaluators then validate the information collected by comparing documents and interviews; and
3. Using validated evidence, evaluators then analyse and interpret the data to answer key questions around the contribution of the initiative to the outcomes observed. This can be done through drafting stories, by creating charts or matrices, or through other data visualisations.

4.2.9 Artificial neural networks

One other method that has been used in other research but has not found its way into evaluations of advocacy initiatives, is using artificial neural networks. We think thre are some strengths associated with this emerging approach that could be used critically to inform the effectiveness of advocacy initiatives. Artificial neural networks are processing devices that can be either algorithms or actual hardware modeled to process information in a way similar to the way it is processed by the human brain. Artificial neural networks are composed of a large number of interconnected processing elements or neurons that work together to process information (University of Wisconsin-Madison 2015). The special feature is that artificial neural networks have the ability to learn from experience and adjust accordingly. They can be used to derive meaning from complex and/or imprecise data and can find patterns and trends that are nearly impossible to be observed through other techniques (Stergiou and Siganos 2016). A well-trained neural network can be used to give future predictions and answer 'what if' questions.

Similar to how children learn to talk and categorise things based on examples from their parents, neural networks get trained to work in certain way based on examples where a defined input should result in its expected outcome. Once an artificial neural network is

presented with a pattern, it uses previous information to make an informed guess on what the pattern might be; depending on how far the result is from the prediction, the neural network adjusts its processes by changing the weights of its assumptions. The disadvantage of using neural networks is that they must be trained carefully with previous information. Additionally, since neural networks create their own pathways to solve problems, the process through which results are achieved can be unpredictable for the user (Stergiou and Siganos 2016).

Artificial neural networks work well in these instances:

- Capturing associations or discovering regularities within a set of patterns;
- Cases where the volume, number of variables or diversity of the data is large;
- When relationships between variables are vaguely understood; and
- When relationships are difficult to describe adequately with conventional approaches. (University of Wisconsin-Madison n.d.)

4.2.10 Qualitative comparative analysis

Qualitative comparative analysis bridges quantitative components of research with the qualitative aspects of case-oriented insight. It does this by bringing together integrative binary logic (e.g. two-way logic affirming whether a condition is true or false) and empirical intensity to qualitative approaches, such as Boolean logic (AND, OR and NOT), to establish meaningful interpretations of data (Marshall 1998; Raab and Stuppert 2014). Qualitative comparative analysis can be employed in relatively small and simple data sets and functions to establish necessity or sufficiency among various conditions of causal pathways in complex observed or hypothetical cases (Ragin 2008; Raab and Stuppert 2014).

Qualitative comparative analysis can be advantageous in the following situations (Befani 2016; Kane et al. 2014):

- Analysing a small number of cases (e.g. 5–30);
- Identifying multiple pathways to an outcome;
- Addressing complex pathways with multiple factors and combinations (e.g. INUS conditions[12]);
- Examining relative contributions of different factors and pathways; and
- Establishing alternate or asymmetric pathways.

Raab and Stuppert (2014) provide a celebrated example of qualitative comparative analysis, examining 39 studies for conditions for effective or ineffective evaluation effects, evaluation context, and aspects of evaluation quality in greater depth related to violence against women and girls.

4.3 Tools

Separate from evaluation approaches in this section, we discuss some tools that evaluators of advocacy programmes may use in addition to or to complement well-

[12] Causal conditions that are insufficient but necessary parts of causal recipes, which are themselves unnecessary but sufficient.

recognised quantitative and qualitative methods, to provide important insights into advocacy approaches and processes.

4.3.1 After-action review

After-action review is a straightforward and flexible tool used to facilitate ongoing learning and organisational improvement by convening a team to discuss an activity, event or project in an open and transparent manner. This tool was developed for a military context, but it is used broadly at corporations and increasingly by activists, advocates and political organisations.

After-action review asks relevant stakeholders the following questions (BetterEvaluation 2015):

- What was anticipated to happen?
- What actually happened?
- Why were there differences?
- What worked?
- What didn't?
- Why?
- What would you do differently next time?[13]

4.3.2 Intense period debrief

Intense period debrief is a tool similar to the after-action review, developed by the Innovation Network. The process is undertaken shortly after an intense period of advocacy activity to engage advocates in evaluative inquiry. Periods of high-intensity activity provide critical opportunities for data collection and learning. However, advocates have little capacity for reflection during such times, often resulting in a lack of data collection or debriefing during this time. This absence of data at such a crucial time in the initiative unfortunately leaves important data gaps. Intense period debrief aims to remedy this problem by convening key actors through focus groups or individual interviews, following a debrief interview protocol to acquire information and data about advocates' recent experiences (Coffman 2015).

Because intense period debrief collects in-depth and real-time information from key actors, it can capture the public mood and political context during the intense window of activity. It aims to answer questions related to what occurred; how campaign members responded to actions and events; and what activities occurred behind closed doors, how those closed-door activities occurred and what their meaning was. It can also collect perspective on the outcome(s) achieved or not achieved and how strategies could have been modified to be more successful. As the method's developers note,

> The idea of the debrief grew out of the need to have a forum that encouraged participation from key groups and individuals engaged in different layers or "spheres of influence" surrounding decision makers. It … [is] particularly useful for providing a way for individuals in the "inner circle" of those spheres … to tell the story of what happened behind the scenes (Coffman 2015 p.7).

[13] This summary is taken, with permission, from an ODI/RAPID publication (Ramalingam 2006). Sources for the original article are Collison and Parcell (2001) and Whiffen (2001).

4.3.3 Policymaker rating tool

The policymaker rating tool is designed to gauge political will or support for a particular advocacy issue or proposal among a defined group of policymakers (e.g. a legislature or council). Developed in response to the perceived inadequacy of indicators commonly used to gauge policymaker support on issues (e.g. number of bills introduced on the issue, number of bill co-sponsors or co-signers, or number of votes for or against specific bills), the policymaker rating tool capitalises on advocates' insider knowledge about individual policymakers' stances on policy issues (Coffman and Reed 2009).

4.3.4 Most significant change

When stakeholders do not agree on the most important outcomes or what success looks like, the most significant change tool can be used to understand and account for different viewpoints. This approach involves generating and analysing personal accounts of change and then determining which are the most significant, and why, in a collective manner.

There are three basic steps in using most significant change:

1. Deciding types of stories that should be collected;
2. Collecting stories and determining which stories are the most significant; and
3. Sharing stories and discussion of values with stakeholders.

In addition to guiding the collection and reporting of stories, most significant change provides guidance for learning from these stories. In particular, it provides a tool to learn about the similarities and differences according to what different groups and individuals value. Most significant change is used primarily for classifying values held by different stakeholders and for identifying intended and unintended outcomes. Hence, it is usually combined with other tools or methods to evaluate an intervention. Most significant change can be helpful in explaining how processes and influential mechanisms and in what situations and contexts change comes about (Dart and Davies 2003).

Table 10 provides a brief summary of the additional methods of advocacy evaluation discussed here.

Table 10: Additional methods and tools for advocacy evaluation

	Keywords	Purposes	Prospective evaluation	Appropriate for long-term nature	Appropriate for changing and highly uncontrolled environment	Pros	Potential Drawbacks
				Methods			
Case study	In-depth research tool of What, Why, How?	Looks at a case in its natural context, examines multiple exposures	Yes	Yes	Yes	Detailed, highlights variables of interest	Unacknowledged bias, poor analysis, not widely applicable, cannot make causal claims
Process tracing	Testing hypothesis, observing avenues of change	Examines the fit of a theory to the intervention or initiative's causal steps	No	No	No	Validates and provides insight into the assumptions and predictions that are part of the identified causal mechanism	Not helpful if looking at a weak hypothesis, can lead to losing the big picture
Contribution analysis	Mapping contribution, testing theory of change	Tests accuracy of theory of change assumptions, identifies Why and How of results	No	No	No	Builds confidence in theory of change's assumptions and predictions, identifies what actions of an intervention affected the outcomes	Looking at a weak hypothesis
General elimination method	Defensible case for initiative, eliminate alternative theories	Rules out alternative explanations for outcomes to reach the most plausible reasoning	No	No	No	Narrows the potential causes influencing outcomes	Need exhaustible list of causes, not possible in some settings
Episode study	Influence of single or comparative episodes	Assesses impact of advocacy or research on policy change	No	No	No	Provides a comprehensive narrative	Inefficient for rare instances, might overlook important factors

	Keywords	Purposes	Prospective evaluation	Appropriate for long-term nature	Appropriate for changing and highly uncontrolled environment	Pros	Potential Drawbacks
System mapping and network diagrams	Visualising a network of goals	Maps different potential roles in advocacy efforts (e.g. lobbying, protesting, advising or advocating)	No	No	No	Provides a visual representation	Supplementary tool only
Outcome mapping	Mapping theories of change	Visualises theory of change and analyses relationships and outcomes	Yes	Yes	Yes	Provides a visual representation, highlights strengths and shortcomings, captures interim outcomes	Potentially long and complex procedures, contribution not attribution
Outcome harvesting	Linking actions to outcomes	Assesses an initiative's contribution to verify connections by working backwards	No	Yes	Yes	Especially useful in complex situations, analyses unintended outcomes	Identifies contribution not attribution, only captures outcomes that the informants are aware of
Artificial neural networks	Learning machine	Generalises unseen observations	Yes	Yes	No	Picks up hidden relationships, approximates any continuous function to any desired degree of accuracy	Vulnerable to potential misclassification, time consuming
Qualitative comparative analysis	Analysis of causal pathways	Pinpoints decisive cross-case patterns	Yes	Yes	Yes	Assesses causation in complex cases, can be used for a small number of cases, flexible and adaptable to different situations	Vulnerable to biases due to selection of cases analysis of small number of cases results in limited generalisation
Tools							

