

EVIDENCE REPORT

No 164

Policy Anticipation, Response and Evaluation

Annotated Bibliography – Evaluating Impact Investing

Justin Flynn and Jeremy Young

January 2016

The IDS programme on Strengthening Evidence-based Policy works across seven key themes. Each theme works with partner institutions to co-construct policy-relevant knowledge and engage in policy-influencing processes. This material has been developed under the Policy Anticipation, Response and Evaluation theme.

The material has been funded by UK aid from the UK Government, however the views expressed do not necessarily reflect the UK Government's official policies.

AG Level 2 Output ID: 616

ANNOTATED BIBLIOGRAPHY – EVALUATING IMPACT INVESTING

Justin Flynn and Jeremy Young

January 2016

This is an Open Access publication distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are clearly credited.

First published by the Institute of Development Studies in January 2016
© Institute of Development Studies 2016

IDS is a charitable company limited by guarantee and registered in England (No. 877338).

Contents

	Abbreviations	2
1	Introduction	3
2	Annotated bibliography	4
Annex	Strengths and weaknesses of existing evidence and evaluation approaches and methods related to SIBs and DIBs: a summary	36
	References	38

Abbreviations

ANDE	Aspen Network for Development Entrepreneurs
CDI	Centre for Development Impact
CSP	corporate social performance
CSR	Corporate Social Responsibility
DFID	Department for International Development
DIB	development impact bond
EIA	environmental impact assessment
ESG	environmental, social and governance
EVAS	Expanded Value Added Statement
GIIN	Global Impact Investing Network
GIIRS	Global Impact Investing Rating System
GRI	Global Reporting Initiative
HMP	Her Majesty's Prison
IDS	Institute of Development Studies
IRIS	Impact Reporting and Investment Standards
M&E	monitoring and evaluation
ON	Omidyar Network
PPI	Progress out of Poverty Index
PSM	propensity score matching
RCT	randomised controlled trial
SIA	social impact assessment
SIB	social impact bond
SROI	social return on investment
UEA	University of East Anglia

1 Introduction

The Centre for Development Impact (CDI) contributes to learning and innovation in the field of impact evaluation, through the use of appropriate, mixed method, and robust evaluation designs. It is a joint initiative between Institute of Development Studies (IDS), Itad and the University of East Anglia (UEA).

'Impact' in impact investing has been touted as bringing a new dimension to investing, going beyond investment's traditional 'risk' and 'return' dimensions, and even becoming a new paradigm in solving society's problems (Social Impact Investment Taskforce 2014a). Yet, there has been a growing body of literature in the past few years calling for increased scrutiny of the sector so that the nature of ventures can be improved and so that a more significant positive change can be brought to the beneficiaries ultimately targeted by impact investments, namely communities, households and individuals (Jackson 2013b). In this sense, evaluation of impact investing must go beyond the function that most often underpins it, namely accountability (Jackson 2013b; Liket, Rey-Garcia and Maas 2014).

We created this annotated bibliography because we wanted to bring the literature on both 'impact investing' and 'evaluation' into one place – and in doing so, further encourage debate on the role of evaluation in understanding and measuring the social and environmental dimensions of such investments. This annotated bibliography is intended for a global audience of business professionals, impact evaluation professionals, students, academics and researchers. It is designed to enable quick and easy access to current and relevant publications in the evaluation of impact investing.

This bibliography was made possible by funding from the IDS programme on Strengthening Evidence-based Policy (funded by UK aid from the UK government), and as part of the CDI grant under the Innovation for Impact Evaluation for Development project (funded by the Rockefeller Foundation).

Searches were made on Web of Science and Google using the following terms: impact invest*,¹ development impact bond, social impact bond*, pay-for-success, social stock exchange, performance measurement, (strategic) philanthropy, M&E (monitoring and evaluation), evaluation/assessment, CSR/CSP (Corporate Social Responsibility/corporate social performance), ratings/measurement, SROI (social return on investment), and performance management. References from articles found were also evaluated for further sources. We also explored publications related to evaluation on websites dedicated to impact investing. Overall, 101 documents were found to relate to the topic, of which 55 documents were deemed relevant (based on reviews of their abstracts/executive summaries, tables of contents, titles, conclusions and general body of text).

The journal articles and other publications are listed in alphabetical order by author. The search criteria for each is presented in brackets after the publication title. These 55 publications are further summarised in *Impact Investments: A Literature Review* (Flynn, Young and Barnett 2015).

¹ The asterisk denotes a boolean search term used to return all relevant entries for keywords starting with 'impact invest' (e.g. impact investor, impact investment, impact investing).

2 Annotated bibliography

1. **Bannick, M. and Hallstein, E. (2012) 'Learning from Silicon Valley', *Stanford Social Innovation Review (Advancing Evaluation Practices in Philanthropy, Sponsored Supplement of the SSIR)* 10.3: 8–11 (*M&E, *Evaluation/assessment, *Strategic) Philanthropy)**

Omidyar Network (ON) is a philanthropic investment firm. It has invested in numerous microfinance projects, and other projects such as low-cost private schools – many of these projects are situated in developing countries. Essentially, the article focuses on ON's approach to monitoring and evaluation (M&E), which is largely shaped by a due diligence process of ventures it invests in. The authors claim that this does not constitute normal evaluation practice, yet they stress the significant benefits of this process. Namely, it helps ON know whether to invest in the organisation or not; it helps to establish the metrics used to evaluate the organisation – the due diligence process thus has an impact on all future M&E of the organisation; and this process helps to build trust with the organisation over time.

In terms of metrics, ON measures *reach*, i.e. how many individuals are reached (e.g. how many users on Wikipedia per month), and *engagement*, i.e. the depth of that interaction (e.g. how long users stay on Wikipedia for per month). ON does not use experimental or quasi-experimental methods to measure impact, rather ON uses the metrics of reach and engagement as proxies for impact. ON also uses operating income as another proxy for impact, and the sustainability of this impact.

Finally, the authors highlight the lessons learned relating to the M&E of philanthropic investments, which are to: *foster partnerships* (for easier M&E and evaluation), *exercise judgement* (in setting the metrics), *be flexible* (to adapt metrics if need be), and *embrace feedback* (carefully look at data returned).

2. **Bice, S. (2013) *Creating Meaningful Corporate Social Responsibility Programs for Rural Australian Mining Communities: Reflections and Improvements on Current Practice*, www.tasa.org.au/wp-content/uploads/2013/11/Bice.pdf (accessed 4 December 2015) (*M&E, *CSR/CSP)**

This article focuses on Corporate Social Responsibility (CSR) activities of mining companies in Australia, and the way these are monitored, evaluated and reported. Though not representative of all corporate CSR efforts, it can give some insights into CSR culture extending beyond mining companies operating in Australia. Also, the article does not give detailed technical insight into evaluation or monitoring and evaluation (M&E) practices used in evaluating CSR initiatives, but rather contextual reasons for why these are often poorly developed or implemented.

The author starts by stating that mining companies face scrutiny from various actors, but do not scrutinise their own CSR activities, leading to several problems in CSR implementation and reporting (including M&E) strategies. One reason for this lack of programme-specific scrutiny is that companies do not engage in meaningful two-way dialogue about the problems or possibilities of their CSR programmes, nor how to shape and improve them, because of their 'instrumental, calculative and rational' *auditing culture*. Thus, feedback loops with the community are not well established, leading to a lack of community feedback, and resulting in poor service provision.

Another problem, according to the author, is that global CSR reporting requirements are for overall company CSR efforts rather than programme-specific efforts, thus providing little impetus for the monitoring of each programme. Further, on-site M&E reporting may not be happening because of disengagement between headquarter-level agenda setting and implementation by managers lower downstream in the management scale; this leads in part to monitoring and evaluating project implementation, but not measuring and understanding social outcomes achieved. Better M&E of each programme would help to improve this, and provide lessons for future CSR practices.

Finally, the article mentions another problem, namely that social impact assessments (SIAs) are often included in environmental impact assessments (EIAs), and do not seem to inform community development programming. Thus, better use of SIAs would better inform CSR practices.

3. Brest, P. and Born, K. (2013) 'Unpacking the Impact in Impact Investing', *Stanford Social Innovation Review*, www.ssireview.org/articles/entry/unpacking_the_impact_in_impact_investing (accessed 4 December 2015) (*Evaluation/assessment – especially the theory of it)

The article makes a theoretical contribution to impact investing assessment by providing an assessment framework. It mentions three types of impact that investments can have: *enterprise impact* ('the social contribution of the enterprise itself, whether from the enterprise's products or from operations'), *investment impact* ('the contribution that investors make to the enterprise's impact by providing capital that it could not otherwise obtain, or by providing it at lower cost') and *nonmonetary impact* ('the contribution of activities besides the financial investment itself to the social impact of an enterprise or, indeed, to the social impact of a sector'). These impacts, most fundamentally, can only represent social impact if an investment 'increases the quantity or quality of the enterprise's social outcomes beyond what would otherwise have occurred' – referring to the principle of *additionality*.

Further, the article states that metrics such as the Global Impact Investing Rating System (GIIRS) and Impact Reporting and Investment Standards (IRIS) focus on measuring outputs, but do not go far enough in assessing outcomes. A quote from the article summarises the existing situation of impact investing evaluation:

The superficiality of the current rating schemes may reflect the expense, complexity, and often indeterminate results of evaluations. *In this respect, impact investing is not far behind philanthropy, which has not yet succeeded in creating a comprehensive system for evaluating the performance of non-profit organizations* [emphasis added]. But the absence of data and analysis makes it difficult for impact investors to assess the social impact of the enterprises in which they invest.

Although the article does not provide concrete calculation tools to assess the various types of impact, it provides clues to each type of possible impact from an investment, allowing investors to estimate the potential impact.

Finally, the article is quite rich in providing a range of useful information on theory behind impact investing, for example, on such concepts as *impact investing*, *non-concessionary investing* (i.e. not conceding on potential financial returns), *concessionary investing*, *mission-related investment* (MRI; non-concessionary), *programme-related investment* (PRI; concessionary), and *social value* (= social benefit/production cost). Lastly, it states that assessing outcomes will continue to require large-scale studies such as those conducted by the MIT Poverty Lab, realistically only funded by foundations or international development organisations.

4. **Brest, P. and Born, K. with responses from Choi, A.; Speirn, S.K.; Rodriguez Arregui, A.; Chu, M. and O'Donohoe, N. (2013) 'When Can Impact Investing Create Real Impact?', *Stanford Social Innovation Review* 11.4: 22–31 (*Evaluation/assessment)**

This article is a shorter version of Brest and Born (2013), but includes responses from experts.

Speirn: This critic agrees to a large extent with the tenets of the article. Indeed, he also describes the shortcomings of the evaluation of impact investments, by stating that it is 'a field that is desperate to have the practice catch up with the rhetoric, and that until now has surrounded a provocative idea with too much wishful and fuzzy thinking'. One contribution the expert does provide is the idea of *learning return*, which is about investing in areas not as well known by an organisation, generating additional expertise otherwise not produced without this new exposure. For example, investing in the technology sector can help to spur expertise in usage of big data in assessments (though not much more information on *big data* is provided).

Rodriguez Arregui and Chu: These critics on the other hand disagree with the authors of the article, stating that outputs, such as producing glasses, are a good measure of impact since they obviously are of benefit to the poor. Getting bogged down with such 'academic questions' disrupts the road to investments in various ventures.

5. **Cave, S.; Williams, T.; Jolliffe, D. and Hedderman, C. (2012) *Peterborough Social Impact Bond: An Independent Assessment – Development of the PSM Methodology*, London: UK Ministry of Justice (*Evaluation/assessment, *SIB)**

This report details the propensity score matching (PSM) methodology used to evaluate the first social impact bond (SIB) ever implemented, namely, the HMP (Her Majesty's Prison) Peterborough SIB. This report is very useful to our review as it covers the main points in the design of the PSM methodology. Concerning the SIB, it was launched in 2010, providing interventions for adult males (aged 18 or over) receiving custodial sentences of less than 12 months ('short-sentence prisoners') and getting discharged from HMP Peterborough. Interventions are flexible to meet the needs of the offender group (though admittedly making it harder to attribute impact to the specific intervention/counselling approach since it varies for each prisoner). The current intervention model is based on pre- and post-release mentoring and connecting prisoners to services in order to help them break the cycle of reoffending.

Concerning measurement, there are three cohorts of approximately 1,000 people each. Each cohort is defined by the amount of time it takes to release 1,000 prisoners, but will not exceed 24 months (thus the size of the cohorts could be less depending on whether 1,000 prisoners are released within 24 months or not). Based on power calculations, each cohort will be matched with a control group sample size in the ratio of 10:1, drawn from the other 35 male local prisons in England and Wales (yet the report states that there are 9,000 control individuals for each cohort). The data used to keep track of prisoners come from the Police National Computer (tracking 'reconviction frequency'), and LIDS (prison database – used to find control group prisoners and elaborate the PSM methodology).

The study used 2008 data to carry out a pilot and test the PSM methodology – matched inmates had similar rates of reconviction: 1.64 reconviction at HMP Peterborough vs 1.71 reconviction of matched prisoners coming from all other male local prisons. However, criticisms have emerged about comparing HMP Peterborough prisoners to others, since it is

argued that HMP Peterborough has a different approach and culture of rehabilitation compared to other prisons, thus making it not completely comparable.²

6. Center for Global Development and Social Finance (2013) *Investing in Social Outcomes: Development Impact Bonds: The Report of the Development Impact Bond Working Group*, Washington DC and London: Center for Global Development and Social Finance (*DIB, *Evaluation/assessment)

This report gives an overview of what development impact bonds (DIBs) are and of the considerations to keep in mind when attempting to implement them. It also maps out existing and developing social impact bonds (SIBs) and DIBs, and provides additional details (yet few relevant to our study) about potential DIB interventions.

