



© 2020 by the author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International (CC BY 4.0) License (<https://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

Citation: Zappalà, G. 2020. Meaningful Evaluation: A Holistic and Systemic Approach to Understanding and Assessing Outcomes. *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, 12:2–3, 1–19. <https://doi.org/10.5130/ccs.v12.i2-3.7034>

ISSN 1837-5391 | Published by UTS ePRESS | <https://epress.lib.uts.edu.au/journals/index.php/mcs>

REFEREED PAPER

Meaningful Evaluation: A Holistic and Systemic Approach to Understanding and Assessing Outcomes

Gianni Zappalà

University of New South Wales, Sydney, Australia
University of Technology Sydney, Sydney, Australia

Corresponding author: Gianni Zappalà, Centre for Social Impact, University of New South Wales, Level 7 Science Engineering Building (E8), High St, Kensington, NSW, 2052.

g.zappala@unsw.edu.au

DOI: <https://doi.org/10.5130/ccs.v12.i2-3.7034>

Article History: Received 22/01/2020; Revised 21/04/2020; Accepted 07/07/2020; Published 12/02/2021

Abstract

The concept and practice of Social Impact and Social Impact assessment has developed and matured over the last decade. Despite this growth, confusion still exists with respect to definitions of social impact as well as which tools and frameworks are most appropriate. Central to the concept and practice of social impact is a focus on outcomes. This paper argues that a discernible albeit problematic pattern and approach has emerged in outcomes measurement within the social impact field. After briefly reviewing some of the key problems with how outcomes measurement is practised, it presents some recent approaches from the evaluation field that attempt to address some of these concerns before introducing a new approach to understanding and evaluating outcomes – Meaningful Evaluation (ME). A pilot Meaningful Evaluation of volunteer ethics teachers in the Primary Ethics program in NSW is used to illustrate the approach. ME combines the Map of Meaning (MoM) with insights from next generation evaluation approaches to understand and assess outcomes. Informed by appreciative inquiry perspectives, it provides a means to bridge the divide between positivist and interpretivist approaches in evaluation. Key to ME is the assumption that it is more likely that immediate outcomes lead to medium and long-term outcomes (changes in behaviour) that are sustainable and lead to impact if participants experience program interventions as ‘meaningful’. Meaning is an important internal outcome that is essential if longer-term external outcomes are to

DECLARATION OF CONFLICTING INTEREST The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. **FUNDING** The author(s) received no financial support for the research, authorship, and/or publication of this article.

occur. It also shifts the focus to capture unintended outcomes, key to developing holistic and systemic rather than linear and mechanistic Theories of Change.

Keywords

Map of Meaning; Meaningful Evaluation; Outcomes Measurement; Social Impact Assessment; Theory of Change

Introduction

The field of Social Impact and Social Impact assessment has developed and matured over the last decade spawning a wide range of specialist consultancies, practitioner networks, University Centres and courses, and publications ([Onyx 2014a; b](#); [OECD 2015](#); [Bloch 2012](#); [Epstein & Yuthas 2014](#)). The development has also occurred with respect to practice within the social sector itself, witnessed by the popular usage of terms such as Theory of Change (ToC), the publication of social impact reports and philanthropic grants that focus on impact ([Callis et al. 2019](#)).

Despite this growth, confusion still exists with respect to definitions of social impact as well as which tools and frameworks are most appropriate ([Onyx 2014b](#); [Grieco 2015](#); [Ormiston 2019](#)). The lack of consensus on the definition of social impact and its assessment is problematic for both theory and practice ([Maas and Liket 2011](#); [Ormiston 2019](#)). The confusion is often due to the different ways various stakeholders within the social sector understand the term ([Wallman-Stokes et al 2013](#); [Ormiston 2019](#)).

Some scholars caution against impact assessment. Apart from the resources required in terms of the money, staff, and expertise to undertake such assessments, the different types of complexity that beset social problems mean that what is trying to be tracked and measured is often elusive, and metrics often suffer from a lack of validity and reliability ([Zappalà 2011](#)). Ebrahim and Rangan's (2010) framework, for instance, suggests that not-for-profit (NFP) organisations should only devote the time and resources to demonstrating their social impact when their programs are underpinned by a complex causal logic and operational strategy, and where an organisation can exercise sufficient control over results to attribute any impacts to its work. In practice, many NFP organisations apply relatively complicated, time consuming and expensive measurement approaches that attempt to quantify and/or monetise impact to relatively simple and straightforward programs, where measuring inputs and outputs and at times outcomes would suffice ([Ebrahim & Rangan 2010, 2014](#); [Gugerty & Karlan 2018](#); [Zappalà 2011](#)).

Nevertheless, a coming together on definitions and key aspects of social impact has emerged ([OECD 2015](#); [Hearn & Buffardi 2016](#)). Central to the concept and practice of social impact is a focus on outcomes. Outcomes, of course, have always been of interest to evaluators, but the growth of a specialist social impact discourse and approach over the last decade, especially promoted by philanthropic and government funders, has led to a rhetoric rather than reality of outcomes measurement ([MacKeith 2011](#)). Indeed, many social purpose organisations question the quality of the outcomes measurement they do undertake ([Callis et al. 2019](#)).

A discernible albeit problematic pattern and approach has emerged in outcomes measurement within the social impact field. This paper briefly reviews some of the key problems with how outcomes measurement is practised, then reviews some recent approaches from the evaluation field that attempt to address some of these concerns before introducing a new approach to understanding and evaluating outcomes – Meaningful

Evaluation (ME). A pilot evaluation of volunteer ethics teachers in the Primary Ethics program in NSW is referred to in order to illustrate aspects of the ME approach ([Zappalà 2019](#))¹.

The main aim of the pilot was to help Primary Ethics better understand the experience of their volunteer teachers in order to suggest potential areas for program improvement and contribute to the development of their broader Theory of Change. While not a key target beneficiary group, the volunteer ethics teachers are key to the program's success and sustainability as they are the delivery vehicle of the Primary Ethics program without whom it would not exist. Understanding the perspective of the ethics teachers was critical for Primary Ethics' goals of scaling the program in terms of volunteer recruitment and retention.

The Problem with Current Approaches to Outcomes

There are at least five problems in current approaches to outcome measurement: 1) An over reliance on the Impact Value Chain; 2) The de-contextualisation of Outcomes; 3) The monetisation of Outcomes; 4) The focus on external Outcomes; and 5) A positivist approach to validity.

OVER RELIANCE ON THE IMPACT VALUE CHAIN

In evaluation, an outcome is the change that occurs in a person or program participant resulting from an intervention. Outcomes usually result from achieving a particular program's outputs. Outcomes measurement refers to the measurement of change resulting from an initiative, program or organisation. Measuring outcomes provides evidence of whether programs and organisations make a difference to program participants ([Flatau et al. 2016](#)). There is generally consensus that the term impact refers to long-term results at the end of an Impact Value Chain (see [Figure 1](#)).