	Keywords	Purposes	Prospective evaluation	Appropriate for long-term nature	Appropriate for changing and highly uncontrolled environment	Pros	Potential Drawbacks
After-action review	Structured review	Facilitates ongoing learning by convening actors to openly discuss an activity/event/project	No	No	Yes	Highlights strengths and shortcomings	Questions need to be planned and executed well
Intense period debrief	Post-initiative inquiry	Engages advocates in evaluative inquiry after periods of high-intensity activity	No	No	No	Captures current mood and context	Vulnerable to respondent bias
Policymaker rating tool	Policymakers' support	Gauges political will or support	No	No	No	Assesses level of support on policies	Can be misleading, doesn't cover context
Most significant change	Collecting stories	Analyses personal accounts of change to determine the most significant ones	No	No	No	Provides context, helps identify most important outcomes for stakeholders	Supplementary tool only

4.4 Discussion

Non-experimental approaches are less likely to provide the same level of certainty as robust experimental evaluation tools and are less structured than experimental and/or quasi-experimental approaches; however, they may offer the flexibility, adaptability and rigor required for more structured mixed methods. They could also be valid alternatives in situations where experimental and quasi-experimental approaches are hardly or not applicable.

Qualitative components of advocacy evaluation sometimes simply consist of a few group discussions and interviews with key informants or ethnographic surveys to understand the pathways to impacts and, in some cases, to monitor programme and evaluation fidelity. Qualitative tools, help to describe possible causal links between an initiative and impacts and can be quite important and useful when used in combination with quantitative approaches, which help to measure the strength of these linkages and their impact.

We compiled a set of methods for advocacy evaluation through literature review and interviews with advocacy experts, which enabled us to address the shortcomings of strictly experimental approaches. However, it is important to recognise that no method is one size fits all. Many experts use several methods or parts of methods to design an evaluation strategy that fits the unique context and goals of the initiative they are evaluating. As many experts interviewed noted, designing an advocacy evaluation plan requires the combination of several methods and tools, and often the flexibility to change which methods or tools are used to best fit the context and needs of the evaluation.

5. Conclusion

Advocacy has long been utilized to influence public policy, it is increasingly being recognized as an important component of social and behavioral change initiatives. Its role has been acknowledged by researchers, programme planners and funding agencies, leading to an increased use of advocacy efforts, often as complements to other interventions. With the development and use of numerous and diverse advocacy techniques, it becomes essential to evaluate these advocacy efforts to understand their efficacy. Our interviews and literature review set out the challenges and considerations for evaluating advocacy. Importantly, the need for flexible and more nuanced approaches emerged.

One of the ways advocacy initiatives can be evaluated is by using established or innovative impact evaluation methodologies. We reviewed published articles that had used impact evaluation methods to evaluate behavioural advocacy programmes and identified factors associated with effective advocacy. Identifying and highlighting the factors contributing to positive outcomes can thus inform future advocacy initiatives (programme planners, funders and beneficiaries) of which programme elements work better than others and can be included in subsequent endeavors. Evaluating advocacy carries myriad questions regarding what else we should look at for advocacy efforts – efficacy versus effectiveness, causal density and other uncharted trajectories of change. Undertakings could range from large, transformational change to slow, incremental

adjustments. Furthermore, observing and defining tipping points carries its own set of assumptions.

However, evaluating advocacy programmes using these methods also proved to have shortcomings. As advocacy occurs in sociopolitical contexts that are continually in flux, these methods may be too rigid in their approach. Advocacy requires evaluation methods that capture the nuances of the contextual dynamics, which operate at different levels as advocacy proceeds. In evaluating complex advocacy initiatives, identifying the theory of change is essential in order to understand and highlight the definition of the work, the stage of progress and the final goal. Thinking about the timing, effort and agency of change likely requires multidisciplinary investigation.

Thus, we have presented an array of advocacy evaluation methods that have greater capability to capture the fluid nature of advocacy. It is important to keep in mind that evaluation of advocacy initiatives may require a combination of evaluation methods. Furthermore, a method that is best suited for one advocacy initiative may not be suitable for another. Evaluation methods must therefore be selected based on the nature of the advocacy and the context of its setting. Therefore, the goals of this report are to provide readers and users with a set of diverse evaluation tools and, ultimately, strengthen advocacy evaluation through their appropriate application.

This study adds to the growing study and practise of advocacy evaluation by:

- Identifying the factors contributing to successful advocacy and shedding light on conditions that enable change as a result of advocacy initiatives through a review of advocacy impact evaluations; and
- Discussing advocacy evaluation methods; identifying promising methodologies and illustrating the circumstances in which those methodologies could be most effective..

There is a need for more research in this area to:

- Identify the characteristics of policy advocacy efforts that may contribute to their success, as evaluations of these initiatives did not make the inclusion criteria for this study;
- Continue investigating the impact of providing information on behaviour change. As we discuss, the information type, channel and provider play a key role in the outcomes of an initiative. Further research could shed light on how different types of information are more or less effective at changing the behaviour or attitudes of a population based on the characteristics of that population. A study identifying the appropriateness of the messages in information campaigns would contribute significantly to the evidence base of what makes advocacy successful; and
- Look deeper at advocacy (for behaviour, attitude or policy change) in each policy sector. This would require qualitative assessments of the limited number of advocacy evaluations in a given sector, but could allow for more specific conclusions on the effectiveness of advocacy – e.g. in health or political participation or even deeper, as in preventative medicine or voting.

Appendix A: Advocacy evaluation resources

Tool or concept	Resources
Bellwether Methodology	Coffman, J and Reed, E, 2009. *Unique Methods in Advocacy Evaluation*. The California Endowment. Available at: http://www.innonet.org/resources/files/Unique_Methods_Brief.pdf.
Case studies	Gienapp, A and Cohen, C, 2011. *Advocacy Evaluation Case Study: The Chalkboard Project*. Washington, DC: Center for Evaluation Innovation. Swanborn, P, 2010. *Case study research: what, why and how?* London: SAGE Publications Ltd. Whelan, J, 2006. *Assessing Advocacy: Extracts from 'Work Justice' Case Study*. The Change Agency. Available at: http://www.thechangeagency.org/campaigners-toolkit/research-projects/advocacy-evaluation/.
Contribution analysis	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). Beer, T and Coffman, J, 2015. *Four Tools for Assessing Grantee Contribution to Advocacy Efforts*. Washington, DC: Center for Evaluation Innovation. Available at: <http://www.evaluationinnovation.org/publications/four-tools-assessing-grantee-contribution-advocacy-efforts>. Kotvojs, F, 2006. *Contribution Analysis: A New Approach to Evaluation in International Development*. Paper presented at the Australian Evaluation Society 2006 International Conference, Darwin, Australia. Larbi, G, Christensen, J, Jackson, P and Ura, K, 2006. *Capacity Development in Bhutan: Capacity Development Outcome Evaluation of Danish Supported Organisations in Bhutan*. Copenhagen: Danida. Available at: http://www.oecd.org/dataoecd/45/9/42218416.pdf. Lemire, S, 2010. *Contribution Analysis: The Promising New Approach to Causal Claims*. Paper presented at European Evaluation Society International Conference, Prague. Available at: <https://www.alnap.org/help-library/contribution-analysis-the-promising-new-approach-to-causal-claims> Mayne, J, 2001. Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Programme Evaluation*, 16, pp.1–24. Mayne, J, 2008. *Contribution Analysis: An Approach to Exploring Cause and Effect*. Institutional Learning and Change Brief No. 7. Rome, Italy: Institutional Learning and Change Initiative. Available at: <http://www.adcoesao.pt/sites/default/files/avaliacao/4_22_contribution_analysis_an_approach_to_exploring_cause_and_effect_maio_2008.pdf>. Mayne, J, 2011. Contribution analysis: addressing cause and effect. In: K Forss, M Marra and R Schwartz, eds. *Evaluating the complex*. New Brunswick, NJ: Transaction Publishers.