Overall, this report does not provide enough in-depth analysis to be of considerable use to our review, but it does raise key considerations in identifying appropriate metrics to assess DIB impacts (even though these considerations are quite basic):

- **Measurability:** For example, using an innovative method such as portable cameras to measure absenteeism (in Udaipur, India);
- **Avoidance of perverse incentives:** There could be too many metrics, detracting from the overall programme goal (as mentioned in entry 1 on the Omidyar Network, above). Also, an example of a perverse incentive includes people getting paid to be treated against tuberculosis (in India), but not taking full medication to prolong treatment. The payment thus became capped at six months, judging that it was enough time to treat individuals;
- **Ability to evaluate success:** It mentions three ways:
 - Randomised controlled trials (RCTs) ‘widely considered to be the most rigorous way of determining that a significant change has occurred’;
 - Live comparison groups (mentions PSM);
 - Establishing a historical baseline. An advantage is that there is no need to exclude programme participants if strong baselines of outcomes are established, tracked, and proven;
- **Potential for independent verification of results:** Assessment needs to be done by a third party (to avoid conflict of interest):
 - One way to assess integrity of evaluation/data is to have a recount by another independent organisation;
 - Or, random spot checks;
 - Or, direct observations by an independent agent.

Finally, the report provides approaches to valuing outcomes (i.e. valuing the appropriate financial return of DIBs), though these guidelines are more relevant for DIB intermediary organisations (organisations bringing together investors, programme implementers, evaluators, etc., and which are responsible for developing the contract and payment framework).

² More on this criticism can be found in a very insightful blog entry of the *Stanford Social Innovation Review*: www.ssireview.org/blog/entry/what_the_first_social_impact_bond_wont_tell_us (accessed 8 December 2015).

7. **Clark, C.; Rosenzweig, W.; Long, D. and Olsen, S. (2004) *Double Bottom Line Project Report: Assessing Social Impact in Double Line Ventures, Methods Catalog* (*Evaluation/assessment, *Ratings/measurement, *SROI, *(Strategic) Philanthropy)**

This is a very oft-cited report, which seems to have drawn considerable attention towards the importance and emergence of the social aspect of 'double bottom line' investments and projects. The report is divided into various sections. First, a background to the emergence of social impact assessment is given (where 'lines between grant-making and investing have begun to blur'). Usefully, a summary of key characteristics (i.e. whether a tool is used for *process* evaluation, *impact* evaluation, or *monetisation*) of the various methods reviewed in the report is provided in a graph at the end of the first section.

Second, the report provides two-page summaries of nine existing methods of social impact assessment at the time of writing. The nine approaches are: (1) theories of change as applied by New Schools Venture Fund; (2) balanced scorecard as applied by New Profit Inc.; (3) Acumen McKinsey Scorecard as applied by Acumen Fund; (4) social return assessment as applied by Pacific Community Ventures; (5) AtKisson Compass Assessment for Investors as applied by Angels with Attitude; (6) ongoing assessment of social impacts as applied by Roberts Enterprise Development Fund (REDF); (7) social return on investment as applied by REDF; (8) benefit-cost analysis as applied by Abt Associates and AmeriCorps; and (9) poverty and social impact analysis as applied by the World Bank. Features in the summaries include backgrounds on the development of various tools, their applications to date, some observations about the tools (such as high feasibility and low implementation costs), and the potential risks to credibility. Further information on which elements of the impact value chain are assessed by the method, assessment purposes (such as scaling, external reporting, etc.), time breakdown and other components are also summarised into easily understandable graphics.

Finally, method details (such as questionnaires used in the application of a method) and examples of method uses can be found in the report's appendix.

8. **Disley, E. and Rubin, J. (2014) *Phase 2 Report from the Payment by Results Social Impact Bond Pilot at HMP Peterborough (Revised)*, London: UK Ministry of Justice (*SIB, *Evaluation/assessment)**

This report is the second process evaluation report of the pilot of *One Service* to the HMP Peterborough social impact bond (SIB) project. The report is based on interactions with 39 individuals involved in programme design, implementation, and ongoing operation, and provides additional information on the HMP Peterborough SIB. Among other things, it provides additional information on the intervention, the rationale behind the metrics used in assessing the intervention, and the tendering process of hiring the evaluation consultants.

Clearer information on the nature of the intervention towards prisoners is provided: *One Service* (the name of the programme) is provided by paid and volunteer care workers, and '*There is no fixed intervention model or single theory of change articulated for the One Service*', meaning that *One Service* uses a mix of counselling approaches specifically catered to inmates in order to reduce the likelihood of reconviction. Also, additional information is provided on how the State calculated the monetary valuation of outcomes.

Concerning metrics, the report states that they have supposedly been designed to prevent 'cherry-picking' (focusing on those most likely to make the programme succeed), and avoid 'parking' (neglecting those whose needs are strongest and/or have the least chance of success). Clarification is provided on the main outcome used to measure the success of the intervention (footnote 13): 'There are various ways in which reoffending can be measured.

Most official measures underestimate the true level of reoffending because only a proportion of crime is reported, detected and/or sanctioned’.

One of the key lessons from the SIB implementation is ‘considering appointing a *performance manager* with some experience in the field and knowledge and understanding of the local area’, to improve implementation.

Concerning the choice of the independent assessors, QinetiQ and the University of Leicester were chosen via a competitive tendering process.

Regarding payment, the outcome of being paid upon 10 per cent of reduction of reconviction (for each of the three cohorts – thus three chances for payment) was calculated based on this number being *statistically significant, based on prior data of reconviction, and using 90 per cent statistical significance, 80 per cent power and 10:1 control group ratio*. Payment was also made if a reduction of 7.5 per cent was found across all tracked released prisoners (n= 3,000 across the three cohorts).

The report has a section on the research approach and limitations (detailing the interviewing process, its pros and cons, from page 16 onwards). *Interestingly, one possible caveat to the findings mentioned is that stakeholders involved in the SIB (in the prison) may have added incentive to make the intervention successful*. Finally, of note is also the fact that the report states that there was no research or evaluation of a completed SIB at the time of writing (2014).

9. Drew, R. and Clist, P. (2015) *Evaluating Development Impact Bonds: A Study for DFID*, London: Department for International Development (DFID) (*Evaluation/assessment, *DIB)

As stated in the executive summary, this study seeks to present concepts for understanding development impact bonds (DIBs). First, these are based on economic theory (i.e. calculation of interest rates, who the payers are, etc.). In essence, DIBs are elaborated and implemented in very similar ways to SIBs. The study then considers issues relating to evaluation of individual DIBs (while listing those currently being developed) and how data from those evaluations can be synthesised effectively into an overall evidence base related to DIBs.

As part of the section on the evaluation of DIBs, several important points are mentioned. First, it is interesting to note that there has been mention of some opposition to DIBs (i.e. Oxfam 2013), because of the diversion of programme funds to interventions with very little evidence base. Second, usefully, the main findings from a literature review on DIBs are presented, basing many of the findings from the literature on social impact bonds (SIBs). Namely, one review covering evaluation methods of SIBs is mentioned (Liebman and Sellman 2013) which provides a list of evaluation strategies, including randomised controlled trials, quasi-experimental methods (such as regression discontinuity design), and non-experimental methods (including comparison to a historical baseline). Very importantly to our study, the report states that ‘if there is no credible approach to establishing a counterfactual outcome for a particular intervention, then a SIB may not be feasible’. It also concludes that a comparison group will not be feasible in most cases. Third, DFID’s expectations of DIB evaluations are listed from point 51 onwards, which can be insightful to our review. Among them, it will be particularly important to build evaluation into DIBs from the earliest stage, so that these can be as informative as possible in the context of the use of a new type of development mechanism/intervention. Fourth, there is discussion, at point 57, of the role of the evaluator as one of the required ‘actors’ in the DIB mechanism development.

Of further interest to our study are the descriptions of the experiences of evaluation of SIBs (since no reports exist of DIB evaluations), categorised according to method (experimental, quasi-experimental, etc.). Further, among other things, it has been reported that evaluation rigour has not been consistently implemented or delivered concerning SIBs (i.e. because of limits to evaluative resources, to ensure enough capital for the programme) – point 67. The findings contained in Drew and Clist’s review of the literature are very usefully summarised in a table, which is presented in the Annex of this Evidence Report (see p.36).

Interestingly, the report also states that simply establishing the existence of impact from comparison with a control group does not constitute a full evaluation exercise – impacts on non-payment outcomes and measures should be taken into account, as well as not simply whether, but how, impact occurred in this intervention. *Attention is also brought to the difference between evaluating the intervention vs the instrument (i.e. SIB/DIB) (point 78).*

The authors also advocate for the generation of evaluation evidence across the use of SIBs/DIBs, where synthesis of individual evaluations would present a more complete picture of the experience of evaluation SIBs/DIBs. The elaboration of a framework to raise the adequate questions to ask in order to conduct generic evaluations as well as synthesise their findings is proposed and detailed (from page 30 onwards). While several benefits of the use of such a framework are given and explained, a major benefit is that it would allow for the distinction between *process and impact evaluations*.

Methods to answer the generated evaluation questions are then listed from point 88 onwards (i.e. using experimental, quasi-experimental and other evaluation methods). A comment about experimental evaluations is that they should be informed by strong theory-based *qualitative analysis* to understand causal mechanisms (there is also comment that there is a great deal of interest in using theory-based methods to evaluation: point 99). Another point is that it seems likely that there will be relatively few situations where experimental methods will be able to be used in the evaluation process – situations of ‘high control’ are rare.

Concerning the use of non-experimental methods, the authors comment that the use of historical baselines as a basis for comparison is likely to be highly problematic, but it does not mean that there is no place at all for them in DIB evaluation. Also, the strategy for accumulating knowledge across DIB initiatives is mentioned, listing the following recommendations: ‘Evaluations should be designed around a common evaluation framework’; ‘Capacity is needed to conduct real-time synthesis’; ‘Periodic, retrospective synthesis exercises are likely to be very valuable’. At the end of that section of the report, lessons learned and key recommendations are provided. These include: that there are more actors involved in DIBs than SIBs (a six-actor framework to be used by DFID is suggested in the report); that rigorous evaluations need to be conducted particularly in the early stage of DIB development and implementation; the use of the generic evaluation framework for DIBs and for the development of evaluation questions would help to streamline evidence and lessons learned from DIBs; and that three ways have actually been suggested to synthesise learning from individual evaluations, i.e. the use of the common framework for evaluations of DIBs, contracting an actor to conduct a real-time synthesis of DIB evaluations, and conducting periodic and retrospective syntheses of learning from evaluations – it is suggested that DFID uses a combination of the three approaches.

Finally, there is an example of a ‘Design for the Impact Evaluation of a DIBs Programme’ in Annex 6, which attempts to provide a generic template for such a design. It is argued, however, that establishing such a template is difficult since DIBs diverge so much on the basis of the context of the intervention.

10. Ebrahim, A. and Rangan, V.K. (2014) 'What Impact? A Framework for Measuring the Scale and Scope of Social Performance', *California Management Review* 56.3: 118–41 (*Ratings/measurement, *Performance management)

This article makes a significant contribution to the theory behind establishing metrics to assess social outcomes. Rather than contributing a specific tool for establishing these metrics, it instead provides a framework for determining the proper ways of measuring outputs, outcomes and impacts. The main tenets are summarised as follows:

The core of our framework for measuring social performance is relatively simple: clarify the operational mission, specify the set of activities to address that mission (scope), and identify the target size of the problem (scale). Yet, such measurement is rare in practice.

Interestingly, outcomes and impacts are defined differently to usual, where *outcomes* refer to lasting changes in the lives of individuals and *impacts* to lasting results achieved at a community or societal level (impact, for example, is often defined as the effect which can only be attributed to a specific intervention). Importantly, the authors state that outcomes can only be measured when two conditions exist, which are uncommon to the social sector, namely: (1) 'when the causal link between outputs and outcomes is well established'; and (2) 'when the range of the integrated interventions needed to achieve outcomes are within the control of the organization'.

Concerning measurement, the authors make a few points. First, they state that the growing emphasis on measuring performance and impact is coming from institutional funders seeking to use their resources more effectively and strategically, and looking to publicly legitimise their allocation decisions. Second, the authors find that funders often expect organisations to measure outcomes and impact themselves without necessarily supporting them to do so. Yet, the authors make the case that it is funding organisations that are better placed to conduct evaluations, since they often have more resources, and can actually conduct more sectoral evaluations, which help to give a better assessment of the impacts of interventions (it is important to keep the authors' specific definition of 'impact' in mind here). Thus, measuring outcomes and impacts for organisations is, according to the authors, not necessary – they need simply to be hypothesised.

Finally, the authors state the growing emphasis on incorporating impact evaluation during programme design and implementation, namely to allow organisations to get real-time feedback for improving their work. They also list a number of performance measurement methodologies, including participatory and integrative tools such as *constituency feedback*, most *significant changes techniques*, and *developmental evaluation*, as well as network-based approaches such as *collective impact* and *outcome mapping* (see Appendix 1 for approaches on social performance assessment) suited to settings of high complexity involving interactions across multiple organisations and sectors.

11. Freireich, J. and Fulton, K. (2009) *Investing for Social & Environmental Impact: A Design for Catalyzing an Emerging Industry*, Monitor Institute (Ratings/measurement)

This early report in the impact investing and measurement/evaluation literature points out some interesting considerations for ensuring the success of the impact investing industry, based in part on good evaluation of social/environmental outcomes/impacts. For example, the report provides some thoughts on how impact investing could fail, namely by making impact investing too easy. By this the authors mean that the definition of social and environmental impact could be used too loosely and end up having virtually no meaning. At best, this outcome would turn this type of investing into a 'feel good' rather than a 'do good'

exercise. Also, another risk would be the ‘greenwashing’ and dilution of standards, reducing effectiveness of impact investing.