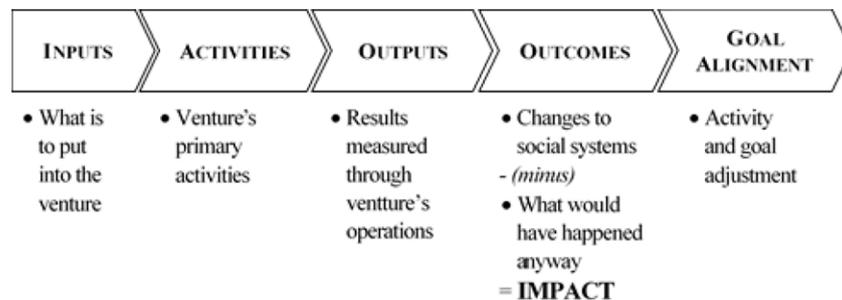


Figure 1. The Impact Value Chain

Source: Clarke et al. ([2004](#))

Impact is distinct from 'outputs' – which are the direct products resulting from the implementation of intervention activities – and from 'outcomes'. Outcomes are sometimes broken down into immediate, medium and long-term changes in the target group(s) who have been engaged in the intervention and which precede, and are usually a pre-condition for, impact. Impact is usually defined either as outcomes less what would have happened without the intervention (as in [Figure 1](#) where outcomes and impact are similar

¹ Primary Ethics, an independent NFP, is the sole approved provider of ethics classes to NSW public schools. Its mission is to 'support children to develop skills in ethical reasoning, critical thinking and respectful discussion'. It has developed an integrated age-appropriate curriculum for all of the Primary School stages (Kindergarten to Year 6) which is delivered free of charge via a network of specially trained and accredited volunteers. Since its establishment, ethics classes are available in almost 500 schools in NSW, with approximately 45,000 children in ethics classes, delivered by a network of almost 2,800-trained volunteers.

kinds of effects), or as significant long-term sustainable change (outcomes and impact are different kinds of effects) ([Peersman et al. 2016](#)).

The 'if-then' causal logic inherent in the Impact Value Chain often oversimplifies and reinforces the view that change is linear, predictable, straightforward and mechanistic. In other words, the problem specification, intervention and any resulting outcomes lie within Simple and/or Complicated rather than Complex domains ([Glouberman & Zimmerman 2002](#); [Kurtz & Snowden 2003](#)). The assumptions underlying the logic as to why a short-term outcome will lead to a medium-term and then a long-term outcome and eventually impact are often not clearly articulated, nor whether the change is sustainable. The aspirations of the program participants that lie beneath the logic model's intended impact are often unrealistic and simplistic given the uncertainty and unpredictability of the real-world context.

THE DE-CONTEXTUALISATION OF OUTCOMES

The second problem of current approaches to outcome measurement is that outcomes are usually identified and analysed outside of their contextual background. Many social impact frameworks (e.g. Social Return on Investment - SROI) describe, analyse and present outcomes as single line items along a spreadsheet or outcomes map, seen outside of their broader holistic context ([Onyx 2014a](#)). The insights from Realistic evaluation, that outcomes will depend on how mechanisms are enacted in any given context by people working with constraints, either are forgotten or ignored. It is not programs per se that make things change but that people embedded in their context (and exposed to programs) do something to activate mechanisms and achieve change ([Pawson & Tilley 1997](#)). Context not only effects how outcomes and impact should be measured, but the 'more contexts vary, the more likely it is that a rigid approach displaces a more insightful one' ([Ruff & Olsen 2016](#)).

A problem in social impact assessment is that 'foundations and non-profits often jump to collecting data and setting up measurement systems without thinking about their surrounding environment (read context)' ([Coffman & Beer 2019](#)). Indeed, to do this requires 'Contextual Intelligence', the 'ability to quickly and intuitively recognize and diagnose the dynamic contextual variables inherent in an event or circumstance', especially in situations characterised by uncertainty, ambiguity, and non-linear causal relations ([Kutz 2008](#), pp. 6-7).

THE MONETISATION OF OUTCOMES

The third problem of current approaches to outcome measurement is the desire to monetise outcomes. Giving a monetary measure to social impact is seen as appropriate and something the sector should strive for, '...as it helps to translate impact assessment results into the common language: the dollar sign!' ([Radeke 2016](#)). This trend to give outcomes a financial value, such as in SROI approaches, gives those doing the analysis the incentive to focus on those outcomes that are more amenable to monetisation while ignoring the qualitative and relational aspects of outcomes ([Onyx 2014b](#)). Several studies have now highlighted the problematic nature of using financial proxies, the risks of commodifying social value, as well as ignoring the qualitative and social aspects of outcomes ([Arvidson et al. 2010](#); [Onyx 2014b](#); [Zappalà & Lyons 2009](#)).

THE FOCUS ON EXTERNAL OUTCOMES

The fourth problem of current approaches to outcome measurement is the tendency to focus on external or so-called 'hard' outcomes that are more easily measured, tangible and quantifiable (e.g. Tenancy for a homeless person, a job for an unemployed person). The trend is to discount so-called 'soft' or internal outcomes that are less tangible and embodied in processes rather than explicit in external outcomes (e.g. Perceptions, attitudes, meaning, personal skills). The term 'soft' devalues internal changes that usually need to occur within people if they are to meet longer-term goals and are usually the external manifestations

of internal changes. Shifts in attitudes, beliefs, and skills are often the necessary pre-conditions (and the missing assumptions in impact value chains) for longer-term sustainable change to occur ([MacKeith 2009](#)).

POSITIVIST APPROACH TO VALIDITY

The fifth problem of many current approaches to outcome measurement is the positivist approach to validity. Positivist approaches usually treat program participants as passive ‘objects’ when collecting and measuring outcomes data rather than valuing the usefulness of the data collection tool for program staff and clients. Such an approach assumes an objective, verifiable truth that is separate from the researcher. The validity of an outcomes measurement tool is about assessing the tool against other ‘objective’ measures and assessing its own consistency (reliability) over time and across users. In contrast, phenomenological (Interpretivist) approaches see human beings not as passive receivers of an objective reality but acknowledges that they interpret and understand their world by creating stories that make sense to them ([MacKeith 2011](#); [Liamputtong 2013](#)). ‘Truth’ is constructed through subjective experience and is not separate from the person having the experience. The subject is seen as an active agent constantly involved in ‘sense-making’ and whose actions will be informed by their evolving understanding.

Other shortcomings of current approaches to outcome measurement include the short time-frames for evaluation (impact usually only occurs in the longer-term), the assumption that all programs should measure impact and the push to develop a standardised approach to measuring value and impact ([Ebrahim & Rangan 2010](#); [Mulgan 2010](#); [Zappalà 2011](#)).

Next Generation Evaluation

Many of the problems of outcome measurement noted above plagued the broader field of evaluation well before the mania with ‘social impact’. A more recent challenge has been responding to the increased complexity of both the environment and the nature of interventions designed to deal with social problems ([Kahane 2004](#); [Kania & Kramer 2013](#); [2014](#)). As was noted previously, a problem of current approaches to outcome measurement is that outcomes are usually specified as resulting from contexts that are primarily simple or complicated in nature. People and social systems, however, usually behave like Complex Adaptive Systems (CAS) ([Patton 2011](#)). Complexity results from the interaction of many adaptive agents who, pursuing their own plans, adapt to the feedback and behaviour of others. Complex systems are characterised by nonlinearity, unpredictability, and instability. They are less amenable to being controlled, they cannot be broken down into more simple parts and emerge through self-organisation. It is difficult to predict the emergence of outcomes from complex environments ([Kania & Kramer 2013](#); [2014](#)). Rogers captured the dilemma of the implications of complexity for evaluation:

‘Life is not simple, but many of the logic models used in programme theory evaluation are. Is it a problem to represent reality as a simple causal model of boxes and arrows, or should the logic models we use address the complexity of life – and if so, how?’ ([Rogers 2008](#), p. 29)

The response of so-called ‘next generation evaluation’ approaches, which are seen as complementary rather than replacing traditional evaluation approaches and methods, was to embrace complexity ([Gopalakrishnan et al. 2013](#); [Preskill et al. 2014](#)). Some approaches have attempted to either build more sophisticated logic models ([Funnel & Rogers 2010](#); [Fujita 2010](#)), introduce Systems Thinking and Systems concepts ([Rogers & Williams 2010](#); [Canty-Waldron 2014](#)), or bring logic models and systems thinking together ([Renger et al. 2011](#); [Renger et al. 2019](#)). Appreciative Inquiry ([Coglan et al. 2003](#)), Participatory Inquiry ([Heron & Reason 1997](#)) and Developmental Evaluation ([Gamble 2008](#); [Dozois et al. 2010](#); [Patton 2011](#)) are also important responses to dealing with greater complexity. Some of the key differences and focus between traditional and so-called next generation evaluation are summarised in [Table 1](#).