Tool or concept	Resources
General Elimination Methodology	Patton, MQ, 2008. Advocacy impact evaluation. *Journal of MultiDisciplinary Evaluation*, 5(9). Scriven, M, 1976. Maximising the power of causal investigations: the Modus Operandi method. In: GV Glass, ed. *Evaluation studies review annual*, Vol. 1. London: SAGE Publications Ltd. Scriven, M, 2005. *Can We Infer Causation from Cross-sectional Data?* Paper presented at a Symposium on the Use of School-Level Data for Evaluating Federal Education Programmes, Washington, DC. Scriven, M, 2008. A summative evaluation of RCT methodology and an alternative approach to causal research. *Journal of MultiDisciplinary Evaluation*, 5(9), pp.11–24.
Outcome harvesting	Wilson-Grau, R and Britt, H, 2012. *Outcome Harvesting*. New York: Ford Foundation.
Outcome mapping	Ambrose, K and Roduner, D, 2009. *A Conceptual Fusion of the Logical Framework Approach and Outcome Mapping*. Outcome Mapping Ideas Paper No. 1, May. Armstrong, J, Carden, F, Coe, A and Earl, S, 2000. International Model Forest Network Secretariat (IMFNS): Outcomes Assessment. Earl, S, Carden, F and Smutylo, T, 2001. *Outcome Mapping: Building Learning and Reflection into Development Programs*. Ottawa, Canada: International Development Research Centre. Earl, S and Carden, F, 2002. Learning from complexity: the International Development Research Centre's experience with outcome mapping. *Development in Practice*, 12(3/4), August. Howard, G, Jeger, M and Wilson-Grau, R, 2011. BioNET 2007–2010: An outcomes evaluation and assessment of the prospects for BioNET to increase its impact on food security, in particular through greater support to plant health systems. Available at : < http://webapp-hq.nl/node/1460> Jones, H and Hearn, S, 2009. *Outcome Mapping: A Realistic Alternative for Planning, Monitoring and Evaluation*. London: Overseas Development Institute (ODI) Background Note, October. Majot, J, Richert, W and Wilson-Grau, R, 2010. Evaluation of Oxfam Novib's Global Programme 2005–2008. Available at: <https://www.oxfamnovib.nl/Redactie/Downloads/Jaarverslagen/Landenevaluaties/GloPro_Eval_2005_2008_Summ.pdf> Outcome Mapping Learning Community website. Available at: <http://www.outcomemapping.ca/> Roduner, D, Schläppi, W and Egli, W, 2008. Logical framework approach and outcome mapping, a constructive attempt of synthesis. *Rural Development News*, 2, pp.1–24. Smutylo, T, 2005. *Outcome mapping: a method for tracking behavioural changes in development programs*. ILAC Brief 7, August.

Tool or concept	Resources
	Van Ongevalle, J, Chipimbi, R and Sibanda, M, n.d. *Analysing Outcome Mapping Monitoring Data: The Case of the Quality Education and Vulnerability Programme in Zimbabwe 2008–2013*. Working Paper. Outcome Mapping Learning Community. Available at: <http://www.outcomemapping.ca/> Van Ongevalle, J, Chipimbi, R and Sibanda, M, 2009. *Monitoring for Impact in a Programme's Sphere of Influence – a Case Study of the Quality Education and Vulnerability Programme in Zimbabwe*. Presented at the African Evaluation Association (AFREA) Conference, 29 March–3 April 2009, Cairo, Egypt. Available at: <http://www.outcomemapping.ca/>
Process tracing	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). Bennett, A, 2008. Process-tracing: a Bayesian perspective. In: JM Box-Steffensmeier, HE Brady and D Collier, eds. *The Oxford handbook of political methodology*. Oxford: Oxford University Press, pp.702–21. Bennett, A, 2010. Process tracing and causal inference. In: H Brady and D Collier, eds. *Rethinking social inquiry: diverse tools, shared standards*. Lanham, MD: Rowman and Littlefield. Checkel, JT, 2006. It's the process, stupid! Process tracing in the study of European and international politics. ARENA Centre for European Studies, University of Oslo. Available at: <http://www.sv.uio.no/arena/english/research/publications/arena-working-papers/2001-2010/2005/wp05_26.pdf> Checkel, JT, 2008. Tracing causal mechanisms. *International Studies Review*, 8(2), pp.362–70. Collier, D, 2011. Understanding process tracing. *Political Science and Politics*, 44(4), pp.823–30. George, AL and Bennett, A, 2005. *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press. Hughes, K and Hutchings, C, 2011. *Can we obtain the required rigour without randomisation? Oxfam GB's non-experimental global performance framework*. 3ie Working Paper 13. Available at: <http://3ieimpact.org/en/publications/working-papers/working-paper-13/> Reilly, R, 2010. Process tracing. In: AJ Mills, G Durepos and E Wiebe, eds. *Encyclopedia of case study research*. London: SAGE Publications Ltd. Wesleyan University, n.d. *A User's Guide to Political Science: Process Tracing*. Available at: <http://govthesis.site.wesleyan.edu/research/methods-and-analysis/analysing-qualitative-data/process-tracing/>
Strategic plausibility	Coe, J and Smith, J, 2015. Making a case for change: the value of strategic plausibility in evaluation. *openDemocracy*. Available at: <https://www.opendemocracy.net/openglobalrights/jeremy-smith-jim-coe/making-case-for-change-value-of-strategic-plausibility-in-eval>

Tool or concept	Resources
General advocacy evaluation resources	BetterEvaluation website. Available at: <http://betterevaluation.org> Center for Evaluation Innovation website. Available at: <www.evaluationinnovation.org>
Neural networks	Stergiou, C and Siganos, D, 2016. *Neural Networks*. London Imperial College. Available at: <https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html> [Accessed 13 May 2016]. University of Wisconsin-Madison, n.d. *A Basic Introduction to Neural Networks*. Available at: <http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html> [Accessed 10 May 2016].
Qualitative comparative analysis	Befani, B, 2016. *Pathways to change: evaluating development interventions with Qualitative Comparative Analysis (QCA)*. Stockholm: EBA. Marshall, G, 1998. Qualitative Comparative Analysis. *A Dictionary of Sociology*. Encyclopedia.com. Available at: <http://www.encyclopedia.com/social-sciences/dictionaries-thesauruses-pictures-and-press-releases/qualitative-comparative-analysis> [Accessed 26 November 2017]. Raab, M and Stuppert, W, 2014. *Review of evaluation approaches and methods for interventions related to violence against women and girls (VAWG)*. DFID. Available at: <https://assets.publishing.service.gov.uk/media/57a089b440f0b652dd00037e/61259-Raab_Stuppert_Report_VAWG_Evaluations_Review_DFID_20140626.pdf> [Accessed 10 November 2015]. Ragin, CC, 2008. *What is qualitative comparative analysis (QCA)?* Department of Sociology and Department of Political Science, University of Arizona. Available at: <http://eprints.ncrm.ac.uk/250/1/What_is_QCA.pdf>

Online appendixes

Online appendix B: Semi-structured interview guide for advocacy evaluation expert interviews

This appendix is only available online and can be accessed from http://www.3ieimpact.org/media/filer_public/2017/12/14/wp29-online-appendix-b.pdf

Online appendix C: List of key experts interviewed

This appendix is only available online and can be accessed from http://www.3ieimpact.org/media/filer_public/2017/12/14/wp29-online-appendix-c.pdf

Online appendix D: Coding sheet for eligible studies

This appendix is only available online and can be accessed from http://www.3ieimpact.org/media/filer_public/2017/12/14/wp29-online-appendix-d.pdf

Online appendix E: Inclusion and exclusion criteria for thematic evidence review

This appendix is only available online and can be accessed from http://www.3ieimpact.org/media/filer_public/2017/12/14/wp29-online-appendix-e.pdf

References

Alexander, J and Bonino, F, 2015. Addressing causation in humanitarian evaluation: a discussion on designs, approaches and examples. *ALNAP Discussion Series: Improving the Quality of EHA Evidence, Method Note,* 4 (January).

Allcott, H, 2011. Social norms and energy conservation. *Journal of Public Economics*, 95(9–10), pp.1082–95.

Ambrose, K and Roduner, D, 2009. *A Conceptual Fusion of the Logical Framework Approach and Outcome Mapping*. Outcome Mapping Ideas Paper No. 1, May.

Armstrong, J, Carden, F, Coe, A and Earl, S, 2000. International Model Forest Network Secretariat (IMFNS): Outcomes Assessment, July.

Ashraf, N, Kelsey, J, and Emir, K, 2013. Information and subsidies: complements or substitutes? *Journal of Economic Behavior & Organization* 88(2013), pp.133–39.

Banerjee, A, Kumar, S, Pande, R and Su, F, 2011. *Do Informed Voters Make Better Choices? Experimental Evidence from Urban India.* Working Paper (November).

Banerjee, A, Duflo, E, Imbert, C, Pande, R, Walton, M and Mohapatra, BP, 2014. *An impact evaluation of information disclosure on elected representatives' performance*. 3ie Impact Evaluation Report 11. New Delhi: International Initiative for Impact Evaluation (3ie).

Banerjee, AV, Banerji, R, Duflo, E, Glennerster, R and Khemani, S, 2010. Pitfalls of participatory programs: evidence from a randomized evaluation in India. *American Economic Journal: Economic Policy*, 2(1), pp.1–30.

Barooah, B, Kaushish, B and Puri, J, 2017. A story of bias: An assessment of the quality of evidence in agricultural risk management by smallholder farmers. Manuscript submitted for publication.

Beach, D and Pedersen, RB, 2013. *Process-tracing methods: foundations and guidelines.* Ann Arbor: University of Michigan Press.

Beer, T and Coffman, J, 2015. *Four Tools for Assessing Grantee Contribution to Advocacy Efforts*. Washington, DC: Center for Evaluation Innovation. Available at: <http://www.evaluationinnovation.org/publications/four-tools-assessing-grantee-contribution-advocacy-efforts>

Beer, T and Reed, ED, 2009. A model for multilevel advocacy evaluation. *The Foundation Review*, 1(3), p.149–161.

Beer, T, 2012*. Best Practices and Emerging Trends in Advocacy Grantmaking*. Washington, DC: Center for Evaluation Innovation*.*

Beer, T, Ingargiola, PS and Beer, MF, 2012. *Advocacy & public policy grantmaking: matching process to purpose*. Denver, CO: The Colorado Trust*.*

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).

Befani, B, 2016. *Pathways to change: evaluating development interventions with Qualitative Comparative Analysis (QCA)*. Stockholm: EBA.