What is needed for success, according to the authors of the report are: ‘better metrics’ so investors know what they are paying for and can validate the achievement of social or environmental outcomes; being able to articulate social value gain relative to the potential sacrifice of social return; impact rating systems and minimum standards to ensure some existence of impact – and common approaches to assessing social/environmental elements of investment; and finally, awareness of potential ‘integrity constraints’, where indicators make organisations look good, but where these organisations are actually treating employees or beneficiaries unfairly (such as lending microcredit at predatory rates).

As such, as an early publication on this topic, the theoretical contributions it offers for really making the most use of evaluation are limited; mostly advocating for use of metrics and standards to ensure at least some levels of impact.

12. Galloway, I. (2014) ‘Using Pay-For-Success to Increase Investment in the Nonmedical Determinants of Health’, *Health Affairs* 33.11: 1897–1904 (*SIB)

Overall, this article provides a good description of social impact bonds (SIBs), or pay-for-success programmes, and of the current extent to which they are being implemented and considered within the United States. It mentions that some SIBs being considered are based on prior evidence of success of similar interventions, such as randomised controlled trials, and feasibility studies (such as the one carried out for the Fresno Health Investment Bond (HIB)), but provides little other information on the type of data relied upon to plan other SIBs. It also mentions that data, especially concerning non-medical interventions, can be hard to get, but that a greater collaboration on the use of data between providers and government can help unlock a correct mapping of outcomes related to various problems and interventions.

13. Godeke, S. and Pomares, R. (2009) *Solutions for Impact Investors: From Strategy to Implementation*, New York: Rockefeller Philanthropy Advisors (*Evaluation/assessment, *Ratings/measurement, *SROI)

In relation to our study, this report mentions the difference between commercial and philanthropic investors – equating philanthropy with generating financial and social/environmental returns, and not just the latter. Also, it provides some rationale on how to conduct impact investing evaluation and measurement. For example, in comparison to Ebrahim and Rangan’s (2014) article on impact measurement, which focuses on developing metrics closely attuned to an organisation’s mission, this report advocates this practice while at the same time developing metrics enabling comparison to other investors’ performances. Yet, it does bring up the fact that not all returns on investment can be compared across the sector, as impact is sometimes region- or intervention-specific. Finally, it mentions two tools used in assessing impact across sectors: the Social Return on Investment (SROI) Calculator Tool (Calvert Foundation) – designed to be measured across communities (but not necessarily more broadly than that); and Impact Reporting and Investment Standards (IRIS) – designed to be measured across a broad range of impact investments.

14. Hardi, P.; Radasci, G. and Schmidt, K. (2012) *Evaluations of CSR Performance and Impact as Seen by Key Actors other than Business*, Working Paper 3, CSR Impact (*Evaluation/assessment, *CSR/CSP)

This report is a systematic review of Corporate Social Responsibility (CSR) performance evaluations in various European countries, including the United Kingdom. It structures reviews of CSR performance evaluations according to organisation-type (including political actors, non-governmental organisations), for-profit organisations, labour unions, and others such as academic institutions) and finds that overall, evaluations of corporate CSR activities and their own business performance (firm-level impacts) are scarce and of little value:

The assessment of the economic, social and environmental outcomes and impacts of CSR activities is sporadic. There are almost no independent measurement tools and quantitative methods applied. If assessments are performed, these are done by reviewing managers' personal observations, or publicly available company communications (such as CR [corporate responsibility] reports).

It mentions, among other things, the limitations of the British *AccountAbility* CSR assessment tool, for example. Nevertheless, a list of stakeholders and CSR assessments is provided in Section 5.

15. Harji, K. (2008a) *Accounting for Social Impact: The Expanded Value Added Statement*, Ottawa: Carleton Centre for Community Innovation, Carleton University (*Evaluation/assessment)

Overall, the tool explained in this document, the Expanded Value Added Statement (EVAS), attempts to measure social value creation in economic terms, where it is usually not counted. Of primary interest is the value of the involvement in the non-profit sector that is generated for volunteers and others employed in a non-profit. Also of interest is the value of volunteers' work (to measure non-profit effectiveness) and the value created by a non-profit to the community. However, no methodology on how to do this is provided (rather the document refers the reader to a book on EVAS to gain greater knowledge of the EVAS methodology, entitled *What Counts: Social Accounting for Nonprofits and Cooperatives*, by Mook *et al.* (2007) – yet this resource is inaccessible via University of Sussex/IDS resource databases and the internet).

Finally, the author states that the EVAS is intended to 'supplement existing financial statements, not replace them'. He also proposes for the EVAS to triangulate findings from social impact assessments from SROI and triple bottom line accounting. Overall, however, it is unclear how widely used or relevant this tool is to evaluation of impact investing.

16. Harji, K. (2008b) *Social Return on Investment*, Ottawa: Carleton Centre for Community Innovation, Carleton University (*Ratings/measurement, *SROI, *CSR/CSP)

This document provides information on social return on investment (SROI), and how it can be used to assess Corporate Social Responsibility (CSR) performance. It introduces the notion of *blended value*, which is that 'all organizations create value that consists of economic, social and environmental value'. Corporations are looking to increase this value to address concerns of wider stakeholder groups, while simultaneously increasing or not compromising shareholder value. SROI is defined as the monetised value of the social value created by an organisation, against the financial value needed to create the monetised social value, and as such utilises cost-benefit analysis techniques to monetise social value. It can be used to assess the 'biggest bang for the buck', and ways to express it include

outputs/US\$ (e.g. number of housing units built per US\$ invested). Among other uses, the author states that SROI is increasingly being used in CSR performance assessment.

Further, the document lists a number of organisations as having developed their own SROI tool, namely, the Roberts Enterprise Development Fund (REDF) and the New Economics Foundation (NEF). Social Capital Partners employs a sustainable livelihoods framework in assessing the organisation's role in creating employment. However, the author states some caveats to the use of SROI in assessing social value: (1) not all existing SROI methods can be applied to assess CSR investments; (2) appropriate data may be difficult to collect if not already being collected; (3) there are difficulties in assigning a dollar amount to various social benefits, such as increased morale/self-esteem from volunteering; (4) it is a subjective exercise, as social value is different for different groups; and (5) causality and correlation can be difficult to attribute – assumptions must therefore be clearly articulated.

17. HM Government (2014) *Qualitative Evaluation of the London Homelessness Social Impact Bond: First Interim Report*, London: Department for Communities and Local Government (*Evaluation/assessment, *SIB)

This report is the first process evaluation of a social impact bond (SIB) being carried out in London to reduce rough sleeping of 836 individuals. It is a qualitative evaluation of the intervention thus far, and is being carried out alongside a quantitative impact and economic evaluation. The qualitative aspect involves three stages: an initial baseline (included in this HM Government report), mid-point collection data, and final data collection, at which time a synthesis evaluation (quantitative + qualitative) of the project is scheduled to be published in 2016. To conduct the qualitative evaluation, various stakeholders were interviewed, including a representative of the City of London, social investors and intermediaries, provider organisations and members of the cohort. The evaluation itself is based on a baseline, with no comparison group.

Some aspects of monitoring reporting are provided, such as monthly data checks with providers. Concerning the data, results of the qualitative evaluation show that some indicators may be somewhat problematic, such as the indicator of two nights spent rough sleeping in a quarter. Indeed, this indicator was not necessarily able to show the extent of the successes being achieved since the evaluation showed that individuals 'slept rough' sporadically and out of choice (e.g. to spend time with a partner), and were measured as sleeping rough even when they lodged at relatives', since no receipts of a hostel were provided as proof of accommodation. The evaluation shows that one way to better assess the success of the intervention would be to compare data on rough sleeping to historical trends, rather than the maximum two nights of rough sleeping per quarter indicator.

Finally, the evaluation shows the challenges associated with implementing SIBs, which sometimes implement innovative interventions, yet for which little evidence of evaluation/success exists:

In modelling PbR [payment by results], commissioner and provider decisions about what can be achieved will ultimately be a matter of judgement without a robust evidence base from tested interventions. As a result, the assessment of risk is complex and some of the risk should be shared with the commissioner [i.e. increasing potential investor returns for 'riskier investments'], if a wide range of providers are to consider PbR viable.

18. Hornsby, A. (2012) *The Good Analyst: Impact Measurement and Analysis in the Social-Purpose Universe*, London: Investing for Good (*Evaluation/assessment, *Ratings/measurement)

This book focuses on advancing the theoretical debate around social impact measurement. It does so by providing a new and independent tool for measuring social impact: the MIAA (Methodology for Impact Analysis and Assessment). This tool is based on a CRI rating (confidence–return–impact), and integrates cost-benefit analysis into it in order to ‘price’ the creation of social value. Another concept added to the MIAA is the impact contribution (i.e. the impact of a specific capital to the organisation’s impact).

Essentially, a full MIAA analysis comprises:

- ‘a non-evaluative mapping operation;
- an evaluative analysis and assessment of confidence;
- an evaluative analysis and assessment of impact (with an impact of contribution bolt-on available as required).’

Page 53 has a summary of the methodology around producing the various outputs from an MIAA analysis, while Chapter 2 of the book provides a detailed explanation of the MIAA methodology.

19. Ingram, V. (2013) *Corporate Social Responsibility: The Role of Public Policy. A Systematic Literature Review of the Effects of Government Supported Interventions on the Corporate Social Responsibility (CSR) Behaviour of Enterprises in Developing Countries*, Den Haag: Policy and Operations Evaluation Department (IOB), Ministry of Foreign Affairs (*Evaluation/assessment, *CSR/CSP)

This study is useful in that it provides information on various evaluations of Corporate Social Responsibility (CSR) performance for firms operating in developing countries which were supported by government. Yet, the study found that:

Most reviewed evidence relates to CSR behaviour in companies and does not refer to the influence of government policies on CSR behaviour or impacts. CSR behaviour was most commonly seen displayed by larger, multinational companies, and in the energy, consumer goods and agricultural sectors.

More importantly to our study, one of the study’s research questions is what the impact of CSR activities was on poverty alleviation for companies supported by government interventions. The answer is that there was a wide variety of impacts, and that these were not measurable to a large extent, since they were conducted mainly with the use of *qualitative methods*. Still the study mentions some ways of assessing impacts of CSR activities on poverty: environmental indices, waste and water emissions; labour and industrial relations, such as wage levels; community involvement defined by things such as goods procured by the local community or investments in community projects; and public disclosure, such as CSR impacts published in reports, benchmarks and websites.

Finally, the review is useful since it is composed of findings from several CSR activity evaluations in developing countries, which can thus be more closely inspected (see Ingram 2013 reference list; also an example of a relevant evaluation was reviewed, under Triodos Facet (2010)).

20. IRIS (2013) *Getting Started with IRIS: How to Select IRIS Metrics for Social and Environmental Performance Measurement*, Global Impact Investing Network (GIIN) (*Ratings/measurement)

This report details what the Impact Reporting and Investment Standards (IRIS) is and how it is meant to be used. Among the various things it measures are the following performances: financial, operational, product, sector, and social and environmental objective performance. IRIS provides metrics in a wide range of sectors, including housing/community facilities, water and waste, financial services, environment, health, energy, education, agriculture and even cross-sector. It is a catalogue of metrics, which does not advocate for one single combination of metrics to be used, yet it is also not an evaluation tool, a data management platform nor a reporting framework. The guide also states that IRIS is only meant to complement other performance measurement tools, and that it should only be one component of an impact measurement programme.

Four steps are detailed in order to find the right metrics to assess social performance ('start with the end in mind'; 'create a "data-wanted list"'; 'organise your portfolio'; and 'choose your metrics'). Further in the guide, it also advises to collaborate with the investee in order to determine the right metrics, as they will be able to notify the data tracing team on the realistic possibility of tracking various chosen metrics.

The example of the metrics chosen to track social performance by an investment fund is provided (page 20 onwards). Finally, a notable advantage of using IRIS is the fact that metrics can be compared across investment portfolios, and across sectors.

21. Jackson, E.T. (2013a) 'Evaluating Social Impact Bonds: Questions, Challenges, Innovations, and Possibilities in Measuring Outcomes in Impact Investing', *Community Development* 44.5: 608–16 (*Evaluation/assessment, *SIB)

This article reviews the social impact bond (SIB) landscape as of the end of 2013; provides background information on SIBs; mentions problems associated with, as well as recommendations for, conducting their evaluations; and it brings up other questions surrounding SIB evaluation. The article first starts by giving good background information on the emergence and relevance of SIBs (in the introduction), and states that evaluation is a prime tool to hold investors and sponsors of SIBs accountable for their declared intentions on realising social outcome targets. The focus of the article is on the United States, United Kingdom, Canada and Australia, where SIBs have been introduced or are being considered for introduction. In terms of language, the author equates them to *pay-for-success financings*.

One of the problems surrounding SIBs is related to SIB evaluation; namely, that it is difficult (according to the author) for SIBs to achieve their very ambitious social outcome targets, partly because only some factors that influence outcomes are related to programme design. Also, the author points out a problem with the HMP Peterborough SIB, i.e. that it was difficult during the interim study to pinpoint exactly which elements of the services delivered were responsible for social outcomes. This and other questions will thus have to be addressed in future evaluations.

Recommendations on conducting evaluations of SIBs focus on putting theory of change at the centre of SIB design while making sure to interrogate it using the tools at the evaluator's disposal, and also while retaining some flexibility in the design and operations of an SIB. They also stress engaging with primary stakeholders (i.e. in a participatory evaluation process), using contribution analysis to assess how SIBs improve and sustain wellbeing, being informed by the Impact Reporting and Investment Standards (IRIS) and Global Impact Investing Rating System (GIIRS), and using comparative studies to assess common performance indicators within and across countries, sectors, SIB design, etc.

Finally, among the points of discussion brought up in the article and which, according to the author, has not yet been taken up by SIB proponents and sponsors, is the importance of looking at the impact of SIBs at the individual level and within households. The impact of neighbourhood and community dynamics on outcomes has also apparently not been looked at closely in the evaluation of SIBs.