Table 1. Key features and focus of traditional and next-generation evaluation

Traditional	Next-generation
Accountability	Learning
Compliance	Capacity building
Formative/summative	Developmental
Logic models	Systems approach
Accounting for impact	Demonstrate learning/adaptation to change
Evaluator as external/objective	Evaluator as facilitator/coach
Pre-determined goals	Shifting goals/adaptation
Fixed plan	Variable/emergent plan
Anticipated outcomes	Unintended outcomes
Individual outcomes	Holistic outcomes

Source: Adapted from various sources cited above

This paper introduces another approach (still in its infancy) consistent with aspects of ‘next generation’ evaluation approaches, termed ‘Meaningful Evaluation’ (ME). The author first developed and used the ME approach in his Graduate Demonstrating Social Impact course at the Centre for Social Impact, University of New South Wales². Subsequently, the ME approach was trialled by several students enrolled in the Graduate Certificate in Social Impact at UNSW as part of their Field Project assignments (2014 – 2016). In 2019, the ME approach was applied in a pilot evaluation to understand the experience of volunteer ethics teachers in the Primary Ethics program in NSW by the Centre for Social Justice & Inclusion at UTS (Zappalà 2019). The evaluation identified the main Pathways along which meaning occurs as well as any imbalances or barriers to meaning. Aspects of this evaluation are highlighted throughout the remainder of this paper to illustrate the approach.

Meaningful Evaluation arose from a desire to address some of the problems with existing outcomes frameworks discussed above and wanting to deal with complex social phenomena in a holistic and systemic manner. Meaningful Evaluation combines the Map of Meaning (MoM) framework with some of the key principles from next generation evaluation approaches. As meaning is a subjective phenomenon, Meaningful Evaluation acknowledges the importance of focusing on the subjective experience of participants and their active participation in any program evaluation, although it is also applicable with more positivist approaches. The next section provides a summary of the MoM framework upon which ME draws.

The Map of Meaning

The Map of Meaning (MoM) is a simple yet profound model originally developed to engage with meaning and purpose in work (Lips Wiersma & Morris 2011). Based and validated on empirical research in cross-national settings, a community of MoM practitioners has developed over the last two-decades, applying the MoM successfully in Counselling, Coaching, Leadership, Organisational and community change, and

² ‘Complexity and Social Impact – Undertaking Meaningful Evaluation’, Unit 11 presentation for Demonstrating Social Impact, CSI, UNSW, 2014.

most recently through the author's own work, in evaluation³. A brief summary of the MoM follows with particular focus on its relevance in an evaluative context.

There are three main elements or dimensions to the Map of Meaning (MoM) (see [Figure 2](#) below). The first element comprises the four Pathways to meaning⁴:

- i) *Developing the Inner Self⁵ (Reflect)*: For an activity, event or participation in a program to be meaningful, it usually enables a degree of self-development and personal growth that comes from inner reflection. Three main sub-themes comprise this Pathway:
 - a. Moral development – the ability to distinguish between different values, perspectives and right and wrong;
 - b. Personal Growth – the ability to learn new aspects of ourselves; and
 - c. Being true to self – the ability to be and act authentically in all situations.
- ii) *Unity with others (Connect)*: Meaning is relational and for an activity, event or participation in a program to be meaningful, there usually needs to be a sense of connection with others. Connection does not necessarily imply conformity, but experiencing unity even when diversity is present. Three main sub-themes comprise this Pathway:
 - a. Working together – the sense of being able to achieve more with the support of others, whether colleagues or co-participants;
 - b. Shared values – the ability to raise and make values public; and
 - c. Belonging – the sense of being part of a larger community or tribe with whom we feel connected.
- iii) *Service to Others (Respect)*: For an event, activity or participation in a program to be meaningful, people need to feel they are serving others through contributing to the well-being of others, including the broader environment. Two main sub-themes comprise this Pathway:
 - a. Making a Positive difference – the ability to improve things for others;
 - b. Serving the needs of humanity and the planet – the need to feel that what we are doing is useful to the wider context, be it a social cause, the global community or the planet.
- iv) *Expressing Full potential (Express)*: An event, activity or participation in a program is meaningful to the extent that it enables people to utilise and demonstrate their unique gifts, talents and skills. In contrast to the 'Developing the inner self' Pathway, which is reflective and generally passive in nature, this Pathway focusses on active, external outcomes. Three main sub-themes comprise this Pathway:
 - a. Creating – the ability to be creative and discover new ways of doing things;
 - b. Achieving – the sense of accomplishment that comes from completing a job or task; and
 - c. Influencing – the power and ability to affect change in something or someone.

³ The author began accredited Map of Meaning Training with Lani Morris in October 2013 and is a Certified Practitioner of the Map of Meaning (www.themapofmeaning.org).

⁴ The word in brackets that follow the Pathway Headings are the author's abbreviated headings for the respective Pathways.

⁵ In a revised version of the Map this Pathway was renamed 'Integrity with Self' although the substance and sub-themes of the Pathway remained the same ([Lips-Wiersma & Morris 2018](#)).

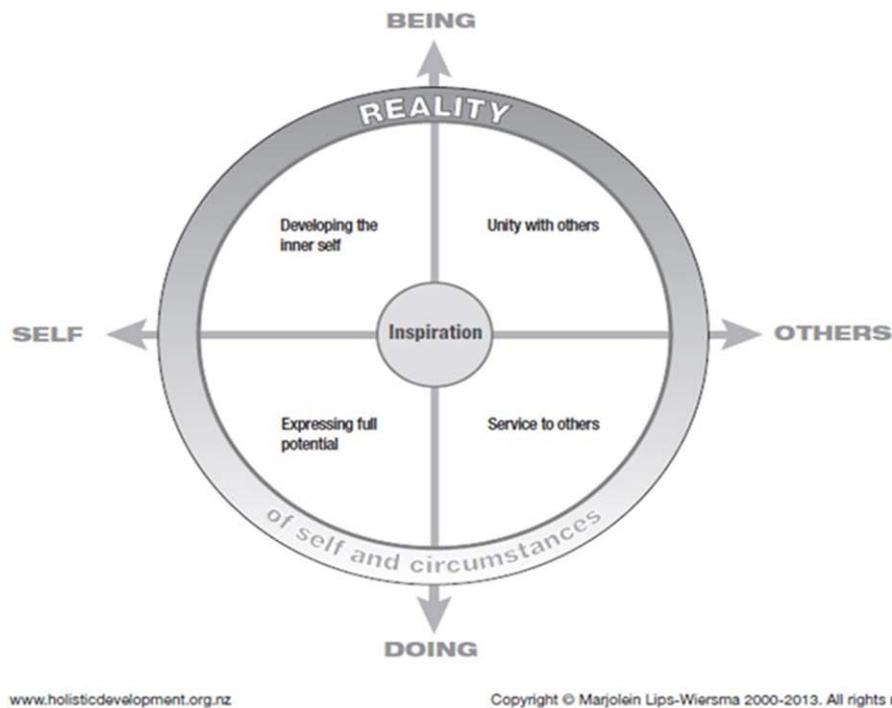


Figure 2. The Map of Meaning

Source: Lips-Wiersma & Morris (2011)⁶

The second element of the MoM emerges from the intersection of two key dimensions inherent in the human search for meaning. On the horizontal plane in [Figure 2](#), the relational tension between the needs of the Self (the need to develop and express ourselves as individuals) and the needs of Others (the need to make a difference and contribute so we feel life has been worthwhile). On the vertical plane in [Figure 2](#), is the ‘Presencing’ tension between the need for Being (the inner need to reflect and contemplate) and the need for Doing (the need to act and do, the outward expression of Being). These tensions need to be in balance to sustain meaning in our work, lives and to participate in any program.