Benneara, L, Tarozzi, A, Pfaff, A, Balasubramanya, S, Ahmed, KM and van Geen, A, 2013. Impact of a randomized controlled trial in arsenic risk communication on household water-source choices in Bangladesh. *Journal of Environmental Economics and Management*, 65(2), pp.225–40.

Bennett, A and Checke, JT, eds., 2015. *Process Tracing: From Metaphor to Analytic Tool*. Cambridge, UK: Cambridge University Press.

Bennett, A, 2008. Process-tracing: a Bayesian perspective. In: JM Box-Steffensmeier, HE Brady and D Collier, eds. *The Oxford handbook of political methodology*. Oxford: Oxford University Press, pp.702–21.

Bennett, A, 2010. Process tracing and causal inference. In: H Brady and D Collier, eds. *Rethinking social inquiry: diverse tools, shared standards*. Lanham, MD: Rowman and Littlefield.

BetterEvaluation, n.d.a *Approaches*. Available at <http://betterevaluation.org/approaches>

BetterEvaluation, n.d.b *Positive Deviance*. Available at: <http://www.betterevaluation.org/en/plan/approach/positive_deviance>

BetterEvaluation, 2014a. *Triangulation.* Available at: <http://betterevaluation.org/evaluation-options/triangulation>

BetterEvaluation, 2014b. *Episode Studies*. Available at: <http://betterevaluation.org/evaluation-options/episode_studies>

BetterEvaluation, 2014c. *Network Diagram*. Available at: <http://betterevaluation.org/evaluation-options/networkmapping>

BetterEvaluation, 2015. *After Action Review*. Available at: <http://betterevaluation.org/evaluation-options/after_action_review>

Bidwell, K, Casey, K and Glennerster, R, 2015. *Debates: The Impact of Voter Knowledge Initiatives in Sierra Leone.* Working Paper, April 2015.

Boaz, A, Fitzpatrick, S and Shaw, B, 2008. *Assessing the impact of research on policy: A review of the literature for a project on bridging research and policy through outcome evaluation*. King's College, University of London and Policy Studies Institute.

Bonell, CP, Hargreaves, J, Cousens, S, Ross, D, Hayes, R, Petticrew, M and Kirkwood, BR, 2011. Alternatives to randomisation in the evaluation of public health interventions: design challenges and solutions. *Journal of Epidemiology and Community Health*, 65(7), pp.582–7.

Borkum, E, Sivasankaran, N, Sridharan, S, Rotz, D, Sethi, S, Manoranjini, M, Ramakrishnan, L and Rangarajan, A, 2015. Evaluation of the Information and Communication Technology (ICT) Continuum of Care Services (CCS) in Bihar. Mathematica Policy Research, Princeton, NJ.

Bowen, HL, 2013. Impact of a mass media campaign on bed net use in Cameroon. *Malaria Journal*, 12(1).

California Endowment, 2009. Advocacy Evaluation Advances Convening Framing Paper. Available at: <http://www.pointk.org/resources/node/687>

Campbell, M and Coffman, J, 2009. Tools to support public policy grantmaking. *The Foundation Review*, 1(3).

Carden, F, 2009. Understanding influence: the episode studies approach. In: *The politics of trade: the role of research in trade policy and negotiation*. Ottawa, Canada: International Development Research Centre (IDRC). Ch.10.

Center for Evaluation Innovation. Available at: www.evaluationinnovation.org

Center for Evaluation Innovation, 2010. Advocacy Evaluation Update, Issue #9 (May). Washington, DC: Center for Evaluation Innovation.

Center for Evaluation Innovation, 2011. Advocacy Evaluation Update, Issue #12 (October). Washington, DC: Center for Evaluation Innovation.

Céspedes, J, Briceño, G, Farkouh, ME, Vedanthan, R, Baxter, J, Leal, M, Boffetta, P, Woodward, M, Hunn, M, Dennis, R and Fuster, V, 2013. Targeting preschool children to promote cardiovascular health: cluster randomized trial. *The American Journal of Medicine*, 126(1), pp.27–35.e3.

Chapman, J and Wameyo, A, 2001. Monitoring and Evaluating Advocacy: A Scoping Study. South Africa. ActionAid.

Chase, C and Do, Q-T, 2012. *Handwashing Behavior Change at Scale: Evidence from a Randomized Evaluation in Vietnam.* World Bank Policy Research Working Papers.

Checkel, JT, 2006. It's the process, stupid! Process tracing in the study of European and international politics. ARENA Centre for European Studies, University of Oslo. Available at: <http://www.sv.uio.no/arena/english/research/publications/arena-working-papers/2001-2010/2005/wp05_26.pdf>

Checkel, JT, 2008. Tracing causal mechanisms. *International Studies Review*, 8(2), pp.362–70.

Chong, A, Karlan, D, Shapiro, J and Zinman, J, 2011. 'Information campaign to reduce plastic bag consumption in supermarkets in Mexico'. Unpublished.

Chong, A, Karlan, D, Shapiro, J and Zinman, J, 2013. (Ineffective) Messages to encourage recycling: evidence from a randomized evaluation in Peru. *World Bank Economic Review*, 28(2), pp.1–27.

Chong, A, De La O, AL, Karlan, D and Wantchekon, L, 2015. Does corruption information inspire the fight or quash the hope? A field experiment in Mexico on voter turnout, choice, and party identification. *The Journal of Politics,* 77(1), pp.55–71.

Citrin, J, Green, D and Levy, M, 2014. The effects of voter ID notification on voter turnout: results from a large-scale field experiment. *Election Law Journal,* 13(2), pp.228–42.

Cluster Munition Coalition, n.d. *Cluster Munition Coalition*. Available at: <http://www.stopclustermunitions.org/en-gb/home.aspx>

Coates, B and David, R, 2002. Learning for change: the art of assessing the impact of advocacy work. *Development in Practice*, 12, pp.3–4*.*

Coe, J and Majot, J, 2013. Monitoring, Evaluation and Learning in NGO Advocacy: Findings from Comparative Policy Advocacy MEL Review Project. Oxfam America.

Coe, J and Mayne, R, 2008. *Is your campaign making a difference?* London: NCVO.

Coe, J and Schlangen, R, 2011. Looking through the right end of the telescope. *Advocacy Evaluation Update: Advocacy in Theory, Methods and Practice*, p. 1- 2. Washington, DC: Center for Evaluation Innovation.

Coe, J and Smith, J, 2015. Making a case for change: the value of strategic plausibility in evaluation. *openDemocracy*. Available at: <https://www.opendemocracy.net/openglobalrights/jeremy-smith-jim-coe/making-case-for-change-value-of-strategic-plausibility-in-eval>

Coffman, J, 2003. *Lessons in evaluating communications campaigns: five case studies*. Washington, DC: Communications Consortium Media Center and Harvard Family Research Project.

Coffman, J, 2007. What's different about evaluating advocacy and policy change? *The Evaluation Exchange*, 8, 2–4.

Coffman, J, 2009. *Overview of Current Advocacy Evaluation Practice.* Washington, DC: Center for Evaluation Innovation. Available at: <http://www.evaluationinnovation.org/sites/default/files/Coffman%20Brief%201.pdf>

Coffman, J, 2010. Monitoring and Evaluating Advocacy: Companion to the Advocacy Toolkit. New York. UNICEF.

Coffman, J, 2011. Revolutionary Lessons About Interim Outcomes. *Advocacy Evaluation Update*, 11 (August). Washington, DC: Center for Evaluation Innovation.

Coffman, J and Reed, E, 2009. *Unique Methods in Advocacy Evaluation*. The California Endowment. Available at: <http://www.innonet.org/resources/files/Unique_Methods_Brief.pdf>

Coffman, J, Hendricks, A, Kaye, J, Kelly, T and Masters, B, 2007. 'The Advocacy and Policy Change Composite Logic Model.' Advanced Practice Institute, conducted at the 58th Annual Conference of the Council on Foundations, Seattle, WA. Available at: <http://www.pointk.org/client_docs/File/advocacy/clm_final.ppt>

Cole, S, and Fernando, AN, 2012. *The Value of Advice: Evidence from Mobile Phone-Based Agricultural Extension*. Harvard Business School Working Paper (November). Cambridge, MA.

Collier, D, 2011. Understanding process tracing. *Political Science and Politics,* 44(4), pp.823–30.

Collison, C and Parcell, G, 2001. *Learning to fly*. Oxford: Capstone.

Creel, AH, Rimal, RN, Mkandawire, G, Böse, K and Brown, JW, 2011. Effects of a mass media intervention on HIV-related stigma: 'Radio Diaries' program in Malawi. *Health Education Research*, 26(3), pp.456–65.

Christiansen, K and Hovland, I, 2003. *The PRSP initiative: multilateral policy change and the role of research*. London: Overseas Development Institute (ODI).

Cugelman, B and Otero, E, 2010. Evaluating Complex Campaigns: Oxfam GB's Climate Change Campaign. Available at: <https://www.interaction.org/sites/default/files/BrianCugelmanEvaOtero.pdf>

Cummins, J, 2015. *Ten things we learnt about behaviour change and sustainability*. Available at: <https://www.theguardian.com/sustainable-business/behavioural-insights/behaviour-change-sustainability-debate>

Cutler, DM, 2004. Behavioral health interventions: what works and why? Critical perspectives on racial and ethnic differences. In: NB Anderson, RA Bulatao and B Cohen, eds. *Health in late life: National Research Council (US) panel on race, ethnicity, and health in later life*. Washington, DC: National Academies Press. Available at: <https://www.ncbi.nlm.nih.gov/books/NBK25527/>

Dart, J and Davies, R, 2003. A dialogical, story-based evaluation tool: the most significant change technique. *American Journal of Evaluation*, 24(2), pp.137–55.