22. Jackson, E.T. (2013b) 'Interrogating the Theory of Change: Evaluating Impact Investing Where it Matters Most', *Journal of Sustainable Finance & Investment* 3.2: 95–110 (*Evaluation/assessment)

This article explores theory of change analysis in depth and makes a strong case for its usage and benefits across the impact investing evaluation sector. In the first part of the article, the author provides an overview of the impact investing industry in general, mentioning, for example, that the most favoured types of investments made thus far are in microfinance institutions and affordable housing schemes, but that this range of investments is broadening. It also mentions the Impact Reporting and Investment Standards (IRIS) and Global Impact Investing Rating System (GIIRS) as measurement tools in the industry, but states that there is also a growing customised and decentralised array of metrics at the level of institutions or organisations – how sector-wide and smaller-scale metrics can interact is an important question. Until recently, the author considered that the industry was 'metrics-rich but data-poor'. Interestingly, the author states that the GIIRS as well as the Nexii both require the organisations that they review to make their theories of change explicit. Also, in relation to the use of various tools, social return on investment and the Expanded Value Added Statement are both mentioned as being able to monetise social value created by an investment which otherwise goes unaccounted for.

Further, several benefits of using theory of change analysis are mentioned, such as the fact that it can be a cost-effective method of framing and informing an evaluation (cheaper than large-scale experimental studies), and it also enables an improving and refining of programme strategies and implementation, while also being able to generate further and even more useful insights when combined with a whole host of other evaluation methods (which are listed in the text).

Finally, the article talks about accountability and learning, where the most effective evaluation strategies aim to balance these two components, but where in reality, most underpinnings of impact evaluation come from the accountability function. It also suggests ways of going forward, emphasising the importance of integrating the perspectives of various stakeholders more greatly into impact evaluations (i.e. using participatory evaluation methods), especially those of primary stakeholders – that is, the ultimate beneficiaries of impact investments in designing, participating and using the impact assessment results. This has two important benefits: it is able to hold funders and investors to greater account for their statements, intentions and practices while also ensuring that the functioning of impact is correctly understood at micro level, i.e. at an individual and intra-household level.

23. Jackson, E.T. and Harji, K. (2013a) *Assessing Impact Investing: Five Doorways for Evaluators*, <http://issuu.com/rockefellerevaluation/docs/impactinvestingthematicbriefs2014> (accessed 4 December 2015) (*Evaluation/assessment, *Ratings/measurement)

This short article describes five ways that evaluators can engage with impact investment assessment. Of interest is also the information on the extent (in resources) of assessment of the Global Impact Investing Rating System (GIIRS) as of mid-2013. The five ways of engaging with impact investment assessment are detailed below:

1. **Through industry-wide systems (IRIS/GIIRS):** The author recommends that evaluation work undertaken in this field should be informed and linked to the industry-wide systems that are IRIS and GIIRS.
 2. **Theory of change:** Used even in assessment of social stock exchanges (see article 26, below). Learning to use this tool should be a priority.
 3. **Policy assessment:** Policy assessment is a specialised area of evaluation. The IIPC (Impact Investing Policy Collaborative) is an international network of scholars, practitioners and policymakers from a dozen countries that examines policies which support impact investing. A policy example: nearly all American states have a law that mandates or encourages pension systems to be invested in geographically targeted economic development or energy-efficient real estate; evaluating the upstream and downstream components of such policies would represent a gateway for evaluators to engage with impact investing assessment.
 4. **Sector-based interventions:** Specialising in certain sectors (such as medical technologies, affordable private education) and evaluating differentially various actors in these sectors – ‘market scalers’, ‘market innovators’ and enterprises and organisations that provide infrastructure (such as credit ratings) and research to the field (such as IDS).
 5. **Outcome-based financing instruments:** Such as social impact bonds (SIBs)/pay-for-success bonds. Upstream and downstream analysis is important, while answering questions such as ‘how will the social outcomes be maintained over time?’ is also important. Authors advocate for using a counterfactual before the start of the SIB to strengthen the analysis.
24. **Jackson, E.T. and Harji, K. (2013b) *Measurement that Matters: Ten Steps for Assessing Social Impact*, <http://issuu.com/rockefellerevaluation/docs/impactinvestingthematicbriefs2014> (*Ratings/measurement)**

This article explains ten steps to assessing impact; the most relevant to our study include:

2. **Test and refine a theory of change:** Important questions to keep in mind: ‘What change are you seeking to make?’, ‘What is your contribution to this vision?’, ‘How will you know when your objectives are achieved?’, and ‘To what extent are the results due to your efforts and not to other factors and actors?’
4. **Enhance utility and relevance:** Basically, the focus should be on developing and using measurement tools/approaches as *a means to an end, rather than as an end itself*. Figuring out what the tool is for will help to develop the tool more effectively. For example, *there is currently an overreliance on measurement data based on quantitative tools such as randomised controlled trials (RCTs)*.
5. **Coordinate standardised and customised approaches:** Social return on investment (SROI) and RCTs have come in vogue, but other approaches can be more useful to certain functions of evaluation. One way to coordinate approaches is to use some indicators, for example, of the Impact Reporting and Investment Standards (IRIS) toolkit. *According to the authors, investors can and should use both standardised and customised evaluation tools*.
7. **Share experiences with peers:** The Acumen Fund’s BACO (Best Available Charitable Option) apparently does this well.

25. **Jackson, E.T. and Harji, K. (2012a) *Unlocking Capital, Activating a Movement: The Final Report of the Strategic Assessment of the Rockefeller Foundation's Impact Investing Initiative*, <http://issuu.com/rockefellerevaluation/docs/impactinvestingscan2012> (accessed 4 December 2015) (*Evaluation/assessment)**

This report presents the findings from an evaluation of an organisation-wide strategy, the Rockefeller Foundation's Impact Investing Initiative, which is a five-year US\$40m global grant-making initiative which aimed to foster a global impact investing movement. Data were collected through an extensive review of the literature on impact investing, open-ended interviews with over 100 leaders in the impact investing field from 11 countries, and participant observations at impact investing conferences and meetings in eight countries. The evaluation was carried out from 2011 to 2013.

Among others, the report details sections on the theory of change of the initiative (Chapter 4) and the section on effectiveness (Chapter 6) which presents the outputs, outcomes and impact of the initiative. The chapter on the theory of change is brief but provides a good example of a theory of change applied to assessing impact investments; namely, it includes a full diagram of the Initiative's theory of change, while also including a table on the Initiative's result statements and indicators.

Chapter 6 reports outputs, outcomes and impacts. Outputs are mainly reported in the form of amounts of money granted to the various organisations composing the Initiative's portfolio. Among outcomes, there is a lot of focus on the evolution of the Global Impact Investing Network (GIIN), Global Impact Investing Rating System (GIIRS) and Impact Reporting and Investment Standards (IRIS), which were all investees of the Initiative. Regarding the IRIS, for example, the report states that it will still require another five to seven years (the report was published in 2012) and more funding to be fully developed. At the same time, more effort is needed to increase its credibility and adoption around the world, as it is still very much a North American-led initiative. There is also some mention of the effectiveness of social stock exchanges, which have been supported by the Rockefeller Initiative, but it is too early to tell what social impacts these have produced. Regarding impact, the report states that there has been evidence of some positive impacts arising from impact investments, but that data demonstrating such impacts are still underdeveloped and insufficient. Operationally testing organisations' own theories of change and collecting better information are two ways further light can be shed on the social impact of investments.

26. **Jackson, E.T. and Harji, K. (2012b) *Accelerating Impact: Achievements, Challenges and What's Next in Building the Impact Investing Industry*, New York: The Rockefeller Foundation (Evaluation/assessment)**

The sections of this document relevant to our study (i.e. on evaluation, highlighted in the document's table of contents) focus on providing an overview of the impact investing sector. The authors state that there is an overall trend currently to find standardised ways of assessing impact across the sector, while at the same stating the need to recognise the use of smaller customised impact assessment tools. The authors describe the Global Impact Investing Rating System (GIIRS) (investor-focused) and the Impact Reporting and Investment Standards (IRIS) (has broader applicability), both of which have emerged as the current leading global initiatives in social impact measurement, while nevertheless pointing out two particularly interesting tools for measuring outputs and outcomes: 'Pulse' (a management information system developed by Acumen Fund integrating the IRIS taxonomy) and social return on investment (SROI). Also, the authors mention that there are two origins for increased demand for measurement standards, namely philanthropic initiatives (making charities more transparent) and ESG metrics (environmental, social and governance).

Useful, however, are the overall aims used for impact measurement mentioned in the text: to improve the nature of a venture (to be more effective) and to prove impact.

The authors then proceed to mention various networks for social impact assessment (e.g. SROI Network, ANDE – the Aspen Network for Development Entrepreneurs, and the Social Impact Analysts Association), as well as various assessment tools (e.g. IRIS, B Impact Ratings System, Progress out of Poverty Index, LEED (Leadership in Energy and Environmental Design), etc.). Among all of these measurement tools, however, organisations risk ‘drowning in data’ trying to use too many of them.

Finally, as in other texts by the same authors, the authors state the importance of impact investments reaching individuals, households and communities – which can be strengthened by solid elaboration of theories of change. Also, the authors conclude by enumerating various ways of improving impact assessment through greater collaboration of impact assessment efforts across the field, improved data utilisation and analysis capacity, balancing the diversity of needs of impact assessment users – *assessments are used in various ways, to different ends!* –, utilising top-down and bottom-up approaches to find ways of triangulating data and adequately attributing sources of impact (to the correct intervention), and increasing the social impact measurement standard, thereby also providing accountability and verification.

27. Jackson, E.T. and Harji, K. (2009) ‘Measuring the Blended Value of CSR and Social Enterprise’, presented at the 2009 Annual CES Conference, Ottawa, 31 May – 3 June 2009 (*Evaluation/assessment, *CSR/*CSP)

This PowerPoint presentation is on ways of measuring blended value, mentioned in Harji’s (2008b) document on social return on investment (SROI). Blended value refers to the financial, social and environmental value created by Corporate Social Responsibility (CSR) and social enterprises. More specifically, the presentation attaches the notion of blended value creation to the context of CSR activities.

In evaluating blended value creation through CSR, the presentation mentions various approaches, including logic models, i.e. theories of change, SROI, and participatory data collection and analysis, while various tools are also mentioned, including ‘Pulse’ (Acumen Fund), which provides portfolio-level assessment; Best Available Charitable Option (BACO); and the Progress out of Poverty Index (Grameen Foundation). An extract is provided on slides 32–33, which provides an example of the use of SROI to assess the CSR efforts of TurnAround Couriers. It mentions that SROI can be resource-intensive and still has issues around feasibility, replication and reporting. Further, in terms of approaches, what makes an effective CSR evaluation according to the authors of the presentation is ‘logic models + social accounting + participatory strategies = effective CSR evaluation’.

In addition, the presentation provides information on the levels at which CSR can have an impact: micro-level (i.e. the direct beneficiaries), meso-level (corporate level, increased reputation), and macro-level (increased job creation in a given area). It also lists the various objectives of the evaluation of CSR, namely management (e.g. organisation sustainability, attract new investment), social investors (e.g. impact of grants, mission alignment and accountability measures) and government programmes/policy (e.g. making the case for investment in an organisation/approach). Finally, some challenges surrounding social impact assessment are mentioned, i.e. fragmentation and *immaturity of approaches*, siloing, comparing apples to oranges, and data availability and collection.

- 28. Jackson, E.T.; MacPherson, N.; Gariba, S. and Ofir, Z. (2012) 'Evaluating Impact Investing: An Emerging Area of Development Evaluation Practice', panel presentation to the 6th biennial conference of the African Evaluation Association, International Conference Centre, Accra, Ghana, January 2012 (*Evaluation/assessment)**

This presentation provides useful insights to our study on several levels. First, it provides a good context and definition of impact investing. Then, it presents features which make measurement approaches compelling, before listing a number of approaches that can be used in impact investing evaluation itself.

Indeed, first, the presentation situates the emergence of evaluation in the impact investing sector in a context of *pressure to reduce Western aid budgets* and the *need and opportunity to lever private sector capital to reduce poverty, fight climate change and increase human wellbeing* (the focus on improving human wellbeing is repeated in Picciotto (2011) and Rodin and MacPherson (2012) – see below). Then, definitions of 'impact investing' are provided, as well as facts and information on the impact investing field such as examples of impact investments. Further, issues to resolve in this field going forward are mentioned, which include understanding and resolving different cultures and practices around the impact investing and development evaluation fields, balancing metrics, data and performance assessment at centralised and decentralised levels, and making the business case for social return on investment (SROI).

Second, the presentation lists the most relevant evaluation methods in impact investing, namely: methods that provide for measurement, accountability and understanding through learning customised to each assignment; methods that can be applied at multiple levels (from individual to the system); methods that distinguish between groups (e.g. for equity issues); methods that deal with systems and emergence; and tools that provide for assessing financial returns and business performance as well as viability.

Finally, the document lists a range of evaluation methods that can be used in impact investing and provides a potential rationale/aim for their use. It also provides references for each method proposed in the presentation. These rationales and methods for impact investment evaluation include, among others: *clarifying the logic and assessing why, under which conditions, etc.* (methods can include theory-based evaluation (theory of change analysis) and realistic evaluation), *making provision for emerging, turbulent or fast-changing contexts* (developmental evaluation is proposed as a method), *understanding and assessing (un)expected impact* (possible methods include 3ie, NONIE, IE4ID, InterAction Guidance, and contribution analysis) and *valuation beyond direct financial return* (includes methods such as SROI, environmental impact assessment and EVAS (Expanded Value Added Statement)).

- 29. Knickman, J.R. and Hunt, K.A. (2007) 'The Robert Wood Johnson Foundation's Approach to Evaluation', in S.L. Isaacs and D.C. Colby (eds), *The Robert Wood Johnson Foundation Anthology: To Improve Health and Health Care Volume XI*, San Francisco: Jossey-Bass (*Evaluation/assessment)**

This document details the Robert Wood Johnson Foundation's (RWJF) (which focuses on health) approach towards evaluation of its philanthropic endeavours. The article starts by giving some background into philanthropic impact assessment, stressing the growing importance that outcomes and effectiveness assessments and evaluations have gained within philanthropy over the past 20 years. The main part of this report describes the four tiers of assessment that RWJF uses to assess its own impact.