The third element of the MoM is the bigger realm within which meaningfulness takes place, the space between the inner and outer circles in [Figure 2](#). At the centre is ‘Inspiration’, what gives people hope and purpose, ideals and visions for the future. At any particular point in time, however, the real-world pressures and circumstances test and limit these ideals. An event, activity or participation in a program is meaningful therefore to the extent that people are authentic and accepting of the context within which they find themselves at any point in time.

The MoM framework suggests that the key to greater meaning lies in achieving a more balanced and holistic state across the four Pathways, addressing the dual ‘relational’ and ‘presencing’ tensions, and keeping true to the values and inspiration within the reality of the current situation.

⁶ Used with permission from the authors and the Map of Meaning International Trust (<http://www.themapofmeaning.org>)

Meaningful Evaluation in Brief

Applying the MoM to evaluation and social impact analysis is a simple yet profound framework to capture complex, emergent, internal and unintended outcomes⁷. It can inform program design; evaluate existing programs; and develop more holistic and systemic rather than linear and mechanistic Theories of Change, hallmarks of best-practice outcomes measurement ([Bagnoli & Megal 2011](#); [Ógáin et al. 2012](#)).

As noted above, evaluators define outcomes as the changes that occur in a person resulting from an intervention as part of participating in a program. Most definitions of social impact are based on a ‘logic’ or ‘impact value chain’, where outcomes are delineated along a spectrum as falling into short, medium and long-term results, with impact defined as significant and sustainable change in long-term outcomes. Assumptions as to why short-term outcomes lead to longer term outcomes often do not capture the complexity and inner drivers that shift human behaviour.

Applying a ME approach is a process to program design and evaluation assumes that:

1. It is more likely that immediate outcomes lead to Medium and Long-term outcomes (changes in behaviour) that are significant and sustainable, that is, achieve social impact, if the experience/s from participating in a program are ‘meaningful’ ([McNeil et al. 2012](#));
2. Meaning is an important internal outcome, needed for longer-term external outcomes to occur. Longer-term change and the accomplishment of so-called ‘hard’ outcomes often requires shifts in attitudes and meaning ([MacKeith 2011](#)). Important aspects of inner change also include intangible and tacit changes embodied in relational processes ([Edwards et al. 2012](#); [Onyx 2014a](#));
3. Capturing meaningful unintended outcomes is just as important as capturing intended outcomes, especially where people did not experience those intended outcomes as meaningful ([Wallman-Stokes et al. 2013](#); [Hearn & Buffardi 2016](#)).

Simply stated, ME puts in play a process that uncovers the extent to which participants in any given program experience the various activities and outcomes of a program as ‘meaningful’ as defined by the MoM framework. By doing so, it enables the identification and assessment of the sustainability of any outcomes experienced, captures the inner process of change and any unintended consequences of the participation or program intervention. The approach of ME is to identify and examine the main sources (Pathways) along which meaning occurs as well as any imbalances or barriers to meaning. The more balanced and extensive is the experience of meaning across a program the more likely that any intended outcomes achieved are long lasting and lead to impact. Even if program participants may not have achieved the intended outcomes, the focus on ‘meaning’ is an important ‘feedback’ loop in a three-legged approach to assessing impact that is more likely to capture unintended outcomes, also important to achieving social impact ([Twersky 2019](#)).

Key issues explored in the ME of Primary Ethics volunteer ethics teachers, for example, included:

Whether and how participation in the Primary Ethics program led ethics teachers to experience a sense of self-development and personal growth (Reflect Pathway)

While participation in the program had led to some moral development for ethics teachers in terms of thinking about ethical issues, it had particularly influenced their thinking, perspectives and skills in a variety of personal and work contexts and in some cases had challenged participants’ own norms and worldviews. In particular, volunteering as an ethics teacher enabled participants to:

- Have a better understanding of different ethical and social perspectives;
- Increase their self-awareness;

⁷ Simple in this context refers to the MoM framework itself rather than the domain being examined. ME can be applied to Simple, Complicated and Complex domains.

- Improve their emotional intelligence;
- Improve their critical thinking and reasoning;
- Improve their people management and relationship skills;
- Improve their ability to assess events and situations;
- Improve and change their approach to parenting.

These were important, albeit *unintended outcomes* of the program, ones that were experienced as ‘meaningful’ by the ethics teachers and therefore ones that were likely to lead to long-term impact in terms of personal and professional development for the volunteer ethics teachers (Zappalà & Smyth 2020). They also provided insights for ensuring the longer-term sustainability of the program for Primary Ethics in terms of improving volunteer recruitment and management. These included:

- Promoting and marketing the potential benefit for volunteer ethics teachers to develop the above skills, increasingly seen as key to success in the modern labour market (FYA 2015);
- Developing a simple template to assist volunteer ethics teachers reflect on their role as teachers and any associated skills and personal growth that may arise from that role; and
- Exploring the possibility of offering a competency-based professional-development training pathway for ethics teachers.

Whether and how participation in the Primary Ethics program enabled a sense of connection and unity via working together, sharing values with others and creating a sense of belonging (Connect Pathway)

The evaluation suggested that while the Connect Pathway was generally a positive source of meaning for ethics teachers, it also contained barriers to meaning, which if not addressed have the possibility of undermining the longer term sustainability of the program. Some volunteers, for instance, did not feel supported in their role as ethics teachers (the ‘Working Together’ aspect of the ‘Connect’ Pathway) due to a lack of support from the school. Some volunteers felt unsupported by Primary Ethics itself, especially where ethics teachers were in schools where the ethics coordinator role was vacant, or where they felt the person acting in that role was ineffective. This meant that the teachers themselves had to take on the coordinator duties and reduced the opportunities for ethics teachers to engage with each other in some formalised way.

There was also a sense of working in isolation among many of the ethics teachers, of being in the front line without ever seeing or feeling connected to other parts of the organisation. Others felt isolated in their role (the ‘Belonging’ aspect of the ‘Connect’ Pathway) due to the infrequent and ad hoc nature of meetings and communications between the ethics teachers in a particular school. The sense of belonging felt by ethics teachers was primarily to the school community rather than to a Primary Ethics community or a ‘community of ethics teachers’. The analysis highlighted a desire on the part of many volunteers for a greater degree of interaction and ‘collegiality’ with other ethics teachers in their school and schools in their local community. These findings highlighted the importance for the program to have an effective ethics coordinator role in schools who could ensure, for example, that formal regular meetings of volunteer ethics teachers occurred.