Dart, J, 2008. Participatory Performance Story Reporting. Clear Horizon. Available at: <http://www.innonet.org/resources/node/438?referer=L3Jlc291cmNlcy9zZWFyY2gvcmVzdWx0cz9tb2RlPWJyb3dzZSZjYXRlZ29yeT00NyZmaWx0ZXIwPSZmaWx0ZXIxPTQ3JnZpZXdtb2RlPXN1bW1hcnk=>

Dart, JJ and Mayne, J, 2005. Performance story. In: S Mathison, ed. *Encyclopedia of evaluation*. Thousand Oaks, CA: Sage Publications, pp.307–08.

Davidson, EJ, 2005. *Evaluation methodology basics.* Thousand Oaks, CA: SAGE Publications.

Devlin-Foltz, D and Molinaro, L, 2010. Champion and Champion-ness: Measuring Efforts to Create Champions for Policy Change. Washington, DC: Center for Evaluation Innovation.

Dinkelman, T and Martinez, C, 2014. Investing in schooling in Chile: the role of information about financial aid for higher education. *The Review of Economics and Statistics*, 96(2), pp.244–57.

Downes, S, 2014. Field of themes: five early lessons from an innovative advocacy approach. *National Civic Review* (Winter).

Duflo, E, Dupas, P, Seban, J and Huillery, E, 2012. *Impacts of School-Based HIV Education on Reported Behavior and Knowledge of Adolescent Girls, Evidence from Cameroon*.Sciences Po.

Dupas, P, 2011. Do teenagers respond to HIV risk information? Evidence from a field experiment in Kenya. *American Economic Journal: Applied Economics,* 3(1), pp.1–34.

Durland, M and Fredericks, K, eds., 2005. *New Directions for Evaluation: Special Issue: Social Network Analysis in Program Evaluation*, 2005(107), pp.1–101.

Earl, S and Carden, F, 2002. Learning from complexity: the International Development Research Centre's experience with outcome mapping. *Development in Practice*, 12(3/4), August.

Earl, S, Carden, F and Smutylo, T, 2001. *Outcome Mapping: Building Learning and Reflection into Development Programs*. Ottawa, Canada: International Development Research Centre.

Epstein, AI, Tumbo, J, Moriaty, K, Vorley, K and di Marco, A, 2014. Formative Evaluation of The Children's Agenda in Tanzania (2010–2013) Final Report. CfTB Education Trust to the United Nations Children's Fund (UNICEF) in Tanzania.

Espinoza-Gómez, F, Hernández-Suárez, CM and Coll-Cárdenas, R, 2002. Educational campaign versus malathion spraying for the control of *Aedes aegypti* in Colima, Mexico. *Journal of Epidemiology and Public Health*, 56(2), pp.148–52.

Essien, EJ, Mgbere, O, Monjok, E, Ekong, E, Holstad, MM and Kalichmanf, SC, 2011. Effectiveness of a video-based motivational skills-building HIV risk-reduction intervention for female military personnel. *Social Science & Medicine*, 72(1), pp.63–71.

FatheIrahman, AI, Omar, M, Awang, R, Cummings, MK, Borland, R and Mohd Samin, AS, 2010. Impact of the new Malaysian cigarette pack warnings on smokers' awareness of health risks and interest in quitting smoking. *International Journal of Environmental Research and Public Health*, 7, pp.4089–99.

Ferraz, C and Finan, F, 2008. Exposing corrupt politicians: the effects of Brazil's publicly released audits of electoral outcomes. *The Quarterly Journal of Economics,* 123(2), pp.703–44.

Frölich, M and Vazquez-Alvarez, R, 2009. HIV/AIDS knowledge and behavior: have information campaigns reduced HIV infection? The case of Kenya. *African Development Review*, 21(1), pp.86–146.

Fryer Jr., RG, 2013. *Information and Student Achievement: Evidence from a Cellular Phone Experiment*. NBER Working Paper #19113 (June).

Galiani, S, Gertler, P and Orsola-Vidal, A, 2012. *Promoting Handwashing Behavior in Peru: The Effect of Large-Scale Mass-Media and Community Level Interventions*. Policy

Research Working Paper No. 6257. Washington, DC: World Bank. Available at: <https://openknowledge.worldbank.org/handle/10986/19928>

Gardner, A and Brindis, C, 2017. *Advocacy and policy change evaluation: theory and practice*. Stanford, CA: Stanford University Press.

Gardner, A, Marshall, N, Geierstanger, S and Quinn, H, 2008. Achieving a Policy Change: Key Strategies and Factors for Success. Creating a Legacy for Change. Clinic Consortia Policy and Advocacy Program Evaluation.

George, AL and Bennett, A, 2005. *Case studies and theory development in the social sciences*. Cambridge, MA: MIT Press.

Gerber, AS and Green, DP, 2005. Do phone calls increase voter turnout? An update. *The Annals of the American Academy of Political and Social Science,* 601(142), pp.142–54.

Gerber, AS, Green, DP and Larimer, CW, 2008. Social pressure and voter turnout: evidence from a large-scale field experiment. *The American Political Science Review,* 102(1), pp.33–48.

Gienapp, A and Cohen, C, 2011. *Advocacy Evaluation Case Study: The Chalkboard Project*. Washington, DC: Center for Evaluation Innovation.

Gillespie, S, Haddad, L, Mannar, V, Menon, P, Nisbett, N and Maternal and Child Nutrition Study Group, 2013. The politics of reducing malnutrition: building commitment and accelerating progress. *The Lancet*, *382*(9891), pp.552–69.

Godlonton, S, Munthali, A and Thornton, R, 2015. Responding to risk: circumcision, information, and HIV prevention. *The Review of Economics and Statistics*, 98(2), pp.333–49.

Green, D and Vasudevan, S, 2015. *The Impact of a Radio Campaign against Vote-Buying: Evidence from Field Experiment Conducted during the 2014 Indian General Election*. Working Paper.

Guiteras, R, Jannat, K, Levine, DI and Polley, T, 2015. *Testing disgust- and shame-based safe water and handwashing promotion in urban Dhaka, Bangladesh*. 3ie Grantee Final Report. Delhi: International Initiative for Impact Evaluation (3ie). Available at: <http://3ieimpact.org/en/evidence/impact-evaluations/details/217/>

Guthrie, K, Louie, J, David, T and Foster, CC, 2005. *The Challenge of Assessing Advocacy: Strategies for a Prospective Approach to Evaluating Policy Change and Advocacy*. The California Endowment.

Howard, G, Jeger, M and Wilson-Grau, R, 2011. BioNET 2007–2010: An outcomes evaluation and assessment of the prospects for BioNET to increase its impact on food security, in particular through greater support to plant health systems. Available at: <https://www.outcomemapping.ca/download/BioNET-Global-Programme-Evaluation.pdf>

Hudson, A, 2002. Advocacy by UK-based development NGOs. *Nonprofit and Voluntary Sector Quarterly*, 31, p.402. Available at: <http://nvs.sagepub.com/content/31/3/402>

Hughes, K and Hutchings, C, 2011. *Can we obtain the required rigour without randomisation? Oxfam GB's non-experimental global performance framework*. 3ie Working Paper 13. Available at: <http://3ieimpact.org/en/publications/working-papers/working-paper-13/>

Humphreys, M and Weinstein, JM, 2012. *Policing Politicians: Citizen Empowerment and Political Accountability in Uganda.* International Growth Centre Working Paper.

Hutchinson, P and Meekers, D, 2012. Estimating causal effects from family planning health communication campaigns using panel data: the 'Your Health Your Wealth' campaign in Egypt. *PLOS One*, 7(9).

International Institute for Environment and Development (IIED), 2017. Process tracing with Bayesian updating. Last modified March 2017. Available at: <http://pubs.iied.org/pdfs/17402IIED.pdf>

Innovation Network, 2007. Advocacy Evaluation Update, 2 (November) Washington, DC: Center for Evaluation Innovation be added?

Innovation Network, n.d. 'Composite Logic Model' (including supplemental materials). Available at: <http://www.pointk.org/index.php?section_id=6&content_id=637>

International Initiative for Impact Evaluation (3ie), 2012. 'Principles for Impact Evaluation.' Available at: <http://www.3ieimpact.org/media/filer_public/2014/01/15/principles_for_impact_evaluation.pdf>

Ishola, AG and Chipps, J, 2015. The use of mobile phones to deliver acceptance and commitment therapy in the prevention of mother–child HIV transmission in Nigeria. *Journal of Telemedicine and Telecare*, 21(8), pp.423–26.

James, S, Reddy, PS, Ruiter, RA, Taylor, M, Jinabhai, CC, Van Empelen, P and Van den Borne, B, 2005. The effects of a systematically developed photo-novella on knowledge, attitudes, communication and behavioural intentions with respect to sexually transmitted infections among secondary school learners in South Africa. *Health Promotion International*, 20(2), pp.157–65.

Jamison, J, Karlan, D and Raffler, P, 2013. Mixed method evaluation of a passive mHealth sexual information texting service in Uganda. *Information Technologies & International Development*, 9(3).

Jensen, R, 2010. Impact of information on the returns to education on the demand for schooling in the Dominican Republic. *Quarterly Journal of Economics*, 125, pp.515–48.