The first tier involves measuring the impact of specific programmes, where preference is given to evaluations based on comparison. An example of the implications of an impact assessment is given (i.e. influencing policy to adopt a certain programme across the United States), and lessons such as the importance of timing in evaluating programmes (e.g. to influence next steps of programme planning) are provided.

The second tier focuses on tracking the impact of programme portfolios, where short-term (one year), medium-term (two to three years), and long-term (three to ten years) goals are evaluated on a yearly basis. An example of programme portfolio assessment is conducting an evaluation of programmes related to decreasing smoking.

The third tier is the 'scorecard', which analyses organisational effectiveness. It is based on a concept for measuring performance in the business world. Yet in the context of RWJF, it is deployed by asking grantees, experts in the field of health and other stakeholders about the performance of RWJF on various issues, namely programme impact, programme development, and customer service, for example.

The fourth tier is about publishing and sharing the results of the Foundation's evaluation. The organisation has over 2,000 published reports, and it reports more concise anthologies aiming to share lessons of organisational practice with specifically targeted health-care professionals.

30. Lankester, K. and Pease, M. (2013) *Collecting Impact Data Using Mobile Technology: A Pilot Project Measuring Client Poverty Levels, Global Impact Investing Network (GIIN) (*Ratings/measurement, *Evaluation/assessment)*

This report provides results from a pilot study aimed at evaluating the success of using various ways of collecting data related to the Progress out of Poverty Index (PPI). It seeks to know whether the use of mobile technology in Kenya is a good way of collecting data on various poverty indicators or not. The experiment is set up such that three cohorts of 100 individuals each (though not 100 individuals were always found for each round of the experiment) are encouraged to fill out the PPI survey via their mobile phones in three different ways (i.e. through airtime bonus³ and advance outreach⁴: cohort 1; advance outreach only: cohort 2; and airtime bonus only: cohort 3). This experiment is conducted three times (thus a total of nine rounds). It is conducted each time on three companies which have differing levels of contact with clients (i.e. light/occasional contact, medium/regular contact, and high/frequent contact). The surveys are sent to clients of these three businesses.

The results indicate that using a combination of advance outreach and airtime bonus is much more effective at getting respondents to fill out the survey than using either of these treatment arms separately. They also found that the relationship between the businesses who were used as intermediaries to give advance outreach and the clients themselves is also very important. The implications are that short message service (SMS) surveys can be envisaged as a useful monitoring or surveying tool, yet they can be improved if context,

³ Airtime bonus: Respondents were sent a short message service (SMS) that asked them to take part in the PPI survey. It explained that they would receive a small direct transfer of phone airtime at the end of the survey if they did so. If they chose to participate, each answer triggered the next question sent via SMS. At the survey's conclusion, respondents received an airtime bonus worth 30 Kenyan Shillings (US\$0.35), the equivalent of 30 SMSs or 10–15 minutes of airtime, to thank the respondent for participating. The airtime bonus amount was chosen based on the experience of one of the implementing partners, which showed that incentives can be fairly small and they increase response rates.

⁴ Advance outreach: Before clients were sent an SMS asking them to fill out the survey, staff members of the companies made the clients aware of the survey and asked them whether they would like to participate. If clients opted to participate, their phone numbers were recorded and they were sent confirmation messages of their participation. Clients were made aware and asked to participate in the survey through usual company-client interaction channels, which did not require additional time.

proper translation and more frequent SMS communication with clients are better understood, implemented and utilised.

- 31. Liao, P.-C.; Xue, J.; Lui, B. and Fang, D. (2014) 'Selection of the Approach for Producing a Weighting Scheme for the CSR Evaluation Framework', *KSCE Journal of Civil Engineering* 19.6: 1549–59 (*CSR/CSP, *Evaluation/assessment, *Performance management)**

This article provides a technical analysis of Corporate Social Responsibility (CSR) evaluation, focusing on determining the most adequate weighting scheme of key performance indicators (KPIs) within an evaluation framework used in the construction industry. The rationale for carrying out this study is that weighting schemes, which according to the authors are crucial to *performance management*, and evaluation metrics are usually developed subjectively and arbitrarily, as a result of a lack of objective learning for CSR management.

The article provides some background information on CSR-related evaluation, mentioning the existence of a Dow Jones Sustainability Index, as well as the most commonly used Domini 400 Social Index and Ethibel Sustainability Index to rate the sustainability of publicly traded companies. Yet the focus of these indices is to stress the financial side of companies considered 'sustainable'. Indeed, at the beginning of the article, the authors shed light on the rationale for companies to implement CSR activities, which is to enhance corporate competitiveness. Yet other paradigms around the rationale for CSR practice are proposed as well.

The methodology compares and assesses the adequacy of the three main weighting processes highlighted in the literature (Analytic Hierarchy Process (AHP), Fuzzy Analytic Hierarchy Process (FAHP), and Analytic Network Process (ANP)). This is done by applying these approaches to the CSR activities of two different companies (case studies), with each weighting exercise being reviewed by experts. After the three approaches are assessed according to sensitivity and robustness to the priority of the metrics, ANP is selected as the more robust and practical weighting scheme. Overall, although the article may provide very specific information about weighting processes, this information may be relevant more broadly in the context of the development of performance management strategies. The article also provides interesting and relevant information on the rationale and functioning of CSR activities.

- 32. Liket, K. and Maas, K. (2015) 'Strategic Philanthropy: Corporate Measurement of Philanthropic Impacts as a Requirement for a "Happy Marriage" of Business and Society', *Business & Society* 1–33 (*(Strategic) Philanthropy, *CSR/CSP)**

This article attempts to explain the factors influencing the use of strategic philanthropy among companies. The authors begin by giving some theoretical background to strategic philanthropy, and particularly reasons why companies engage in it. Beyond the theoretical contributions of the article on strategic philanthropy, however, our review has little to gain from it, since it provides very little information on how social impact related to companies' activities is measured – indeed the authors actually state that 'more efforts are needed to reveal the methods and indicators used by companies in measuring this impact' – and it only links strategic philanthropy with the presence of corporate social performance measurement.

33. Liket, K.C.; Rey-Garcia, M. and Maas, K.E.H. (2014) 'Why Aren't Evaluations Working and What to Do About it a Framework for Negotiating Meaningful Evaluation in Nonprofits', *American Journal of Evaluation* 35.2: 171–88 (*Evaluation/assessment)

This article makes the case that non-profits are increasingly engaged in the evaluation of their organisations, programmes and projects (due in large part to greater concerns of accountability), but that the utilisation of these evaluations is often low with many organisations finding themselves 'drowning in data' that do not contribute to their decision-making. As such, the article attempts to shed some light on proper ways and approaches to effectively develop evaluation methodologies useful to non-profits.

The article breaks down and explains problems around each of the components of an evaluation methodology, such as specifying the evaluation purpose, posing the right research question and choosing the right evaluation design (e.g. having a problematic theory of change that may answer how to evaluate a project, but not what to evaluate in that project), and then applies key lessons (in effect composing an evaluation framework) to a case study evaluation.

The main recommendation of the article is for greater participation of organisations of stakeholders in the evaluation process so that they may contribute to the design and evaluation questions that really assess what the organisation is interested in. Otherwise, evaluations have little chance of being useful to organisations. Overall, this article makes some relevant theoretical contributions to the evaluation process of non-profits.

34. McKinsey & Company (2010) *Learning for Social Impact: What Foundations Can Do*, White Paper, London: McKinsey & Company (*Evaluation/assessment)

This paper essentially suggests ways to improve impact assessment within the impact investing sector, in part by contributing a rationale to conducting impact assessment. It brings up several points of contention in existing evaluation practice of non-profit investments, such as evaluation for learning rather than strictly for 'accountability'. In fact, the authors state that evaluation should be used for both.

The paper is divided into three sections. The first focuses on the rationale for designing and conducting impact assessment of social investments. The second describes the emerging trends in measuring social impact – notably, it mentions various social impact assessment tools and networks, yet none of which are new (i.e. B Impact Ratings System, GIIRS, IRIS, Pulse, GIIN and ANDE), as well as the TRASI (Tools and Resources for Assessing Social Impact), which lists over 150 different tools for assessing social impact. The third proposes ways of moving the sector forward, namely by suggesting that the sector agrees on a limited number of common metrics in order to enable greater sharing of information and streamlining of results.

35. Moore, K.A.; Walker, K. and Murphey, D. (2011) 'Performance Management: The Neglected Step in Becoming an Evidence-Based Program', in L. Weiss and C. Collins (eds), *Leap of Reason: Managing to Outcomes in an Era of Scarcity*, Washington DC: Venture Philanthropy Partners (*Performance management)

This chapter discusses the relevance of performance management in general, and not necessarily in relation to activities emanating from impact investments.

One of the main arguments that it makes is that too much impact assessment is currently focused on randomised controlled trials or quasi-experimental methods, while an important

tool useful in informing randomised controlled trials or other evaluations is being ignored. Indeed, the authors state that whole programme evaluations would be premature without the implementation of a *performance management* system. Although the authors admit that most organisations implement a performance management system, they do not take full advantage of it. As such, six different components to comprehensive performance management are provided (including ‘the capacity to collect and analyse information on an ongoing basis’ and ‘a willingness to modify programs to address challenges’). Before concluding, the authors mention a case study of a performance management tool specifically developed for an education-related programme (this tool does not seem transferable to other, especially non-education, programmes but the case study still shows innovation in the development of a performance management system).

36. Morino, M. (2011) *Leap of Reason: Managing to Outcomes in an Era of Scarcity*, Washington DC: Venture Philanthropy Partners (*Evaluation/assessment)

On the whole, the book (including the ‘essays’, written by various other experts and practitioners) explores many different aspects of social impact assessment, yet these aspects are explored in a more reflective (and not very deep) manner, offering little in the way of concrete tools to assess social impact. Concerning the first part of the book (the part written by the author), it is essentially a think piece on how to foster a culture of learning from programme implementation, performance management, and evaluation to improve programme implementation and design. Among other things, the author (like many others) emphasises the role that social impact assessment should play in increasing the benefits of programme activities to society (intended beneficiaries), rather than simply being used as an accountability tool – where measurement has become an end in itself.

Of potential interest is a glossary of terms that the book provides in the first chapter, and a listing of the efforts/resources that have emerged around improving social impact assessment (yet no list is provided of particular tools that are being used in assessing social impact).

37. Newmark, T.E. and Pena, M. (2012) *Portfolio for the Planet: Lessons from 10 Years of Impact Investing*, Abingdon and New York: Earthscan (*M&E)

The main section of interest of this book (pages 111–15) details a monitoring and evaluation (M&E) tool of an environment-focused impact investment fund, called EcoEnterprises Fund. The M&E tool (questionnaire) is included in the appendix of the book (Appendix A). There are essentially three components tied to the rationale and design of the M&E, i.e. so it can be used in the following ways according to the authors: ‘as a checklist and guide for investment officers and consultants to use during due diligence and monitoring visits; as an addendum to legal agreements that defined the organisation’s [EcoEnterprises Fund’s] environment and social expectations; and as a template for portfolio companies to help streamline their reporting to us [the Fund] and to certifying agencies’.

One approach the organisation attempted to use in designing the tool was to look to existing certification schemes to borrow indicators to track social and environmental performance and impact, in order to ‘hold companies to the same standard’; even though, it found only a few indicators on biodiversity conservation. As such, in addition to drawing on some indicators from existing cross-sector assessment schemes, interestingly and usefully, the book highlights that the organisation sought to compliment these indicators by developing an individually tailored M&E programme for each portfolio company with indicators reflecting their particular sectors and environmental and social objectives. It also established indicators for which data collection was manageable, trying not to demand too much data from businesses.

Finally, the book states the frequency with which the M&E tool is used (twice a year), and by whom (once by a staff member and once by an external party). It also underlines the refining of the M&E tool after a few years (i.e. two years) as irregularities in reporting (e.g. giving different answers for the number of employees depending on the respondent) have been smoothed out. Lastly, it mentions a Shared Impact Assessment Methodology and Toolbox, developed by FAST (Finance Alliance for Sustainable Trade).

38. Olsen, S. and Galimidi, B. (2008) *Catalog of Approaches to Impact Measurement: Assessing Social Impact in Private Ventures, Social Venture Technology Group* (*Ratings/measurement)

This report focuses on providing information on assessments of impact investments in private and non-profit firms. The report is particularly relevant to our study as it provides a catalogue of 25 different approaches used in assessing social impact of investments (although the report has the limitation of providing older approaches, some of which, however, are definitely still in use). These approaches are detailed, each including among other things a brief description of their approach, methodology, scope of analysis, and applications to date (summarised in 1.5–2 pages).

Additionally, the approaches are categorised in a table defined according to the functions of impact assessment approaches; i.e. *screening* (e.g. ratings systems), *summarising results* (assessment systems), and *ongoing tracking* (management systems). That table is presented on page 19. Also, a table of the stages of impact assessment is given on page 18, defining *implied impact* (storytelling + internal data analysis), *proven impact* (external data analysis + experimental analysis), and *optimised impact* (*proven impact* + interrelationship with financial performance). The stages do not imply chronology of impact assessment within an organisation, but rather their credibility. These stages are mentioned in 'the ways to measure impact' from the Tides (2011) document.