Whether and how participation in the Primary Ethics program provided ethics teachers the opportunities to serve and assist others (Respect Pathway)

The evaluation found that the sense of giving something back through either making a positive difference to the children they taught or through contributing to a more informed and self-aware society was a, if not the most, significant source of meaning for the volunteer ethics teachers. Closely related to this, the program also enabled volunteers to feel that they were making a material difference to the children (the ‘Influence’ aspect of the ‘Express’ Pathway). Comparative analysis of the data related to the meaning Pathways (an integral part of the ME approach) illustrated the strength of meaning and satisfaction that

volunteers derived from seeing how the children in their classes changed through their engagement with the curriculum content. Key findings included:

- Ethics teachers' belief that they were making a positive difference to children's learning outcomes;
- A recognition that the benefits of participating in the program were cumulative and more apparent as children progressed through the program;
- That the program changed how the children considered and evaluated issues by developing more critical thinking and reasoning skills; and
- That the program developed the children's confidence to express themselves in class discussions.

Although an assessment of whether participation in ethics classes leads to improved student learning outcomes was not an objective of the evaluation, the findings from the 'Respect' Pathway provided supportive evidence that Primary Ethics were achieving some of the key intended outcomes of the program, namely, supporting students to become critical thinkers with a strong capacity for questioning and inquiry. In other words, volunteer ethics teachers derived their main source of meaning in areas consistent with the main aims of the program. These findings also contribute to the literature exploring the relationship between the teaching of philosophical ethics and the cognitive, educational or learning outcomes that may result ([Zappalà & Smyth 2020](#)).

Whether and how participation in the Primary Ethics program enabled the expression of ethics teachers' full potential through creating, achieving, influencing others (Express Pathway)

Similar to the 'Connect' Pathway, the analysis of data relating to the 'Create' Aspect of the 'Express' Pathway highlighted that it was both a positive source of meaning as well as a potential barrier to meaning. Ethics teachers experienced creativity in their role through the delivery of the curriculum; managing classroom behaviour and modifying the teaching script without altering the content. Nevertheless, there was a vocal and strong desire from ethics teachers for greater flexibility and autonomy in how they delivered the curriculum, in particular, to enable them to respond to the different contexts and situations that arise in the classroom and for a greater variety of class activities.

While most ethics teachers understood the need to ensure the delivery of a consistent high-quality curriculum for all students, the ME approach revealed a strength of sentiment with respect to this aspect of 'meaningful activity' that led to recommendations to introduce elements of flexibility to the content and delivery of the ethics curriculum without undermining program consistency, quality and accountability. People's need and ability to express themselves through unique and creative ways is a key part of finding an activity meaningful, and the findings suggested that the relative lack of flexibility and autonomy in curriculum delivery had a detrimental effect on some of the volunteers' overall satisfaction and indeed saw some exit the program as a result.

What inspired people to become volunteer ethics teachers?

As the subjects of the evaluation in this case were volunteers, the Inspiration element of the MoM closely related to people's motivations for becoming ethics teachers for the program. Motivations and inspiration included a strong belief in the importance of secular ethical reasoning and critical thinking skills for their children's education as well as the contribution it makes to a healthy and vibrant society. The opportunity to teach their own children in ethics classes for potentially all of the primary school years was also a key motivating factor.

How was the inspiration mediated/constrained by the reality or boundaries of the program and how it affected the overall experience of meaning?

The ‘Reality of Self and Circumstances’ (or context) of the program was the reality of the classroom, in particular the short time frame of the weekly lesson (30 – 45 minutes), the culture of the school, and the challenges of behaviour management for a group of around twenty primary school aged children. These factors provided the boundaries and context within which a volunteer’s motivations were contained. The key constraint that emerged from the analysis was disruptive behaviour in the classroom, ethics teachers’ ability to manage it and the demoralising effects it had on the class as well as volunteers’ own sense of satisfaction and meaning. Most volunteers accepted this reality and accommodated their motivations and inspiration within it. For many, the skills of keeping class discipline and managing disruptive behaviour improved with their experience as a teacher (which was a potential source of personal growth), and as such, had a less deleterious effect on the meaningfulness of their overall experience. In some cases, behaviour management was a deal breaker for volunteers, this ‘reality’ dampening their initial inspiration and motivations, leading them to question the value of continuing with the teaching, dissatisfaction and ultimately exiting the program.

In summary, using the ME approach to explore and evaluate the ethics teacher component of the Primary Ethics program was able to paint a rich and nuanced picture of their experience, that was able to capture elements of the complexity and tensions that existed in their role and avoid many of the problems with understanding outcomes identified previously. The ME uncovered the extent to which participants experienced their role as ethics teachers as ‘meaningful’, and by doing so, enabled the identification and assessment of those outcomes that were most likely to be sustainable because they were the most meaningful, aspects of the inner process of change and unintended outcomes from participating as volunteer ethics teachers. Barriers to meaning were also identified which enabled the pinpointing of potential fault lines in the program that could hinder the longer-term sustainability and impact of the program.

Using a linear logic model framework, for instance, would most likely have missed many of the insights captured, such as personal growth, issues around connection with others, and the importance of feeling a degree of creativity. Indeed, ME based on the MoM is explicitly circular and holistic in nature, with the different Pathways and the component Aspects mutually interacting and reinforcing each other in multivalent ways. The ME approach was able to explicitly contextualise the volunteer ethics teachers’ experience by examining how the activities and process of teaching ethics were constrained by the reality of the program circumstances, show how teachers adapted and responded to their context, and highlight the implications this had for how they experienced meaning. The focus was as much on the inner process of change, how meaning often comes from the intangible and tacit aspects embodied in relational processes as it was on outer manifestations of program delivery. Finally, the evaluation captured meaningful unintended outcomes, such as the personal and professional changes that occurred within ethics teachers themselves.

Measuring Meaning

The process of how participants experience activities as meaningful as defined by the MoM leads to the issue of how meaning is measured. Informed by appreciative inquiry perspectives, ME provides a means to bridge the divide between positivist and interpretivist approaches to evaluation, and has an open approach with respect to how data on meaning is collected. That is, ME does not prescribe any one method or Inquiry paradigm for collecting or analysing data on the issues listed above but rather can be adapted to suit the principles of the inquiry paradigm adopted by the researcher or the nature of the evaluation undertaken.

Under a positivist paradigm, for example, one can use a structured questionnaire that aims to capture and measure the extent of meaning experienced along the four Pathways as well as tensions. The authors of the MoM have developed and used such a questionnaire to measure meaningful work for example

([Lips-Wiersma & Wright 2012](#)). In the Primary Ethics ME, for instance, previous data sources including story transcripts and open-ended comments from exit surveys of volunteer teachers were coded using a framework informed by this questionnaire and adapted to the program context. Even though these data were in different formats and collected for different purposes, analysing them through the MoM framework cast initial insights into which Pathways were most salient in providing volunteers with a meaningful experience, and alternatively, which experiences reduced meaning. This also assisted with the development of the Interview protocol as well as subsequent analysis of the interview transcripts.

Constructivist paradigms of inquiry are concerned with how individuals interpret, understand and experience their social world. They explore the meanings and interpretations people give to their subjective experiences. Methodologically, a qualitative approach describes and analyses from the point of view of research participants. Under this paradigm, for example, using the MoM framework to inform interview and focus group protocols or guide Most Significant Change (MSC) analysis is flexible, responsive and generates rich descriptions of phenomena ([Liamputtong 2013](#); [Mason 1996](#)).

In the Primary Ethics evaluation, for instance, primary data came from face-to-face in-depth interviews with volunteer ethics teachers, using an Interview Protocol based on the MoM and informed by the previous coding⁸. Although a structured interview schedule guided the interviews, the researcher was responsive to and explored issues raised by the interview participants.⁹ Interview transcripts were coded and analysed using NVivo with the MoM providing the primary framework for analysis. As outlined in the previous section, the analysis of the combined data enabled the identification of the key sources as well as barriers to meaning for the volunteer ethics teachers, which enabled the identification of key outcomes (intended and unintended) and the making of several recommendations for program improvement.