Jimenez, E and Puri, J, 2017. The 'wicked' cases of education and climate change: the promise and challenge of theory-based impact evaluations. In: RD van den Berg, I Naidoo, and SD Tamondong, eds. *Evaluation for agenda 2030: providing evidence on progress and sustainability*. Exeter, UK: IDEAS.

Johns Hopkins Bloomberg School of Public Health's Center for Communication Programs, 1999. 'A' Frame for Advocacy. Johns Hopkins Center for Communication Programs and USAID. Available at: <http://www.innonet.org/resources/node/61>

Jones, H and Hearn, S, 2009. *Outcome Mapping: A Realistic Alternative for Planning, Monitoring and Evaluation*. London: Overseas Development Institute (ODI) Background Note, October.

Jones, H, 2011. *A Guide to Monitoring and Evaluating Policy Influence*. London: ODI Background Note.

Kahneman, D, Knetsch, JL and Thaler, RH, 1991. Anomalies: the endowment effect, loss aversion, and status quo bias. *Journal of Economic Perspectives*, 5(1), pp.193–206.

Kalichman, SC, Simbayi, LC, Vermaak, R, Cain, D, Smith, G, Mthebu, J and Jooste, S, 2008. Randomized trial of a community-based alcohol-related HIV risk-reduction intervention for men and women in Cape Town, South Africa. *Annals of Behavioral Medicine*, 36(3), pp.270–9.

Karlan, D and Wood, DH, 2015. *The Effect of Effectiveness: Donor Response to Aid Effectiveness in a Direct Mail Fundraising Experiment*. Yale University Economic Growth Center Discussion Paper 1038.

Kazemi, A, Ehsanpour, S and Nekoei-Zahraei, NS, 2012. A randomized trial to promote health belief and to reduce environmental tobacco smoke exposure to pregnant women. *Health Education Research*, 27(1), pp.151–9.

Keating, J, Meekers, D and Adewuyi, A, 2006. Assessing effects of a media campaign on HIV/AIDS awareness and prevention in Nigeria: results from the VISION Project. *BMC Public Health*, 6(123).

Kellstedt, P and Whitten, G, 2013. *The fundamentals of political science research, Second Edition*. Cambridge University Press, MA.

Klugman, B, 2010*. Evaluating Social Justice Advocacy: A Values Based Approach*. Washington, DC: Center for Evaluation Innovation.

Kotvojs, F, 2006. *Contribution Analysis: A New Approach to Evaluation in International Development*. Paper presented at the Australian Evaluation Society 2006 International Conference, Darwin, Australia.

Kremer, M, Miguel, E and Thornton, R, 2009. Incentives to learn. *Review of Economics and Statistics*, 91(3), pp.437–56.

Larbi, G, Christensen, J, Jackson, P and Ura, K, 2006. *Capacity Development in Bhutan: Capacity Development Outcome Evaluation of Danish Supported Organisations in Bhutan*. Copenhagen: Danida. Available at: <http://www.oecd.org/dataoecd/45/9/42218416.pdf>

Lemire, S, 2010. *Contribution Analysis: The Promising New Approach to Causal Claims.* Paper presented at European Evaluation Society International Conference, Prague 2010. Available at: <https://www.alnap.org/help-library/contribution-analysis-the-promising-new-approach-to-causal-claims>

Lempert, T, 2009. *An Advocate's Perspective: Ten Considerations for Effective Advocacy Evaluation*. Children Now. Available at: <http://www.innonet.org/resources/files/Lempert_10_Considerations.pdf>

Lindquist, E, 2001. 'Discerning Policy Influence: Framework for a Strategic Evaluation of IDRC-Supported Research.' School of Public Administration, University of Victoria.

Lloyd Laney, M, 2003. 'Advocacy Impact Assessment Guidelines.' Communications and Information Management Resource Centre (CIMRC). Available at: <http://www.innonet.org/resources/files/Advocacy_Impact_Assessment_Guidelines.pdf>

Luo, R, Shi, Y, Zhang, L, Zhang, H, Miller G, Medina, A and Rozelle, S, 2012. The limits of health and nutrition education: evidence from three randomized-controlled trials in rural China. *CESifo Economic Studies*, 58.

Lynn, J, 2014. 'Assessing and Evaluating Change in Advocacy Fields.' Washington, DC: Center for Evaluation Innovation. Available at: <http://www.innonet.org/resources/files/Spark-Evaluating_Change_In_Advocacy_Fields.pdf>

Majot, J, Richert, W and Wilson-Grau, R, 2010. Evaluation of Oxfam Novib's Global Programme 2005–2008.

Mantzaria, E, Vogt, F, Shemilt, I, Weid, Y, Higgins, JP and Marteau, TM, 2015. Personal financial incentives for changing habitual health-related behaviors: A systematic review and meta-analysis. *Preventive Medicine*, 75(June), pp.75–85. Marshall, G, 1998. Qualitative Comparative Analysis. *A Dictionary of Sociology*. Encyclopedia.com. Available at: http://www.encyclopedia.com/social-sciences/dictionaries-thesauruses-pictures-and-press-releases/qualitative-comparative-analysis [Accessed: 26 November 2017].

Martin, A, Culey, C and Evans, S, 2005. Make Poverty History Campaign Evaluation. Firetail. Available at: <https://www.oecd.org/dev/pgd/38405846.pps>

Masters, B, 2009. Evaluating Policy Change and Advocacy: The Funder's Perspective. Washington, DC: Center for Evaluation Innovation.

Mayne, J, 2001. Addressing attribution through contribution analysis: using performance measures sensibly. *Canadian Journal of Programme Evaluation*, 16, pp.1–24.

Mayne, J, 2008. *Contribution Analysis: An Approach to Exploring Cause and Effect*. Institutional Learning and Change Brief No. 7. Rome, Italy: Institutional Learning and Change Initiative. Available at: <https://www.betterevaluation.org/resources/guides/contribution_analysis/ilac_brief>

Mayne, J, 2011. Contribution analysis: addressing cause and effect. In: K Forss, M Marra and R Schwartz, eds. *Evaluating the complex*. New Brunswick, NJ: Transaction Publishers. pp.53–96.

Mayoux, L, 2003. *Advocacy for Poverty Eradication and Empowerment: Ways forward for Advocacy Impact Assessment*. WISE Development Limited. Available at: <http://www.innonet.org/resources/node/73>

McConnell, M, 2012. *Between Intention and Action: An Experiment on Individual Savings*. J-PAL Working Paper.

Mill, JS, 1893. *A system of logic, ratiocinative and inductive: Being a connected view of the principles of evidence and the methods of scientific investigation*. New York: Harper & Brothers.

Miller, G, Luo, R, Zhang, L, Sylvia, S, Shi, Y, Foo, P, Zhao, Q, Martorell, R, Medina, A, Rozelle, S and Farnsworth, HF, 2012. Effectiveness of provider incentives for anaemia reduction in rural China: a cluster randomized trial. *British Medical Journal*, 345, p.4809.

Miranda, J, Sabet, S and Brown, A, 2016. Is impact evaluation still on the rise? 11 August 2016. *Evidence Matters: Improving development policy and practice: Blog*. Available at: <http://blogs.3ieimpact.org/is-impact-evaluation-still-on-the-rise/>

NIMH Collaborative HIV/STD Prevention Trial Group, 2010. Results of the NIMH collaborative HIV/sexually transmitted disease prevention trial of a community popular opinion leader. *Journal of Acquired Immune Deficiency Syndromes*, 54(2).

O'Neil, G and Goldschmid, P, 2013. Oxfam's GROW Campaign: Mid-point External Evaluation. Oxfam Wise Research and Evaluation (Owl RE). Available at: <https://www.oxfam.org/en/research/oxfams-grow-campaign-mid-point-external-evaluation>

O'Neil, G, Garay, J, Goldschmid, P, Lienart, L, Roper, L, Ruano, GE, Sarma, N and Schaepers, AP, 2015. *Oxfam's Global Leaders Empowered to Alleviate Poverty (LEAP) Evaluation – Final Report*. Oxfam America Evaluation Report. Available at: <https://www.oxfamamerica.org/static/media/files/LEAP_Final_Evaluation.pdf>

Oreopoulos, P and Dunn, R, 2013. Information and college access: evidence from a randomized field experiment. *The Scandinavian Journal of Economics*, 115(1), pp.3–26.

Outcome Mapping Learning Community. Available at: <http://www.outcomemapping.ca/>

Oxfam Great Britain, 2012. 'Influencing of Policy and Public Management Program: Impact Assessment Summary Report: Reducing Poverty by Improving Public Policy and Public Administration in Bolivia.'

Oxfam Great Britain, 2013. 'Promoting Women's Rights across Africa: Raising Her Voice – Pan Africa Effectiveness Review.'

Paden, CM, 2011. *Civil rights advocacy on behalf of the poor*. Philadelphia, PA: University of Pennsylvania Press.

Padwe, J, 2015. 'Bringing Innovation to Scale: A Synthesis of Learning from Four Projects of the Economic Innovations Incentives Fund.' Oxfam America Evaluation Report.

Paluck, EL, 2009. Reducing intergroup prejudice and conflict using the media: a field experiment in Rwanda. *Journal of Perspectives in Social Psychology*, 96(3).

Panagopoulos, C, and Green, DP. 2011. Spanish-language radio advertisements and Latino voter turnout in the 2006 congressional elections: field experimental evidence. *Political Research Quarterly,* 64(3), pp.588–99.

Pande, R, 2011. Can informed voters enforce better governance? Experiments in low-income democracies. *Annual Review of Economics*, 3, pp.215–37.