39. Picciotto, R. (2011) *Evaluating Development Philanthropy in a Changing World*, Commissioned Paper, Brighton: IDS, The Resource Alliance and The Rockefeller Foundation (*Strategic) Philanthropy, *Evaluation/assessment)

This paper details the steps to take to develop philanthropy evaluation. The first part, however, provides an overview of the changing aid and philanthropic landscape, stating that as public expenses into aid are under increased pressure due to the recession, private giving is rising, leading to an 'emerging aid architecture'. One of the main points of the paper and this emerging aid architecture is that, according to the author, international development philanthropies should cooperate more closely with other development partners (including developing countries' governments) and be more deeply integrated into global development initiatives, which would be in line with the new global development paradigm of improving human wellbeing. The article also mentions the various evaluation approaches linked to the different dimensions of human wellbeing (material – cost-benefit analysis, relational – participatory evaluation, and perceptual – empowerment evaluation).

In the latter half of the paper, the author proposes various ways of improving evaluation practice itself within development philanthropy. One suggestion he makes is to go beyond the anecdotal evidence of successful projects and social enterprises tied to development philanthropy, and rigorously and independently evaluate these interventions. The author then advocates for greater use of participatory methods to be included in philanthropic evaluation in order to empower beneficiaries of philanthropy. One method of improving the involvement of participatory methods would be to make better use of mobile technologies in developing countries; these would provide better real and real-time information, as well as increase the speed and amount of data which could be processed. Finally, other measures to improve

philanthropic development evaluation would be to include management better in the design of evaluation (since some board members themselves reported too often rubber-stamping pre-approved and independently designed evaluations), advocate for more systematic integration of evaluation within project design so as to avoid costs of conducting complex and expensive *ex-post* evaluations which few organisations on the ground (i.e. non-governmental organisations) can afford, and pilot testing indicators before deploying them in monitoring and evaluation systems.

40. Porter, M.E.; Hills, G.; Pfitzer, M.; Patscheke, S. and Hawkins, E. (n.d.) *Measuring Shared Value: How to Unlock Value by Linking Social and Business Results*, Boston/Geneva/Mumbai/San Francisco/Seattle/Washington DC: FSG (*Evaluation/assessment, *Ratings/measurement, *SROI)

The report makes the link between social performance and its effect on economic value for the business. It states that businesses measure social and environmental impacts, and that there is a movement to add sustainability measures to financial reports; yet crucially to the report, there is still a missing framework to link social and business performance to each other. One example of creation of shared value is through 'redefining productivity in the value chain', namely by reducing water use (social result) and thereby reducing logistical and operating costs, and improving profitability.

One benefit from the shared value measurement process is that strategic priorities inform the focus and extent of shared value measurement, and the data and insights from the value measurement inform refinement of the shared value strategy – this constitutes an ongoing feedback loop. The shared value measurement process (whose goal is to unlock further value creation) is described in four steps, and case studies of the implementation of these four steps are also provided further in the text. One tool which is mentioned in assessing organisations' environmental, social and governance (ESG) performance is the Global Reporting Initiative (GRI), allegedly used by 3,500 organisations in more than 60 countries.

Different types of shared value measurement are then listed (sustainability, impact assessment, reputation and compliance) as well as examples of what can be measured (e.g. efficiency in the use of input factors). Further, various ways are proposed to establish a link between social and environmental impacts of an organisation and market value, namely correlating ESG performance and the stock performance of a company; monetising environmental externalities of companies/organisations; and using the social return on investment approach. Finally, various challenges are mentioned regarding measuring shared value, such as the lag between social value and financial value creation.

As such, this guide can be interesting for thinking around ways to link and measure financial value from social value creation, yet few concrete tools are proposed in measuring and establishing the value of social performance.

41. PwC (2013) *Measuring and Managing Total Impact: A New Language for Business Decisions*, PwC (*Evaluation/assessment, *Ratings/measurement)

This report presents a new way of accounting for impacts entering the mainstream investing and business world. Indeed, the report recognises various types of impacts which have changed the business landscape, namely *social, environmental, tax* and *economic* impacts. Together, these constitute total impact, and the report advises on how to measure and manage them (TIMM – Total Impact Measurement and Management). Table 2 (on page 24) mentions the rationales and general guidelines to the measurement of each impact, while few concrete frameworks and indicators are provided (except for tax impact, where reference to a 'Total Tax Contribution Methodology' is given). Still, a five-step process to applying the

TIMM is given on page 26, and more importantly, the TIMM is applied in a detailed manner to a hypothetical example (pages 28–29). Finally, the last chapter (before the conclusion) provides information on the existing use of the TIMM, and how to bring this approach further into the mainstream.

42. Rockefeller Philanthropy Advisors (n.d.) *Assessing Impact*, New York: Rockefeller Philanthropy Advisors (*Evaluation/assessment, *Ratings/measurement, *(Strategic) Philanthropy)

Overall, this is quite a basic introduction to philanthropic investment impact assessment – the report describes the main approaches used in assessing impact (*evaluation*, *return on investment* and *systems thinking*), and provides examples of case studies to illustrate the use of each approach. Each approach is summarily described, providing a brief rationale for their use as well as possible forms they can take (e.g. evaluation can take the form of *goals vs outcomes*, *process evaluation*, *experimental design*, and *cluster evaluation*). This report is of little use to our review since it provides little information on developing evaluation designs effectively for each approach, nor does it provide much information on new and insightful approaches used to carry out philanthropic impact assessment.

43. Rodin, J. and MacPherson, N. (2012) ‘Shared Outcomes’, *Stanford Social Innovation Review (Advancing Evaluation Practices in Philanthropy, Sponsored Supplement of the SSIR) 10.3: 12–15* (*Evaluation/assessment)

This article draws on Robert Picciotto’s 2011 article (see entry 39 above) on philanthropic evaluation in the development context, and particularly his call for greater collaboration on the matter of philanthropic evaluation and increased focus on improving human wellbeing. In January 2012, at a conference in Ghana, a group of developing-country evaluators highlighted five strategies to increase foundations’ and development agencies’ influence in contributing to meaningful social transformation. While the article describes these five strategies, it provides the Rockefeller Foundation’s response to this ‘call to action’ and further describes the five strategies that the Foundation has adopted in its approach toward philanthropic evaluation. The Rockefeller Foundation’s strategies correspond closely to those outlined by the development-country evaluators, and are: (1) engaging stakeholders to develop shared outcomes; (2) expanding capacity through use of non-staff monitoring and evaluation specialists to partner with grantees; (3) sharing knowledge through learning forums and communities of practice; (4) strengthening developing country evaluation practice and ownership of results; and (5) developing innovative methods and approaches to evaluation and learning. Regarding monitoring and evaluation (M&E), the authors state that the Foundation provides a network of ‘critical friends’ in the M&E sector by awarding grants to evaluation specialists in developed and developing countries to make up for a lack of resources attributed to M&E in developing countries. Regarding the point about communities of practice, the article mentions various communities of practice in the developing world, such as the South East Asia Community of Practice in Evaluating Climate Change Resilience (SEA Change). And, regarding new methods and approaches to evaluation, the article usefully points out the example of Ushahidi, an open-source crowdsourcing project enabling individuals to directly report acts of violence through their mobile devices, granting them a place within the evaluative process and thus making this process more democratic.

The article concludes by suggesting four ways to reshape evaluation practice, namely so that grantees and partners in developing countries can better respond to global change and to fulfil their missions and goals more effectively. This, in turn is inspired by an emphasis on the value of evaluation being judged by ‘its usefulness to improve outcomes for target beneficiaries’, a call oft-repeated in the emerging literature on social impact assessment.

44. Rubin, J. (2011) *Evaluation of the Social Impact Bond: Lessons from Planning and Early Implementation at HMP Peterborough, Cambridge and Brussels: RAND Europe (Evaluation/assessment, SIB)*

This document gives a very brief (three pages) and summarised account of the first (process) evaluation of the HMP Peterborough social impact bond (SIB). It provides background information on the SIB, and of most interest to us, highlights the key features of the evaluation, including the use of the chosen outcomes to measure the impact of the intervention. It also mentions the processes to keep in mind when developing an SIB evaluation (i.e. the process of finding a robust outcome measure agreed upon by all stakeholders was time-consuming, which should be kept in mind in the development of future SIBs) and the payment scheme.

In relation to outcome measures, a brief comment is made on their selection. Namely, they enable:

- outcomes to be attributed to the SIB-funded intervention;
- reduced incentives to 'cherry pick' and work with offenders who are easiest to engage;
- ensured statistical significance.

45. Rudd, T.; Nicoletti, E.; Misner, K. and Bonsu, J. (2013) *Financing Promising Evidence-Based Programs: Early Lessons from the New York City Social Impact Bond, New York and Oakland: MDRC (Evaluation/assessment, SIB)*

This report describes the evaluation (but not its findings as they have not yet been released) of a social impact bond (SIB) whose aims are to reduce recidivism among 16–18-year-old inmates at Rikers Prison in New York City (the payment target is set at a 10 per cent reduction in the rate of recidivism). The report has three main sections of interest to us. The first one describes the evaluation method which has been selected to assess the SIB, the second describes the evaluation method's selection process, and the third provides theoretical considerations in choosing an evaluation method for an SIB.

First, regarding the evaluation method used in this SIB, it is a quasi-experimental method, comparing a cohort of adolescents receiving the intervention with a historical cohort that did not receive the intervention. It uses propensity score matching to control for individual-level differences, and to track recidivism in 19–20-year-olds (who are not eligible for the intervention) to control for systemic changes that may influence the outcomes (e.g. if arrest patterns or detention practices change). Figure 2.5 on page 22 shows the timeline of the project and evaluation in clear fashion – adolescent offenders are tracked for two years after their release.

Second, regarding the selection process of the evaluation, the method was chosen through the collaboration of the City of New York, Bloomberg Philanthropies and MDRC (the 'intermediary'). Tasked with identifying the most appropriate evaluation method, MDRC first conducted a literature review to see which interventions worked. From the review, it was realised that 'MRT [Moral Reconciliation Therapy] was the most studied of all interventions examined and was ultimately selected because the available evidence suggested it could best accommodate the unique operational challenges of high turnover on Rikers Island.' A pilot of the intervention was also conducted, which included the training of a cohort of teachers and their carrying out of MRT, but there is no mention of what the results of the pilot were. Presumably, the implementation of the intervention was deemed adequate during the pilot, and was allowed to continue into full implementation. Interestingly, MDRC also monitors the programme, both its participation and smooth operation.

Third, considerations in selecting an evaluation approach for an SIB mention the inclusion of a reliable comparison or control group – indeed the report states that it is the only way to know that the government is paying for success (thus value for impact is essential to an SIB). *SIBs also need to be expansive and thus relatively large-scale in order to both enable cost-savings and to have enough statistical power to detect impact.* Again, the emphasis on the importance of evaluation rigour is reiterated in the case of SIBs, as accountability and learning are both particularly important in this context, and since a proper evidence base needs to be built up, especially since some interventions are innovative and not necessarily proven to work (but based on strong theoretical bases). Finally, some potential risks are mentioned in the evaluation of SIBs, namely that evaluation may suffer as the focus may be on raising sufficient capital to make the investment possible (mentioned in Drew and Clist 2015), *and that there may be political pressure on the need to show the success of SIBs, therefore reducing the incentive to engage in serious evaluation.*

46. Salazar, J.; Husted, B.W. and Biehl, M. (2012) ‘Thoughts On the Evaluation of Corporate Social Performance through Projects’, *Journal of Business Ethics* 105.2: 175–86 (*Evaluation/assessment, *CSR/CSP)

This is a very insightful article on the evaluation of corporate social performance. It is divided into three broad sections, described below.

The first section provides a theoretical overview of Corporate Social Responsibility (CSR) and corporate social performance (CSP), defining the former as ‘policies and practices of corporations that reflect business responsibility for some of the wider societal good’ (Matten and Moon 2008: 405) and the latter as ‘the results of those policies and practices’. It is interesting to understand the rationale for shedding light on CSP. According to the authors, the impact of CSR on social outcomes has been ‘disturbingly understudied’; indeed, this has not been the focus of CSR and CSP for corporations (rather the focus has been on linking CSP and financial performance), since as one manager stated ‘the need is so great, anything we do is helpful.’ There has thus been little preoccupation with the actual extent of impacts that CSR activities have on targeted beneficiaries. Further, the authors additionally remark on the poorly developed state of CSP regarding social outcomes, stating that there is essentially no agreement on ‘what’ types of outcomes should be measured, and ‘how’ they should be measured. Because the authors argue that the primary reason to measure social performance is to improve social outcomes, they advocate for Senian principles in measuring the ‘what’; namely indicators tied to human development (income, health and education). Regarding the ‘how’, the authors advocate for pre- and post-testing and the use of quasi-experimental methods, since using experimental methods is not always possible yet quasi-experimental methods are preferable to cross-sectional methods.

The second section describes the evaluation of Patrimonio Hoy (PH), a CSR initiative by the Mexican cement company Cemex, which offers microcredit for do-it-yourself housing construction in low-income areas. Until this study, there had been no studies of the actual social outcomes of the initiative, despite the high visibility and widespread knowledge of the project. Because no evaluation had been included in the project design, the evaluation relied on propensity score matching methodology to establish comparison groups (treatment group n= 62; control group n= 175). The measured indicators can be seen in the appendix (Table 2), and the results of the case study are given on page 181.

Finally, the last section provides a discussion and conclusion. In terms of discussion, the PH case study illustrates that corporations are more interested in being well perceived as good corporate citizens than actually alleviating poverty, as proper goals, indicators and measures of social outcomes were not even developed along with the CSR activities.

47. Social Finance (2014) ‘Peterborough Social Impact Bond Reduces Reoffending by 8.4%; Investors on Course for Payment in 2016’, press release, 7 August 2014, London: Social Finance Ltd (*Evaluation/assessment, *SIB)

This document is a press release announcing the results from the first cohort from the HMP Peterborough social impact bond (SIB). The most notable result is the fact that there was a reduction of 8.4 per cent in reoffending rates for the first 1,000 inmates released from prison two years after their release (no result of the outcome for other cohorts has yet been released). Further information on the propensity score matching (PSM) design is given in this press release, namely that the PSM was based on 36 characteristics.