The validity of an outcomes measurement tool in constructivist approaches is about assessing the tool against the subjective experience of workers and service users and about the usefulness of the tool in addressing practical problems ([MacKeith 2009](#)). Participatory Assessment and Measurement (PAM) approaches such as Outcomes Star, for instance, treat participants as active agents and validate their experience and perception of the change process. A person receiving a service is an active agent in their own life, inner change is key to achieving outer change, and data on change is collected collaboratively by participant and program worker (i.e. measurement is 'co-produced'). The process of participating in an assessment can itself result in change (through engagement and reflection) ([MacKeith 2011](#)).

Ethics teachers in the evaluation, for instance, often remarked at the end of the formal interview how the process and nature of the questions had led them to reflect and identify issues about themselves and the program that they had not previously thought of (at least not consciously). Several sent follow-up emails to the author with further information they had recalled as they reflected on the questions post interview. This process of ongoing 'emergence' and inner reflection has been found to occur in other cases where the MoM has been used as a framework ([Lips-Wiersma & Morris 2011](#); [2018](#)).

Similar to the Outcomes Star model, standard measures of 'distance travelled' for each meaning Pathway or the relationship between program outcomes sought and each Pathway can be developed to quantify any relative change in meaning for program participants (see [Figure 3](#) & [Table 2](#)). In this way each point on the scale is defined for each meaning Pathway with the scales constructed around a model of change appropriate for the program being evaluated. In this approach the Pathways are discussed between case workers and/

⁸ See Zappalà ([2019](#)) for further details on the characteristics of the volunteer teachers interviewed.

⁹ All interviews were conducted in person at a location convenient to the volunteer, usually the school where they taught ethics, their home or a local café. Interview duration was between 40 and 70 minutes, with an average interview length of 48 minutes; interviews were recorded and professionally transcribed.

or evaluators and participants and plotted on the MoM to provide an overview of the ‘meaningfulness’ of outcomes.

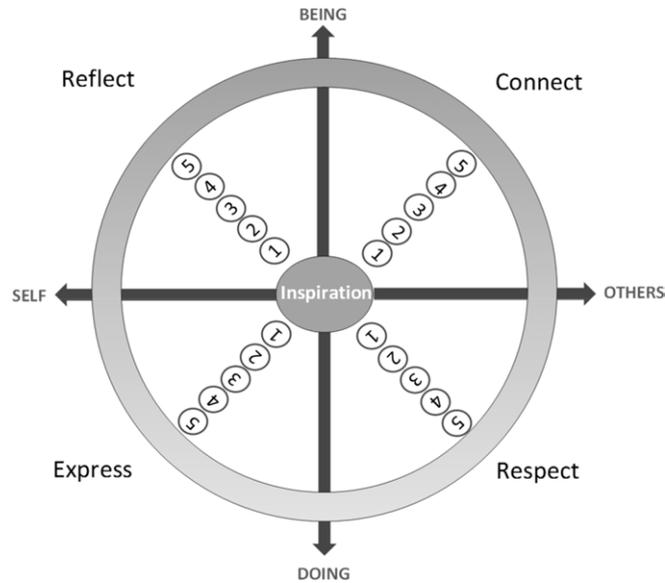


Figure 3. Distance travelled and Meaningful Evaluation

The process can be repeated longitudinally to provide a dynamic analysis of change as well as a point in time snapshot. These data can also be aggregated across all participants to provide program level outcomes. This approach not only enables the collection of richer and deeper data on the sources of meaning, it allows the knowledge about meaning to be ‘co-produced’ between evaluator (and/or case worker) and participant (MacKeith 2011). Meaningful Evaluation is also consistent with Participatory Inquiry paradigms, which favour collaborative forms of action and co-operative inquiry (Heron & Reason 1997).

Table 2. Using Distance travelled with Meaningful Evaluation

Distance travelled	Map of Meaning Pathway & relationship to program outcome sought
1 (static)	No change in the particular Pathway as a result of the program
2 (Awareness)	Person recognizes the need or opportunity that the program provides to change in that particular Pathway
3 (Some Change)	Person can recognize and takes concrete steps to make changes in that particular pathway
4 (Tangible results)	The person is able to demonstrate & report positive changes in that Pathway and link it to the program’s activities
5 (Significant/sustained change)	The person reports longer sustained changes in that Pathway as a result of the program

Conclusion

Next generation approaches in evaluation have made significant inroads in dealing with the increased complexity and richness of social systems. Many of them go some way to addressing the problems noted with current trends in outcome measurement, namely: an over reliance on the impact value chain, the fragmentation and monetisation of outcomes, the focus on external outcomes and a positivist approach to validity. This paper introduced another approach consistent with next generation models to understanding and measuring outcomes – Meaningful Evaluation.

Meaningful Evaluation is a simple yet profound framework to capture complex, emergent, internal and unintended outcomes. It can inform program design; evaluate existing programs; and develop more holistic and systemic rather than linear and mechanistic Theories of Change. The aim of ME is to examine the extent to which participants in any given program experience the various activities and outcomes of a program as ‘meaningful’ as defined by the MoM framework. Meaningful Evaluation seeks to identify the main sources (Pathways) along which meaning occurs as well as identifying any imbalances or barriers to meaning.

The process of ME identifies outcomes that are significant and sustainable, that is, achieve social impact by uncovering the ‘meaningfulness’ from participating in a program. Meaning is an important internal outcome, needed for longer-term external outcomes to occur, and capturing meaningful unintended outcomes is just as important as capturing intended outcomes, especially where people did not experience those intended outcomes as meaningful.

Informed by appreciative inquiry perspectives, ME provides a means to bridge the divide between positivist and interpretivist approaches, and has an open approach with respect to data collection and analysis on meaning. While a practitioner community has developed around applying the Map of Meaning across a range of areas, its application to evaluation and social impact assessment is relatively recent and thus requires further testing and refinement¹⁰. Initial findings are promising, and we invite others who have similar concerns with some of the directions of outcomes measurement practice to consider the Meaningful Evaluation approach.

Acknowledgements

I thank Mitra Gusheh at the Centre for Social Justice & Inclusion, University of Technology Sydney, and the two anonymous referees for their comments on an earlier version of this paper; and Primary Ethics and Mitra Gusheh for agreeing to pilot the approach to evaluate the role of volunteers within the program. I also thank Marjolein Lips-Wiersma and Lani Morris for their support and encouragement in this work and allowing me to adapt the Map of Meaning for use as an evaluation and social impact framework.

References

- Adams, S., Flatau, P., Zaretzy K., McWilliam, D. & Smith, J. 2015, ‘Measuring the difference we make: The state-of-play of outcomes measurement in the community sector in Western Australia’, Bankwest Foundation Social Impact Series, No. 3, Bankwest Foundation, Western Australia. https://www.csi.edu.au/media/uploads/BankwestSocialImpactSeries-measuring-outcomes_aO5S3au.pdf
- Arvidson, M., Lyon, F., McKay, S. & Moro, D. 2010, ‘The ambitions and challenges of SROI’, Working Paper 49, Third Sector Research Centre, University of Birmingham. <https://www.birmingham.ac.uk/Documents/college-social-sciences/social-policy/tsrc/working-papers/working-paper-106.pdf>

10 The Centre for Social Justice & Inclusion at the University of Technology, Sydney is facilitating a Meaningful Evaluation Community of Practice with several evaluations using the ME approach planned.