Pandey, P, Goyal, S and Sundararaman, V, 2009. Community participation in public schools: impact of information campaigns in three Indian states. *Education Economics,* 17(3), pp.355–75.

Patil, SR, Arnold, BF, Salvatore, AL, Briceno, B, Ganguly, S, Colford Jr, JM and Gertler, PJ, 2014. The effect of India's total sanitation campaign on defecation behaviors and child health in rural Madhya Pradesh: a cluster randomized controlled trial. *PLOS Medicine*, 11(8).

Pattanayak, S, Dickinson, KL, Yang, J-C, Patil, SR, Praharaj, P and Poulos, C, 2007. *Promoting Latrine Use: Midline Findings from a Randomized Evaluation of a Community Mobilization Campaign in Bhadrak, Orissa*. Research Triangle Institute Working Paper, No. 07_02.

Patton, MQ, 2007. *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: SAGE Publications Ltd.

Patton, MQ, 2008. Advocacy impact evaluation. *Journal of MultiDisciplinary Evaluation*, 5(9), March.

Patton, MQ, 2011. *Developmental evaluation*. New York: The Guilford Press.

Pelletier, D, Haider, R, Hajeebhoy, N, Mangasaryan, N, Mwadime, R and Sarkar, S, 2013. The principles and practices of nutrition advocacy: evidence, experience and the way forward for stunting reduction. *Maternal and Child Nutrition*, 9(S2), pp.83–100.

Pellini, A, 2013. '52 weeks of BetterEvaluation: Week 33: Monitoring policy influence part 2 - like measuring smoke?' (August). Available at: <http://betterevaluation.org/blog/m-and-e-policy-influence>

Pritchett, L, 2013. Impact Evaluation as a Learning Tool for Development Effectiveness. IDB. Available at: <https://www.cgdev.org/sites/default/files/Impact%20Evaluation%20as%20a%20LearningTool%20for%20Development%20Effectiveness%20Pritchett.pdf>

Raab, M and Stuppert, W, 2014. Review of evaluation approaches and methods for interventions related to violence against women and girls (VAWG). DFID. Available at: <https://assets.publishing.service.gov.uk/media/57a089b440f0b652dd00037e/61259-Raab_Stuppert_Report_VAWG_Evaluations_Review_DFID_20140626.pdf> [Accessed 10 November 2015].

Ragin, CC, 2008. *What is qualitative comparative analysis (QCA)?* Department of Sociology and Department of Political Science, University of Arizona. Available at: <http://eprints.ncrm.ac.uk/250/1/What_is_QCA.pdf>

Ramalingam, B, 2006. *Tools for Knowledge and Learning*. ODI Toolkit. London: Overseas Development Institute (ODI).

Rankin, S, 2013. *Extractive Industries Program in Cambodia (2009–2012)*. Oxfam America Evaluation Report.

Redstone Strategy Group, 2013. *Helping Think Tanks Measure Impact*. Available at: <http://www.redstonestrategy.com/wp-content/uploads/2013/09/2013-09-30-IDRC-Helping-think-tanks-measure-impact.pdf>

Reilly, R, 2010. Process tracing. In: AJ Mills, G Durepos and E Wiebe, eds. *Encyclopedia of case study research*. London: SAGE Publications Ltd.

Reinikka, R and Svensson, J, 2004. Fighting corruption to improve schooling: evidence from a newspaper campaign in Uganda. *Journal of the European Economic Association,* 3(2/3).

Reisman, J, Gienapp, A and Stachowiak, S, 2007.
'A Guide to Measuring Advocacy and Policy.' Organizational Research Services. Available at: <http://www.aecf.org/m/resourcedoc/aecf-aguidetomeasuringpolicyandadvocacy-2007.pdf>

Roduner, D, Schläppi, W and Egli, W, 2008. Logical framework approach and outcome mapping: a constructive attempt of synthesis. *Rural Development News*, 2, pp.1–24.

Rogers, P, 2014. *Overview: Strategies for Causal Attribution*. Methodological Briefs: Impact Evaluation No. 6. Florence, Italy: UNICEF Office of Research.

Roper, L, 2011. 'Summary of Two Evaluations of the Program to Prevent Gender Violence in El Salvador.' Oxfam America Evaluation Report.

Schlangen, R, 2014. 'Monitoring and Evaluation for Human Rights Organizations: Three Case Studies.' Washington, DC: Center for Evaluation Innovation.

Scriven, M, 1976. Maximising the power of causal investigations: the Modus Operandi method. In: GV Glass, ed. *Evaluation studies review annual*, Vol. 1. London: SAGE Publications Ltd.

Scriven, M, 2005. *Can We Infer Causation from Cross-sectional Data?* Paper presented at a Symposium on the Use of School-Level Data for Evaluating Federal Education Programmes, Washington, DC.

Scriven, M, 2008. A summative evaluation of RCT methodology and an alternative approach to causal research. *Journal of MultiDisciplinary Evaluation*, 5(9), pp.11–24

Smiltneek, E, n.d. *Suffrage Movement.* Learning to Give. Available at: <https://www.learningtogive.org/resources/suffrage-movement>

Smith, EA, Palen, L-A, Caldwell, LL, Flisher, AJ, Graham, JW, Mathews, C, Wegner, L and Vergnani, T, 2008. Substance use and sexual risk prevention in Cape Town, South Africa: an evaluation of the HealthWise program. *Prevention Science*, 9(4), pp.311–21.

Smutylo, T, 2005. *Outcome mapping: a method for tracking behavioural changes in development programs*. ILAC Brief 7, August.

Stachowiak, S, 2013. *Pathways for Change: 10 Theories to Inform Advocacy and Policy Change Efforts*. Seattle, WA, and Washington, DC: ORS Impact and Center for Evaluation Innovation.

Start, D and Hovland, I, 2004. Tools for Policy Impact: A Handbook for Researchers. Research in Policy and Development/ODI.

Stedman-Bryce, G, 2013. *Effectiveness Review: Health for All: Towards Free Universal Health Care in Ghana*. Oxfam Great Britain.

Stergiou, C and Siganos, D, 2016. *Neural Networks*. London Imperial College. Available at: <https://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html> [Accessed 13 May 2016].

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.* DFID Working Paper 38. Available at: <http://r4d.dfid.gov.uk/pdf/outputs/misc_infocomm/DFIDWorkingPaper38.pdf>

Swanborn, P, 2010. *Case study research: what, why and how?* London: SAGE Publications Ltd.

Tearfund, 2002. 'Advocacy Toolkit: Practical Action in Advocacy, Section C5 Evaluation.' Available at: <http://www.innonet.org/resources/files/PracticalEvaluation.pdf>

Teles, S and Schmitt, M, 2011. The Elusive Craft of Evaluating Advocacy. *Stanford Social Innovation Review* in May 2011: Retrieved from http://www.ssireview.org/

Thornton, R, Godlonton, S, Chinkhumba, J and Pierotti, R, 2014. *Scaling up male circumcision service provision: results from a randomised evaluation in Malawi.* 3ie Impact Evaluation Report 13. New Delhi: International Initiative for Impact Evaluation (3ie). Available at: <http://www.3ieimpact.org/en/publications/3ie-impact-evaluation-reports/3ie-impact-evaluations/3ie-impact-evaluation-report-13/>

Tibbett, S, 2010. 'Food for Thought: Save the Children's Influencing of the UK Department for International Development on Nutrition – Evidence and Lessons.' *Save the Children.*

Tibbett, S and Stalker, C, 2014. 'Enough Food for Everyone: IF Campaign Evaluation.' *The Advocacy Hub.*

Tiwari, A, Fong, DYT, Yuen, KH, Yuk, H, Pang, P, Humphreys, J and Bullock, L, 2010. Effect of an advocacy intervention on mental health in Chinese women survivors of intimate partner violence: a randomized controlled trial. *JAMA*, 304(5).

Tsui, J, 2014. Week 18: is there a 'right' approach to establishing causation in advocacy evaluation? 6 May 2014. *Josephine Tsui: Blog.* Available at: <http://betterevaluation.org/blog/m-and-e-advocacy-eval>

Tsui, J, Hearn, S and Young, J, 2014. *Monitoring and evaluation of policy influence and advocacy*. ODI Working Paper 395.

UNICEF, 2010. Advocacy toolkit. *A guide to influencing decisions that improve children's lives*. Available at: <https://www.unicef.org/evaluation/files/Advocacy_Toolkit.pdf>

UK Collaborative on Development Sciences (UKCDS), n.d. *Evaluating the Impact of Research Programmes – Approaches and Methods*. Available at: <http://www.ukcds.org.uk/sites/default/files/content/resources/Evaluating%20the%20Impact%20of%20Research%20Programmes%20-%20Approaches%20and%20Methods.pdf>

UN OCHA, n.d. What is the Cluster Approach? *Humanitarian Response*. Available at: <https://www.humanitarianresponse.info/en/about-clusters/what-is-the-cluster-approach>

United States Agency for International Development (USAID), 2016. Development evaluation for USAID. Available at: <https://socialimpact.com/wp-content/uploads/2017/06/DEPA_MERL-web-Lookbook.pdf>

University of Wisconsin-Madison, n.d. *A Basic Introduction to Neural Networks*. Available at: <http://pages.cs.wisc.edu/~bolo/shipyard/neural/local.html> [Accessed 10 May 2016].