The brief also sheds more light on the debate around the particularity of HMP Peterborough, where the Independent Assessor (IA) had noted *a lowered likelihood of reoffending among released prisoners at HMP Peterborough based on a dry run in 2012 using 2008 data, and based on overall data from 2007 to 2013*. However, interim reports by the Ministry of Justice (MOJ) consistently suggested the reverse, yet the IA still states that the potential effects of being released from a certain prison was not controlled for in this SIB. As such, it was determined that the MOJ and Social Finance will seek independent advice on whether to refine the PSM methodology when evaluating the second cohort.

48. Social Impact Investment Taskforce (2014b) *Measuring Impact: Subject Paper of the Impact Measurement Working Group*, Social Impact Investment Taskforce (*Ratings/measurement)

This report from the Social Impact Investment Taskforce attempts to shape the framing of impact investing measurement into the future. Among other things, it strongly advocates for conventions and commonality in impact measurement, indeed to harness and leverage the power of measurement in impact investing initiatives around the world. The importance of this, according to the report, is that impact measurement guidelines would help organisations hold themselves accountable to those they serve, as well as attract more capital by determining the positive impacts of an investment on society and the environment. While the first chapter provides an overview of the state of impact measurement today, the report then focuses on the various guidelines of good impact measurement practice that it proposes, which are targeted towards the global impact investing measurement community.

The guidelines are predicated on four phases of impact measurement (*plan, do, assess, review*), and seven steps (effectively constituting the seven guidelines) found within these phases of impact measurement. The seven guidelines are: set goals, develop framework and select metrics, collect and store data, validate data, analyse data, report data, make data-driven investment management decisions. Unfortunately, although several of these guidelines would seem particularly relevant to our review (e.g. develop framework and select metrics), concrete ways of developing and implementing each guideline are not detailed. Still, Chapter 3 provides five examples of case studies of impact investments and their measurements, providing some but still limited insights into the general implementation of the measurement guidelines. Nevertheless, of particular interest to our review is the description of the measurement framework described around the New York State social impact bond.

Finally, likely trends to affect impact measurement of impact investments, and data qualities important to forming a convention around impact investing measurement are proposed and detailed in the final chapters of the report. Moreover, various resources of interest to our review are mentioned and listed in the report, such as the ‘promising existing initiatives’ to impact measurement which are listed on page 25, the comprehensive glossary of terms found in Appendix 1, and the other existing frameworks and metrics listed in Appendix 5.

49. The Centre for Social Impact (2011) *Report on the NSW Government Social Impact Bond Pilot*, Sydney: The Centre for Social Impact (CSI) (*Evaluation/assessment, *Ratings/measurement, *SIB, *SROI)

This report does not actually provide much insight into methodologies used in a social impact bond (SIB); rather it reports on what infrastructure is necessary to implement SIBs in New South Wales. It provides information on the various methodologies necessary to implement an SIB, namely cost-benefit, cost-effectiveness and/or social return on investment analyses (to assess costs and risks of payment and of an intervention). It also provides information on methodologies to use to evaluate the pilot of an SIB; the authors advocate for the use of a well-developed quasi-experimental study. Finally, it provides input on suggested auditing and performance reporting bodies, namely experts that can carry out evaluation, performance measurement, social impact measurement, etc.

50. Tides (2011) *Assessing Social Impacts of Investments: Developing Effective Evaluation Systems to Measure Social Returns*, Tides (*Evaluation/assessment, *Ratings/measurement)

This report offers several suggestions on establishing systems to evaluate impact investments specifically. It begins by providing steps on planning for the establishment of an impact assessment system. Though most of these steps are familiar to evaluation experts, it does include *calculating the expected return* (expected return = (benefit x likelihood of success)/cost) as one step which seems particularly relevant to the assessment of *impact investments*. Steps on 'conditions to support an impact assessment system' are then provided (such as researching what data and systems already exist).

In terms of 'ways to measure impact', this report refers to various stages of impact assessment (*implied impact, proven impact, optimised impact*), yet no further information on putting these into practice is offered. However, the report does include financial performance (defined as cost-effectiveness) calculation (cost per impact = measured impact/investments made by all sources to realise the impact) in the measurement of impact.

Further, theoretical recommendations on formulating impact assessment systems are made, such as making sure these systems are culturally relevant, as well as flexible to amendment (including the sets of indicators, outputs, outcomes), particularly in the case of innovative and experimental programmes. Coordination of programme implementation and evaluation to avoid siloing is also encouraged.

Usefully, examples of seven organisation/company impact assessment systems are then presented along with short descriptions (the links to their websites are also included). These include familiar systems such as the Global Impact Investing Rating System and Impact Reporting and Investment Standards, but also include others such as environmental, social and governance (ESG); Global Reporting Initiative; Independent Sector; International Finance Corporation; and New Philanthropy Capital.

Finally, the report mentions challenges in using impact assessment systems, such as the fact that measurements which are put in place can reflect what a funder specifically required, yet proxy information and investment in evaluation can improve an organisation's evaluation capacity. Interestingly, there is also mention of Mulago Foundation's impact assessment design model, which really focuses on what an organisation tries to do (e.g. describing in eight words or less an organisation's mission), and advocates using a minimal (sometimes even only one – the best) set of indicators to document impact, etc. This design model could prove insightful to our review.

51. Triodos Facet (2010) *Final Report: Evaluation PSOM/PSI 1999-2009 and MMF*, Zeist: Triodos Facet (*Evaluation/assessment)

This document reports the findings of an evaluation of two programmes (yet three components) related to impact investment supported by the Dutch government (in the form of subsidies – this constitutes another form of government-supported impact investment):

- The PSOM/PSI (Private Sector Investment programme), which aimed to reduce poverty by stimulating sustainable investments (by Dutch and foreign entrepreneurs) in innovative businesses in selected developing countries;
- The MMF (Matchmaking Facility), aimed at matching entrepreneurs in developing countries with Dutch firms to develop a sustainable relationship and hopefully lead to joint investment or another form of trading relationship;
- The third component of the evaluation is the assessment of the performance (efficiency and effectiveness) of the implementing agency – EVD (agency of the Dutch Ministry of Economic Affairs for international business and cooperation).

The research questions posed by the evaluation are found on pages 2–3 of the report, on themes related to ‘efficiency’, ‘effectiveness’, ‘relevance’ and ‘additionality’. The methodology (Section 1.3) used in the evaluation employed a mix of quantitative (gathering data based on indicators for inputs, outputs, outcomes and impact) and qualitative methods (telephone interviews and desk research, staff interviews, field visits to projects, etc.), all based on a sampling of projects.

Findings on the evaluation for each specific programme are provided in detail in the report and are summarised in the conclusion. Interestingly, most of the results point to effectiveness (e.g. whether a significant amount of jobs have been created as a result of the PSI or not) and efficiency (e.g. regarding the costs of implementation by EVD).

52. Tuan, M.T. (2008) *Measuring and/or Estimating Social Value Creation: Insights Into Eight Integrated Cost Approaches*, Bill & Melinda Gates Foundation (*Ratings/measurement, *(Strategic) Philanthropy)

This paper aims to inform effective ways of integrating measuring and/or estimating social value creation in philanthropic activities. It does so by describing and analysing eight approaches which seek to measure and/or estimate social value creation within philanthropy. Most of the new methodologies reviewed draw on concepts from cost-effectiveness and cost-benefit analysis, yet the author states that there is ‘**no perfect methodology**’ to measuring/estimating social value creation. The author further states that at the time the report was written (2008), social programme evaluation (and cost-benefit analysis) lacked maturity. In a 2008 paper, one researcher from RAND attempted to explain this lack of maturity in cost-benefit analysis by the fact that they rarely and do not fully capture the full range of social benefits, and that they are thus not monetised.

The eight reviewed approaches listed on pages 10–13 are of interest (organisations tied to the approaches are mentioned, and the approaches themselves are explained and usefully summarised in two paragraphs), but they are essentially all based on the calculation of ‘expected return’ → $ER = (\text{outcome or benefit} \times \text{probability of success}) / \text{cost}$. Further, information on technical issues and limitations, including assumptions, discount rates, timeframes, shadow prices, interdependencies and value judgments is then provided.

Finally, the report also includes a good analysis of the ‘big picture issues’ in measuring and/or estimating social value, providing ample description of issues such as: inconsistent use of language; lack of common measures in the social sector; lack of quality data on

impact, outcomes, outputs and costs; lack of incentives for transparency; unintended consequences; inadequate utilisation; and cost of measurement.

Issues and implications are then summarised in Section 6 (from page 24 onwards).

53. Twersky, F. and Lindblom, K. (2012) *Evaluation Principles and Practices: An Internal Working Paper*, The William and Flora Hewlett Foundation (*Evaluation/assessment)

This report outlines The Hewlett Foundation's evaluation principles and practices, i.e. their strategy, which focuses on OFG (outcome-focused grant-making). Although it may not prescribe a concrete tool that may be reproduced to evaluate impact, it does provide some thinking around principles of strong evaluation.

The report starts by briefly describing the Foundation's seven principles of evaluation (including 'leading with a purpose', i.e. knowing why something is being evaluated and how the results will be used, and 'evaluation is a learning process', i.e. evaluations help test the logical or illogical assumptions along the theory of change and provide new knowledge). The report further emphasises the importance of theories of change to their strategy, as well as having a flexible evaluation framework (it seems that much of the newer literature emphasises this point), which requires revisiting over time. And, further, another issue of interest to our review is the report's mention of 'strategy evaluation', which evaluates multiple interventions or clusters within the same strategy. According to the authors, this is *relatively new* to the field of evaluation, and as such there is no prescribed way of aggregating these interventions or clusters together. Timing is also noted as a key issue.

In selecting methods, the Foundation believes that the use of multiple methods makes for the strongest evaluations, notably due to the possibility of triangulation. However, the Foundation does recognise that the essence of a good evaluation involves some comparison. Other criteria for success are mentioned, such as establishing a strong link between the implementer and the evaluator, for example, to strengthen data collection, particularly if problems in that area arise.

Of interest also are the guidelines which detail the crafting of a request for proposal (RFP) for an evaluator – the report states that the organisation would ideally choose an evaluator who is strong technically, has subject matter expertise, is pragmatic and communicates well verbally and in writing.

Finally, the report also mentions the Foundation's evaluation of re-granting organisations (i.e. organisations that act as granting intermediaries between the Foundation and small organisations on the ground), posing such evaluation questions as 'Do these intermediaries add value to grantees or are they simply middlemen?' A case study is provided with the International Development Research Centre as the intermediary/re-granting organisation. Lastly, Appendix C provides an evaluation planning tool used by the Foundation to conduct evaluations, which may be of interest.

54. Warner, M.E. (2013) 'Private Finance for Public Goods: Social Impact Bonds', *Journal of Economic Policy Reform* 16.4: 303–19 (*Evaluation/assessment, *Performance management, *SIB)

This article is particularly relevant to our review. It focuses on the context for using social impact bonds (SIBs), debates around performance management and their comparison with service provision through public–private partnerships (PPPs).

Essentially, the findings of the article are that the rationale for using SIBs may be undermining developmental evaluation, namely because the payment system allows little room for programme improvement based on results from performance management. Also, evaluation may be compromised because some SIB schemes, such as the one at HMP Peterborough, cede the control of grantee selection and evaluation to private investors, creating the possibility for collusion. Service provision and learning from interventions may thus be compromised in comparison to PPPs.

Concerning the design of performance management and the overall evaluation design, the author states that evaluation methods of SIBs may focus on traditional positivistic approaches that maintain a control and experimental group. Yet, the author points to a vibrant debate where critical theorists are arguing for increased acknowledgement of perception and context, calling for more collaborative and constructivist approaches to research. Finally, another problem with SIBs is that their design usually allows little experimentation of metrics, due to the need for rigid metrics which can measure tangible and payable outcomes. As such, this limitation somewhat defeats the purpose of SIBs, which are supposed to challenge traditional approaches to service delivery.

55. Woelfel, T.; Thornley, B.; Ren, J. and Sirull, B. (2014) *Invest Northwest Eight-year Social Performance Review: A Comprehensive Case Study on the Impact of a Northwest Area Foundation Mission-Related Investment*, InSight at Pacific Community Ventures (*Evaluation/assessment)

This document reports on Invest Northwest's mission-related investment strategy, which refers to wanting to advance the Foundation's goals, while also receiving a competitive financial return. The document is interesting because it provides a concrete example of the use of Impact Reporting and Investment Standards (IRIS) metrics to track the social performance of investments. Social outcomes that are measured by the Foundation include employment and employment growth among the 11 companies in its investment portfolio, as well as employee wages and job quality.

As part of the results detailed in the conclusion, the Foundation does mention that establishing causality for increased job numbers, for example, is difficult – this may very well point to a weakness of the IRIS, yet the authors assert that the investment of capital has been clearly vital for many Invest Northwest portfolio companies.