- Bagnoli, L. & Megali, C. 2011, 'Measuring Performance in Social Enterprises', *Nonprofit and Voluntary Sector Quarterly*, vol. 40, no. 1, pp. 149-165. <https://doi.org/10.1177/0899764009351111>
- Bloch, B. 2012, 'Everyone's story counts: measuring social impact in the Not-for-profit sector, an overview', *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, vol. 4, no. 3, pp. 1-17. <https://doi.org/10.5130/ccs.v4i3.2623>
- Callis, Z., Seivwright, A. and Flatau, P. 2019, 'Outcomes Measurement in the Australian Community Sector: A National Report Card', Bankwest Foundation Social Impact Series, No. 10, Bankwest Foundation, Western Australia. https://www.csi.edu.au/media/Social_Impact_Series_10.pdf
- Canty-Waldron, J. 2014, 'Using Systems Thinking to create more impactful social policy', *Journal of Future Studies*, vol. 19, no. 2, pp. 61-86.
- Clark, C., Rosenzweig, W., Long, D. & Olsen, S. 2004, *Double bottom line project report: Assessing social impact in double bottom line ventures. Methods catalog*. Center for Responsible Business, Paper 13, Working Paper Series, University of California, Berkeley. <https://escholarship.org/uc/item/80n4f1mf>
- Coffman, J. & Beer, T. 2019, 'Why NFPs need to ask powerful questions' Probono Australia, 9 May 2019. <https://probonoaustralia.com.au/news/2019/05/why-nfps-need-to-ask-powerful-questions/>
- Coghlan, A.T., Preskill, H. & Catsambas, T.T. 2003, 'An overview of appreciative inquiry in evaluation', *New Directions for Evaluation*, vol. 100, pp. 5-22. <https://doi.org/10.1002/ev.96>
- Dozois, E., Langlois, M. & Blanchet-Cohen, N. 2010, *DE 201: A Practitioner's Guide to Developmental Evaluation*, Montreal: The J.W. McConnell Family Foundation & International Institute for Child Rights and Development. <https://www.mcconnellfoundation.ca/assets/Media%20Library/Publications/DE%20201%20EN.pdf>
- Ebrahim, A. & Rangan, V. K. 2010, 'The limits of nonprofit impact: a contingency framework for measuring social performance', Harvard Business School Working Paper 10-099, Cambridge, MA. https://www.hbs.edu/faculty/Publication%20Files/10-099_0b80d379-8e88-4992-9e8f-4b79596b1ff9.pdf
- Ebrahim, A., & Rangan, V. K. 2014, 'What impact?: A framework for measuring the scale and scope of social performance', *California Management Review*, vol. 56, no. 3, pp. 118-141. <https://doi.org/10.1525/cmr.2014.56.3.118>
- Edwards, M., Onyx, J., Maxwell, H. & Darcy, S. 2012, 'Meso level Social Impact: Meaningful indicators of community contribution', *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, vol. 4, no. 3, pp. 18-37. <https://doi.org/10.5130/ccs.v4i3.2576>
- Epstein, M. J., & Yuthas, K. 2014, *Measuring and Improving Social Impacts: A Guide for Nonprofits, Companies, and Impact Investors*, Berrett-Koehler Publishers, San Francisco.
- Flatau, P., Adams, S. & Seivwright, A. 2016, 'The practice, prevalence and funding of outcomes measurement in the community sector: Findings from the Outcomes Measurement in the Western Australian Community Sector Survey', Bankwest Foundation, Social Impact Series, No. 5, Bankwest Foundation, Western Australia. https://www.csi.edu.au/media/uploads/Bankwest_Foundation_and_CSI_Social_Impact_Series_Issue_5.pdf
- Fujita, N. (ed.) 2010, *Beyond Logframe: Using Systems Concepts in Evaluation*, Tokyo, Foundation for Advanced Studies in International Development (FASID). https://www.fasid.or.jp/english/files/web_library/h21-3.pdf
- Funnell, S., & Rogers, P. 2010, *Purposeful Program Theory*, Jossey-Bass, San Francisco.
- Foundation for Young Australians (FYA) 2015, *The New Work Order*, FYA, Melbourne. <http://www.fya.org.au/wp-content/uploads/2015/08/fya-future-of-work-report-final-lr.pdf>
- Gamble, J. 2008, *A Developmental Evaluation Primer*, The J.W. McConnell Family Foundation, Montreal. <https://mcconnellfoundation.ca/wp-content/uploads/2017/07/A-Developmental-Evaluation-Primer-EN.pdf>

- Glouberman, S., & Zimmerman, B. 2002, *Complicated and Complex Systems: What Would Successful Reform of Medicare Look Like?* Discussion Paper no. 8. Commission on the Future of Health Care in Canada, Ottawa. <https://www.alnap.org/system/files/content/resource/files/main/complicatedandcomplexsystems-zimmermanreport-medicare-reform.pdf>
- Gopalakrishnan, S., Preskill, H. & Lu, S. J. 2013, *Next Generation Evaluation: Embracing Complexity, Connectivity, and Change*, FSG, Boston. <https://www.fsg.org/publications/next-generation-evaluation>
- Grieco, C. 2015, *Assessing Social Impact of Social Enterprises Does One Size Really Fit All?*, Springer, Heidelberg. <https://doi.org/10.1007/978-3-319-15314-8>
- Gugerty, M. K., & Karlan, D. 2018, 'Ten Reasons Not to Measure Impact – and What to Do Instead', *Stanford Social Innovation Review*, Summer, pp. 39-47. https://ssir.org/articles/entry/ten_reasons_not_to_measure_impact_and_what_to_do_instead
- Hearn, S., & Buffardi, A. L. 2016, 'What is impact?', A Methods Lab publication, Overseas Development Institute, London.
- Heron, J., & Reason, P. 1997, 'A participatory inquiry paradigm', *Qualitative Inquiry*, vol. 3, no. 3, pp. 274-294. <https://doi.org/10.1177/107780049700300302>
- Kahane, A. 2004, *Solving Tough Problems – An Open Way of Talking, Listening, and Creating New Realities*, Berrett-Koehler, San Francisco.
- Kania, J., & Kramer, M. 2013, 'Embracing emergence: How collective impact addresses complexity', *Stanford Social Innovation Review*, January, pp. 1-7. https://ssir.org/articles/entry/social_progress_through_collective_impact#
- Kania, J., Kramer, M. & Russell, P. 2014, 'Strategic Philanthropy for a Complex World', *Stanford Social Innovation Review*, Summer, 2014, pp. 25-37. https://ssir.org/articles/entry/strategic_philanthropy
- Kriev, G., Munscher, R. & Mulbert, K. 2013, Social Return on Investment (SROI): State-of-the-art and Perspectives – A meta-analysis of practice in Social Return on Investment (SROI) studies published 2002-2012, Centre for Social Investment, Heidelberg University. https://archiv.ub.uni-heidelberg.de/volltextserver/18758/1/CSI_SROI_Meta_Analysis_2013.pdf
- Kurtz, C. F., & Snowden, D. F. 2003, 'The new dynamics of strategy: Sense-making in a complex and complicated world', *IBM Systems Journal*, vol. 42, no. 3, pp. 462-483. <https://doi.org/10.1147/sj.423.0462>
- Kutz, M. R. 2008, 'Contextual Intelligence – An emerging competency for Global Leaders', *Regent Global Business Review*, vol. 2, no. 2, pp. 5-8.
- Liamputtong, P. 2013, *Qualitative Research Methods*, Oxford University Press Australia and New Zealand, South Melbourne.
- Lips-Wiersma, M., & Wright, S. 2012, 'Measuring the meaning of meaningful work: Development and validation of the Comprehensive Meaningful Work Scale (CMWS)', *Group & Organization Management*, vol. 37, no. 5, pp. 655-685. <https://doi.org/10.1177/1059601112461578>
- Lips-Wiersma, M., & Morris, L. 2011, *The Map of Meaning – A Guide to Sustaining our Humanity in the World of Work*, Greenleaf Publishing, Sheffield.
- Lips-Wiersma, M., & Morris, L. 2018, *The Map of Meaningful Work – A Practical Guide to Sustaining Our Humanity* (Revised & updated 2nd ed.) Routledge, London.
- Maas, K. & Liket, K. 2011, 'Social impact measurement: Classification of methods', in Burritt, R. et al. (eds) *Environmental Management Accounting and Supply Chain Management*, Springer, New York, pp. 171-202.
- MacKeith, J. 2009, 'Why hard-nosed agencies measure soft outcomes', *SITRA Bulletin – The magazine for housing with care and support*, February, No. 221, pp. 10-11.