Van Ongevalle, J, Chipimbi, R and Sibanda, M, 2009. *Monitoring for Impact in a Programme's Sphere of Influence – a Case Study of the Quality Education and Vulnerability Programme in Zimbabwe*. Presented at the African Evaluation Association (AFREA) Conference, 29 March–3 April 2009, Cairo, Egypt. Available at: <http://www.outcomemapping.ca/>

Van Ongevalle, J, Chipimbi, R and Sibanda, M, n.d. *Analysing Outcome Mapping Monitoring Data: The Case of the Quality Education and Vulnerability Programme in Zimbabwe 2008–2013*. Working Paper. Available at: <http://www.outcomemapping.ca/>

Vivalt, E, 2017 How Much Can We Generalize from Impact Evaluations? (New York University, under review). Available at: <http://evavivalt.com/wp-content/uploads/2014/12/Vivalt_JMP_latest.pdf>

Wantchekon, L and Fujiwara, T, 2013. Can informed public deliberation overcome clientelism? Experimental evidence from Benin. *American Economic Journal: Applied Economics,* 5(4), pp.241–55.

WaterAid, 2001. 'Advocacy: What's it all about? A guide to advocacy in the water and sanitation sector.' Available at: <http://www.wsp.org/Hygiene-Sanitation-Water-Toolkit/Resources/Readings/WaterAidAdvocacy-all-sections.pdf>

WaterAid, n.d. 'The Advocacy Sourcebook.' Available at: <https://www.wateraid.org/~/media/Publications/advocacy-sourcebook.ashx>

Wen, X, Chen, W, Gans, K, Colby, S, Ciyong, L, Liang, C and Ling, W, 2010. Two-year effects of a school-based prevention programme on adolescent cigarette smoking in Guangzhou, China: a cluster randomized trial. *International Journal of Epidemiology*, 39(3), pp.860–76.

Wesleyan University, n.d. *A User's Guide to Political Science: Process Tracing*. Available at: <http://govthesis.site.wesleyan.edu/research/methods-and-analysis/analyzing-qualitative-data/process-tracing/>

Westin, RA, 2012. Americans' unwillingness to pay taxes before the American Revolution: an uncomfortable legacy. *J. Juris*, 13, p.11.

Whelan, J, 2006. *Assessing Advocacy: Extracts from 'Work Justice' Case Study.* The Change Agency. Available at: <http://www.pointk.org/resources/files/assessing_advocacy.pdf>

Whelan, J, 2008. *Advocacy Evaluation Review and Opportunities*. The Change Agency. Available at: <http://www.pointk.org/resources/files/Whelan_Advocacy_Evaluation.pdf>

Whiffen, P, 2001. Seizing Learning Opportunities at Tearfund, *Knowledge Management Review 4*, December. Available at: <www.library.nhs.uk/knowledgemanagement>

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

White, H and Sabarwal, S, 2014a. *Modelling*. Methodological Briefs: Impact Evaluation No. 13. Italy: UNICEF Office of Research.

White, H and Sabarwal, S, 2014b. *Quasi-experimental Design and Methods*. Methodological Briefs: Impact Evaluation No. 8. Italy: UNICEF Office of Research.

Whitehurst, GJ, Stuit, D, Graves, C and Doan, S, 2013. Measuring the Influence of Education Advocacy: The Case of Louisiana's School Choice Legislation. Brown Center on Education Policy at Brookings.

Wilson-Grau, R and Britt, H, 2012. *Outcome Harvesting*. Ford Foundation.

Wu, Z, Luo, W, Sullivan, S, Rou, K, Lin, P, Sgdsg, F and Ming, Z, 2007. Evaluation of a needle social marketing strategy to control HIV among injecting drug users in China. *AIDS*, 21, Suppl. 8, pp.S115–22.

Yazdani, R, Vehkalahti, MM, Nouri, M, and Murtomaa, H, 2009. School-based education to improve oral cleanliness and gingival health in adolescents in Tehran, Iran. *International Journal of Paediatric Dentistry*, 19(4), pp.274–81.

Young, SD, Konda, K, Caceres, C, Galea, J, Sung-Jae, L, Salazar, X and Coates, T, 2011. Effect of a community popular opinion leader HIV/STI intervention on stigma in urban, coastal Peru. *AIDS and Behavior*, 15(5), pp.930–7.

Zandniapour, L and Brennan, K, 2010. Advocacy evaluation and rigor: how to be methodically rigorous in the field of advocacy evaluation. Advocacy Evaluation Update, 10. Washington, DC: Center for Evaluation Innovation, p.1.

Other publications in the 3ie working paper series

The following papers are available from
http://www.3ieimpact.org/en/publications/workingpapers/

3ie evidence gap maps: a starting point for strategic evidence production and use, 3ie Working Paper 28. Snilstveit, B, Bhatia, R, Rankin, K and Leach, B (2017)

Examining the evidence on the effectiveness of India's rural employment guarantee act, *3ie Working Paper 27*. Bhatia, R, Chinoy, SL, Kaushish, B, Puri, J, Chahar, VS and Waddington, H (2016)

Power calculation for causal inference in social science: sample size and minimum detectable effect determination, 3ie Working Paper 26. Djimeu, EW and Houndolo, DG (2016)

Evaluations with impact: decision-focused impact evaluation as a practical policymaking tool, 3ie Working Paper 25. Shah, NB, Wang, P, Fraker, A and Gastfriend, D (2015)

Impact evaluation and policy decisions: where are we? A Latin American think-tank perspective, 3ie Working Paper 24. Baanante, MJ and Valdivia, LA (2015)

What methods may be used in impact evaluations of humanitarian assistance? 3ie Working Paper 22. Puri, J, Aladysheva, A, Iversen, V, Ghorpade, Y and Brück, T (2014)

Impact evaluation of development programmes: experiences from Viet Nam, 3ie Working Paper 21. Nguyen Viet Cuong (2014)

Quality education for all children? What works in education in developing countries, 3ie Working Paper 20. Krishnaratne, S, White, H and Carpenter, E (2013)

Promoting commitment to evaluate, 3ie Working Paper 19. Székely, M (2013)

Building on what works: commitment to evaluation (c2e) indicator, 3ie Working Paper 18. Levine, CJ and Chapoy, C (2013)

From impact evaluations to paradigm shift: A case study of the Buenos Aires Ciudadanía Porteña conditional cash transfer programme, *3ie Working Paper 17*. Agosto, G, Nuñez, E, Citarroni, H, Briasco, I and Garcette, N (2013)

Validating one of the world's largest conditional cash transfer programmes: A case study on how an impact evaluation of Brazil's Bolsa Família Programme helped silence its critics and improve policy, *3ie Working Paper 16*. Langou, GD and Forteza, P (2012)

Addressing attribution of cause and effect in small n impact evaluations: towards an integrated framework, 3ie Working Paper 15. White, H and Phillips, D (2012)

Behind the scenes: managing and conducting large scale impact evaluations in Colombia, *3ie Working Paper 14*. Briceño, B, Cuesta, L and Attanasio, O (2011)

Can we obtain the required rigour without randomisation? 3ie Working Paper 13. Hughes, K and Hutchings, C (2011)

Sound expectations: from impact evaluations to policy change, 3ie Working Paper 12. Weyrauch, V and Langou, GD (2011)

A can of worms? Implications of rigorous impact evaluations for development agencies, 3ie Working Paper 11. Roetman, E (2011)

Conducting influential impact evaluations in China: the experience of the Rural Education Action Project, *3ie Working Paper 10*. Boswell, M, Rozelle, S, Zhang, L, Liu, C, Luo, R and Shi, Y (2011)

An introduction to the use of randomised control trials to evaluate development interventions, *3ie Working Paper 9*. White, H (2011)

Institutionalisation of government evaluation: balancing trade-offs, *3ie Working Paper 8.* Gaarder, M and Briceño, B (2010)

Impact evaluation and interventions to address climate change: a scoping study, *3ie Working Paper 7*. Snilstveit, B and Prowse, M (2010)

A checklist for the reporting of randomised control trials of social and economic policy interventions in developing countries, *3ie Working Paper 6.* Bose, R (2010)

Impact evaluation in the post-disaster setting, 3ie Working Paper 5. Buttenheim, A (2009)

Designing impact evaluations: different perspectives, contributions, 3ie Working Paper 4. Chambers, R, Karlan, D, Ravallion, M and Rogers, P (2009) [Also available in Spanish, French and Chinese]

Theory-based impact evaluation, *3ie Working Paper 3*. White, H (2009) [Also available in French and Chinese]

Better evidence for a better world, *3ie Working Paper 2.* Lipsey, MW (ed.) and Noonan, E (2009)

Some reflections on current debates in impact evaluation, *3ie Working Paper 1*. White, H (2009)

This working paper adds to the growing study and practice of advocacy evaluation by examining the challenges associated with evaluating advocacy initiatives. The authors identify factors associated with successful advocacy interventions that have been measured using rigorous impact evaluations. Most advocacy evaluation rely on a range of non-experimental designs, methods and tools. This paper summarises them in a toolkit section of the paper. It is designed to be used by evaluators and advocacy practitioners to address some of the challenges in measuring the impacts of advocacy initiatives targeting a change in policy, processes, knowledge, values, or social norms. This paper highlights that evaluating advocacy actions requires a combination of methods, and that any single method is too limited to evaluate the dynamic and multi-level, multi-actor nature of advocacy initiatives.

Working Paper Series

International Initiative for Impact Evaluation
202-203, Rectangle One
D-4, Saket District Centre
New Delhi – 110017
India

3ie@3ieimpact.org
Tel: +91 11 4989 4444

www.3ieimpact.org