Annex Strengths and weaknesses of existing evidence and evaluation approaches and methods related to SIBs and DIBs: a summary

<p>Overall approaches to evaluation</p>	<p>There is a strong commitment to evaluation of SIBs/DIBs from many stakeholders.</p> <p>Some forms of SIB require robust measures of outcome for payment to occur – as a result, SIBs/DIBs may contribute to a greater focus on robust data collection, including through evaluation.</p> <p>Clear expectations of evaluations of SIBs and DIBs are beginning to emerge.</p>	<p>Several manuals on SIBs do not cover issues of evaluation in much detail if at all.</p> <p>Expectations of what evaluations of SIBs and DIBs might produce are widely divergent. It is extremely unlikely that evaluations could meet all these expectations without extremely high levels of resourcing.</p> <p>The need for impact evaluations is not recognised by all SIB/DIB stakeholders and advocates. There is a documented tendency to scale back on evaluations where resources are tight.</p> <p>The value added of a full impact evaluation (over and above the results themselves) is downplayed by some, even in designs where payment is not automatically the same as truly desirable performance.</p>
<p>Evidence base</p>	<p>Some evidence is emerging from experience of SIBs. This is likely to be very relevant to DIBs.</p> <p>Some lessons are beginning to emerge regarding DIB design.</p> <p>Some efforts have been made to synthesise learning from across experience of SIBs. There are plans to do similar work on DIBs.</p>	<p>Much of the evidence about SIBs to date is about process and based on collated opinion.</p> <p>Most of the evidence is through a positive lens, i.e. from advocates of impact bonds. Negative evidence is more limited although there is some. Much of the criticism of impact bonds is un-evidenced opinion.</p> <p>There is little information on development of markets for SIBs, although there is some.</p> <p>There is as yet no evidence base specific to DIBs as there has, as yet, been no implementation experience.</p>

<p>Methods for individual evaluations</p>	<p>There has been some work on categorising available methods for evaluation of SIBs and DIBs.</p> <p>At least one randomised controlled trial of an SIB is underway (in New York).</p> <p>Several SIBs (e.g. Peterborough, Rikers Island) use matched comparison groups as a basis for making payments.</p>	<p>Categorisation of available evaluation methods is incomplete. Other methods could include individual randomised controlled trials, a range of quasi-experimental approaches and the use of theory-based approaches.</p> <p>There has been relatively little use of experimental or quasi-experimental approaches and more reliance on interviews as the main means of primary data collection.</p> <p>Many of the reports of evaluations and reviews lack clear descriptions of methods which would mean that they would be likely to be downgraded in terms of any assessment of the quality/strength of evidence presented. They might be excluded from some forms of data synthesis, e.g. systematic reviews.</p>
<p>Methods for synthesising learning</p>	<p>Efforts to synthesise learning from experience of SIBs has begun.</p> <p>The value of having capacity available to conduct real-time synthesis of evidence from evaluations has been recognised for both SIBs (e.g. the evaluation of the Commissioning Better Outcomes portfolio in the UK) and for DIBs (e.g. the call to establish a Community of Practice).</p> <p>There are a number of reports of SIB experience and evaluations which could be interpreted as a willingness to share information openly and transparently.</p>	<p>SIBs and DIBs are being developed and implemented in very different ways and in different contexts. The applicability of evidence and lessons learned from a SIB/DIB in one context to SIBs/DIBs in other contexts is unclear. Simplistic methods intended to aggregate learning across SIBS/DIBs are unlikely to work.</p> <p>There is reluctance among some stakeholders to talk openly about their experiences and plans. Reasons are complex but include concerns about biasing procurement processes and giving away valuable intellectual property. This reluctance or 'secrecy' is a major potential barrier to synthesising evaluation evidence into a body of evidence.</p>

Source: Drew and Clist (2015).

References

- Bannick, M. and Hallstein, E. (2012) 'Learning from Silicon Valley', *Stanford Social Innovation Review (Advancing Evaluation Practices in Philanthropy, Sponsored Supplement of the SSIR)* 10.3: 8–11
- Bice, S. (2013) *Creating Meaningful Corporate Social Responsibility Programs for Rural Australian Mining Communities: Reflections and Improvements on Current Practice*, www.tasa.org.au/wp-content/uploads/2013/11/Bice.pdf (accessed 4 December 2015)
- Brest, P. and Born, K. (2013) 'Unpacking the Impact in Impact Investing', *Stanford Social Innovation Review*, www.ssireview.org/articles/entry/unpacking_the_impact_in_impact_investing (accessed 4 December 2015)
- Brest, P. and Born, K. with responses from Choi, A.; Speirn, S.K.; Rodriguez Arregui, A.; Chu, M. and O'Donohoe, N. (2013) 'When Can Impact Investing Create Real Impact?', *Stanford Social Innovation Review* 11.4: 22–31
- Cave, S.; Williams, T.; Jolliffe, D. and Hedderman, C. (2012) *Peterborough Social Impact Bond: An Independent Assessment – Development of the PSM Methodology*, London: UK Ministry of Justice
- Center for Global Development and Social Finance (2013) *Investing in Social Outcomes: Development Impact Bonds: The Report of the Development Impact Bond Working Group*, Washington DC and London: Center for Global Development and Social Finance
- Clark, C.; Rosenzweig, W.; Long, D. and Olsen, S. (2004) *Double Bottom Line Project Report: Assessing Social Impact in Double Line Ventures, Methods Catalog*, New York: The Rockefeller Foundation
- Disley, E. and Rubin, J. (2014) *Phase 2 Report from the Payment by Results Social Impact Bond Pilot at HMP Peterborough (Revised)*, London: UK Ministry of Justice
- Drew, R. and Clist, P. (2015) *Evaluating Development Impact Bonds: A Study for DFID*, London: Department for International Development (DFID)
- Ebrahim, A. and Rangan, V.K. (2014) 'What Impact? A Framework for Measuring the Scale and Scope of Social Performance', *California Management Review* 56.3: 118–41
- Flynn, J.; Young, J. and Barnett, C. (2015) *Impact Investments: A Literature Review*, Centre for Development Impact, Brighton: IDS
- Freireich, J. and Fulton, K. (2009) *Investing for Social & Environmental Impact: A Design for Catalyzing an Emerging Industry*, Monitor Institute
- Galloway, I. (2014) 'Using Pay-For-Success to Increase Investment in the Nonmedical Determinants of Health', *Health Affairs* 33.11: 1897–1904

- Godeke, S. and Pomares, R. (2009) *Solutions for Impact Investors: From Strategy to Implementation*, New York: Rockefeller Philanthropy Advisors
- Hardi, P.; Radasci, G. and Schmidt, K. (2012) *Evaluations of CSR Performance and Impact as Seen by Key Actors other than Business*, Working Paper 3, CSR Impact
- Harji, K. (2008a) *Accounting for Social Impact: The Expanded Value Added Statement*, Ottawa: Carleton Centre for Community Innovation, Carleton University
- Harji, K. (2008b) *Social Return on Investment*, Ottawa: Carleton Centre for Community Innovation, Carleton University
- HM Government (2014) *Qualitative Evaluation of the London Homelessness Social Impact Bond: First Interim Report*, London: Department for Communities and Local Government
- Hornsby, A. (2012) *The Good Analyst: Impact Measurement and Analysis in the Social-Purpose Universe*, London: Investing for Good
- Ingram, V. (2013) *Corporate Social Responsibility: The Role of Public Policy. A Systematic Literature Review of the Effects of Government Supported Interventions on the Corporate Social Responsibility (CSR) Behaviour of Enterprises in Developing Countries*, Den Haag: Policy and Operations Evaluation Department (IOB), Ministry of Foreign Affairs
- IRIS (2013) *Getting Started with IRIS: How to Select IRIS Metrics for Social and Environmental Performance Measurement*, Global Impact Investing Network (GIIN)
- Jackson, E.T. (2013a) 'Evaluating Social Impact Bonds: Questions, Challenges, Innovations, and Possibilities in Measuring Outcomes in Impact Investing', *Community Development* 44.5: 608–16
- Jackson, E.T. (2013b) 'Interrogating the Theory of Change: Evaluating Impact Investing Where it Matters Most', *Journal of Sustainable Finance & Investment* 3.2: 95–110
- Jackson, E.T. and Harji, K. (2013a) *Assessing Impact Investing: Five Doorways for Evaluators*, <http://issuu.com/rockefellerevaluation/docs/impactinvestingthematicbriefs2014> (accessed 4 December 2015)
- Jackson, E.T. and Harji, K. (2013b) *Measurement that Matters: Ten Steps for Assessing Social Impact*, <http://issuu.com/rockefellerevaluation/docs/impactinvestingthematicbriefs2014>, p.9
- Jackson, E.T. and Harji, K. (2012a) *Unlocking Capital, Activating a Movement: The Final Report of the Strategic Assessment of the Rockefeller Foundation's Impact Investing Initiative*, <http://issuu.com/rockefellerevaluation/docs/impactinvestingscan2012> (accessed 4 December 2015)
- Jackson, E.T. and Harji, K. (2012b) *Accelerating Impact: Achievements, Challenges and What's Next in Building the Impact Investing Industry*, New York: The Rockefeller Foundation
- Jackson, E.T. and Harji, K. (2009) 'Measuring the Blended Value of CSR and Social Enterprise', presented at the 2009 Annual CES Conference, Ottawa, 31 May – 3 June 2009

Jackson, E.T.; MacPherson, N.; Gariba, S. and Ofir, Z. (2012) 'Evaluating Impact Investing: An Emerging Area of Development Evaluation Practice', panel presentation to the 6th biennial conference of the African Evaluation Association, International Conference Centre, Accra, Ghana, January 2012

Knickman, J.R. and Hunt, K.A. (2007) 'The Robert Wood Johnson Foundation's Approach to Evaluation', in S.L. Isaacs and D.C. Colby (eds), *The Robert Wood Johnson Foundation Anthology: To Improve Health and Health Care Volume XI*, San Francisco: Jossey-Bass

Lankester, K. and Pease, M. (2013) *Collecting Impact Data Using Mobile Technology: A Pilot Project Measuring Client Poverty Levels*, Global Impact Investing Network (GIIN)

Liao, P.-C.; Xue, J.; Lui, B. and Fang, D. (2014) 'Selection of the Approach for Producing a Weighting Scheme for the CSR Evaluation Framework', *KSCE Journal of Civil Engineering* 19.6: 1549–59

Liket, K. and Maas, K. (2015) 'Strategic Philanthropy: Corporate Measurement of Philanthropic Impacts as a Requirement for a "Happy Marriage" of Business and Society', *Business & Society* 1–33

Liket, K.C.; Rey-Garcia, M. and Maas, K.E.H. (2014) 'Why Aren't Evaluations Working and What to Do About it a Framework for Negotiating Meaningful Evaluation in Nonprofits', *American Journal of Evaluation* 35.2: 171–88

McKinsey & Company (2010) *Learning for Social Impact: What Foundations Can Do*, White Paper, London: McKinsey & Company

Moore, K.A.; Walker, K. and Murphey, D. (2011) 'Performance Management: The Neglected Step in Becoming an Evidence-Based Program', in L. Weiss and C. Collins (eds), *Leap of Reason: Managing to Outcomes in an Era of Scarcity*, Washington DC: Venture Philanthropy Partners

Morino, M. (2011) *Leap of Reason: Managing to Outcomes in an Era of Scarcity*, Washington DC: Venture Philanthropy Partners

Newmark, T.E. and Pena, M. (2012) *Portfolio for the Planet: Lessons from 10 Years of Impact Investing*, Abingdon and New York: Earthscan

Olsen, S. and Galimidi, B. (2008) *Catalog of Approaches to Impact Measurement: Assessing Social Impact in Private Ventures*, Social Venture Technology Group

Picciotto, R. (2011) *Evaluating Development Philanthropy in a Changing World*, Commissioned Paper, Brighton: IDS, The Resource Alliance and The Rockefeller Foundation

Porter, M.E.; Hills, G.; Pfitzer, M.; Patscheke, S. and Hawkins, E. (n.d.) *Measuring Shared Value: How to Unlock Value by Linking Social and Business Results*, Boston/Geneva/Mumbai/San Francisco/Seattle/Washington: FSG

PwC (2013) *Measuring and Managing Total Impact: A New Language for Business Decisions*, PwC

Rockefeller Philanthropy Advisors (n.d.) *Assessing Impact*, New York: Rockefeller Philanthropy Advisors

- Rodin, J. and MacPherson, N. (2012) 'Shared Outcomes', *Stanford Social Innovation Review (Advancing Evaluation Practices in Philanthropy, Sponsored Supplement of the SSIR)* 10.3: 12–15
- Rubin, J. (2011) *Evaluation of the Social Impact Bond: Lessons from Planning and Early Implementation at HMP Peterborough*, Cambridge and Brussels: RAND Europe
- Rudd, T.; Nicoletti, E.; Misner, K. and Bonsu, J. (2013) *Financing Promising Evidence-Based Programs: Early Lessons from the New York City Social Impact Bond*, New York and Oakland: MDRC
- Salazar, J.; Husted, B.W. and Biehl, M. (2012) 'Thoughts On the Evaluation of Corporate Social Performance through Projects', *Journal of Business Ethics* 105.2: 175–86
- Social Finance (2014) 'Peterborough Social Impact Bond Reduces Reoffending by 8.4%; Investors on Course for Payment in 2016', press release, 7 August 2014, London: Social Finance Ltd
- Social Impact Investment Taskforce (2014a) *Impact Investment: The Invisible Heart of Markets*, Social Impact Investment Taskforce
- Social Impact Investment Taskforce (2014b) *Measuring Impact: Subject Paper of the Impact Measurement Working Group*, Social Impact Investment Taskforce
- The Centre for Social Impact (2011) *Report on the NSW Government Social Impact Bond Pilot*, Sydney: The Centre for Social Impact (CSI)
- Tides (2011) *Assessing Social Impacts of Investments: Developing Effective Evaluation Systems to Measure Social Returns*, Tides
- Triodos Facet (2010) *Final Report: Evaluation PSOM/PSI 1999-2009 and MMF*, Zeist: Triodos Facet
- Tuan, M.T. (2008) *Measuring and/or Estimating Social Value Creation: Insights Into Eight Integrated Cost Approaches*, Bill & Melinda Gates Foundation
- Twersky, F. and Lindblom, K. (2012) *Evaluation Principles and Practices: An Internal Working Paper*, The William and Flora Hewlett Foundation
- Warner, M.E. (2013) 'Private Finance for Public Goods: Social Impact Bonds', *Journal of Economic Policy Reform* 16.4: 303–19
- Woelful, T.; Thornley, B.; Ren, J. and Sirull, B. (2014) *Invest Northwest Eight-year Social Performance Review: A Comprehensive Case Study on the Impact of a Northwest Area Foundation Mission-Related Investment*, InSight at Pacific Community Ventures



Brighton BN1 9RE

T +44 (0)1273 606261

F +44 (0)1273 621202

E ids@ids.ac.uk

www.ids.ac.uk



UKaid

from the British people