- MacKeith, J. 2011, 'The development of the Outcomes Star: A participatory approach to assessment and outcome measurement', *Housing Care and Support: A Journal on Policy, Research and Practice*, vol. 14, no. 3, pp. 98-106. https://www.outcomesstar.org.uk/wp-content/uploads/A_Development-of-the-Outcomes-Star-Published-in-Housing-Care-and-Support.pdf
- Mason, J. 1996, *Qualitative Researching*, Sage, London.
- McNeil, B., Reeder, N. & Rich, J. 2012, *A Framework of Outcomes for Young People*. The Young Foundation, London. <https://www.youngfoundation.org/publications/framework-of-outcomes-for-young-people/>
- Mulgan, G. 2010, 'Measuring Social Value', *Stanford Social Innovation Review*, Summer, pp. 38-43. https://ssir.org/articles/entry/measuring_social_value
- OECD 2015, *Policy Brief on social impact measurement for social enterprises Policies for social entrepreneurship*, Organisation for Economic Cooperation and Development and European Union. <https://www.oecd.org/social/PB-SIM-Web-FINAL.pdf>
- Ógáin, E. N., Lumley, T. & Pritchard, D. 2012, *Making an Impact; Impact Measurement Among Charities and Social Enterprises*, New Philanthropy Capital, London. <https://www.thinknpc.org/resource-hub/making-an-impact/>
- Onyx, J. 2014a, 'A theoretical model of social impact', *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, vol. 6, no. 1, pp. 1-18. <https://doi.org/10.5130/ccs.v6i1.3369>
- Onyx, J. 2014b, 'The politics of social impact: 'value for money' versus 'active citizenship'? *Cosmopolitan Civil Societies: An Interdisciplinary Journal*, vol. 6, no. 2, pp. 69-78. <https://doi.org/10.5130/ccs.v6i2.3923>
- Ormiston, J. 2019, 'Blending practice worlds: Impact assessment as a transdisciplinary practice', *Business Ethics: A European Review*, vol. 28, no. 4, pp. 423-440. <https://doi.org/10.1111/beer.12230>
- Patton, M. Quinn 2011, *Developmental Evaluation: Applying Complexity Concepts to Enhance Innovation and Use*, Guildford Press, New York.
- Pawson, R. and Tilley, N. 1997, 'An introduction to scientific realist evaluation', in Chelimsky, E. and Shadish, W. (eds.), *Evaluation for the 21st Century: A Handbook*, Sage, London, pp. 405-418.
- Peersman, G., Rogers, P., Guijt, I., Hearn, S., Pasanen, T. & Buffardi, A. L. 2016, 'When and how to develop an impact oriented monitoring and evaluation system', A Methods Lab publication, Overseas Development Institute, London. <https://www.odi.org/sites/odi.org.uk/files/odi-assets/publications-opinion-files/10327.pdf>
- Preskill, H., & Beer, T. 2012, *Evaluating Social Innovation*, FSG, Boston. <https://www.fsg.org/publications/evaluating-social-innovation>
- Preskill, H., & Gopal, S., Mack, K. & Cook, J. 2014, *Evaluating Complexity: Propositions for Improving Practice*, FSG, Boston. <https://www.fsg.org/publications/evaluating-complexity>
- Radeke, J. 2016, 'Impact Assessment: Corporate Sustainability 2.0', *Stanford Social Innovation Review*, June 23. https://ssir.org/articles/entry/impact_assessment_corporate_sustainability_2.0
- Renger, R., Atkinson, L., Renger, J. & Hart, G. 2019, 'The connection between logic models and systems thinking concepts', *Evaluation Journal of Australasia*, vol. 19, no. 2, pp. 79-87. <https://doi.org/10.1177/1035719X19853660>
- Renger, R., Wood, S., Williamson, S. & Krapp, S. 2011, 'Systemic evaluation, impact evaluation and logic models', *Evaluation Journal of Australasia*, vol. 11, no. 2, pp. 24-30. <https://doi.org/10.1177/1035719X1101100204>
- Rogers, P.J. & Williams, B. 2010, 'Using Systems Concepts in Evaluation' in Fujita, N. (ed.) *Beyond Logframe: Using Systems Concepts in Evaluation*, Foundation for Advanced Studies in International Development (FASID), Tokyo. https://www.fasid.or.jp/english/files/web_library/h21-3.pdf

- Rogers, P. 2008, 'Using Programme Theory to evaluate complicated and complex aspects of interventions', *Evaluation*, vol. 41, no. 1, pp. 29-48. <https://doi.org/10.1177/1356389007084674>
- Ruff, K. & Olsen, S. 2016, 'The next frontier in social impact measurement isn't measurement at all', *Stanford Social Innovation Review*, May 10. https://ssir.org/articles/entry/next_frontier_in_social_impact_measurement
- Ryan, P. & Lyne, I. 2008, 'Social enterprise and the measurement of social value: methodological issues with the calculation and application of the social return on investment', *Education, Knowledge and Economy*, vol. 2, no. 3, pp. 223-237. <https://doi.org/10.1080/17496890802426253>
- Twersky, F. 2019, 'Time for a Three-Legged Measurement Stool' *Stanford Social Innovation Review*, Winter. https://ssir.org/articles/entry/time_for_a_three_legged_measurement_stool
- Wallman-Stokes, C., Hovde, K., McLaughlin, C. & Rosqueta, K. 2013, 'What we talk about when we talk about impact', Center for High Impact Philanthropy; Women Moving Millions. September. https://www.impact.upenn.edu/wp-content/uploads/2016/2014/12/What_Are_We_Talking_About_When_We_Talk_About_Impact.pdf
- Wood, C., & Leighton, D. 2010, *Measuring Social Value – the Gap between Policy and Practice*, Demos, London. https://www.demos.co.uk/files/Measuring_social_value_-_web.pdf
- Zappalà, G., & Lyons, M. 2009, 'Recent approaches to measuring social impact in the Third sector: An overview', Background paper 5, Centre for Social Impact, UNSW, July.
- Zappalà, G. 2011, 'Solving social problems & demonstrating impact: A tale of two typologies', Briefing Paper No. 5, Centre for Social Impact, UNSW, February. https://www.csi.edu.au/media/uploads/Solving_Social_Problems_and_Demonstrating_Impact_-_February_2011.pdf
- Zappalà, G. 2019, *Primary Ethics Volunteer Teachers Pilot Evaluation: A Meaningful Evaluation Approach*, Centre for Social Justice & Inclusion, University of Technology Sydney.
- Zappalà, G., & Smyth, C. 2020, 'Student and Teacher outcomes from participating in a Philosophy for Children program: Volunteer ethics teachers' perspectives', submitted to *Journal of Philosophy in Schools*.